Studien zum Physik- und Chemielernen

M. Hopf und M. Ropohl [Hrsg.]

392

Anna Weißbach

Reflexion von Physikunterricht

Entwicklung und Validierung eines Testinstruments mit Assessment-Feedback zur Reflexionsfähigkeit Physik-Lehramtsstudierender

λογος

Studien zum Physik- und Chemielernen

Herausgegeben von Martin Hopf und Mathias Ropohl

Diese Reihe im Logos Verlag Berlin lädt Forscherinnen und Forscher ein, ihre neuen wissenschaftlichen Studien zum Physik- und Chemielernen im Kontext einer Vielzahl von bereits erschienenen Arbeiten zu quantitativen und qualitativen empirischen Untersuchungen sowie evaluativ begleiteten Konzeptionsentwicklungen zu veröffentlichen. Die in den bisherigen Studien erfassten Themen und Inhalte spiegeln das breite Spektrum der Einflussfaktoren wider, die in den Lehr- und Lernprozessen in Schule und Hochschule wirksam sind.
Die Herausgeber hoffen, mit der Förderung von Publikationen, die sich mit dem Physik- und Chemielernen befassen, einen Beitrag zur weiteren Stabilisierung der physik- und chemiedidaktischen Forschung und zur Verbesserung eines an den Ergebnissen fachdidaktischer Forschung orientierten Unterrichts in den beiden Fächern zu leisten.

Martin Hopf und Mathias Ropohl

Studien zum Physik- und Chemielernen

Band 392

Anna Weißbach

Reflexion von Physikunterricht

Entwicklung und Validierung eines Testinstruments mit Assessment-Feedback zur Reflexionsfähigkeit Physik-Lehramtsstudierender

Logos Verlag Berlin



Studien zum Physik- und Chemielernen
Martin Hopf und Mathias Ropohl [Hrsg.]

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.d-nb.de abrufbar.



Dieses Werk ist lizenziert unter der Creative Commons Lizenz CC BY-SA (https://creativecommons.org/licenses/by-sa/4.0/). Die Bedingungen der Creative-Commons-Lizenz gelten nur für Originalmaterial. Die Wiederverwendung von Material aus anderen Quellen (gekennzeichnet mit Quellenangabe) wie z. B. Schaubilder, Abbildungen, Fotos und Textauszüge erfordert ggf. weitere Nutzungsgenehmigungen durch den jeweiligen Rechteinhaber.

Dissertation zur Erlangung des akademischen Grades eines Doktors (Dr. rer.nat.) am Fachbereich 1 der Universität Bremen, vorgelegt im November 2024 von Anna Weißbach.
1. Gutachter: Prof. Dr. Christoph Kulgemeyer
2. Gutachter: Prof. Dr. Josef Riese
Kolloquium am 13.01.2025

Logos Verlag Berlin GmbH 2025

ISBN 978-3-8325-5977-9
ISSN 1614-8967
DOI 10.30819/5977

Logos Verlag Berlin GmbH
Georg-Knorr-Str. 4, Geb. 10
D-12681 Berlin

Tel.: +49 (0)30 / 42 85 10 90
Fax: +49 (0)30 / 42 85 10 92
https://www.logos-verlag.de

Reflexion von Physikunterricht

Entwicklung und Validierung eines Testinstruments mit Assessment-Feedback zur Reflexionsfähigkeit Physik-Lehramtsstudierender

Dissertationsschrift zur Erlangung des akademischen Grades
Dr. rer. nat.
am Fachbereich 1 der Universität Bremen

Vorgelegt von Anna Weißbach

Erstgutachter:	Prof. Dr. Christoph Kulgemeyer
Zweitgutachter:	Prof. Dr. Josef Riese
Eingereicht am:	05.11.2024
Tag des Kolloquiums:	13.01.2025

Um der Druckfassung zu genügen wurden in der vorliegenden Arbeit redaktionelle Änderungen vorgenommen. Insbesondere wurden die ursprünglich separat digital angehängten Transkripte (Kapitel O und P) in den Anhang integriert und Verweise im Text dahingehend angepasst. Der Anhang der vorliegenden Arbeit ist in der Open-Access-Publikation der Arbeit unter https://doi.org/10.30819/5977 einzusehen.

Im Kontext des Projekts entstandene Veröffentlichungen:

1. Kulgemeyer, C., Kempin, M., Weißbach, A., Borowski, A., Buschhüter, D., Enkrott, P., Reinhold, P., Riese, J., Schecker, H., Schröder, J., & Vogelsang, C. (2021). Exploring the impact of pre-service science teachers' reflection skills on the development of professional knowledge during a field experience. *International Journal of Science Education*, *43*(18), 3035–3057. https://doi.org/10.1080/09500693.2021.2006820
2. Weißbach, A., & Kulgemeyer, C. (2022b). Reflexion von Physikunterricht - ein Online-Assessment mit Feedback. In S. Habig & H. van Vorst (Hrsg.), *Unsicherheit als Element von naturwissenschaftsbezogenen Bildungsprozessen* (S. 756–759). Universität Duisburg-Essen
3. Weißbach, A., & Kulgemeyer, C. (2022a). Reflexion von Physikunterricht - ein Online Assessment mit Feedback. *PhyDid B - Didaktik Der Physik - Beiträge Zur DPG-Frühjahrstagung*, 203–210
4. Kulgemeyer, C., Riese, J., Vogelsang, C., Buschhüter, D., Borowski, A., Weißbach, A., Jordans, M., Reinhold, P., & Schecker, H. (2023). How authenticity impacts validity: Developing a model of teacher education assessment and exploring the effects of the digitisation of assessment methods. *Zeitschrift für Erziehungswissenschaft*, 601–625. https://doi.org/10.1007/s11618-023-01154-y
5. Weißbach, A., & Kulgemeyer, C. (2023c). Reflexion von Physikunterricht: Ein Online-Assessment mit Feedback. In H. van Vorst (Hrsg.), *Lernen, Lehren und Forschen in einer digital geprägten Welt* (S. 861–864). Universität Duisburg-Essen
6. Weißbach, A., & Kulgemeyer, C. (2023a). Physikunterricht reflektieren: Eine Lernumgebung für die Lehramtsausbildung. In L. Mientus, C. Klempin & A. Nowak (Hrsg.), *Reflexion in der Lehrkräftebildung* (S. 241–247). Universitätsverlag Potsdam

7. Weißbach, A., & Kulgemeyer, C. (2023b). Reflection on Physics Teaching: An Online-Assessment Offering Automated Feedback. https://www.researchgate.net/publication/385172169_Reflection_On_Physics_Teaching_An_Online-Assessment_Offering_Automated_Feedback
8. Weißbach, A., & Kulgemeyer, C. (2024b). Reflexionsfähigkeit von Physiklehramtsstudierenden: Ein Online-Self-Assessment mit Feedback. *Unterrichtswissenschaft*, *52*(1), 135–161. https://doi.org/10.1007/s42010-024-00195-4
9. Weißbach, A., & Kulgemeyer, C. (2024c). Reflexionsfähigkeit: Validierungsstudien zu einem Test mit Feedback. In H. van Vorst (Hrsg.), *Frühe naturwissenschaftliche Bildung* (S. 266–269). Universität Duisburg-Essen
10. Weißbach, A., & Kulgemeyer, C. (2024a). Reflection on Physics Teaching - A comparison of a performance assessment and a multiple-choice assessment. https://www.researchgate.net/publication/379483629_Reflection_on_Physics_Teaching_-A_comparison_of_a_performance_assessment_and_a_multiple-choice_assessment
11. Weißbach, A., & Kulgemeyer, C. (2024d). Wie hängen Unterrichtsreflexion und Noticing zusammen? https://www.researchgate.net/publication/384925967_Wie_hangen_Unterrichtsreflexion_und_Noticing_zusammen

1 Kurzzusammenfassung

Die Fähigkeit, Unterricht zu reflektieren, ist für Lehrkräfte von zentraler Bedeutung. Sie dient der Entwicklung von Unterrichtsqualität und der Professionalisierung von Lehrkräften. Da sich insbesondere die Reflexionsfähigkeit Studierender allerdings als nur gering ausgeprägt erweist, besteht grundsätzlich Bedarf zu ihrer Förderung. Gegenstand der vorliegenden Arbeit ist die Entwicklung und Evaluation eines Tests zur Erfassung der Fähigkeit zur Unterrichtsreflexion (angehender) Physiklehrkräfte mit Feedback, das als Ausgangspunkt für die fundierte Selbsteinschätzung der Reflexionsfähigkeit dienen und nach positiver Evaluation für weitere Untersuchungen zur Struktur und Entwicklung der Reflexionsfähigkeit herangezogen werden soll. Um das Feedback zeitnah bereitstellen zu können, ist ein Kompromiss zwischen der möglichst authentischen Abbildung dieser Fähigkeit und einer hohen Standardisierung des Tests herzustellen. Die Standardisierung ermöglicht zwar die automatisierte Auswertung des Reflexionstests, ist allerdings mit einer Einschränkung der Handlungsfreiheit gegenüber realen Reflexionssituationen verbunden. Im entwickelten Testinstrument werden in Anlehnung an eine kollegiale Reflexionssituation geskriptete Videovignetten eines fiktiven Mitpraktikanten in Multiple-Choice-Aufgaben reflektiert.

Kern der Evaluation des Reflexionstests ist die Untersuchung der Frage, inwieweit die Interpretation der Testwerte des Reflexionstests einerseits als Maß für die Reflexionsfähigkeit und andererseits als Ausgangspunkt für einen weiteren Professionalisierungsprozess der (angehenden) Physiklehrkräfte als gültig angenommen werden kann. Die Entwicklung der Validitätsargumentation erfolgt in vier Schritten: Zunächst wird die angemessene Abbildung des Konstrukts der Reflexionsfähigkeit im Test evaluiert. Anschließend wird überprüft, ob bei der Bearbeitung der Aufgaben inhaltlich zur Unterrichtsreflexion passende Überlegungen angestellt werden; dazu werden die kognitiven Prozesse während der Testbearbeitung in Think-Aloud-Interviews analysiert. Außerdem werden auf Grundlage von $N = 192$ vorliegenden Bearbeitungen die zugewiesenen Testwerte und Zusammenhänge zu externen (u. a. demografischen) Variablen untersucht. Da die Teilnehmenden ein Feedback erhalten, ist darüber hinaus im Sinne der konsequentiellen Validität sicherzustellen, dass sie dieses angemessen interpretieren. Hierzu wird eine Interviewstudie durchgeführt.

Insgesamt legen die Ergebnisse nahe, dass die Testwerte valide als Maß für die Reflexionsfähigkeit interpretiert werden können. Das Konstrukt der Unterrichtsreflexion

ist - wenn auch unter Einschränkung der Handlungsfreiheit durch die verwendeten Multiple-Choice-Aufgaben - in relevanten und repräsentativen Aspekten in den Testaufgaben abgebildet und die Überlegungen der Teilnehmenden während der Bearbeitung des Instruments beziehen sich überwiegend auf den dargestellten Unterricht. Die Testwerte zum Reflexionstest basieren auf einer reliablen Skala, erscheinen geeignet zur Differenzierung zwischen den Teilnehmenden und weisen Zusammenhänge u. a. zum Studienfortschritt und zur Unterrichtsreflexion in einem authentischeren Setting auf. Lediglich ein Zusammenhang zum fachdidaktischen Professionswissen konnte nicht festgestellt werden. Zuletzt erscheint auch das Feedback angemessen, um den Teilnehmenden die fundierte Selbsteinschätzung der Reflexionsfähigkeit zu ermöglichen.

Zusätzlich erweist sich der Test als dazu geeignet, in einer latenten Profilanalyse Stufen der Reflexionsfähigkeit zu unterscheiden und einen Beitrag zur Klärung des Verhältnisses der Fähigkeiten zu Noticing und Unterrichtsreflexion zu leisten: Die Fähigkeiten erscheinen zwar unterscheidbar, aber zusammenhängend, wobei gelingendes Noticing eine Voraussetzung für bedeutsame Unterrichtsreflexion zu sein scheint und die Fähigkeit zum Noticing zur Entwicklung der Reflexionsfähigkeit beitragen könnte.

2 Abstract

Reflection on teaching is a crucial skill for teachers. It is assumed to be beneficial to teaching quality and professional development. However, various studies reveal rather low reflection skills of (prospective) teachers. Accordingly, there is a need to promote prospective teachers' reflection skills. This work presents the development and evaluation of an assessment tool of (prospective) physics teachers' reflection skills including feedback, that is intended to serve as a basis for a well-found self-assessment of their reflection skills and further explorative analysis regarding the structure and development of reflection skills. In order to provide timely feedback, a compromise must be found between an authentic representation of reflection on teaching and standardization of the assessment. Although standardization enables an automation of the examination and scoring process, it is associated with restricted authenticity compared to real reflection. The developed assessment tool simulates a peer reflection with a fictitious fellow intern on his teaching which is presented in scripted video vignettes.

The key question regarding the evaluation of the assessment on reflection skills concerns the investigation of the extent to which the test scores can be interpreted as a measure of reflection skills on the one hand and as a starting point for further professional development on the other. A validity argument is developed in four steps: First, the well-founded representation of reflection skills in the assessment is verified. Subsequently, the developed tasks are evaluated as evoking reflections on teaching. For this purpose, a think-aloud study is conducted to analyze the cognitive processes. In addition, the test scores and correlations with external variables are analyzed (including but not limited to demographics) based on $N = 192$ available cases. Since the (prospective) teachers receive feedback, in terms of consistent validity, it must be ensured, that they understand and interpret this feedback appropriately. To gain insight into their interpretations, an interview study is conducted.

Overall, the results suggest the valid interpretation of the test scores as a measure of (prospective) teachers' reflection skills. Relevant aspects of reflection skills are represented in the assessment - albeit with restricted authenticity - and the participants' cognitive processes while working on the assessment predominantly relate to the teaching presented in video vignettes. The test scores are based on a reliable scale, appear to be suitable for differentiating between the participants and, amongst others, correlate to the individual progress in teacher training

and a performance assessment measuring reflection skills in a more authentic setting. No correlation to the professional knowledge could be established. Finally, the feedback appears to be appropriate for enabling (prospective) teachers to self-assess their reflection skills.

Additionally, the assessment proves to be suitable for identifying different levels of reflection skills in a latent profile analysis and it contributes to the theoretical clarification of the relationship between reflection skills and noticing skills. Both skills appear to be distinguishable but closely interrelated. Moreover, successful noticing appears to be a prerequisite for meaningful reflection on teaching and noticing skills might contribute to the development of reflection skills.

Inhaltsverzeichnis

1 Kurzzusammenfassung **V**

2 Abstract **VII**

3 Ausgangslage **1**

4 Ziele und Aufbau der Arbeit **5**
4.1 Ziele der Arbeit 5
4.2 Einbettung der vorliegenden Arbeit in den Projektzusammenhang *ProfiLe-P* 6
4.3 Aufbau der vorliegenden Arbeit 7

I Theoretischer Hintergrund **9**

5 Professionelles Wissen und professionelle Handlungskompetenz **11**
5.1 Wirkungsmodelle schulischen Lernens 11
5.2 Professionelles Wissen 13
5.2.1 Bestandteile des professionellen Wissens 13
5.2.2 Pädagogisches Wissen 16
5.2.3 Fachwissen 19
5.2.4 Fachdidaktisches Wissen 21
5.2.5 Empirische Erkenntnisse zur Bedeutung von fachlichem und fachdidaktischem Wissen für die Gestaltung und Qualität von Unterricht 25
5.2.6 Zur Zusammenhangsstruktur der Bereiche des Professionswissens 27
5.3 Professionelle Handlungskompetenz 28
5.4 Zum Verhältnis von professionellem Wissen und Handeln 31
5.5 Zwischenfazit: Professionelles Wissen und Handeln von Physiklehrkräften 34

6 Zur Messung von professionellem Wissen und Handeln **37**
6.1 Gütekriterien empirischer Forschung 38
6.1.1 Objektivität 38

6.1.2 Reliabilität 39
6.1.3 Validität 42
6.1.4 Nebengütekriterien 57
6.1.5 Akzeptanz 59
6.2 Klassifikation unterschiedlicher Testtypen nach Miller (1990) . . . 63
6.3 Methoden zur Erfassung des professionellen Wissens und Handelns entlang Millers Klassifikation 67
6.3.1 Wissens- und Kompetenztests 68
6.3.2 Handlungstests 69
6.3.3 Performanztests 71
6.4 (Video-)Vignetten-Tests 74
6.5 Zwischenfazit: Messung von professionellem Wissen und Handeln . 78

7 Reflexion von Physikunterricht 79
7.1 Vorläufige Definition: Der gemeinsame Kern des Reflexionsbegriffs 79
7.2 Zur Bedeutung von Reflexion für (angehende) Lehrkräfte 79
7.3 Zum vielfältigen Verständnis des Reflexionsbegriffs 82
7.4 Reflexionskompetenz, Reflexionsfähigkeit und Reflexionsbereitschaft als Personenmerkmale 86
7.4.1 Reflexionskompetenz, Reflexionsfähigkeit und Reflexionsbereitschaft 86
7.4.2 Zum Begriff der Reflexivität 87
7.5 Modellierung des Reflexionsbegriffs 88
7.5.1 Prozessmodelle 88
7.5.2 Stufenmodelle 91
7.5.3 Weitere Kriterien zur Bewertung der Qualität von Reflexionen 93
7.5.4 Reflexionsmodell nach Nowak, Kempin, Kulgemeyer und Borowski (2019) 94
7.6 Reflexion, Noticing und professionelle Unterrichtswahrnehmung . 97
7.7 Empirische Ergebnisse der Erfassung der Reflexionskompetenz . . 99
7.7.1 Selbstauskünfte zur Reflexionsbereitschaft und -fähigkeit . 99
7.7.2 Verfahren zur Messung der Reflexionsfähigkeit 101
7.7.3 Reflexionsperformanztest von Kempin, Kulgemeyer und Schecker (2018) 105
7.8 Ansätze zur Förderung der Reflexionsfähigkeit 113
7.9 Zwischenfazit: Reflexion von Physikunterricht 118

8 Feedback 121
8.1 Feedback-Modell von Hattie und Timperley (2007) 122
8.2 Empfehlungen für effektives Assessment-Feedback 124
8.3 Zum Verhältnis von Feedback zum und der Reflexion von Unterricht 125
8.4 Zwischenfazit: Feedback 126

9 Konsequenzen für die Entwicklung der Lernumgebung zur Reflexion von Physikunterricht sowie für weitere Analysen zur Reflexionsfähigkeit 129
9.1 Bestimmung und Modellierung des Begriffs der Reflexionsfähigkeit für diese Arbeit . . . 129
9.1.1 Definition von Unterrichtsreflexion . . . 129
9.1.2 Fähigkeit zur Unterrichtsreflexion . . . 131
9.1.3 Abgrenzung zum Begriff der (Unterrichts-)Analyse . . . 131
9.1.4 Modellierung von Unterrichtsreflexion . . . 132
9.2 Anforderungen an die zu entwickelnde Lernumgebung . . . 132
9.3 Schlussfolgerungen für die explorative Untersuchung der Reflexionsfähigkeit und ihrer Entwicklung . . . 134

II Entwicklung der Lernumgebung zur Reflexion von Physikunterricht 137

10 Reflexionstest 139
10.1 Aufbau des Reflexionstests . . . 139
10.2 Entwicklung und Ausgestaltung des Reflexionstests . . . 145
10.2.1 Auswahl der Videovignetten . . . 145
10.2.2 Entwicklung der Aufgaben . . . 146
10.2.3 Bewertung der Antwortoptionen . . . 148
10.3 Bestimmung der Testwerte des Reflexionstests . . . 149
10.3.1 Bestimmung der Testwerte zur Unterrichtsreflexion und zum Noticing . . . 150
10.3.2 Diskussion weiterer Bepunktungsvorschriften . . . 150
10.4 Erläuterungen zur Konzeption des Reflexionstests auf Grundlage des Reflexionsperformanztests von Kempin, Kulgemeyer und Schecker (2018) . . . 152

11 Assessment-Feedback 155
11.1 Aufbau des Assessment-Feedbacks . . . 155
11.2 Aufbau des ergänzenden Selbstlernmaterials . . . 157
11.3 Assessment-Feedback für Dozierende . . . 158

III Validierungsstudien zur entwickelten Lernumgebung 161

12 Das Interpretations-Nutzungs-Argument zum Reflexionstest (FF 1) 163
12.1 Indikatoren zum Teil 1 des INA: Übersetzung Reflexionsfähigkeit in Testinstrument . . . 165
12.2 Indikatoren zum Teil 2 des INA: Übersetzung Testinstrument in Verhalten . . . 165

12.3 Indikatoren zum Teil 3 des INA: Übersetzung Verhalten in Testwerte 166
12.4 Indikatoren zum Teil 4 des INA: Übersetzung Testwerte in Assessment-Feedback . . . 169
12.5 Bemerkungen zur Auswahl der Indikatoren für die Validitätsargumentation . . . 170

13 Weitere eingesetzte Testinstrumente 175
13.1 Fragebogen zur Demografie . . . 175
13.2 Items zur Motivation während der Testbearbeitung . . . 176
13.3 Fragebogen zur Akzeptanz der Lernumgebung . . . 176
13.4 Test zum fachdidaktischen Wissen . . . 180

14 Zusammenfassende Übersicht über die durchgeführten Untersuchungen und die Stichprobe 183
14.1 Übersicht über die Stichprobe . . . 183
14.2 Übersicht über die im Rahmen der Validierung durchgeführten Untersuchungen . . . 185

15 Untersuchung der Anforderung an Teil 1 des INA: Die Aufgaben im Reflexionstest bilden relevante und repräsentative Inhalte der Fähigkeit zur Unterrichtsreflexion präzise ab 191
15.1 Evaluation der Abbildung der Fähigkeit zur Unterrichtsreflexion im Reflexionstest (Indikator 1.1) . . . 191
15.2 Evaluation der Bewertung der Antwortoptionen (Indikator 1.2) . . 196
15.3 Beurteilung des Teils 1 des INA: Übersetzung Unterrichtsreflexion in Testinstrument . . . 197

16 Untersuchung der Anforderung an Teil 2 des INA: Die Proband*innen zeigen bei der Bearbeitung des Reflexionstests konstruktrelevantes Verhalten 199
16.1 Evaluation der kognitiven Validität des Reflexionstests (Indikator 2.1) . . . 199
16.1.1 Die Methode der Think-Aloud-Interviews . . . 200
16.1.2 Umsetzung der Think-Aloud-Interviews . . . 202
16.1.3 Aufbereitung der Think-Aloud-Interviews . . . 203
16.1.4 Kodierung der Think-Aloud-Interviews . . . 204
16.1.5 Ergebnisse . . . 207
16.2 Evaluation der Akzeptanz des Testverfahrens (Indikator 2.2) . . . 214
16.2.1 Evaluation des Akzeptanzfragebogens im Reflexionstest . . 214
16.2.2 Auswertung der Kommentare zum Testinstrument . . . 216
16.2.3 Evaluation des Fragebogens zur Motivation während der Bearbeitung des Reflexionstests . . . 218

16.3 Beurteilung des Teils 2 des INA: Übersetzung Testinstrument in Verhalten 219

17 Untersuchung der Anforderung an Teil 3 des INA: Das (Ankreuz-)-Verhalten der Proband*innen während der Testbearbeitung wird adäquat in Testwerte übersetzt 223

17.1 Evaluation der Differenzierungsfähigkeit des Reflexionstests (Indikator 3.1) 224
17.1.1 Analyse der vorliegenden Verteilung 224
17.1.2 Analyse der Lösungshäufigkeiten 226
17.1.3 Bewertung des Indikators 3.1 227
17.2 Evaluation der Reliabilität der verwendeten Skalen (Indikator 3.2) 227
17.2.1 Betrachtung des Testwerts zur Reflexionsfähigkeit 228
17.2.2 Betrachtung der Testwerte für die Teilfähigkeiten zum Bewerten und Vorschlagen von Alternativen 228
17.2.3 Bewertung des Indikators 3.2 229
17.3 Evaluation der Struktur der Testwerte (Indikator 3.3) 230
17.4 Evaluation des Zusammenhangs zur Reflexion in authentischeren Situationen (Indikator 3.4) 231
17.4.1 Studiendesign 232
17.4.2 Ergebnisse 232
17.4.3 Bewertung des Indikators 3.4 234
17.5 Evaluation des Zusammenhangs zu Lerngelegenheiten (Indikator 3.5) 236
17.5.1 Vorgehen bei den Analysen 236
17.5.2 Zusammenhänge zu demografischen Variablen (Quasi-Längsschnitt) 237
17.5.3 Untersuchung der Testwerte im Längsschnitt über ein Semester 242
17.5.4 Bewertung des Indikators 3.5 242
17.6 Evaluation des Zusammenhangs zum fachdidaktischen Wissen (Indikator 3.6) 245
17.6.1 Studiendesign 245
17.6.2 Ergebnisse 245
17.6.3 Bewertung des Indikators 3.6 246
17.7 Beurteilung des Teils 3 des INA: Übersetzung Verhalten in Testwerte 247

18 Untersuchung der Anforderung an Teil 4 des INA: Das Assessment-Feedback stellt eine geeignete Form der Rückmeldung der Testwerte dar 251

18.1 Evaluation des Assessment-Feedbacks (Indikatoren 4.1 und 4.2) . 251
18.1.1 Pilotierung des Assessment-Feedbacks 252
18.1.2 Evaluation des weiterentwickelten Assessment-Feedbacks . 272
18.1.3 Ergänzende Erkenntnisse aus dem Akzeptanzfragebogen . 293
18.1.4 Bewertung der Indikatoren 4.1 und 4.2 293

18.2 Beurteilung des Assessment-Feedbacks vor dem Hintergrund von Kriterien für effektives Feedback (Indikator 4.3) 295
18.3 Beurteilung des Teils 4 des INA: Übersetzung Testwerte in Assessment-Feedback 298

19 Validitätsargumentation zum Reflexionstest (Beantwortung der Forschungsfrage 1) 301
19.1 Zusammenfassende Bewertung des Teils 1 des INA: Übersetzung Reflexionsfähigkeit in Testinstrument 302
19.2 Zusammenfassende Bewertung des Teils 2 des INA: Übersetzung Testinstrument in Verhalten 304
19.3 Zusammenfassende Bewertung des Teils 3 des INA: Übersetzung Verhalten in Testwerte 305
19.4 Zusammenfassende Bewertung des Teils 4 des INA: Übersetzung Testwerte in Assessment-Feedback 309
19.5 Bewertung sonstiger Gütekriterien 310
19.6 Fazit 314
19.7 Ausblick 315

IV Untersuchungen zur Reflexionsfähigkeit und ihrer Entwicklung 321

20 Ziele und Forschungsfragen der Untersuchungen 323

21 Analyse latenter Profile von Reflexionsfähigkeit 327
21.1 Latente Profilanalyse 327
21.2 Ergebnisse der latenten Profilanalyse 328
21.3 Beschreibung der Profile 329
21.4 Interpretation der Profile als Fähigkeitsstufen 331
21.5 Charakterisierung der Fähigkeitsstufen über inhaltliche Merkmale der Aufgaben 332
21.6 Charakterisierung der Fähigkeitsstufen über demografische Variablen 334
21.7 Untersuchung von Zusammenhängen zwischen Testmotivation sowie Bearbeitungsdauer und der Zuordnung zu Fähigkeitsstufen 336
21.8 Erste Erweiterung des Modells zum Zusammenhang von Noticing und Unterrichtsreflexion (Beantwortung der Forschungsfrage 2) . . 337

22 Zum Zusammenhang von Noticing und Reflexion 339
22.1 Ergebnisse 340
22.1.1 Voruntersuchungen zu den Testwerten zur Noticingfähigkeit 340
22.1.2 Prüfung der Hypothese 1 340
22.1.3 Prüfung der Hypothese 2 341

22.1.4 Prüfung der Hypothese 3 341
22.2 Zweite Erweiterung des Modells zum Zusammenhang von Noticing und Unterrichtsreflexion (Beantwortung der Forschungsfrage 3) . . 344

23 Zum Entwicklungszusammenhang der Fähigkeiten zum Noticing und zur Reflexion 347
23.1 Statistische Analysen im Cross-Lagged-Panel-Design 348
23.2 Ergebnisse 349
23.2.1 PLS-SEM für alle vorliegenden Prä-Post-Datensätze 349
23.2.2 PLS-SEM für die Teilstichprobe der Personen in den oberen drei Fähigkeitsstufen 349
23.2.3 Cross-Lagged-Panel-Analysen mit partiellen Korrelationen 351
23.3 Dritte Erweiterung des Modells zum Zusammenhang von Noticing und Unterrichtsreflexion (Beantwortung der Forschungsfrage 4) . . 353

24 Zusammenfassung und Fazit zu den explorativen Untersuchungen zur Reflexionsfähigkeit 355
24.1 Stufen der Reflexionsfähigkeit 355
24.2 Zum Zusammenhang von Noticing und Reflexion 356
24.3 Noticingfähigkeit als Voraussetzung für die Entwicklung der Reflexionsfähigkeit 356
24.4 Erweitertes Modell zum Zusammenhang von Noticing und Unterrichtsreflexion 357
24.5 Limitationen der explorativen Untersuchungen 358

V Schlussbemerkungen 361

25 Beitrag der Arbeit für die Erfassung und Förderung der Fähigkeit zur Unterrichtsreflexion 363

Literaturverzeichnis 365

Abbildungsverzeichnis 403

Tabellenverzeichnis 407

Abkürzungsverzeichnis 411

VI Anhang **415**

Der Anhang ist Bestandteil der Open-Access-Publikation dieser Arbeit. Diese ist unter dem folgenden Link abrufbar: https://doi.org/10.30819/5977

A Aufgaben im Reflexionstest **417**
A.1 Zu Videovignette 1 . . . 417
A.2 Zu Videovignette 2 . . . 419
A.3 Zu Videovignette 3 . . . 420
A.4 Zu Videovignette 4 . . . 423
A.5 Zu Videovignette 5 . . . 426
A.6 Zu Videovignette 6 . . . 430
A.7 Zu Videovignette 7 . . . 433
A.8 Zur globalen Reflexion . . . 436

B *R*-Skript zur Erstellung der Rückmeldungen **439**
B.1 Auswertung und Erstellung der Diagramme . . . 439
B.2 Erstellung der Rückmeldungen . . . 446

C Zugang zum Selbstlernmaterial **453**

D Fragebogen zur Demografie **455**

E Items zur Motivation während der Testbearbeitung **459**

F Vollständiger Akzeptanzfragebogen **461**

G Manual zur Auswertung der Think-Aloud-Interviews **463**
G.1 Transkriptionsleitfaden . . . 463
G.2 Segmentierung der Think-Aloud-Protokolle . . . 464
G.3 Kodierung der Think-Aloud-Protokolle . . . 465
G.3.1 Übersicht über die Kategorien . . . 466
G.3.2 Kodierschema . . . 473

H Ergänzende Auswertung der Think-Aloud-Protokolle zum Teil 2 des INA: Kodierungen von Bewertungen und Alternativen zu den entsprechenden Aufgaben **477**

I Vollständige statistische Untersuchungen zum Teil 3 des INA **479**
I.1 Indikator 3.2: Übersicht über Kennwerte für die Reliabilität (α_{Cronbach}) und Trennschärfen . . . 479
I.2 Indikator 3.4 . . . 480
I.2.1 Ergebnisse des Kolmogorow-Smirnow-Tests . . . 480
I.2.2 Levene-Tests und ANOVAs . . . 481

I.3 Indikator 3.5 . . . 482
I.3.1 Indikator 3.5: Quasi-Längsschnitt . . . 482
I.3.2 Indikator 3.5: Längsschnitt . . . 489
I.4 Indikator 3.6 . . . 490

J Interviewleitfäden und Beispielrückmeldungen zur Evaluation des Assessment-Feedbacks 491
J.1 Interviewleitfaden zur Pilotierung . . . 491
J.1.1 Einleitung/Einverständnis . . . 491
J.1.2 Fragen 1: Eigene Rückmeldung . . . 492
J.1.3 Fragen 2: Beispielrückmeldung . . . 492
J.1.4 Fragen 3: Bereitstellung der Rückmeldung . . . 492
J.1.5 Abschluss . . . 493
J.2 Beispielrückmeldung zur Pilotierung . . . 494
J.3 Interviewleitfaden zur Erhebung mit Kohorte B . . . 497
J.3.1 Einleitung/Einverständnis . . . 497
J.3.2 Fragen 1: Eigene Rückmeldung . . . 497
J.3.3 Fragen 2: Beispielrückmeldung . . . 497
J.3.4 Fragen 3: Selbstlernmaterial . . . 498
J.3.5 Abschluss . . . 498
J.4 Beispielrückmeldung zur Erhebung mit Kohorte B . . . 499
J.5 Interviewleitfaden zur Erhebung mit Kohorte C . . . 504
J.5.1 Einleitung/Einverständnis . . . 504
J.5.2 Vorab: Mehrfache Testbearbeitung . . . 504
J.5.3 Fragen 1: Eigene Rückmeldung . . . 504
J.5.4 Fragen 2: Beispielrückmeldung . . . 505
J.5.5 Fragen 3: Selbstlernmaterial . . . 505
J.5.6 Fragen 4: Zur wiederholten Testbearbeitung . . . 506
J.5.7 Abschluss . . . 506
J.6 Beispielrückmeldung zur Erhebung mit Kohorte C . . . 507

K Ergebnisse der Akzeptanzbefragung zum Assessment-Feedback und zum Selbstlernmaterial 513

L Vollständige statistische Untersuchungen zur Profilanalyse (Forschungsfrage 2) 515
L.1 Gruppenunterschiede in Testwerten (gesamt und Teilfähigkeiten Bewerten und Vorschlagen von Alternativen) . . . 515
L.2 Gruppenunterschiede in Aufgabenscores . . . 516
L.2.1 Aufgaben zu Vignette 1 . . . 516
L.2.2 Aufgaben zu Vignette 2 . . . 516
L.2.3 Aufgaben zu Vignette 3 . . . 517
L.2.4 Aufgaben zu Vignette 4 . . . 518

L.2.5 Aufgaben zu Vignette 5 . . . 519
L.2.6 Aufgaben zu Vignette 6 . . . 520
L.2.7 Aufgaben zu Vignette 7 . . . 521
L.2.8 Aufgaben zur globalen Reflexion . . . 522
L.3 Gruppenunterschiede in demografischen Variablen . . . 523
L.4 Gruppenunterschiede in Testmotivation und Bearbeitungsdauer . 524

M Vollständige statistische Untersuchungen zum Zusammenhang zwischen Noticing und Unterrichtsreflexion 525
M.1 Zu Hypothese 1 . . . 525
M.2 Zu Hypothese 2 . . . 525
M.3 Zu Hypothese 3 . . . 526

N Vorausgehende statistische Untersuchungen zum Entwicklungszusammenhang zwischen den Fähigkeiten zu Noticing und Unterrichtsreflexion 529
N.1 Gesamtstichprobe ($N = 41$) . . . 529
N.2 Teilstichprobe der oberen drei Profile ($N = 37$) . . . 530

O Segmentierte und kodierte Transkripte der Think-Aloud-Interviews zur Prüfung der Anforderung an Teil 2 des INA (Untersuchung kognitiver Prozesse) 531
O.1 Interview 1: LD1 . . . 531
O.2 Interview 2: LD2 . . . 577
O.3 Interview 3: LD3 . . . 616
O.4 Interview 4: LD4 . . . 649
O.5 Interview 5: LD5 . . . 689
O.6 Interview 6: LD6 . . . 729
O.7 Interview 7: LD7 . . . 766

P Transkripte der Interviews zur Prüfung der Anforderung an Teil 4 des INA (Evaluation der Rückmeldung) 789
P.1 Interviews zur Pilotierung (Kohorte A) . . . 789
P.1.1 Interview 1: A1 . . . 789
P.1.2 Interview 2: A2 . . . 796
P.1.3 Interview 3: A3 . . . 804
P.1.4 Interview 4: A4 . . . 810
P.1.5 Interview 5: A5 . . . 814
P.1.6 Interview 6: A6 . . . 818
P.2 Interviews zur 1. Evaluation (Kohorte B) . . . 823
P.2.1 Interview 1: B1 . . . 823
P.2.2 Interview 2: B2 . . . 826
P.2.3 Interview 3: B3 . . . 831

P.2.4 Interview 4: B4 . . . 834
P.2.5 Interview 5: B5 . . . 840
P.2.6 Interview 6: B6 . . . 846
P.3 Interviews zur 2. Evaluation (Kohorte C) . . . 851
P.3.1 Interview 1: C1 . . . 851
P.3.2 Interview 2: C2 . . . 855
P.3.3 Interview 3: C3 . . . 858
P.3.4 Interview 4: C4 . . . 863
P.3.5 Interview 5: C5 . . . 870

Q *R*-Skript für die Bestimmung der Beurteilerübereinstimmung zur Prüfung der Anforderung an Teil 2 des INA 877
Q.1 Vorbereitend: Pakete einfügen und Einlesen der Datei mit der Übersicht über die Kodierungen . . . 877
Q.2 Vorbereitung der Kreuztabellen zur Bestimmung von κ_{Cohen} . . . 877
Q.3 Bestimmung von κ_{Cohen} für die Kategorien . . . 879

R *R*-Skript für die statischen Auswertungen zur Prüfung der Anforderung an Teil 3 des INA 881
R.1 Vorbereitend: Pakete einfügen und Aufbereitung der Daten aus dem Reflexionstest . . . 881
R.2 Zu Indikator 3.1: Untersuchung der Verteilung und Lösungshäufigkeiten . . . 888
R.2.1 Prüfung der Normalverteilung . . . 888
R.2.2 Betrachtung der Lösungshäufigkeiten für die Gesamtstichprobe sowie Bachelor- und Masterstudierende . . . 889
R.3 Zu Indikator 3.2: Untersuchung der Reliabilität und Trennschärfen 891
R.3.1 Über alle Aufgaben . . . 891
R.3.2 Jeweils über die Aufgaben zu Bewertungen und Alternativen 892
R.4 Zu Indikator 3.3: Explorative Faktorenanalyse . . . 893
R.5 Zu Indikator 3.4: Vergleich mit dem Reflexionsperformanztest . . 894
R.5.1 Aufbereiten der Daten des Reflexionsperformanztests . . . 894
R.5.2 Prüfen von Voraussetzungen und Skalengüte . . . 897
R.5.3 Berechnung von Korrelationen zwischen den Tests . . . 900
R.5.4 Untersuchung von Gruppenunterschieden (Masterstudierende (1), Grundschullehramtsstudierende (2), Bachelorstudierende (3)) . . . 900
R.6 Indikator 3.5: Zusammenhänge zu Lerngelegenheiten/demografischen Variablen . . . 901
R.6.1 Quasi-längsschnittliche Betrachtungen: Zusammenhänge zwischen Gesamtscore/Teilergebnissen und Indikatoren für Lerngelegenheiten . . . 901
R.6.2 Längsschnittliche Betrachtungen . . . 910

R.7 Indikator 3.6: Zusammenhang zum fachdidaktischen Wissen . . . 915
R.7.1 Aufbereiten der Daten aus FDW-Test . . . 915
R.7.2 Zusammenfügen der Daten aus dem Reflexions- und FDW-Test für beide MZPs . . . 915
R.7.3 Berechnung der Kennzahlen (M, SD, Test auf Normalverteilung, Cronbachs-Alpha) . . . 916
R.7.4 Berechnung der Korrelationen für die Gesamt-, Prä- und Post-Stichprobe . . . 919
R.7.5 Berechnung der Tests auf Normalverteilung sowie der Korrelationen zur Reflexionsfähigkeit für die Facetten des FDW-Tests . . . 919

S *R*-Skript für die statistischen Auswertungen der explorativen Untersuchungen zur Reflexionsfähigkeit 923
S.1 Vorbereitend: Pakete einfügen und Aufbereitung der Daten aus dem Reflexionstest . . . 923
S.2 Zur Profilanalyse . . . 930
S.2.1 Berechnung der Modelle . . . 930
S.2.2 Beschreibung der Profile . . . 931
S.2.3 Längsschnittliche Betrachtungen . . . 932
S.3 Charakterisierung der Fähigkeitsstufen über inhaltliche Merkmale der Aufgaben . . . 936
S.4 Charakterisierung der Fähigkeitsstufen über demografische Variablen 942
S.4.1 Ausbildungsphase . . . 942
S.4.2 Fachsemester . . . 943
S.4.3 Umfang der Erfahrungen im Pflichtpraktikum . . . 943
S.4.4 Testmotivation . . . 943
S.4.5 Bearbeitungsdauer . . . 943

T *R*-Skript für die statistischen Auswertungen zum Zusammenhang von Reflexion und Noticing 945
T.1 Vorbetrachtungen zum Testwert zur Noticingfähigkeit . . . 945
T.2 Hypothese 1: Korrelation zwischen den Testwerten der Fähigkeiten zum Noticing und zur Reflexion . . . 945
T.3 Hypothese 2: Unterschiede in den Testwerten der Fähigkeiten zum Noticing zwischen den Profilen . . . 946
T.4 Hypothese 3: Betrachtung der Abhängigkeit der Aufgabenscores abhängig von der Anwahl in den Noticing-Aufgaben in den Profilen 947
T.4.1 Bestimmung der zusätzlichen Reflexionsscores $\text{Reflexionsscore}_{\text{angewählt im Noticing}}$ und $\text{Reflexionsscore}_{\textit{nicht} \text{ angewählt im Noticing}}$ . . . 947
T.4.2 Histogramme zu beiden Reflexionsscores . . . 950
T.4.3 Ergänzende statistische Auswertung . . . 951

U *R*-Skript für die statistischen Auswertungen zum Entwicklungszusammenhang von Reflexion und Noticing (CLP-Design) 955
U.1 PLS-SEM 955
U.1.1 Gesamtstichprobe 955
U.1.2 Obere drei Fähigkeitsstufen 956
U.2 (Partielle) Korrelationen 957
U.2.1 Gesamtstichprobe 957
U.2.2 Obere drei Fähigkeitsstufen 958

3 Ausgangslage

Für das Lernen von Schüler*innen gilt das Handeln von Lehrkräften als zentraler Einflussfaktor, denn sie sind die Organisator*innen von Lerngelegenheiten im Unterricht (Hattie, 2009, S. 108ff.). Gleichzeitig sind Lehrkräfte mit vielfältigen, komplexen Anforderungen und der hohen Situativität des Unterrichts konfrontiert, der stets Herausforderungen bereithält, auf die Lehrkräfte im Rahmen ihrer Ausbildung nicht explizit vorbereitet werden können. Daraus ergibt sich eine grundlegende Notwendigkeit für Lehrkräfte, ihr (professionelles) Handeln im Unterricht zu evaluieren, anzupassen und sich weiter zu professionalisieren. Hierzu soll die Reflexion des Unterrichts dienen. Reflexion gilt dementsprechend als eine Kernaufgabe von Lehrkräften (Sekretariat der Kultusministerkonferenz, 2004) und wird als „Schlüsselbedingung für die Verbesserung des eigenen Unterrichts“ (Helmke, 2015, S. 116) verstanden. Auch für die persönliche Weiterentwicklung von Lehrkräften wird sie vorausgesetzt, indem ihr zugesprochen wird, zwischen Praxiserfahrungen und theoretischem Wissen zu vermitteln (z. B. von Aufschnaiter, Hofmann, Geisler und Kirschner, 2019; McAlpine et al., 1999).

Gleichzeitig wurde wiederholt festgestellt, dass (angehende) Lehrkräfte vorwiegend auf niedrigen Qualitätsniveaus reflektieren und im Rahmen der Reflexion kaum alternative Handlungsoptionen bzw. Konsequenzen ableiten (z. B. für Physik-Lehramtsstudierende von Kulgemeyer et al. (2021) oder Szogs, Große und Korneck (2020) und darüber hinaus u. a. von Wyss (2013, S. 168) oder Hatton und Smith (1995)) und das übereinstimmend, obwohl der Reflexionsbegriff auch über die genannten Studien hinaus nicht einheitlich verwendet wird und verschiedene Studien demnach auf unterschiedlichen theoretischen Grundlagen basieren. Um die positiven Effekte, die der Reflexion von Unterricht zugesprochen werden, auszunutzen, kommt der Förderung der Reflexionsfähigkeit (d. h. der kognitiven personalen Bedingungen einer Lehrkraft, die sie zur Reflexion befähigen) vor diesem Hintergrund eine besondere Bedeutung zu. Während eine ausgeprägte Fähigkeit zur Reflexion von Unterricht bei Lehrkräften im Beruf z. B. durch die Entwicklung alternativer Vorgehensweisen bei der Erklärung eines Sachverhalts oder der Demonstration eines Experiments unmittelbar zu Verbesserungen der eigenen Unterrichtsqualität beitragen kann, gilt dies für Lehramtsstudierende analog zunächst nur in Praxisphasen. Doch auch darüber hinaus erscheint Reflexion gewinnbringend für Studierende: Ihre Funktion zur Relationierung von Theorie und Praxis (z. B. Meier, Vogelsang, Watson und Schaper, 2022) kann dazu beitragen,

aus beobachtetem Unterricht bzw. in universitären Veranstaltungen diskutierten Unterrichtsbeispielen zu lernen und die Verzahnung von theoretischem Wissen und (beobachteter) Unterrichtspraxis voranzutreiben. Passend dazu ist der Einsatz von Reflexionen bereits im Studium verbreitet (Hilzensauer, 2017, S. 14). Insofern erscheint die Förderung der Reflexionsfähigkeit (angehender) Lehrkräfte auch über Praxisphasen hinaus schon im Studium bedeutsam.

Die Förderung der Reflexionsfähigkeit ist aus praktischer und theoretischer Perspektive allerdings mit verschiedenen Herausforderungen verbunden: Zunächst bedarf es zur zielgerichteten Förderung aus praktischer Perspektive einer vorausgehenden Diagnostik bzw. Messung des Fähigkeitsstandes der (angehenden) Lehrkräfte sowie eines darauf aufbauenden Feedbacks zum Fähigkeitsstand, um vorliegende Entwicklungsbedarfe aufzuzeigen. Beides basiert auf der validen Messung der Reflexionsfähigkeit. Bislang werden für objektive Messungen der Reflexionsfähigkeit vorrangig schriftliche Reflexionen herangezogen, die anschließend unter entsprechend hohem personellen Aufwand nach verschiedenen Qualitätskriterien bewertet werden. Eine zeitnahe Rückmeldung der Fähigkeiten ist so allerdings kaum umzusetzen. Außerdem liegen trotz der inzwischen jahrzehntelangen Betonung der Bedeutung von Reflexion für Lehrkräfte weiterhin Forschungsdesiderate vor. Einerseits ist bislang offen, welche Zusammenhänge tatsächlich zwischen der Ausprägung der Reflexionsfähigkeit bei Lehrkräften und der Gestaltung von Unterricht sowie dem Lernen von Schüler*innen festgestellt werden können (Häcker, 2022). Andererseits liegen kaum Erkenntnisse zur Entwicklung der Reflexionsfähigkeit und dazu notwendiger Rahmenbedingungen sowie der „Wirkung spezifischer Lerngelegenheiten“ (Reintjes & Kunze, 2022, S. 10) auf die Entwicklung der Reflexionsfähigkeit vor (s. auch Leonhard und Abels (2017)). Insbesondere ist auch empirisch nicht abschließend geklärt, ob hierarchische Stufen der (Entwicklung der) Reflexionsfähigkeit unterschieden werden können (van Beveren, Roets, Buysse & Rutten, 2018), sowie auf Grundlage welcher Kriterien eine Entwicklung festgestellt bzw. Entwicklungsstufen beschrieben werden können (Riel, 2022, S. 83f.). Diesbezügliche Erkenntnisse könnten zur Auswahl angemessener Fördermaßnahmen beitragen, insbesondere zu deren Anlage und zum Umfang anzustrebender Entwicklungen.

In Anbetracht der beschriebenen Herausforderungen und Forschungsdesiderate nimmt diese Arbeit zunächst die valide Messung und Förderung der Fähigkeit zur Unterrichtsreflexion in den Blick. Zur Messung der Reflexionsfähigkeit wird eine Lernumgebung bestehend aus einem Testinstrument mit dazugehörigem Assessment-Feedback entwickelt und die Interpretation der Testwerte als Maß für die Reflexionsfähigkeit und Ausgangspunkt für weiterführende Professionalisierungsprozesse umfassend validiert. Anschließend werden explorative Untersuchungen zur Struktur und Entwicklung der Reflexionsfähigkeit durchgeführt, die einen Einblick insbesondere in Zusammenhänge zwischen den Fähigkeiten zur Unterrichtsreflexion und zum Noticing ermöglichen. Auf dieser Grundlage

werden abschließend weiter zu prüfende Hypothesen über die Zusammenhänge beider Fähigkeiten und zur Förderung der Reflexionsfähigkeit abgeleitet, um einen Beitrag zur Klärung der offenen Fragen in diesem Bereich beizutragen.

4 Ziele und Aufbau der Arbeit

In diesem Kapitel werden die Ziele der vorliegenden Arbeit vorgestellt (Kapitel 4.1). Zusätzlich erfolgt eine Einordnung der Arbeit in den zugrundeliegenden Projektzusammenhang *ProfiLe-P* (Kapitel 4.2). Dem folgt ein Überblick über den Aufbau dieser Arbeit (Kapitel 4.3).

4.1 Ziele der Arbeit

Ziel dieser Arbeit ist zunächst die Entwicklung und Evaluation einer Lernumgebung zur Reflexion von Physikunterricht, die ein Testinstrument zur Erfassung der Reflexionsfähigkeit (angehender) Physiklehrkräfte und dazugehöriges Assessment-Feedback umfasst. Damit soll Teilnehmenden eine fundierte Einschätzung ihrer Reflexionsfähigkeit ermöglicht und ein Bedarf für die individuelle Weiterentwicklung aufgezeigt werden. Zur Herstellung einer gewissen Authentizität in Bezug auf die Reflexion von Unterricht werden die Teilnehmenden im Reflexionstest in eine fiktive kollegiale Beratungssituation versetzt: Ein Mitpraktikant bittet um Feedback zu seinem Unterricht, der in Videovignetten dargestellt ist. Um außerdem die zeitnahe Bereitstellung von Assessment-Feedback zu ermöglichen, wird bei der Entwicklung des Tests auf ein geschlossenes Antwortformat zurückgegriffen.

Der Evaluation der Lernumgebung wird folgende Forschungsfrage zugrunde gelegt:

> **Forschungsfrage 1:** Inwieweit kann die folgende Testwertinterpretation gerechtfertigt werden?
>
> *Die Testwerte des Reflexionstests können (a) als Maß für die Reflexionsfähigkeit und (b) als Ausgangspunkt für einen weiteren Professionalisierungsprozess aufgefasst werden.*

Zentral für die Beantwortung dieser Forschungsfrage ist die Sicherstellung der Validität der Testwertinterpretation. Im Sinne des argumentbasierten Ansatzes nach Kane (z. B. 1992; 2013) wird dazu eine Validitätsargumentation entwickelt, die den Prozess nachvollzieht, der notwendig wird, um ausgehend von der Zieldomäne der Fähigkeit zur Unterrichtsreflexion über die Entwicklung von Testaufgaben,

das Verhalten der Teilnehmenden bei der Bearbeitung und die Zuweisung der Testwerte die (adäquate) Interpretation des erhaltenen Assessment-Feedbacks durch die Teilnehmenden zu ermöglichen. Dazu werden die genannten Produkte der Übersetzungsschritte (z. B. Testaufgaben oder das Assessment-Feedback) vor dem Hintergrund bestimmter Anforderungen evaluiert.

Darüber hinaus werden nach Abschluss der Validierungsstudien und auf Grundlage der vorliegenden Daten weiterführende, explorative Analysen zur Struktur und Entwicklung der Reflexionsfähigkeit durchgeführt. Ziel dieser Analysen ist die Ableitung von Hypothesen zu beiden Aspekten, um zur Entwicklung der Theorie in diesem Bereich und der zielgerichteten Förderung der Reflexionsfähigkeit beizutragen. Dabei werden zunächst latente Profile der Reflexionsfähigkeit untersucht, die einen Ausgangspunkt für die zielgerichtete Förderung der Reflexionsfähigkeit bieten könnten:

> **Forschungsfrage 2:** Welche latenten Profile der Reflexionsfähigkeit lassen sich auf Grundlage der Bearbeitungen des Reflexionstests identifizieren?

Außerdem wird die Fähigkeit zum Noticing, dem Erkennen von bedeutsamen Aspekten im Unterrichtsgeschehen, als Ausgangspunkt für die Reflexion des Unterrichts näher betrachtet. In realen Reflexionssituationen stellt das Noticing insofern eine Voraussetzung für die Reflexion dar als lediglich solche Aspekte reflektiert werden können, die der reflektierenden Person auffallen. Auf Grundlage der vorliegenden Daten zum Reflexionstest, in dem die beiden Fähigkeiten zum Noticing und zur Unterrichtsreflexion getrennt erfasst werden, werden zunächst Zusammenhänge zwischen beiden Fähigkeiten untersucht:

> **Forschungsfrage 3:** Wie hängen die Fähigkeiten zum Noticing und zur Unterrichtsreflexion zusammen?

Darauf aufbauend wird anschließend die Entwicklung beider Fähigkeiten in den Blick genommen und eine mögliche wechselseitige Beeinflussung dieser untersucht:

> **Forschungsfrage 4:** Wie beeinflussen die Fähigkeiten zum Noticing und zur Unterrichtsreflexion wechselseitig ihre Entwicklung?

4.2 Einbettung der vorliegenden Arbeit in den Projektzusammenhang *ProfiLe-P*

Ziel der Projektreihe *ProfiLe-P* (*Professionskompetenz im Lehramtsstudium Physik*) ist, zu einer evidenzbasierten Gestaltung von Physik-Lehramtsstudiengängen

beizutragen. Im Rahmen der ersten beiden Projektphasen *Profile-P* (Kulgemeyer et al., 2012) und *Profile-P+* (Vogelsang, Riese, Kulgemeyer & Borowski, 2017) wurden daher Modelle und Testinstrumente zum fachbezogenen Professionswissen (physikalisches fachliches und fachdidaktisches Wissen) und professionellen Handeln (Erklären physikalischer Sachverhalte sowie Planung und Reflexion von Physikunterricht) entwickelt.

Im Projekt *Profile-P-Transfer* wurde darauf aufbauend der Transfer der entwickelten Instrumente in die Lehrpraxis angestrebt. Ausgangspunkt für diese Arbeit war das Ziel, das vorliegende Testinstrument zur Reflexion von Physikunterricht für einen nachhaltigen Transfer zu adaptieren. Hierzu bestand Bedarf für die Entwicklung einer technisch weitgehend automatisierten Kombination aus Testinstrument und Assessment-Feedback, um die Testergebnisse zeitnah so zu präsentieren, dass sie niedrigschwellig nutzbar sind und gleichzeitig hohen (v. a. personellen) Aufwand in der Durchführung und Auswertung des Instruments vermeiden. Die Entwicklung und Evaluation dieses Testinstruments und Assessment-Feedbacks sind Gegenstand der vorliegenden Arbeit.

4.3 Aufbau der vorliegenden Arbeit

Die vorliegende Arbeit ist in vier Teile untergliedert: Auf die Darlegung der theoretischen Grundlagen (Teil I) und die Vorstellung der entwickelten Lernumgebung aus Reflexionstest und Assessment-Feedback (Teil II) folgen deren theorie- und evidenzbasierte Evaluation und Bewertung im Rahmen der Validierungsstudien (Teil III). Auf dieser Grundlage werden abschließend weiterführende explorative Untersuchungen zur Entwicklung der Reflexionsfähigkeit (Teil IV) dargestellt.

Zunächst wird in Teil I ein Überblick über relevante theoretische Grundlagen und bisherige empirische Erkenntnisse gegeben. Dabei wird zunächst auf das professionelle Wissen und Handeln von Lehrkräften eingegangen (Kapitel 5); diese gelten im Rahmen von Wirkungsmodellen schulischen Lernens generell als Grundlage für gutes unterrichtliches Handeln der Lehrkräfte. Das professionelle Wissen wird im weiteren Verlauf der Arbeit als Bezugsrahmen für die Reflexion und die damit verbundene fundierte Einschätzung von Unterricht herangezogen. Anschließend werden Grundlagen zur Messung von professionellem Wissen und Handeln diskutiert (Kapitel 6), die im Rahmen der Entwicklung der Lernumgebung zur Reflexion von Physikunterricht sowie dessen Evaluation Berücksichtigung finden und insbesondere das der Evaluation zugrundeliegende Verständnis des Validitätsbegriffs beinhalten. Die dann folgende theoretische Diskussion des Reflexionsbegriffs (Kapitel 7) fokussiert insbesondere auf die Schwierigkeit eines einheitlichen Begriffsverständnisses und gibt einen Überblick über verschiedene Verwendungen des Reflexionsbegriffs sowie Modellierungen des Reflexionsprozesses bzw. der Qualität von Reflexionen. Auch werden empirische Ergebnisse zur Reflexionsfähigkeit (ange-

hender) Lehrkräfte vorgestellt, die auf einen Bedarf zur Förderung dieser Fähigkeit hindeuten, dem u. a. mit Feedback begegnet wird. Dazu werden in Kapitel 8 die für die Entwicklung und Bewertung des Assessment-Feedbacks relevanten theoretischen Grundlagen zu Feedback beschrieben. Abschließend werden in Kapitel 9 das dieser Arbeit zugrundeliegende Verständnis des Begriffs der Unterrichtsreflexion und der Reflexionsfähigkeit konkretisiert und die Konsequenzen beschrieben, die sich aus der Theorie für die Entwicklung und Evaluation der Lernumgebung zur Reflexion von Physikunterricht sowie die explorativen Untersuchungen zur Struktur und Entwicklung der Reflexionsfähigkeit ergeben.

In Teil II wird das entwickelte Testinstrument zur Erfassung der Reflexionsfähigkeit (Kapitel 10) mit zugehörigem Assessment-Feedback (Kapitel 11) vorgestellt. Zur Entwicklung des Testinstruments wurden der im Rahmen des vorangegangenen Projekts *ProfiLe-P+* entwickelte Performanztest zur Messung der Reflexionsfähigkeit sowie damit gewonnene empirische Erkenntnisse zugrunde gelegt.

Kernstück dieser Arbeit bildet Teil III. Hier wird die entwickelte Lernumgebung aus Testinstrument und Assessment-Feedback hinsichtlich der Validität der Testwertinterpretation untersucht und bewertet. Anschließend an die theoretischen Grundlagen zur validen Erfassung professionellen Handelns von Lehrkräften wird als Bezugsrahmen für die Validierung das Interpretations-Nutzungs-Argument festgelegt, das den Rahmen für die Evaluation der Lernumgebung bietet und Anforderungen beinhaltet, die für die valide Interpretation der Testwerte des Reflexionstests (a) als Maß für die Reflexionsfähigkeit sowie (b) als Ausgangspunkt für einen weiteren Professionalisierungsprozess erfüllt sein sollten (s. Kapitel 12). Entlang des formulierten Bezugsrahmens wird die entwickelte Lernumgebung schrittweise theorie- und evidenzbasiert untersucht (Kapitel 15 bis 18); auch die jeweils zur Anwendung kommenden Untersuchungsmethoden werden hier beschrieben. In Kapitel 19 werden die Erkenntnisse zur vollständigen Validitätsargumentation zusammengefasst, die die Beurteilung der Gültigkeit der Testwertinterpretation sowie einen Ausblick beinhaltet.

Anschließend werden in Teil IV auf Grundlage der vorliegenden Daten zum Reflexionstest Ergebnisse explorativer Untersuchungen zur Entwicklung der Reflexionsfähigkeit berichtet. Die Untersuchungen werden entlang eines Modells zum Zusammenhang von Noticing und Unterrichtsreflexion strukturiert (Kapitel 20) und umfassen eine latente Profilanalyse zur Reflexionsfähigkeit (Kapitel 21) sowie die Untersuchung von (Entwicklungs-)Zusammenhängen der Fähigkeiten zum Noticing und zur Reflexion (Kapitel 22 und 23).

Teil I

Theoretischer Hintergrund

5 Professionelles Wissen und professionelle Handlungskompetenz von Physiklehrkräften

In diesem Kapitel steht die Bedeutung des Wissens und adäquaten Handelns von Lehrkräften für die gelingende Unterrichtsgestaltung im Fokus. Ausgehend von einfachen Wirkungsmodellen schulischen Lernens, die den Einfluss der in der Lehrkräfteausbildung gewonnenen Fähigkeiten modellieren (Kapitel 5.1), wird ein Überblick über theoretische Grundlagen zum professionellen Wissen (Kapitel 5.2) und dem damit verbundenen professionellen Handeln von Lehrkräften gegeben (Kapitel 5.3), bevor auf die Modellierung des Verhältnisses zwischen professionellem Wissen und Handeln eingegangen wird (Kapitel 5.4).

5.1 Wirkungsmodelle schulischen Lernens

Als Planer*innen und Gestalter*innen des Unterrichts spielen Lehrkräfte eine zentrale Rolle für die Schaffung von Lerngelegenheiten für Schüler*innen. Plausibel erscheint dabei eine Abhängigkeit der Unterrichtsqualität vom Wissen und von sonstigen Voraussetzungen (z. B. Einstellungen) der Lehrkräfte. Terhart (2012) formuliert in diesem Sinne ein heuristisches Modell der Wirkungskette der schulischen Bildung (s. obere Zeile in Abbildung 5.1), das die Ausbildung der Lehrkräfte als Ausgangspunkt für ihr Handeln (insbesondere die Unterrichtsgestaltung) versteht, die somit das Lernen der Schüler*innen bestimmt. Neuweg (2014) differenziert in dieser Wirkungsstruktur die Rolle der Lehrkräftebildung weiter aus und unterscheidet zwischen angebotenen Lerngelegenheiten sowie angeregten Lernprozessen auf Seiten der Lehrkräfte und dem von ihnen erlangten Wissen.

Diez (2010) formuliert einen strukturell ähnlichen Prozess ausgehend von der Lehrkräfteausbildung über das erlernte Wissen und weitere Dispositionen von Lehrkräften und ihr unterrichtliches Handeln bis hin zum Schüler*innenlernen (s. Abbildung 5.1 inklusive grauer Ergänzungen), geht aber nicht von einer unmittelbaren Wirkung, sondern - abgeschwächt - von der Beeinflussung des jeweils

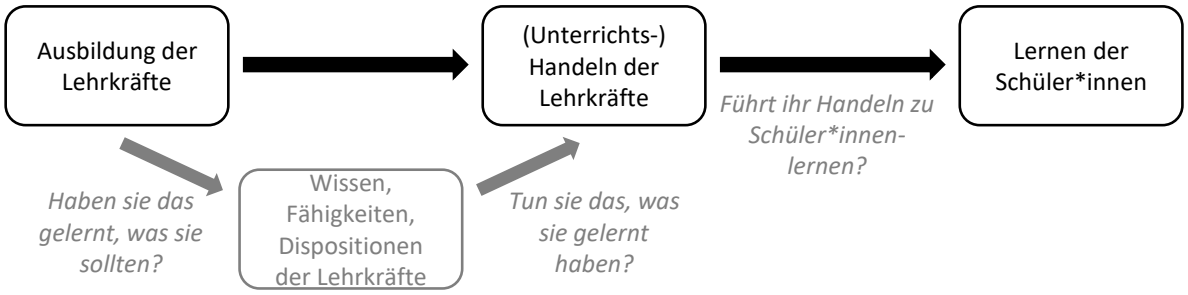


Abbildung 5.1: Wirkungskette der schulischen Bildung nach Terhart (2012, S. 7) in schwarz sowie (mit grauen Ergänzungen) Modell des Einflusses der Lehrkräftebildung auf das Schüler*innenlernen nach Diez (2010, S. 442) (deutsche Übersetzung angelehnt an Terhart (2012, S. 7))

nächsten Schritts aus (Diez, 2010; Terhart, 2012). Damit schafft er Raum für die Berücksichtigung diverser weiterer Einflussfaktoren (auf Seiten der Lernenden z. B. die Intelligenz, Motivation oder Unterstützung durch Eltern oder auf struktureller Ebene z. B. die Klassengröße (Bromme, 1997)). Solche Erweiterungen der modellierten Wirkungszusammenhänge erscheinen notwendig, auch, da die unmittelbaren Wirkungszusammenhänge empirisch nicht bestätigt werden konnten (Bartels, 2018, S. 11). Umfassendere Berücksichtigung finden äußere Einflussfaktoren im Rahmen von sogenannten Angebots-Nutzungs-Modellen. Ein inzwischen etabliertes Modell, das die Nutzung des Angebots *Unterricht* abhängig u. a. von Kontextfaktoren wie der Klassenzusammensetzung oder von individuellen Lernvoraussetzungen versteht, stellt Helmke (2015, S. 71ff.) vor.[1]

Empirische Untersuchungen in diesem Bereich orientieren sich zwar häufig an Wirkungs- bzw. Einflussmodellen der Lehrkräfte(-ausbildung) auf die Unterrichtsgestaltung und den Lernertrag der Schüler*innen, können diese aber nicht in ihrer Komplexität berücksichtigen, sondern fokussieren jeweils auf einzelne Bestandteile (z. B. das fachdidaktische Wissen von Lehrpersonen (Gramzow, 2015)) oder ausgewählte Zusammenhänge, bspw. zwischen dem Professionswissen und der Unterrichtsqualität (z. B. Vogelsang, 2014). Auf das Professionswissen und dessen Bestandteile wird in den nächsten Abschnitten näher eingegangen. Die Wirkungs- bzw. Angebots-Nutzungs-Modelle können dann dazu genutzt werden, die Erkenntnisse in einen größeren (Wirkungs-)Zusammenhang einzuordnen (Seidel, 2014).

[1]Während mit der Trennung zwischen Angebot und Nutzung des Unterrichts die aktive Rolle der Schüler*innen im Lernprozess betont wird, sehen Kohler und Wacker (2013) allerdings die Gefahr, dass Lehrkräfte sich nicht oder nur in geringerem Ausmaß als verantwortlich für das Erreichen eines Lernertrags verstehen.

5.2 Professionelles Wissen

Mit der Klassifikation des Lehrberufs als Profession geht die zentrale Bedeutung einer geteilten, für die Profession konstitutiven Wissensgrundlage, des *professionellen Wissens* (auch *Professionswissen*), einher (Schützeichel, 2020; Combe und Helsper, 2002; Stichweh, 1996). Das Professionswissen umfasst das Wissen, „über das die Angehörigen einer bestimmten Profession verfügen sollten“ (Riese et al., 2015, S. 57). Übertragen auf das Unterrichten als Kerntätigkeit des Lehrberufs umfasst es jenes Wissen, das für eine erfolgreiche Durchführung des Unterrichts als spezifisch angesehen wird und diese auszeichnet (Cauet, Liepertz, Borowski und Fischer, 2015; Riese et al., 2015; Kirschner, 2013, S. 6). Es ist also „domänenspezifisch“ (Baumert & Kunter, 2006, S. 483) und umfasst „alle theoretisch fundierten Wissensbestandteile, die im Rahmen der Ausbildung und unterrichtlichen Praxis von Lehrkräften erworben werden können“ (Borowski et al., 2010, S. 342).

5.2.1 Bestandteile des professionellen Wissens

Zur Charakterisierung des Professionswissens schlägt Shulman (1986; 1987) eine einschlägige Ausdifferenzierung zuletzt in folgende sieben Kategorien vor:

- „content knowledge;
- general pedagogical knowledge, with special reference to those broad principles and strategies of classroom management and organization that appear to transcend subject matter;
- curriculum knowledge, with particular grasp of the materials and programs that serve as ‚tools of the trade‘ for teachers;
- pedagogical content knowledge, that special amalgam of content and pedagogy that is uniquely the province of teachers, their own special form of professional understanding;
- knowledge of learners and their characteristics;
- knowledge of educational contexts, ranging from the workings of the group or classroom, the governance and financing of school districts, to the character of communities and cultures; and
- knowledge of educational ends, purposes, and values, and their philosophical and historical grounds.“ (Shulman, 1987, S. 8)

Bromme (1997) übernimmt in seiner Topologie des Professionswissens fünf der von Shulman (1987) aufgeführten Kategorien und unterscheidet *fachliches Wissen*, *pädagogisches Wissen*, *curriculares Wissen*, *fachspezifisch-pädagogisches Wissen* und die *Philosophie des Schulfachs*. Die verschiedenen Wissensbestände weisen Überschneidungen auf (Bromme, 1997). Als zentrales Charakteristikum des Pro-

fessionswissens von Lehrkräften benennt er dazu passend „[d]ie Verschmelzung von Kenntnissen unterschiedlicher Herkunft“ (Bromme, 1997, S. 198), d. h. der unterschiedlichen Wissensbestände. Die Integration unterschiedlicher Wissensbestände kann durchaus unbewusst passieren (Bromme, 1992).

Baumert und Kunter (2006, S. 482) differenzieren im Rahmen ihres Modells der professionellen Handlungskompetenz zwischen den zentralen Wissensarten *pädagogisches Wissen*, *Fachwissen*, *fachdidaktisches Wissen* sowie dem *Organisationswissen*, das dem *knowledge of educational contexts* nach Shulman (1987) entspricht (Cauet, 2016, S. 13), und dem für die „Kommunikation mit Laien“ (Baumert & Kunter, 2006, S. 482) relevanten *Beratungswissen*. Für die Diskussion der beiden letztgenannten Wissensarten schränken sie allerdings ein, dass diese „sinnvoll nur im Rahmen institutionalisierter und sozial verteilter Wissensbestände behandelt werden können“ (Baumert & Kunter, 2006, S. 482).

Im deutschsprachigen Raum hat sich eine Dreiteilung des Professionswissens in *pädagogisches Wissen* (entsprechend dem *general pedagogical knowledge* nach Shulman (1987)), *Fachwissen* (entsprechend Shulmans (1987) *content knowledge* oder auch *subject matter knowledge*), und *fachdidaktisches Wissen* (häufig synonym verwendet zu Shulmans (1987) *pedagogical content knowledge (PCK)*, z. B. von Baumert und Kunter (2006))[2] weitgehend durchgesetzt (z. B. Cauet, Liepertz, Borowski und Fischer, 2015; Lenske et al., 2015; Riese, 2009, S. 73ff.; Baumert und Kunter, 2006). Die Trennbarkeit dieser drei Dimensionen des Professionswissens konnte z. B. von Kirschner (2013, S. 76) und Riese und Reinhold (2012) empirisch bestätigt werden. Überdies weist diese Struktur hohe Übereinstimmung mit jener der universitären Lehramtsausbildung auf, die ebenfalls pädagogische, fachwissenschaftliche und fachdidaktische Anteile umfasst (Gramzow, 2015, S. 27). Auch diese Arbeit folgt der etablierten Struktur des Professionswissens; dessen drei

[2]Die Gleichsetzung von PCK und fachdidaktischem Wissen ist aufgrund konzeptioneller Unterschiede inhaltlich problematisch, wie Gramzow, Riese und Reinhold (2013) ausführen, da i. d. R. konzeptionelle Unterschiede in den Rahmenmodellen bestehen - z. B. bezüglich des Verhältnisses von PCK und fachdidaktischem Wissen zum Fachwissen und pädagogischen Wissen. Gramzow, Riese und Reinhold (2013) beschreiben allerdings, dass das fachdidaktische Wissen das PCK inkludiert und zusätzlich bspw. Wissensbestände umfasst, die nicht unmittelbar für die Unterrichtspraxis relevant sind. Beide Konzeptualisierungen können also zumindest als verwandt verstanden werden, sodass sie im Folgenden nicht dezidiert unterschieden werden.

zentrale Facetten und deren Bedeutung für die Qualität von Unterricht[3] werden in den Kapiteln 5.2.2 bis 5.2.5 näher charakterisiert.

Das Professionswissen kann (bzw. dessen einzelnen Bereiche können) überdies in Bezug auf die kognitiven Anforderungen bzw. die Handlungsnähe, die mit der Anwendung des Wissens verbunden sind, weiter ausdifferenziert werden. Stender, Brückmann und Neumann (2015) beschreiben die Gemeinsamkeit dieser Ausdifferenzierungen als Unterscheidung „zwischen zwei Polen von Wissenstypen [...] - einem an der Universität erworbenen (eher theoretischen) Wissen und einem durch die aktive Auseinandersetzung mit konkreten Unterrichtssituationen entwickelten“ (Stender, Brückmann & Neumann, 2015, S. 123). Verbreitet ist demgemäß eine Unterteilung in *deklaratives* und *prozedurales Wissen* (z. B. Baumert und Kunter, 2006; Bromme, 1997).[4] Mit deklarativem Wissen ist das Wissen „über Begriffe, Fakten und Sachverhalte“ (Dittrich, 2020, S. 48; z. B. auch Bromme, 1992, S. 43) gemeint, das verbalisierbar ist, aber „nicht automatisch in der Praxis angewendet werden kann“ (Dittrich, 2020, S. 48). Es wird auch als „Faktenwissen“ (Seifert, Hilligus & Schaper, 2009) oder „konzeptuelle[s] Wissen“ bezeichnet (Dittrich, 2020, S. 48). Das prozedurale Wissen (auch „Prozess- und Handlungswissen“ (Dittrich, 2020, S. 48) oder „Können“ (Baumert & Kunter, 2006, S. 481)) ist das handlungsleitende Wissen, das z. B. Wissen über Lösungsstrategien für bestimmte Aufgaben beinhaltet bzw. situativ angemessenes Handeln ermöglicht (Dittrich, 2020, S. 48; Bromme, 1992, S. 43).[5] Andere Autoren schlagen Ausdifferenzierungen in drei Bereiche vor, die als *Reproduzieren*, *Verstehen und Anwenden* sowie *Beurteilen*

[3] Unterricht kann als gut oder erfolgreich bewertet werden, wenn die mit ihm verbundenen Ziele (i. d. R. ein Kompetenzerwerb auf fachlicher Ebene durch die Schüler*innen sowie nicht fachspezifische Bildungsziele (z. B. Demokratiebildung oder Bildung für nachhaltige Entwicklung (Kultusministerkonferenz, 2020) und motivationale Ziele) erfüllt werden (Cauet, 2016, S. 24f.; Vogelsang, 2014, S. 142ff.; Gräsel und Göbel, 2011). Zur Beschreibung der Qualität von Unterricht werden unterschiedliche Kriterien herangezogen (z. B. Helmke, 2015, S. 168ff.). Unterschieden wird aufbauend auf den Basismodellen für das Unterrichten von Oser und Baeriswyl (2001) bspw. zwischen *Sicht-* und *Tiefenstruktur* des Unterrichts, wobei letztere bedeutsamer für Lernerfolge der Schüler*innen ist (Kunter & Voss, 2011). Verbreitet ist eine Differenzierung von (mindestens) drei Bereichen oder Dimensionen der Unterrichtsqualität: *Klassenführung*, *konstruktive Unterstützung* und *kognitive Aktivierung* (z. B. Kunter und Voss, 2011; Cauet, 2016, S. 27ff.). Cauet (2016, S. 25ff.) und Vogelsang (2014, S. 151ff.) liefern einen Überblick über unterschiedliche Modelle zur Bewertung der Qualität von (Physik-)Unterricht.

[4] Teilweise ergänzt um eine dritte Komponente, das *konditionale Wissen*, welches das Wissen über die (Begründung der) Angemessenheit einer Handlung und damit verbundener Rahmenbedingungen beschreibt (Lenske et al., 2015; Tepner et al., 2012). Zusammenhänge zu anderen Klassifikationen verschiedener Wissensformen (u. a. dem impliziten und expliziten Wissen) beschreibt Dittrich (2020, S. 47ff.) überblicksartig.

[5] Eine ausführlichere Unterscheidung in Faktenwissen (Kenntnis von isolierten Fakten), Konzeptwissen (Kenntnis von komplexeren Konzepten, z. B. Theoriewissen), Prozedurales Wissen (Wissen über Methoden und deren adäquater Einsatz) und Metakognitives Wissen (u. a. Wissen über Lösungsstrategien, Aufgaben und Kontexte) schlagen z. B. Anderson et al. (2001, S. 27ff.) vor.

und Analysieren zusammengefasst werden können (Riese, 2009, S. 78f.; Seifert, Hilligus und Schaper, 2009). Auch hier nimmt die Komplexität der kognitiven Anforderung zu.

5.2.2 Pädagogisches Wissen

Das *pädagogische Wissen*[6] ist zentraler Bestandteil des Professionswissens von Lehrkräften (z. B. Dittrich, 2020, S. 56; Baumert und Kunter, 2006; Shulman, 1987) und gilt damit neben fachspezifischem Wissen als eine „Voraussetzung für erfolgreiches Handeln im Unterricht“ (Voss & Kunter, 2011, S. 199). Es umfasst im Allgemeinen das „fächerübergreifende Wissen um Strategien und Mittel zur Genese und Erhaltung lernförderlicher Bedingungen und zur Optimierung von Lehr-Lern-Situationen“ (Lenske et al., 2015, S. 227) und wird damit - im Gegensatz zu fachdidaktischem Wissen und Fachwissen - nicht als fachgebunden verstanden (Lenske et al., 2015; Voss, Kunina-Habenicht, Hoehne und Kunter, 2015; König und Blömeke, 2009).[7] In Bezug auf konkrete Definitionen lässt sich kein einheitliches Verständnis des pädagogischen Wissens in der Literatur ausmachen (Dittrich, 2020, S. 56). Frühere Beschreibungen des pädagogischen Wissens fokussieren häufig auf die Klassenführung (Voss & Kunter, 2011), so umfasst bspw. nach Shulman (1987) das pädagogische Wissen fachunabhängige Prinzipien und Strategien der Klassenführung und -organisation. Zeitlich nachfolgende (theoretische) Ausdifferenzierungen und Operationalisierungen des pädagogischen Wissens enthalten zwar überwiegend auch „die Effizienz des Klassenführung“ (Voss, Kunina-Habenicht, Hoehne und Kunter, 2015, S. 192, s. auch Dittrich, 2020, S. 58), sind allerdings nicht unbedingt auf Belange des Unterrichtens beschränkt,

[6] Auch (mit jeweils etwas unterschiedlicher definitorischer Grundlage) z. B. *(generelles) pädagogisch-psychologisches Professionswissen* (Voss und Kunter, 2011; Lenske et al., 2015), *allgemeines pädagogisches Wissen* (Baumert & Kunter, 2006), *erziehungswissenschaftliches Wissen*, (Seifert, Hilligus & Schaper, 2009) oder *bildungswissenschaftliches Wissen* (Lohse-Bossenz et al., 2018)

[7] Obwohl die Fachunabhängigkeit des pädagogischen Wissens übereinstimmend vertreten wird, beschreiben Voss und Kunter (2011) ein „Ineinandergreifen von fachspezifischen und fachunabhängigen Wissensbeständen“ (Voss & Kunter, 2011, S. 198) u. a. in Bezug auf die Anregung von Lernprozessen bei Schüler*innen, für die neben fachunabhängigem Wissen bspw. „Wissen über fachliche Aufgaben und deren Potenzial“ (Voss & Kunter, 2011, S. 199) benötigt wird. Auch zeigen Lohse-Bossenz et al. (2018), dass insbesondere zu Beginn des Vorbereitungsdienstes Niveauunterschiede zwischen Lehrkräften, die (nur) mathematisch-naturwissenschaftliche Fächer studierten, und denen, die (nur) sprach- und gesellschaftswissenschaftliche Fächer studierten, in den Bereichen Diagnostik/Evaluation (zugunsten der Gruppe mit mathematisch-naturwissenschaftlichen Fächern) und Bildungstheorie sowie Lernen/Entwicklung existierten (zugunsten der Gruppe mit sprach- und gesellschaftswissenschaftlichen Fächern). Für die unterrichtsnahen Bereiche Unterrichtsgestaltung und Lernen/Entwicklung sind allerdings keine Unterschiede feststellbar (Lohse-Bossenz et al., 2018).

sondern beinhalten darüber hinaus z. B. die Unterrichtsplanung als Facette (Voss, Kunina-Habenicht, Hoehne und Kunter, 2015; König und Blömeke, 2009).[8]

Stark rezipiert (z. B. von Lohse-Bossenz et al., 2018; Voss, Kunina-Habenicht, Hoehne und Kunter, 2015; König und Blömeke, 2009; Riese, 2009, S. 42) sind die vier von Baumert und Kunter (2006) vorgeschlagenen Facetten des pädagogischen Wissens, die später (Kunter, Baumert & Blum, 2011) ausdifferenziert und um eine fünfte Facette ergänzt wurden:

1. Konzeptuelles bildungswissenschaftliches Grundlagenwissen
2. Allgemeindidaktisches Konzeptions- und Planungswissen
3. Wissen über Unterrichtsführung und Orchestrierung von Lerngelegenheiten
4. Fachübergreifende Prinzipien des Diagnostizierens, Prüfens und Bewertens
5. Methodische Grundlagen empirischer Sozialforschung

Baumert und Kunter (2006) stellen diese Ausdifferenzierung als Systematisierung von „konsensfähigen Vorschlägen“ (Baumert & Kunter, 2006, S. 484) der Literatur vor. Eine empirische Prüfung dieser Facetten steht allerdings noch aus.

König und Blömeke (2009) schlagen eine Operationalisierung des pädagogischen Wissens in fünf Dimensionen entlang „beruflicher Anforderungen“ (König & Blömeke, 2009, S. 504) an Lehrkräfte vor. Die *Strukturierung von Unterricht* beschäftigt sich mit der Planung und Analyse sowie der curricularen Strukturierung von Unterricht (König & Blömeke, 2009). Die Facette *Motivierung* umfasst die „Leistungsmotivation“ (König & Blömeke, 2009, S. 505) der individuellen Lernenden sowie „Motivationsstrategien im Unterricht“ (König & Blömeke, 2009, S. 505). Kenntnisse zur Umsetzung von Differenzierungsmaßnahmen und methodischer Vielfalt werden der Dimension *Umgang mit Heterogenität* zugerechnet und die letzten beiden Dimensionen fokussieren die effiziente *Klassenführung* und angemessene *Leistungsbeurteilung* u. a. unter Berücksichtigung relevanter Gütekriterien und Vermeidung von „Urteilsfehlern“ (König & Blömeke, 2009, S. 507). Außerdem unterscheiden König und Blömeke (2009) zwischen *deklarativem Wissen* („Wissen, dass...“ (König & Klemenz, 2015, S. 251)) und dem für konkrete Handlungen relevantem *prozeduralem Wissen* („Wissen, wie...“ (König & Klemenz, 2015, S. 251)). Diese Modellierung des pädagogischen Wissens ist auch Grundlage für die Erfassung desselben mit einem Testinstrument im Rahmen von *TEDS-M*[9] und darauf aufbauenden Studien (z. B. zum internationalen Ver-

[8] Eine Übersicht über verschiedene Konzeptualisierungen des pädagogischen Wissens und dessen Operationalisierung zur Erfassung mit Testinstrumenten stellen Voss, Kunina-Habenicht, Hoehne und Kunter (2015) dar, die von Dittrich (2020, S. 56f.) um neuere Beschreibungen ergänzt wird.

[9] *Teacher Education and Development Study in Mathematics*: Ein Projekt zum internationalen Vergleich der Wirksamkeit der Ausbildung von Mathematiklehrkräften; eine umfangreiche Darstellung der Ergebnisse liefern Blömeke, Kaiser und Lehmann (2010).

gleich der Lehrkräfteausbildung (König & Klemenz, 2015) oder zur Entwicklung eines Kompetenzniveaumodells des pädagogischen Wissens (Felske et al., 2020)). Sie unterscheiden dabei zwischen den kognitiven Anforderungen *Erinnern*, *Verstehen/Analysieren* und *Kreieren/Generieren von Handlungsoptionen*, die mit zunehmender Praxisnähe einhergehen (König & Klemenz, 2015). Im Rahmen einer Rasch-Modellierung erweist sich die Annahme einer mehrdimensionalen Struktur des pädagogischen Wissens einem eindimensionalen Modell statistisch überlegen, was zusätzlich durch geringe Korrelationen der Dimensionen untereinander gestützt wird (König & Blömeke, 2009).

Im Rahmen von *COACTIV-R*[10] wird das pädagogisch-psychologische Wissen von Lehrkräften in den vier Facetten *Wissen über effiziente Klassenführung*, *Wissen über Unterrichtsmethoden*, *Wissen über Leistungsbeurteilung* sowie *Wissen über Lernprozesse und individuelle Besonderheiten von Schüler*innen* im Vorbereitungsdienst erfasst (Voss & Kunter, 2011). Auch hier kann die Angemessenheit der mehrdimensionalen Modellierung empirisch bestätigt werden (Voss & Kunter, 2011).

Seifert, Hilligus und Schaper (2009) entwickeln vor dem Hintergrund der inhaltlichen Ausgestaltung der Paderborner Lehramtsausbildung ein zweidimensionales Modell der Struktur des pädagogischen Wissens, das in seinen Inhaltsbereichen (erste Dimension) *Erziehung und Bildung*, *Unterricht und Allgemeine Didaktik* und *Schulentwicklung und Gesellschaft* alle in den KMK-Standards für die Lehrerbildung (Sekretariat der Kultusministerkonferenz, 2004) enthaltenen Kompetenzbereiche umfasst. In einer zweiten Dimension sind die kognitiven Anforderungen *Wissen reproduzieren, Verstehen*, *Reflektieren, Anwenden* und *Urteilen, Bewerten, Entscheiden* in einer hierarchischen Struktur abgebildet (Seifert, Hilligus & Schaper, 2009).

Dittrich (2020, S. 58) fasst die in den verschiedenen Definitionen des pädagogischen Wissens aufgeführten Inhaltsbereiche in *planerische Wissensbereiche* (z. B. Classroom-Management), *interaktive Wissensbereiche* (z. B. Wissen um Methoden oder zum Umgang mit Heterogenität) und *theoretische Wissensbereiche* (z. B. Wissen zum Diagnostizieren und Bewerten) zusammen.

Hinsichtlich der tatsächlichen Bedeutung des pädagogischen Wissens für gelingenden bzw. qualitätsvollen Unterricht gibt Riese (2009) zu bedenken, dass die unterschiedlichen Facetten des pädagogischen Wissens „in unterschiedlicher Entfernung zur praktischen Berufs- und Unterrichtstätigkeit von Lehrkräften stehen, wobei mit zunehmendem Abstand vom Unterricht die Handlungsrelevanz schwieriger nachzuweisen sein dürfte“ (Riese, 2009, S. 44). Dennoch lässt sich empirische Evidenz

[10] *Cognitive Activation in the Classroom: The Orchestration of Learning Opportunities for the Enhancement of Insightful Learning in Mathematics - Referendariatsstudie*: Professionswissen von Lehrkräften, kognitiv aktivierender Mathematikunterricht und die Entwicklung von mathematischer Kompetenz (Ergänzungsstudie zum Referendariat); ein Forschungsprojekt zur Untersuchung der Entwicklung der Professionalität von Lehrkräften im Vorbereitungsdienst

für die Bedeutung des pädagogischen Wissens für die Unterrichtspraxis finden. So berichten Voss, Kunina-Habenicht, Hoehne und Kunter (2015) auf Grundlage von Daten aus *COACTIV-R*, das pädagogische Wissen sei „prädiktiv für die spätere Effizienz der Klassenführung sowie die konstruktive Lernunterstützung“ (Voss, Kunina-Habenicht, Hoehne & Kunter, 2015, S. 209). Zwischen pädagogischem Wissen und der kognitiven Aktivierung im Unterricht wird kein Zusammenhang festgestellt (Voss et al., 2014).[11] Positive Zusammenhänge zwischen dem pädagogischen Wissen (erfasst mit dem Instrument aus *TEDS-M*) und der Klassenführung, der Klarheit der Instruktionen und de Beziehung zwischen der Lehrkraft und den Schüler*innen bestätigen auch Untersuchungen von König und Pflanzl (2016), die die Unterrichtsqualität von Mathematikunterricht basierend auf Einschätzungen der Schüler*innen messen. Auch für die mit einem videobasierten Testinstrument erfasste Klassenführung kann der Zusammenhang zum pädagogischen Wissen bestätigt werden (König & Kramer, 2016). Für den Bereich der Physik kann Cauet (2016, S. 173ff.) einen Zusammenhang zwischen dem pädagogischen Wissen der Lehrkräfte und dem Fachwissen der von ihnen unterrichteten Schüler*innen (nach Kontrolle u. a. des Vorwissens der Schüler*innen) nachweisen. Auch können - im Widerspruch zu den Ergebnissen von Voss et al. (2014) - u. a. Zusammenhänge zwischen dem pädagogischen Wissen und der kognitiv aktivierenden Unterrichtsgestaltung gefunden werden (Cauet, 2016, S. 188; Vogelsang, 2014, S. 479). Auch wenn die Befundlage bezüglich der Zusammenhänge des pädagogischen Wissens zur kognitiven Aktivierung als Dimension der Unterrichtsqualität (möglicherweise aufgrund unterschiedlicher genutzter Testinstrumente[12]) nicht eindeutig ist, kann die Bedeutung des pädagogischen Wissens für gelingenden Unterricht also nicht nur theoretisch begründet, sondern auch empirisch beobachtet werden.

5.2.3 Fachwissen

Das Fachwissen von (angehenden) Physiklehrkräften umfasst ihr Verständnis physikalischer Konzepte (Woitkowski & Riese, 2017); es ist domänenspezifisch und gilt als notwendige Voraussetzung für das Gelingen von Unterricht (z. B. Krauss et al., 2008; Baumert und Kunter, 2006). Bestandteil des Fachwissens muss

[11] Voss et al. (2014) vermuten hier eine erhöhte Bedeutung des fachdidaktischen Wissens anstelle des pädagogischen, schließen aber auch die Wahl des Testinstruments nicht als Ursache aus.

[12] Brühwiler et al. (2017) finden bei der Untersuchung dreier weiterer Instrumente zur Erfassung des pädagogischen Wissens keine signifikanten Zusammenhänge zwischen ihnen. Variierende Ergebnisse abhängig von der Schwerpunktsetzung bei der Entwicklung des jeweiligen Instruments sind also durchaus möglich. Somit kann an dieser Stelle nicht eindeutig von einem Zusammenhang zwischen pädagogischem Wissen und kognitiver Aktivierung der Schüler*innen im Unterricht als Merkmal der Unterrichtsqualität ausgegangen werden. Da aber in beiden Untersuchungen ein Zusammenhang zu einem Bestandteil des Professionswissens (zum pädagogischen bzw. fachdidaktischen Wissen) gefunden wurde, kann dies zumindest als Hinweis auf dessen Bedeutung für die Unterrichtsqualität verstanden werden.

das Wissen sein, das Lehrkräfte unterrichten; es ist allerdings nicht ausreichend, um die Anforderungen zu bewältigen, „die sich Lehrkräften bei der Vorbereitung und Durchführung des Unterrichts stellen“ (Baumert & Kunter, 2006, S. 485). Shulman (1986) fordert sogar, „that the subject matter content understanding of the teacher be at least equal to that of his or her lay colleague, the mere subject matter major“ (Shulman, 1986, S. 9), da Lehrkräfte nicht nur verstehen müssten, *dass* etwas so ist, wie es ist, sondern zusätzlich, *weshalb*.

Aufbauend auf Modellierungen des Fachwissens von Mathematiklehrkräften in den Projekten COACTIV[13] (Krauss et al., 2008) und MT21[14] (Blömeke, Kaiser & Lehmann, 2008) schlägt Riese (2009, S. 76ff.) eine dreidimensionale Modellierung der physikalischen Fachkompetenz vor, die neben den verschiedenen physikalischen Inhaltsbereichen (für die Mechanik z. B. Impuls oder Energie (Riese, 2009, S. 77f.)) zwischen den drei explizit nicht als hierarchisch zu verstehenden kognitiven Aktivitäten *Reproduzieren*, *Verstehen* und *Beurteilen* unterscheidet (Riese, 2009, S. 78f.) und eine Unterteilung des (physikalischen) Fachwissens in drei Niveaustufen vornimmt, die mit zunehmender fachlicher Schwierigkeit bzw. einem zunehmenden Abstraktionsgrad einhergehen (Woitkowski und Borowski, 2017; Riese, 2009, S. 79f.):

- *Schulwissen* ist das Wissen, das im schulischen Physikunterricht gemäß curricularen Vorgaben erlernt werden soll (Riese et al., 2015).
- *Vertieftes Schulwissen*[15] beschreibt die Betrachtung von Inhalten des schulischen Physikunterrichts „von einem höheren Standpunkt aus“ (Riese, 2009, S. 80). Dazu gehört u. a. das Wissen um unterschiedliche Strategien zum Lösen einer Aufgabe oder die adäquate Adaption einer Aufgabe für eine bestimmte Zielgruppe (Riese et al., 2015).[16]
- *Universitäres Wissen* umfasst die „Inhalte, die an Universitäten für Abschlussprüfungen im Fach vorausgesetzt werden“ (Riese et al., 2015, S. 60), die „nicht mehr Teil des Schulcurriculums“ (Riese, 2009, S. 80) sind.

[13] *Cognitive Activation in the Classroom: The Orchestration of Learning Opportunities for the Enhancement of Insightful Learning in Mathematics*: Professionswissen von Lehrkräften, kognitiv aktivierender Mathematikunterricht und die Entwicklung von mathematischer Kompetenz; ein Forschungsprojekt zur Untersuchung der Entwicklung der Professionalität von Mathematik-Lehramtsstudierenden

[14] *Mathematics Teaching in the 21st Century*: Eine internationale Vergleichsstudie zur Wirksamkeit der Mathematiklehrerausbildung, in der u. a. das mathematische, mathematikdidaktische und pädagogische Wissen von (angehenden) Mathematiklehrkräften erfasst wurde

[15] Riese (2009, S. 80) benennt diese Stufe zunächst lediglich *vertieftes Wissen*, betont aber durchgängig die Nähe zu Inhalten des Schulwissens (Riese, 2009, S. 80; Riese und Reinhold, 2010). Vergleichbar verwenden Baumert und Kunter (2011) die Bezeichnung „profundes mathematisches Verständnis des Hintergrunds des Schulstoffs“ (Baumert & Kunter, 2011, S. 37).

[16] Eine Übersicht über die von verschiedenen Autoren beschriebenen charakteristischen Merkmale des vertieften (Schul-)Wissens liefern Woitkowski und Borowski (2017).

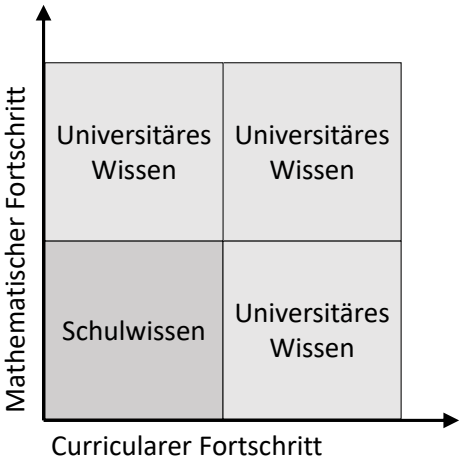


Abbildung 5.2: Abgrenzung von Schulwissen und universitärem Wissen abhängig vom curricularen und mathematischen Fortschritt nach Enkrott (2021, S. 45)

Empirische Evidenz für die Modellierung des Fachwissens mit diesen drei trennbaren Facetten[17] legt Enkrott (2021, S. 70f.) vor. Obwohl Enkrott (2021, S. 44) die Facetten des Fachwissens explizit nicht als Niveaustufen versteht, geht das universitäre Wissen gegenüber dem Schulwissen mit höherem curricularem Fortschritt der jeweiligen Fachinhalte als auch mit höherem mathematischen Fortschritt einher (s. Abbildung 5.2; Enkrott, 2021, S. 44). Das vertiefte Schulwissen ist nicht über ein zugehöriges Curriculum definiert, geht über schulisches Wissen hinaus (Woitkowski, 2015, S. 123) und wird i. d. R. nicht explizit in schulischen oder universitären Veranstaltungen thematisiert (Enkrott, 2021, S. 46). Riese (2009) beschreibt allerdings, dass Aufgaben, die das vertiefte Wissen betreffen, „auf Grund der Nähe zum Schulcurriculum prinzipiell noch von sehr guten Schülern gelöst werden können“ (Riese, 2009, S. 80). Das vertiefte Schulwissen kann also gleichzeitig Teile des universitären wie auch des schulischen Wissens betreffen. Es ist außerdem davon auszugehen, dass es mit zunehmendem curricularen Fortschritt „en passant“ (Woitkowski & Borowski, 2017, S. 64) erlernt wird.

5.2.4 Fachdidaktisches Wissen

Das *fachdidaktische Wissen* von Physiklehrkräften stellt die dritte zentrale Komponente ihres Professionswissen dar. Shulman beschreibt das verwandte PCK als „knowledge *for teaching*“ (Shulman, 1986, S. 9), das spezifisch und konstitutiv

[17] In Abgrenzung zur Modellierung der Niveaustufen des Fachwissens von Riese (2009, S. 79), nach welchem Schulwissen, vertieftes Wissen und universitäres Wissen mit zunehmender Schwierigkeit einhergehen, fasst Enkrott (2021) diese drei Facetten als nicht hierarchisch geordnete „trennbare Dimensionen eines übergeordneten Konstrukts“ (Enkrott, 2021, S. 44) auf.

für (Physik-)Lehrkräfte ist und sie von Fachspezialisten (hier: Physikern) unterscheidet (Magnusson, Krajcik und Borko, 1999; Shulman, 1987). Es umfasst jenes Wissen, das es Lehrkräften ermöglicht, „fachliche Gegenstände zu strukturieren, darzustellen, zu erklären und zu vernetzen“ (Riese und Reinhold, 2010, S. 170), um diese für Lernende aufzubereiten. Umfangreiches fachdidaktisches Wissen setzt also notwendigerweise Fachwissen voraus (Krauss et al., 2008; Baumert und Kunter, 2006), wobei unklar ist, ob hier (auch) universitäres Wissen notwendig oder Schulwissen ausreichend ist (Riese & Reinhold, 2010).

Magnusson, Krajcik und Borko (1999) entwickeln ausgehend von Shulmans (1986, 1987) Konzeptualisierung ein international etabliertes und häufig adaptiertes Modell des PCK für naturwissenschaftliche Fächer (Chan und Hume, 2019, S. 43; Kirschner et al., 2016). Nach diesem Modell (s. Abbildung 5.3, schwarze Pfeile) beinhaltet das PCK Wissen über curricular vorgegebene Ziele und (erprobte) Lehrgänge (*Knowledge of Science Curricula*), Wissen über Bedarfe der Schüler*innen im Lernprozess sowie deren (ggf. fachlich falsche) Vorstellungen und Schwierigkeiten (*Knowledge of Students' Understanding of Science*), Wissen über die Bewertung (der Ergebnisse) naturwissenschaftlicher Lernprozesse (*Knowledge of Assessment of Scientific Literacy*), Wissen über fachspezifische Instruktionsstrategien und themenspezifische Repräsentationsformen (*Knowledge of Instructional Strategies*) sowie - übergeordnet und mit diesen Facetten in Wechselwirkung stehend - Wissen und Überzeugungen zu Zielen des naturwissenschaftlichen Unterrichts (*Orientations towards Teaching Science*). Abgrenzend dazu konzeptualisieren Park und Chen (2012) das PCK mit denselben fünf Facetten (s. Abbildung 5.3, durchgezogene Pfeile), betonen allerdings die Einflüsse aller Facetten aufeinander. Sie definieren PCK als „integration of two or more components in the pentagon model“ (Park & Chen, 2012, S. 928).

Im deutschsprachigen Raum liegen Konzeptionen des fachdidaktischen Wissens z. B. für naturwissenschaftliche Fächer aus dem Projekt *ProwiN*[18] (Tepner et al., 2012) vor, explizit für die Physik von Riese (2009, S. 82f.) und für die Mathematik aus dem Projekt *COACTIV* (Krauss et al., 2008, S. 234). Insgesamt liegt für das fachdidaktische Wissen keine konsensuale Modellierung vor (Chan und Hume, 2019, S. 53f.; Gramzow, Riese und Reinhold, 2013). Unterschiede lassen sich in Bezug auf die Dimensionalität der Modellierung finden, die berücksichtigten (inhaltlichen) Facetten des fachdidaktischen Wissens und die berücksichtigten Wissensarten, wie dem deklarativen, prozeduralen und konditionalen Wissen (Gramzow, Riese und Reinhold, 2013; Tepner et al., 2012).[19]

[18] *Professionswissen von Lehrkräften in den Naturwissenschaften*: Ein Projekt zur Untersuchung des Einflusses der Bestandteile des Professionswissens auf die Unterrichtsgestaltung und den Kompetenzzuwachs bei Schüler*innen in den Fächern Physik, Chemie und Biologie (Borowski et al., 2010).

[19] Umfangreiche Übersichten über die in unterschiedlichen Modellierungen berücksichtigten Facetten des fachdidaktischen Wissens (bzw. PCK) finden sich z. B. in Kirschner et al. (2016), Cauet (2016, S. 18) und Gramzow (2015, S. 79)

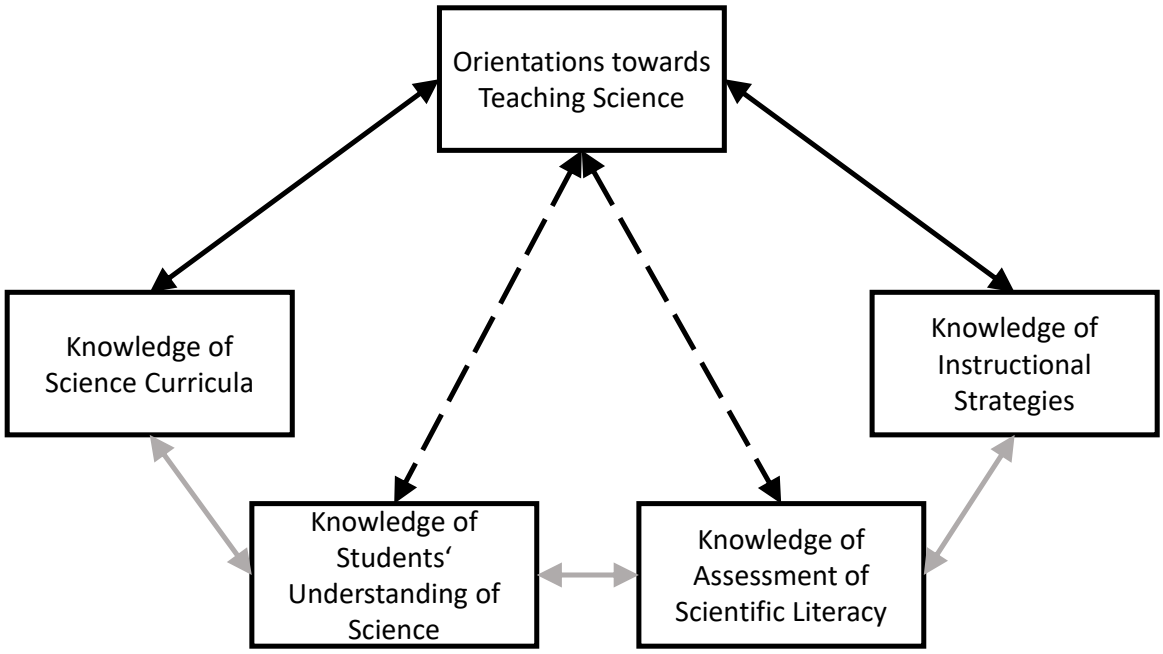


Abbildung 5.3: Komponenten des PCK nach Magnusson, Krajcik und Borko (1999, S. 99) (schwarze gestrichelte und durchgezogene Pfeile beschreiben eine gegenseitige Beeinflussung) sowie die darauf aufbauende Modellierung des PCK von Park und Chen (2012, S. 925) (durchgezogene graue und schwarze Pfeile markieren die gegenseitige Beeinflussung der Facetten).

Gramzow, Riese und Reinhold (2013) stellen eine zweidimensionale Modellierung des fachdidaktischen Wissens angehender Physiklehrkräfte vor, die Grundlage für die Messung des fachdidaktischen Wissens im Projekt *ProfiLe-P* ist (Riese et al., 2015). Während eine Dimension abstrakt die verschiedenen physikalischen Inhaltsbereiche abdeckt, unterscheidet die zweite die folgenden Facetten des fachdidaktischen Wissens (s. Abbildung 5.4; Gramzow, Riese und Reinhold, 2013, S. 20ff.): Die Facette...

- *Instruktionsstrategien* umfasst Wissen über ebendiese sowie zu deren Umsetzung und Bewertung.
- *Schülervorstellungen*[20] umfasst das Wissen über Schülervorstellungen zu den physikalischen Inhalten sowie zum Umgang mit diesen.
- *Experimente und Vermittlung eines angemessenen Wissenschaftsverständnisses* umfasst „sowohl das Wissen über Arten, Funktionen und die Bedeutung von Experimenten [...] als auch Wissen über das Vorgehen beim Experimentieren“ (Gramzow, Riese & Reinhold, 2013, S. 23) sowie Wissen bezüglich der Entwicklung eines angemessenen Wissenschaftsverständnisses auf Seiten der Lernenden.
- *Kontext und Interesse* umfasst Wissen über spezifische Interessen der Lernenden sowie zum Einsatz von Interesse fördernden Kontexten auch unter Berücksichtigung der unterschiedlichen Interessen von Mädchen und Jungen.
- *Curriculum, Bildungsstandards und Ziele von Physikunterricht* umfasst Wissen über (curricular vorgegebene) Ziele des Unterrichts und deren Umsetzung.
- *(Digitale) Medien* umfasst Wissen zu angemessenen Einsatzmöglichkeiten dieser.
- *Fachdidaktische Konzepte* umfasst Wissen um Methodenkonzeptionen und Lerntheorien.
- *Aufgaben* umfasst Wissen über die Konzeption und den Einsatz von Aufgaben sowie deren lernpsychologische und didaktische Funktionen.

Zur Entwicklung von Testitems zum fachdidaktischen Wissen schlägt Gramzow (2015, S. 99f.) ein auf vier Facetten und den Themenbereich Mechanik reduziertes Modell des fachdidaktischen Wissens vor. Gleichzeitig wird eine dritte Dimension *Kognitive Anforderungen* mit den Ausprägungen *reproduzieren*, *anwenden* und *analysieren* ergänzt. Mit einem auf Grundlage dieses Itementwicklungsmodells entwickelten Tests konnten Riese, Gramzow und Reinhold (2017) das physikdidaktische Wissen Lehramtsstudierender erfassen, wobei sie Hinweise dafür fanden, dass

[20] Der Begriff der Schülervorstellungen wird für fachlich korrekte sowie inkorrekte „themenbezogene Vorstellungen ebenso wie themenübergreifende Denkrahmen bzw. Schemata“ (Schecker & Duit, 2018, S. 12) der Lernenden zur Physik verwendet und in Einklang mit bisheriger Literatur aus diesem Forschungsfeld nicht geschlechtsneutral formuliert.

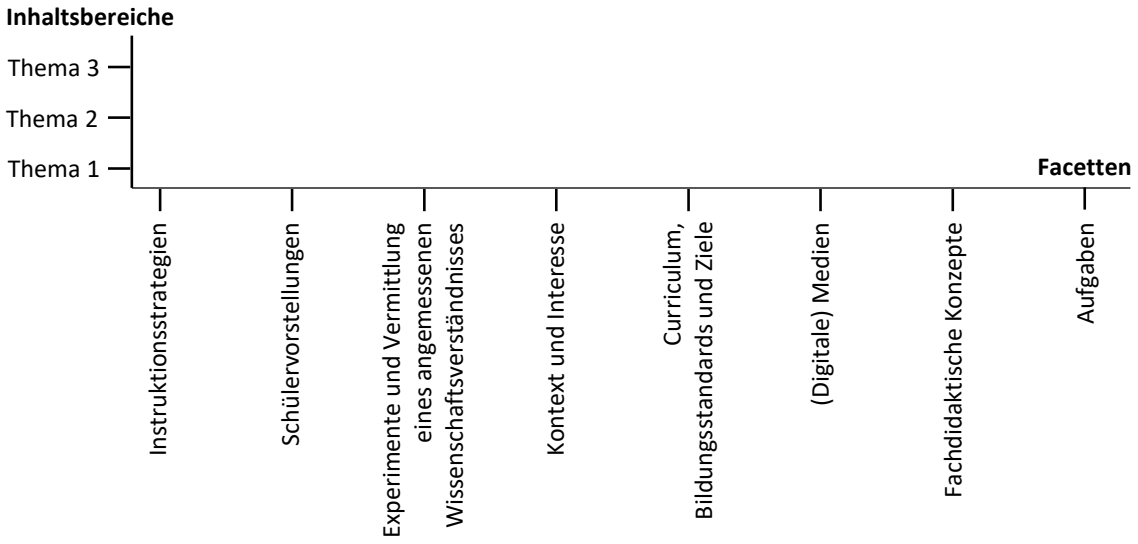


Abbildung 5.4: Modell des fachdidaktischen Wissens nach Gramzow, Riese und Reinhold (2013, S. 21)

die berücksichtigten Facetten Schülervorstellungen, Fachdidaktische Konzepte, Experimente und Instruktionsstrategien empirisch trennbar erfasst werden können - zumindest sofern ein gewisses „Basiswissen“ (Riese, Gramzow & Reinhold, 2017, S. 108) vorliegt.

5.2.5 Empirische Erkenntnisse zur Bedeutung von fachlichem und fachdidaktischem Wissen für die Gestaltung und Qualität von Unterricht

Aufgrund der teilweise verbundenen empirischen Befunde zu Zusammenhängen zwischen dem (erfolgreichen) unterrichtlichen Handeln von Lehrkräften und dem fachlichen und fachdidaktischen Wissen werden diese für beide Wissensbereiche hier zusammengefasst betrachtet.[21]

Während insbesondere für das Fach Mathematik durchaus positive Befunde zu Zusammenhängen u. a. zwischen dem Fachwissen von Lehrkräften und den Fachwissensleistungen von Schüler*innen konstatiert werden können (überblicksartig bei Baumert und Kunter, 2006 oder Lipowsky, 2006), sind Zusammenhänge zwischen dem vorliegenden physikalischen Fachwissen einer Lehrkraft und ihrer Unterrichtsgestaltung bislang nur in geringem Umfang belegt. Grundsätzlich ermöglicht ein umfangreicheres Fachwissen Physiklehrkräften z. B. den flexibleren Umgang

[21] Einen ausführlicheren Überblick über vorliegende Befunde für Mathematik und naturwissenschaftliche Fächer geben Vogelsang und Reinhold (2013) und Gramzow, Riese und Reinhold (2013). Hier werden insbesondere aktuellere Befunde ergänzt.

mit auftretenden (fachlichen) Schwierigkeiten von Schüler*innen im Unterricht und den Einbezug bzw. die Adaption verschiedener Experimente und Beispiele (Hashweh, 1987). Auch Ma (1999) stellt - wiederum für Mathematiklehrkräfte - eine mittelbare Bedeutung des Fachwissens für die Unterrichtsgestaltung fest, da „das tatsächlich im Unterricht verfügbare fachdidaktische Handlungsrepertoire von Lehrkräften weitgehend von der Breite und Tiefe ihres konzeptuellen Fachverständnisses abhängt“ (Baumert & Kunter, 2006, S. 492).

Bezüglich des Unterrichtserfolgs stellt Cauet (2016, S. 188) unter Verwendung der im Projekt *ProwiN* entwickelten Testinstrumente fest, dass die Ausprägung des Fachwissens von Lehrkräften nicht zur Varianzaufklärung in den Fachwissensleistungen der Schüler*innen beiträgt (wie es gemäß der einfachen Wirkungsmodelle schulischen Lernens zu erwarten wäre). Auch Liepertz und Borowski (2019) berichten über ausbleibende Einflüsse des Fachwissens der Lehrkraft auf den Lernerfolg von Schüler*innen. In der PLUS-Studie[22] können zwar keine direkten Zusammenhänge zwischen dem Fachwissen einer Lehrkraft und der Unterrichtsqualität festgestellt werden (betrachtet wurde die Sequenzierung von Lernprozessen sowie die inhaltliche Sachstruktur des Unterrichts); es bestehen aber Hinweise bezüglich eines Einflusses des Fachwissens auf die Leistung der Schüler*innen (mediiert durch die Unterrichtsqualität) (Ohle, Fischer & Kauertz, 2011).

Vogelsang (2014) und Cauet (2016) untersuchen Zusammenhänge zwischen dem fachlichen und fachdidaktischen Wissen (erfasst mit dem *Paderborner Instrument* von Riese (2009) (Vogelsang, 2014) bzw. den im Projekt *ProwiN* entwickelten Instrumenten (Cauet, 2016)) und Merkmalen der Unterrichtsqualität.[23] Während beide Autor*innen keine signifikanten Zusammenhänge zwischen dem fachdidaktischen Wissen und Merkmalen der Unterrichtsqualität und keine signifikanten Einflüsse auf die Fachwissensleistungen der Schüler*innen feststellen können (Cauet, 2016, S. 188; Vogelsang, 2014, S. 479), unterscheiden sich die Ergebnisse in Bezug auf das Fachwissen: Cauet (2016) berichtet signifikante Zusammenhänge zwischen dem Fachwissen von Physiklehrkräften und einer kognitiv aktivierenden Unterrichtsgestaltung; Vogelsang (2014, S. 479) beobachtet hier keinen Zusammenhang, zu anderen Merkmalen der Unterrichtsqualität (Classroom Management und Adaptivität) tendenziell sogar negative Zusammenhänge.

Positive Befunde bezüglich der Bedeutung von fachlichem und fachdidaktischem Wissen können für Physiklehrkräfte für separat erfasste Fähigkeiten wie die Erklärfähigkeit (hier erscheint das Fachwissen mediiert über das fachdidaktische Wissen

[22] *Professionswissen von Lehrkräften, naturwissenschaftlicher Unterricht und Zielerreichung im Übergang von der Primar- zur Sekundarstufe*; untersucht wurde der Unterricht von Lehrkräften im Primarbereich zum physikbezogenen Fachinhalt der Aggregatzustände

[23] Cauet (2016) fokussiert hier die kognitiv aktivierende Unterrichtsgestaltung, die auch von Vogelsang, 2014 einbezogen wird. Sie nutzt dazu eine adaptierte Version des Kategoriensystems von Vogelsang, 2014 (Cauet, 2016, S. 222), sodass die Ergebnisse besonders gut in Beziehung gesetzt werden können.

Einfluss bedeutsam (Kulgemeyer & Riese, 2018)) bzw. in komplexitätsreduzierten, durch Vignetten dargestellte Settings (Riese, 2009, S. 181) festgestellt werden. Sadler et al. (2013) berichten bei vorhandenem Fachwissen der Lehrkräfte einen signifikanten Zusammenhang zwischen der Fähigkeit, verbreitete Fehlvorstellungen zu erkennen, und dem Lernerfolg von Schüler*innen.

Die insgesamt uneindeutige Befundlage bezüglich der Bedeutung des fachlichen und fachdidaktischen Wissens von Lehrkräften für die gelingende Unterrichtsgestaltung zeugt also weiterhin von Forschungsbedarf (Vogelsang & Cauet, 2017). Ausbleibende beobachtbare Zusammenhänge zwischen dem Professionswissen von Lehrkräften und Merkmalen der Unterrichtsqualität könnten ihre Ursache in der Komplexität unterrichtlicher Situationen haben, die in standardisierten Instrumenten nicht abgebildet werden kann (Vogelsang & Rehfeldt, 2021). Auch ein Einfluss der unterschiedlichen eingesetzten Instrumente und ggf. mangelnden Handlungsvalidität kann hier (ebenso wie für das pädagogische Wissen) nicht ausgeschlossen werden (Vogelsang & Reinhold, 2013).

5.2.6 Zur Zusammenhangsstruktur der Bereiche des Professionswissens

Beim Vergleich der in verschiedenen Modellierungen des fachdidaktischen Wissens (bzw. PCK) berücksichtigten Facetten (eine Übersicht bieten z. B. Kirschner et al. (2016)) fällt auf, dass einige Konzeptualisierungen das Fachwissen sowie pädagogisches Wissen und Wissen über Kontexte als Facetten des fachdidaktischen Wissens verstehen (z. B. Hashweh, 2005), während die meisten anderen Autor*innen diese explizit nicht als Teil des fachdidaktischen Wissens ansehen, sondern das Fachwissen sowie i. d. R. pädagogisches und fachdidaktisches Wissen als nicht hierarchisch geordnete Bereiche des Professionswissens ansehen (z. B. Gramzow, 2015, S. 30f.; Tepner et al., 2012). Gleichzeitig weisen auch einige der bereits berichteten Ergebnisse zur Bedeutung des Professionswissens für den (Physik-)Unterricht auf einen engen Zusammenhang zwischen fachlichem und fachdidaktischem Wissen hin (z. B. Kulgemeyer und Riese, 2018; Sadler et al., 2013). Dies eröffnet die Frage nach der Zusammenhangsstruktur von fachdidaktischem, pädagogischem und fachlichem Wissen.

In diesem Sinne schlägt Gess-Newsome (1999) die Unterscheidung eines integrativen und eines transformativen Modells für das PCK vor: während das integrative Modell das fachdidaktische Wissen als Schnittmenge von Fachwissen, pädagogischem Wissen und dem Wissen über Kontexte (auch Organisationswissen, Baumert und Kunter, 2006) beschreibt, wird das fachdidaktische Wissen im Rahmen des transformativen Modells als eine eigenständige Domäne des Professionswissens konzeptualisiert, die auf Grundlage selbiger drei Wissensbereiche entsteht. Neumann, Kind und Harms (2019) beobachteten zwei, sich quasi unabhängig voneinander

entwickelnde, Forschungslinien entsprechend diesen beiden Modellierungen, wobei überwiegend Operationalisierungen gemäß dem transformativen Modell gefunden werden können (Chan & Hume, 2019, S. 43).

Auch empirische Befunde deuten darauf hin, dass Fachwissen, fachdidaktisches und pädagogisches Wissen (angehender) Physiklehrkräfte zwar eng zusammenhängen, aber empirisch trennbare Konstrukte sind und somit eigenständige Dimensionen des Professionswissens darstellen (Neumann, Kind und Harms, 2019; Sorge, Stender und Neumann, 2019[24]; Kirschner et al., 2016; Riese und Reinhold, 2012).

5.3 Professionelle Handlungskompetenz

Vorliegendes Professionswissen gilt als eine Voraussetzung für gelingenden Unterricht (Lipowsky, 2006). Zur empirischen Untersuchung des Professionswissens wird dieses daher üblicherweise als Bestandteil der professionellen Handlungskompentenz operationalisiert (z. B. Rutsch et al., 2018; Cauet, 2016, S. 9ff.; Riese, 2009, S. 25ff.; Baumert und Kunter, 2006). Übereinstimmend wird beschrieben, dass eine kompetente Person in der Lage ist, Probleme, die sich in ihrem professionellen Umfeld ergeben, adäquat und situationsspezifisch zu lösen (Klieme und Leutner, 2006; Frey, 2006; Weinert, 2002b). Professionelle Handlungskompetenz kann demnach als Disposition bzw. Ressource für adäquates Handeln in professionellen Situationen verstanden werden und zeigt sich in diesem (Klieme & Hartig, 2007). Stark rezipiert ist dafür der Kompetenzbegriff nach Weinert (2002b) (z. B. in den Projekten *COACTIV* (Kunter, Klusmann & Baumert, 2009) und *TEDS-M* (König & Blömeke, 2009)), der daher auch dieser Arbeit zugrunde gelegt wird. Weinert (2002b) versteht (professionelle Handlungs-)Kompetenzen als

> „die bei Individuen verfügbaren oder durch sie erlernbaren kognitiven Fähigkeiten und Fertigkeiten, um bestimmte Probleme zu lösen, sowie die damit verbundenen motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten um die Problemlösungen in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können“ (Weinert, 2002b, S. 27f.).

Diese Definition umfasst also neben den für professionelles Handeln notwendigen kognitiven Dispositionen[25] (dem „funktionalen Aspekt“ (Vogelsang, 2014,

[24] Sorge, Kröger, Petersen und Neumann (2019) untersuchen Zusammenhänge zwischen dem Fachwissen, pädagogischen Wissen und dem „topic-specific professional knowledge“ - dem Teil des fachdidaktischen Wissens, das angehende Lehrkräfte im Rahmen ihrer Ausbildung erlernen sollen - anstelle des fachdidaktischen Wissens (Sorge, Kröger, Petersen & Neumann, 2019).

[25] Diese umfassen nach Weinert (2002b) die „fachliche und fachübergreifende Kompetenzen“ (Weinert, 2002b, S. 28), was vor dem Hintergrund seiner eigenen Definition, die neben kognitiven auch affektiv-motivationale Dispositionen berücksichtigt, missverständlich erscheint.

S. 30)) auch affektiv-motivationale Dispositionen[26] (den „intentionalen Aspekt“ (Vogelsang, 2014, S. 30)).[27] Die professionelle Handlungskompetenz ist also Handlungsressource für professionsspezifisches Handeln und kann als kognitive und affektiv-motivationale Bestandteile beinhaltende Disposition für dieses Verstanden werden, das den Angehörigen einer Profession gemein ist (z. B. Vogelsang, 2014, S. 29f.; Weinert, 2002b). Die professionelle Handlungskompetenz ist stets auf einen bestimmten (professionellen) Kontext bezogen, der angibt, *wofür* eine Person kompetent, d. h. in der Lage ist, professionell zu handeln (Klieme & Hartig, 2007, S. 17). Die Definition einer Kompetenz bedarf daher einer präzisen Analyse und Beschreibung der professionellen Handlungssituation, in der sie zum Einsatz kommt (Blömeke, Gustafsson & Shavelson, 2015). Ausdruck findet die professionelle Handlungskompetenz wiederum in beobachtbarem adäquaten Handeln der Person in Situationen, in denen die professionelle Handlungskompetenz „funktional“ (Vogelsang, 2014, S. 29) ist (der sogenannten *Performanz*[28]).

Die Betrachtung der professionellen Handlungskompetenz von Lehrkräften im Rahmen der Untersuchung von Wirkungszusammenhängen wie zwischen dem Professionswissen und der Unterrichtsqualität ist bedeutsam, da sie im Rahmen der Lehrkräfteausbildung als erlernbar gilt (Blömeke, Gustafsson und Shavelson, 2015; Klieme und Hartig, 2007) und dementsprechend angenommen wird, dass sie unterschiedliche Ausprägungen annehmen kann (Blömeke, Gustafsson und Shavelson, 2015; Neumann, 2013).

Aufbauend auf Weinerts (2002b) Verständnis des Kompetenzbegriffs sowie dem Modell der professionellen Handlungskompetenz von Baumert und Kunter (2006), das neben dem Professionswissen individuelle Überzeugungen, motivationale Orientierungen und selbstregulative Fähigkeiten berücksichtigt, präsentieren Riese und Reinhold (2010) ein Strukturmodell der professionellen Handlungskompetenz von Lehrkräften (s. Abbildung 5.5). Sie unterscheiden auf einer ersten Ebene zwischen dem Professionswissen und „damit verbundenen motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten einer Lehrkraft“ (Riese & Reinhold, 2012,

[26] In Anlehnung an das Kompetenzmodell von Blömeke, Gustafsson und Shavelson (2015) werden die von Weinert (2002b) als motivationale, volitionale und soziale Dispositionen ausdifferenzierten nicht-kognitiven Dispositionen (Bereitschaften) als affektiv-motivationale Dispositionen zusammengefasst. Unterscheidungsmerkmal der etablierten begrifflichen Trennung kognitiver Dispositionen (Fähigkeiten) und nicht-kognitiver Dispositionen (Bereitschaften) ist „der meinende und bewertende Charakter von Bereitschaften [...] im Gegensatz zum (angenommenen) ‚rationaleren‘ Wissen“ (Vogelsang, 2014, S. 29).

[27] Bei der Definition und Untersuchung der professionellen Handlungskompetenz von Lehrkräften beschränken sich z. B. Klieme und Leutner (2006) auf die Untersuchung kognitiver (Leistungs-)Dispositionen. Dazu verweisen sie auf den Hinweis Weinerts (2001), kognitive und affektive Dispositionen der professionellen Handlungskompetenz getrennt zu erfassen.

[28] Voss, Kunina-Habenicht, Hoehne und Kunter (2015) kritisieren im Rahmen der Diskussion des pädagogischen Wissens die häufig unzureichende Trennung zwischen beobachtbarer Performanz und professioneller Handlungskompetenz als Disposition für diese im Rahmen von empirischen Studien.

S. 113). In einer zweiten Ebene werden beide Bereiche weiter ausdifferenziert: das Professionswissen in die oben ausgeführten Bestandteile Fachwissen, fachdidaktisches und pädagogisches Wissen und die motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten in *Belief Systems* („Überzeugungen und Vorstellungen zum Fach, zum Lehren und Lernen und zur Schule“ (Riese & Reinhold, 2012, S. 113)) und *Motivationale Orientierungen* (z. B. Selbstwirksamkeitserwartungen (Riese & Reinhold, 2010, S. 171)).

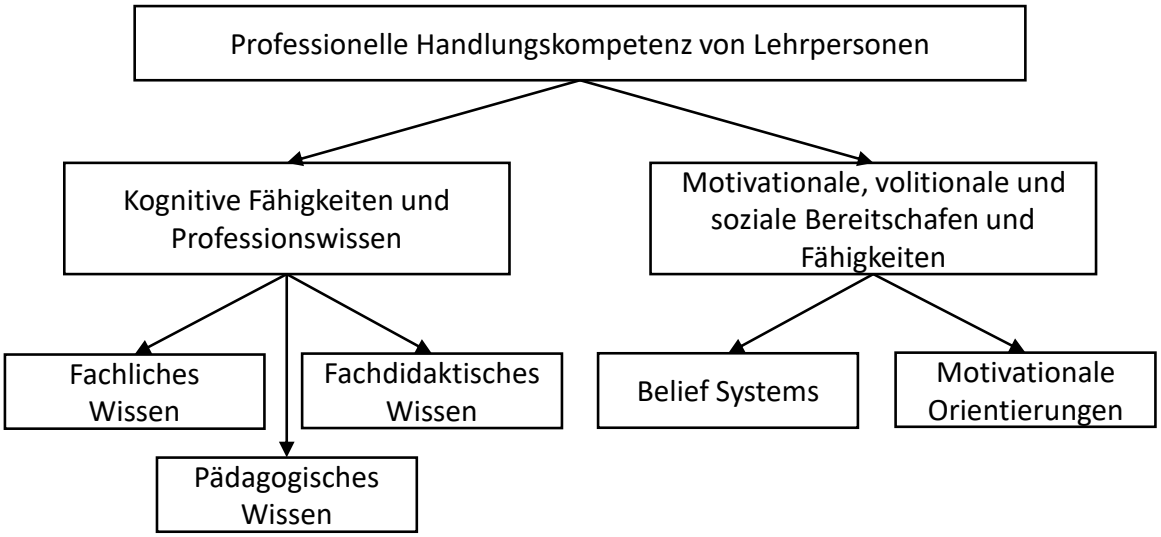


Abbildung 5.5: Modell der professionellen Handlungskompetenz nach Riese und Reinhold (2010, S. 170)

Während der Kompetenzbegriff nach Weinert (2002a) und somit auch das darauf aufbauende Modell der professionellen Handlungskompetenz die für eine Handlung notwendigen Fähigkeiten, Fertigkeiten und Bereitschaften für das Handeln, nicht aber die tatsächliche Handlung umfasst, schlagen Blömeke, Gustafsson und Shavelson (2015) eine Modellierung von Kompetenz als Kontinuum vor, an dessen Ende die Performanz, also beobachtbares Verhalten, steht (s. Abbildung 5.6). Affektiv-motivationale und kognitive Fähigkeiten - die Bestandteile der professionellen Handlungskompetenz im Modell von Riese und Reinhold (2010) - stellen grundlegende Dispositionen für ein Set an situationsspezifischen Fähigkeiten (Wahrnehmen, Interpretieren und Entscheiden) dar, die ihrerseits das beobachtbare Verhalten bestimmen (Blömeke, Gustafsson & Shavelson, 2015).[29]

[29]Vor dem Hintergrund der Kritik von Voss, Kunina-Habenicht, Hoehne und Kunter (2015) erscheint diese Modellierung zunächst kontraproduktiv: Eine präzise Differenzierung zwischen Kompetenz und Performanz wird insofern weiter erschwert, als dass Performanz nun als ein Teil der Kompetenz zu verstehen ist und somit keine klare Trennung möglich ist. Übertragen

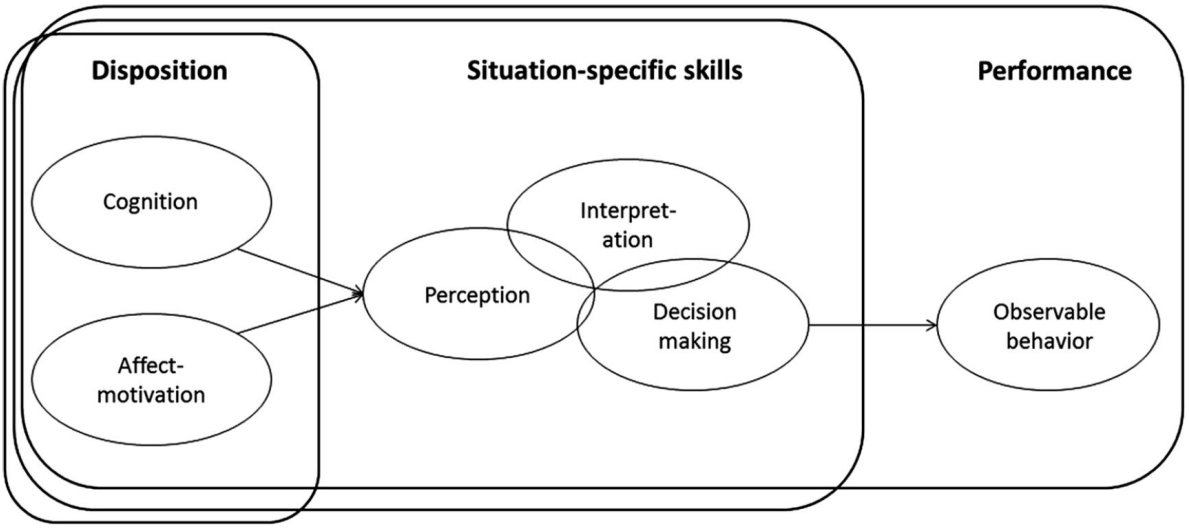


Abbildung 5.6: Modell von Kompetenz als Kontinuum von Blömeke, Gustafsson und Shavelson (2015, S. 7)

5.4 Zum Verhältnis von professionellem Wissen und Handeln

Die verschiedenen Modellierungen der professionellen Handlungskompetenz sprechen dem Professionswissen jeweils eine zentrale Rolle als Bestandteil (Baumert und Kunter, 2006; Riese und Reinhold, 2010) bzw. Disposition (Blömeke, Gustafsson & Shavelson, 2015) zu. Die Modellierung von Kompetenz als Kontinuum, in welcher das Professionswissen die Performanz nur mittelbar und mediiert durch situationsspezifische Fähigkeiten beeinflusst (Blömeke, Gustafsson & Shavelson, 2015), berücksichtigt, dass von vorhandenem Professionswissen nicht automatisch auf adäquates Unterrichtshandeln geschlossen werden kann. Vielmehr muss davon ausgegangen werden, dass Lehrkräfte aus verschiedenen Gründen (z. B. aufgrund fehlender Ressourcen) nicht so unterrichten (können), wie sie es gelernt haben (Diez, 2010; Wahl, 2001). Insbesondere ist davon auszugehen, dass adäquates (unterrichtliches) Handeln, also solches von *Könnern* bzw. Experten, nicht explizit theoriegeleitet, sondern intuitiv und damit zumindest teilweise implizit und situativ ist (Neuweg, 2015a, S. 25ff.).

Zum Zusammenhang zwischen professionellem Wissen und Handeln von Lehrkräften formulieren Vogelsang et al. (2019) die in Abbildung 5.7 dargestellte grobe Professionalisierungskette. Handlungsroutinen werden demnach auf Grundlage universitär erworbenen Wissens ausgebildet und in der Praxis ggf. angepasst und

auf die Modellierung kann die Kritik auf die Forderung einer Differenzierung zwischen Dispositionen (als Voraussetzung für professionelles Handeln und Teil der professionellen Handlungskompetenz) und Performanz (als professionelles Handeln und ebenfalls Teil der professionellen Handlungskompetenz) bezogen werden.

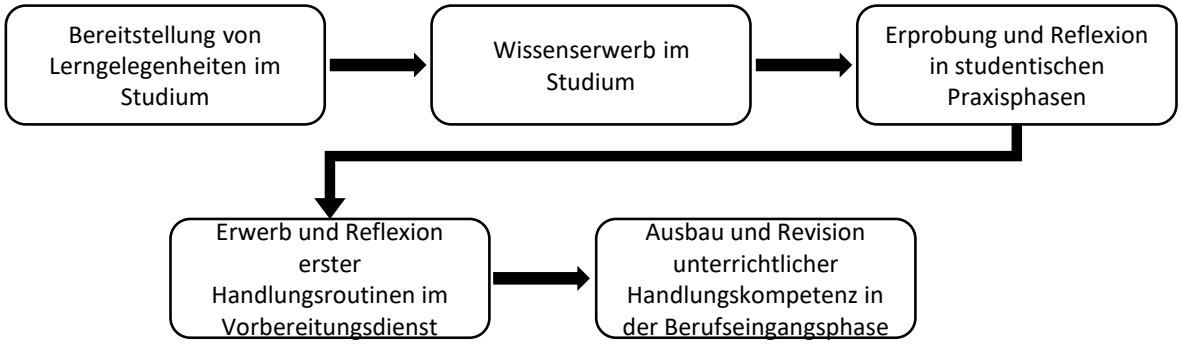


Abbildung 5.7: Vereinfachte Professionalisierungskette von Lehrkräften nach Vogelsang et al. (2019, S. 474)

etabliert. Für die Entwicklung der Handlungsroutinen und darauf aufbauend der Handlungskompetenz auf Grundlage des erworbenen Wissens müssen Transformationsprozesse angenommen werden, da nicht davon ausgegangen werden kann, dass das theoretische Wissen vollständig als Handlungsressource genutzt wird (Vogelsang & Reinhold, 2013).[30]

Ein Modell zur Transformation von Professionswissen in Handeln stellt Stender (2014, S. 33ff.) vor (s. Abbildung 5.8) und validiert es für die Planung von Unterricht (Stender, Brückmann & Neumann, 2015): Universitär erworbenes (deklaratives) Professionswissen muss zunächst ohne konkreten Praxisbezug repliziert und angewendet (*interpretiert*) werden. Anschließend erfolgt die Phase der *Kompilierung*. Hier werden durch die Anwendung des Wissens Handlungsskripte generiert, die wiederum im Rahmen des wiederholten Einsatzes im Unterricht einer *Feinabstimmung* bezüglich ihrer Passung zur jeweiligen Unterrichtssituation unterliegen (Stender, 2014, S. 34). Die Feinabstimmung ermöglicht die Etablierung und fortlaufende Anpassung der Unterrichtsskripte, sie basiert auf der „Reflexion der Entscheidungen, also [der] kritische[n] Auseinandersetzung mit den Konsequenzen der Planung und des Handelns“ (Stender, 2014, S. 39).[31] Zusätzlich zum deklarativen Professionswissen sind motivationale und volitionale Dispositionen der Lehrkräfte zu berücksichtigen sowie ihre selbstregulativen Fähigkeiten, die die Transformation des Professionswissens in adäquates Handeln (*Können*) begünstigen oder ihr im Wege stehen können (Stender, 2014, S. 35).

[30] Der Berücksichtigung solcher Transformationsprozesse ist die verbreitete (implizite) Annahme der unmittelbaren Anwendbarkeit des Professionswissen gegenüberzustellen (Vogelsang & Rehfeldt, 2021).

[31] Unter Bezugnahme auf das Modell der Transformation von Professionswissen (Stender, 2014, S. 33ff.) schreiben Jeschke, Lindmeier und Heinze (2021) der Reflexion eine Mediatorrolle für die Entwicklung professioneller Handlungskompetenz (bezeichnet als *aktionsbezogene Kompetenz*) aus (deklarativem) Professionswissen zu und bestätigen diese empirisch. Auch das Refined Consensus Model sieht eine Weiterentwicklung des personal bzw. enacted PCK

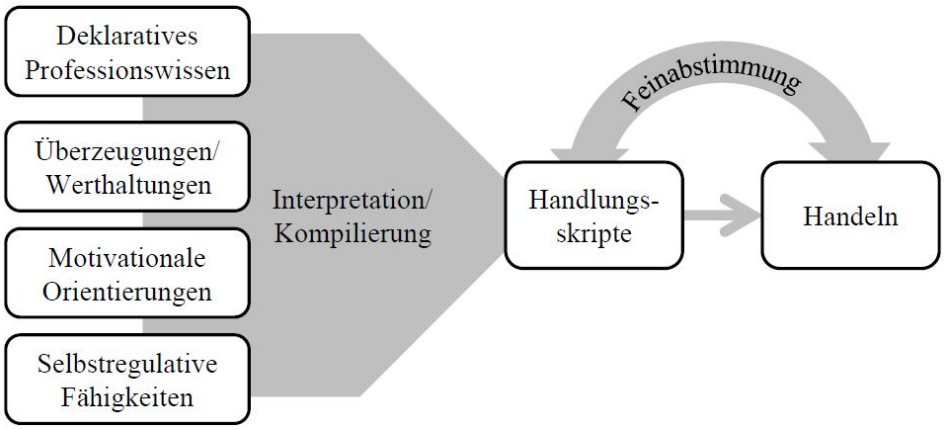


Abbildung 5.8: Modell der Transformation von Professionswissen in Handeln von Stender (2014, S. 33)

Mit dem *Refined Consensus Model* wird eine Modellierung des fachdidaktischen Wissens vorgeschlagen, die sowohl die Zusammenhangsstruktur der verschiedenen Bereiche des Professionswissens als auch darüber hinaus das Zusammenspiel dieser beim Unterrichten berücksichtigt (s. Abbildung 5.9; Carlson et al., 2019). Das fachdidaktische Wissen wird hier als transformiertes Wissen verstanden, es entsteht also auf Grundlage anderer Bereiche des Professionswissens (Tepner & Sumfleth, 2019). Berücksichtigt werden dabei das Fachwissen (*Content Knowledge*) sowie die als fachunabhängig verstandenen Bereiche pädagogisches Wissen (*Pedagogical Knowledge*) und Wissen zum Bewerten (*Assessment Knowledge*), Wissen zur Strukturierung von Lernprozessen (*Curricular Knowledge*) und Wissen zu Voraussetzungen der Lernenden (*Knowledge of Students*) (Carlson et al., 2019). Alle diese Bereiche des Professionswissens (auch das fachdidaktische Wissen) werden im Refined Consensus Model nicht als individuelle Ressource verstanden, sondern als geteilter Wissensbestand einer (Fach-)Community. Carlson et al. (2019) präzisieren dieses kollektive fachdidaktische Wissen daher als *collective PCK* (*cPCK*). Eine Teilmenge davon bildet das *personal PCK* (*pPCK*). Es umfasst alles fachdidaktische Wissen, das einer einzelnen Lehrkraft zur Verfügung steht - gefiltert durch Kontextbedingungen (*Learning Context*), wie curriculare Vorgaben, spezifische Rahmenbedingungen des Unterrichts oder Eigenschaften der Lerngruppe (Carlson et al., 2019). Basierend auf (eigenen) unterrichtlichen Erfahrungen sowie dem Austausch mit anderen Lehrkräften, Wissenschaftler*innen etc. entwickelt sich pPCK dynamisch weiter (Carlson et al., 2019). Damit wird auch der Zusammenhang zwischen dem professionellen Wissen und Handeln modelliert: Werden konkrete Unterrichtssituationen geplant, durchgeführt oder reflektiert, nutzen Lehrkräfte wiederum nur die dafür relevante Teilmenge ihres pPCK - das situative *enacted PCK* (*ePCK*) (Carlson et al., 2019), das im Unterrichtskontext alle Bereiche des

durch Reflexion vor (Sorge, Kröger, Petersen & Neumann, 2019). Auf den Reflexionsbegriff und dessen Bedeutung für das Unterrichten wird in Kapitel 7 näher eingegangen.

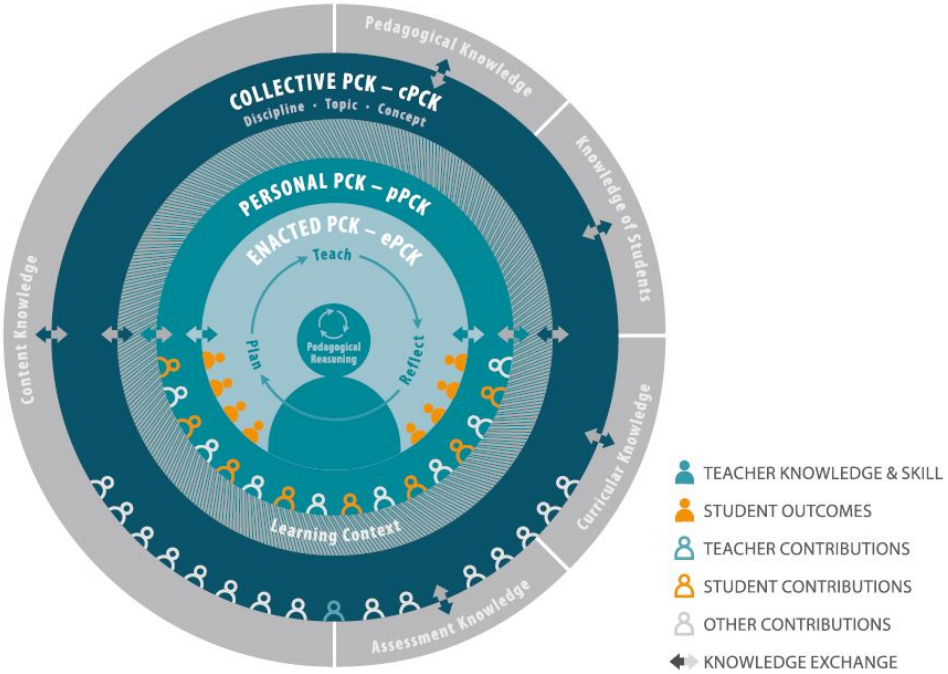


Abbildung 5.9: Refined Consensus Model of PCK von Carlson et al. (2019, S. 84)

Professionswissens integriert (Tepner & Sumfleth, 2019). So können bspw. Schülervorstellungen zu unterschiedlichen physikalischen Inhaltsbereichen Teil des pPCK sein, im Mechanikunterricht werden aber z. B. Schülervorstellungen zur Elektrodynamik nicht berücksichtigt, sie sind also nicht Teil des ePCK. Die in Abbildung 5.9 zwischen allen Ringen des Modells eingezeichneten Pfeile kennzeichnen die Möglichkeit einer Entwicklung des jeweiligen Wissens auf Grundlage des anderen (ohne die zugrundeliegenden Mechanismen näher zu modellieren (Carlson et al., 2019)). So ist z. B. das pPCK Grundlage für situatives Handeln gemäß dem ePCK, während das pPCK durch die Reflexion spezifischer Erfahrungen des Unterrichts bereichert wird (Carlson et al., 2019).

5.5 Zwischenfazit: Professionelles Wissen und Handeln von Physiklehrkräften

In Wirkungsmodellen schulischen Lernens wird das Handeln von Lehrkräften bzw. die adäquate Gestaltung von Unterricht als Grundlage für das Lernen von Schüler*innen verstanden. Auch im Rahmen von ausdifferenzierteren Angebots-Nutzungs-Modellen, nach denen weitere individuelle oder strukturelle Kontextfaktoren (z. B. sozioökonomische Hintergründe von Schüler*innen oder die Unterrichtszeit) als Gelingensbedingungen für Lernerfolge bei Schüler*innen zu berücksichtigen sind, wird das professionelle Handeln (als Ausdruck der professionellen Handlungs-

kompetenz) der Lehrkräfte als bedeutsamer Einflussfaktor auf den Ertrag von Unterricht modelliert (s. Kapitel 5.1). Als zentrale Disposition für professionelles Handeln gilt - neben affektiv-motivationalen Bestandteilen - das Professionswissen (s. Kapitel 5.3). Dieses wird üblicherweise in die drei Bestandteile des fachlichen, fachdidaktischen und pädagogischen Wissens ausdifferenziert (s. Kapitel 5.2), die zwar als eng zusammenhängend, aber (empirisch begründet) eigenständige Bereiche des Professionswissens verstanden werden.

Wie empirische Untersuchungen zeigen, kann vorliegendes Professionswissen nicht automatisch in eine qualitativ hochwertige Unterrichtsgestaltung übersetzt werden; das Verhältnis von Professionswissen und professionellem Handeln erscheint komplexer. Neben einem Einfluss affektiv-motivationaler Dispositionen wird der Reflexion eine zentrale Bedeutung für die Vermittlung zwischen dem Professionswissen und konkreten Handlungssituationen beigemessen (s. Kapitel 5.4), indem sie als Mechanismus zur Generierung von Wissen auf Grundlage der Handlungssituationen verstanden wird (auf die Bedeutung der Relevanz von Reflexion für (angehende) Lehrkräfte wird in Kapitel 7.2 näher eingegangen).

6 Zur Messung von professionellem Wissen und Handeln

In der Literatur finden sich vielfältige methodische Herangehensweisen an die Messung von professionellem Wissen und Handeln. Grundlegend unterschieden werden kann zwischen *objektiven* und *subjektiven* sowie *proximalen* und *distalen* Messungen (Kunter & Klusmann, 2010). Während subjektive Messungen auf Selbsteinschätzungen bzw. Selbstauskünften beruhen, basieren objektive Messungen auf der Erfassung externer Kriterien (z. B. Testergebnisse) (Kunter & Klusmann, 2010). Proximale Messungen basieren auf der direkten Erfassung von Indikatoren für die interessierende Kompetenz mit einem Testverfahren; distale Messungen basieren hingegen auf Indikatoren, die das Konstrukt nicht unmittelbar messen (auch *indirekte Indikatoren* (Kunter & Klusmann, 2010)) wie z. B. Noten, erreichte Abschlüsse oder die Kennziffern zur Berufserfahrung, auf deren Grundlage man sich aber Rückschlüsse über die interessierende Kompetenz erhofft (Rehm und Bölsterli, 2014; Kunter und Klusmann, 2010).[32] Darüber hinaus unterscheiden z. B. Vogelsang (2014, S. 92) und Dollny (2011, S. 42) in diesem Zuge zwischen qualitativen und quantitativen erhobenen Daten. Da diese Unterscheidung nicht spezifisch für die Messung von professionellem Wissen und Handeln ist, wird sie hier nicht weiter thematisiert.

Die bislang getroffenen Unterscheidungen der Herangehensweisen an die Messung von professionellem Wissen und Handeln beziehen sich vorrangig auf die Art und Weise, wie Professionswissen, aber insbesondere auch professionelles Handeln erfasst werden kann. Ergänzend dazu wird im Folgenden ein Konzept zur Klassifikation verschiedener Testtypen nach Miller (1990) vorgestellt, das Tests zum Professionswissen und zur professionellen Handlungskompetenz in Bezug auf die Authentizität bzw. Realitätstreue der Testsituation differenziert (Kapitel 6.2). Anschließend werden Möglichkeiten zur methodischen Umsetzung der Unterschied-

[32] Während Kunter und Klusmann (2010) lediglich für objektive Messungen zwischen distalen und proximalen Indikatoren unterscheiden, erweitern Rehm und Bölsterli (2014) die Ausdifferenzierung auch auf subjektive Messungen und unterscheiden damit bspw. objektiv aus Zeugnissen erfasste Noten von Selbstauskünften über solche Noten. Eine solche Unterscheidung erscheint insbesondere dann sinnvoll, wenn nicht davon ausgegangen werden kann, dass Noten im Rahmen von Selbstauskünften wahrheitsgetreu wiedergegeben werden; entweder, da befragte Personen sich nicht erinnern oder sich einen Vorteil davon versprechen, falsche Angaben zu machen.

lichen Testtypen beschrieben (Kapitel 6.3) und (Video-)Vignettentests als für diese Arbeit besonders relevante Methode näher beschrieben (Kapitel 6.4). Vorab wird auf Gütekriterien empirischer Forschung eingegangen, die als Rahmung für die Evaluation jeglicher Tests dienen (Kapitel 6.1).

6.1 Gütekriterien empirischer Forschung

Die Beurteilung der Qualität eines Tests[33] erfolgt entlang verschiedener Gütekriterien. Klassischerweise wird hierbei zwischen den Hauptgütekriterien *Objektivität*, *Reliabilität* und *Validität* und verschiedenen Nebengütekriterien unterschieden (Bühner, 2011, S. 58; Pospeschill, 2010, S. 17; Lienert und Raatz, 1998, S. 7). Lienert und Raatz (1998, S. 11ff.) benennen hier lediglich die *Normierung* (bzw. *Eichung* (Moosbrugger & Kelava, 2012)), *Vergleichbarkeit*, *Ökonomie* und *Nützlichkeit*. Pospeschill (2010, S. 28ff.) ergänzt die *Skalierung*[34], *Zumutbarkeit*, *Unverfälschbarkeit*, *Fairness* und *Transparenz*. Zusätzlich betont Himme (2007) die Bedeutung der *Generalisierbarkeit* der Ergebnisse. Die aufgeführten Haupt- und Nebengütekriterien werden im Folgenden näher beschrieben. Bedeutsam ist in dieser Arbeit auch die *Akzeptanz* (teilweise auch *Augenscheinvalidität*, s. Kapitel 6.1.3) des entwickelten Tests, da z. B. eine Ablehnung des Instruments die Bearbeitung und die Ergebnisinterpretation beeinflussen kann (Moosbrugger und Kelava, 2012; Pospeschill, 2010, S. 27).[35]

6.1.1 Objektivität

Die *Objektivität* beschreibt das Maß an „Standardisierung des gesamten Testvorganges“ (Fisseni, 2004, S. 47), das für einen Test erreicht wird. Ein Test ist dann objektiv, wenn die Untersuchung - unabhängig von der untersuchenden Person - zu denselben Ergebnissen führt (Lienert & Raatz, 1998, S. 7). Üblicherweise wird die Objektivität entsprechend der durchlaufenen Phasen der Testdurchführung, -auswertung und Ergebnisinterpretation in die drei Aspekte *Durchführungs-*, *Auswertungs-* und *Interpretationsobjektivität* ausdifferenziert (Moosbrugger und Kelava, 2012; Pospeschill, 2010, S. 18; Lienert und Raatz, 1998, S. 8):

[33] Hierbei ist unter *Test* nicht nur ein verwendetes Testinstrument, sondern das vollständige wissenschaftlich fundierte, standardisierte, diagnostische „Verfahren zur Untersuchung eines Persönlichkeitsmerkmals“ (Lienert & Raatz, 1998, S. 1) zu verstehen, es werden also auch die Auswertung und auf Grundlage der Auswertung erlangte Schlussfolgerungen berücksichtigt (Pospeschill, 2010, S. 375).

[34] Das Gütekriterium der *Skalierung* wird von Bühner (2011, S. 67f.) als viertes Hauptgütekriterium erfasst. Das erscheint aber vor dem Hintergrund weiterer Beschreibungen der Gütekriterien (z. B. Pospeschill, 2010, S. 17; Lienert und Raatz, 1998, S. 7) nicht konsensfähig.

[35] Über die Validierung hinaus ist eine positive Einstellung gegenüber dem Test Voraussetzung für einen perspektivisch angestrebten Praxistransfer (vgl. Kollmann, 1998, S. 42).

- Ein hohes Maß an *Durchführungsobjektivität* stellt sicher, dass ein erzieltes Testergebnis unabhängig von der Testleitung ist. Das wird durch eine hohe Standardisierung der Durchführungsbedingungen erreicht (Moosbrugger & Kelava, 2012): Rahmenbedingungen, wie die zeitliche Dauer des Tests, Hilfsmittel, etc. sollten in einem Testmanual festgelegt und personenunabhängig sein, was üblicherweise mit einer Reduktion der Interaktion zwischen Testleitung und Teilnehmenden „auf ein unumgängliches Minimum“ (Lienert & Raatz, 1998, S. 8) einhergeht bzw. die Berücksichtigung und Interpretation der Interaktion erfordert (Fisseni, 2004, S. 48f.).
- *Auswertungsobjektivität* wird erreicht, wenn die Auswertung der gegebenen Antworten unabhängig von der auswertenden Person zu gleichen Testergebnissen führt, d. h. gleichwertigen Antworten gleiche Testwerte zugewiesen werden bzw. bei einer kategorialen Auswertung die Einordnung in Kategorien unabhängig von der auswertenden Person ist (Moosbrugger und Kelava, 2012; Lienert und Raatz, 1998, S. 8). Dafür bedarf es möglichst eindeutiger Auswertungsregeln, die für Multiple-Choice-Formate mit Hilfe einer Musterlösung bzw. Schablonen unkompliziert erstellt werden können; für offene Antwortformate sollten Beispielantworten zur Orientierung ergänzt werden (Moosbrugger und Kelava, 2012; Bühner, 2011, S. 59f.; Lienert und Raatz, 1998, S. 8). Die Berechnung von Übereinstimmungsmaßen (z. B. Cohens κ (Wirtz & Caspar, 2002, S. 56ff.)) liefert Informationen über das Ausmaß der Übereinstimmung der von unterschiedlichen Personen durchgeführten Auswertungen (*Interraterreliabilität*[36]) und ermöglicht so eine Überprüfung der Auswertungsobjektivität (Moosbrugger & Kelava, 2012).
- *Interpretationsobjektivität* liegt vor, wenn verschiedene Personen aus vorliegenden Testergebnissen identische Schlüsse ziehen (z. B. bezüglich der Einordnung der Testergebnisse in eine Vergleichsgruppe) (Bühner, 2011, S. 60). Auch die Interpretation der Ergebnisse kann mit Hilfe von Erläuterungen bzw. Durchführungsmanualen standardisiert werden (Moosbrugger & Kelava, 2012).

6.1.2 Reliabilität

Die *Reliabilität* (auch *Zuverlässigkeit*) eines Tests gibt an, wie präzise ein interessierendes Merkmal gemessen werden kann bzw. mit welchem Fehler der erhaltene Messwert behaftet ist (Moosbrugger und Kelava, 2012; Lienert und Raatz, 1998,

[36] Wirtz und Caspar (2002, S. 33ff.) differenzieren die *Beurteilerübereinstimmung*, die quantifiziert, in welchem Ausmaß verschiedene Personen Testantworten identisch bewerten und die *Beurteiler- bzw. Interraterreliabilität*, die Informationen über das Ausmaß ähnlicher Beurteilungen liefert. Die Beurteilungsübereinstimmung könnte als Spezialfall der Interraterreliabilität formuliert werden - zur Berechnung werden allerdings unterschiedliche Koeffizienten genutzt (Wirtz & Caspar, 2002, S. 40ff.).

S. 9; Fisseni, 2004, S. 49). Voraussetzung für die Bestimmung der Reliabilität ist die Standardisierung „der Messung, also [der] Testsituation, [der] Testdurchführung und [der] Testauswertung“ (Schermelleh-Engel & Werner, 2012, S. 120) - also Objektivität. Ausgehend von der Annahme der klassischen Testtheorie, dass jeder Messwert aus einem wahren Wert und einem Messfehler zusammengesetzt ist, wird das Maß der Reliabilität formal als das Verhältnis der Varianz der wahreren Werte zur Varianz der beobachteten Werte definiert (Schermelleh-Engel und Werner, 2012; Fisseni, 2004, S. 50). Die Schätzung[37] der Reliabilität erfolgt dann unabhängig vom zu messenden Merkmal durch die Bestimmung eines Korrelationsmaßes, des Reliabilitätskoeffizienten $0 \leq r \leq 1$, wobei ein Wert von Eins die vollkommene Abwesenheit von Messfehlern bedeutet (die Wiederholung des Tests unter denselben Rahmenbedingungen führt dann zu identischen Ergebnissen) und ein Wert von Null ausschließlich auf Messfehler zurückzuführen ist (ein wiederholt gleiches Ergebnis wäre also rein zufällig) (Moosbrugger und Kelava, 2012; Schermelleh-Engel und Werner, 2012). Zur Bestimmung des Reliabilitätskoeffizienten liegen vier verbreitete Methoden vor:

- Die *(Test-)Retest-Reliabilität* gibt die Korrelation zwischen den Testergebnissen an, die sich ergeben, wenn derselbe Test wiederholt zu zwei verschiedenen Zeitpunkten von denselben Personen bearbeitet wird (z. B. Lienert und Raatz, 1998, S. 9). Dabei wird vorausgesetzt, dass das Merkmal (der wahre Wert) sowie äußere Einflüsse (die sich in der Varianz der Messfehler niederschlagen) stabil bleiben. Dass diese Annahme nicht tragfähig ist, sondern von systematischen und unsystematischen Veränderungen z. B. durch Lerngelegenheiten oder dem Auftreten von Erinnerungseffekten auszugehen ist, muss bei der Wahl des Retest-Intervalls berücksichtigt werden: Ein längeres Intervall reduziert den Einfluss durch Erinnerungseffekte, lässt aber mehr Spielraum für übrige Übungseffekte und andersherum (Schermelleh-Engel und Werner, 2012; Pospeschill, 2010, S. 22; Fisseni, 2004, S. 53).
- Zur Bestimmung der *Paralleltest-Reliabilität* werden vergleichbare Tests zur Messung desselben Merkmals entwickelt (Paralleltests, die das Merkmal von Interesse mit gleicher Genauigkeit erfassen und in Bezug auf Mittelwerte und Varianzen jeweils vergleichbare Items enthalten) und von der gleichen Personengruppe bearbeitet. Anschließend wird die Korrelation zwischen den Testergebnissen beider Instrumente bestimmt (Schermelleh-Engel und Werner, 2012; Pospeschill, 2010, S. 22; Lienert und Raatz, 1998, S. 9). Auch hier können Einflüsse aufgrund des Zeitintervalls zwischen dem Einsatz beider Instrumente nicht ausgeschlossen werden (Schermelleh-Engel & Werner, 2012).

[37] Eine Berechnung ist nicht möglich, da aus den Messwerten die jeweiligen wahren Werte und Messfehler nicht zu rekonstruieren sind (Schermelleh-Engel & Werner, 2012).

- Die *Split-Half-Reliabilität* kann bei längeren Testinstrumenten Anwendung finden. Um sie zu bestimmen, wird ein Testinstrument von einer Personengruppe bearbeitet und anschließend in zwei in Bezug auf die Merkmalserfassung gleichwertige Hälften geteilt. Beide Testhälften können nun als Paralleltests aufgefasst und ihre Testergebnisse korreliert werden (Bühner, 2011, S. 60f.; Lienert und Raatz, 1998, S. 9f.). Zur Testhalbierung liegen unterschiedliche Verfahren vor, die z. B. statistische Kennwerte der Testitems oder die abwechselnde Zuordnung der Aufgaben zu den Testhälften vorsehen („Methode der Itemzwillinge“ bzw. „Odd-Even-Methode“ (Schermelleh-Engel & Werner, 2012, S. 129)), die die Gleichwertigkeit beider Hälften allerdings nicht per se garantieren, sodass unterschiedliche Teilungen des Instruments zu unterschiedlichen Reliabilitätskoeffizienten führen (Schermelleh-Engel und Werner, 2012; Pospeschill, 2010, S. 23).[38]
- Die *interne Konsistenz* stellt eine Erweiterung der *Split-Half-Reliabilität* dar[39]: Unter der Voraussetzung, dass alle Testitems das interessierende Merkmal erfassen, werden sie jeweils als separate Testelemente aufgefasst und es wird der Korrelationskoeffizient α_{Cronbach} bestimmt, der die jeweiligen Itemvarianzen mit der Varianz des gesamten Testinstruments in Beziehung setzt; es ergibt sich ein umso höherer Koeffizient, je höher die Korrelationen zwischen den einzelnen Items sind (Schermelleh-Engel und Werner, 2012; Pospeschill, 2010, S. 23; Lienert und Raatz, 1998, S. 10).

Die Retest-Reliabilität unterscheidet sich grundlegend von den weiteren Methoden der Reliabilitätsschätzung, da nur hier alle Testitems wiederholt bearbeitet werden. Die übrigen berichteten Methoden entsprechen sogenannten *Item-Sampling-Modellen*, denen die Annahme zugrunde liegt, dass mit einer jeweils unterschiedlichen Auswahl von Items das gleiche Merkmal erfasst werden kann (Fisseni, 2004, S. 52).

Auch wenn die verschiedenen Methoden zur Schätzung der Reliabilität leicht unterschiedliche Werte liefern können (Lienert & Raatz, 1998, S. 10), wird übergreifend ein Reliabilitätskoeffizient von $0,7$ als „ausreichend“ (Himme, 2007, S. 377) bzw. „akzeptabel“ (Pospeschill, 2010, S. 21) angesehen und Werte ab über $0,8$ als „gut“ (Pospeschill, 2010, S. 21). Für die Bewertung der Höhe eines bestimmten Reliabilitätskoeffizienten ist außerdem die Komplexität des erfassten Merkmals zu berücksichtigen, insbesondere kann ein niedrigerer Korrelationskoeffizient zuguns-

[38] Da die Teilung des Testinstruments in zwei Hälften mit einer Reduktion der Items einhergeht, Reliabilitätskoeffizienten allerdings mit zunehmender Anzahl homogener Items steigen, unterschätzt der bestimmte Korrelationskoeffizient die Reliabilität des Testinstruments. Dem kann mit Hilfe der Spearman-Brown-Korrektur entgegengewirkt werden (Schermelleh-Engel und Werner, 2012; Pospeschill, 2010, S. 23).

[39] Einige Autoren fassen auch die *Split-Half-Reliabilität* als Form der internen Konsistenz auf (z. B. Bühner, 2011, S. 60f.; Himme, 2007; Lienert und Raatz, 1998, S. 9f.).

ten einer größeren inhaltlichen Bandbreite in Kauf genommen werden (Schecker, 2014).

6.1.3 Validität

Die Bedeutung der Validität als (Haupt-)Gütekriterium eines Tests ist in der Literatur unumstritten (z. B. Leuders, 2014; Crooks, Kane und Cohen, 1996). Weniger Einigkeit besteht in Bezug auf das zugrundeliegende Konzept bzw. Verständnis der Validität. Hier lassen sich im Rahmen der historischen Entwicklung zwei verschiedene Konzeptualisierungen finden.[40] Im klassischen Verständnis der Validität wird diese als *Eigenschaft des Testinstruments* begriffen: Ein Testinstrument ist dann valide, wenn es dasjenige Konstrukt misst, welches es messen soll und das Testergebnis nicht durch andere Merkmale beeinflusst wird (z. B. Bühner, 2011, S. 61ff.; Lienert und Raatz, 1998, S. 10; Moosbrugger und Kelava, 2012). In der zweiten Hälfte des letzten Jahrhunderts gewann eine zweite Interpretation an Bedeutung (Newton, 2012), die sich heute weitgehend durchgesetzt hat (Wolming & Wikström, 2010): Das Verständnis der Validität als Eigenschaft der Testwertinterpretation unter Berücksichtigung des Verwendungszwecks (z. B. American Educational Research Association, American Psychological Association und National Council on Measurement in Education, 2014; Messick, 1989; Kane, 1992, Schaper, 2014). Soll also bspw. die Reflexionsfähigkeit Studierender erfasst werden, ist nicht der Test selbst als (nicht) valide zu bewerten, sondern es wird die Interpretation der Testwerte evaluiert (z. B. *Die Testwerte sind Indikatoren für die Ausprägung der Reflexionsfähigkeit der Studierenden.*), wobei eine bestimmte Nutzungsabsicht der Ergebnisse zu berücksichtigen ist (z. B. *Auf Grundlage der Testwerte sollen Förderempfehlungen für die Entwicklung der Reflexionsfähigkeit ausgesprochen werden.*). Auch in den *Standards for Educational and Psychological Testing* der American Educational Research Association (AERA), American Psychological Assiciation (APA) und des National Council on Measurement in Education (NCME) wird auf dieses Verständnis rekurriert:

> „Validity refers to the degree to which evidence and theory support the interpretations of test scores for proposed uses of tests. [...] It is the interpretations of test scores for proposed uses that are evaluated, not the test itself. When test scores are interpreted in more than one way [...], each intended interpretation must be validated." (American Educational Research Association, American Psychological Association & National Council on Measurement in Education, 2014, S. 11)

Stobart (2009) fasst diese Entwicklung zusammen als Transformation „from treating validity as measurement of a fixed property towards seeing it as an argument

[40] Rekonstruktionen der historischen Entwicklung des nachfolgend vorgestellten holistischen Validitätskonzepts liefern Shaw und Crisp (2011) und Kane (2013, S. 4ff.).

about the appropriateness of the inferences drawn from the results and the consequences of these inferences. Even a well-constructed test becomes invalid if results are misunderstood or misused“ (Stobart, 2009, S. 162). Mit dem letzten Satz in diesem Zitat stellt Stobart (2009) die übergeordnete Relevanz der Interpretation der Testwerte für die Beurteilung der Validität heraus, wählt aber insofern eine ungünstige Formulierung als er dem *Test(instrument)* im Falle einer unangemessenen Interpretation die Validität abspricht. Solche und weitere inkonsistente Verwendungen des Validitätsbegriffs kritisiert Newton (2012). Während auch die American Educational Research Association, American Psychological Association und National Council on Measurement in Education (2014) den Ausdruck *Validität eines Tests* explizit als falsch bewerten, hält Kane (2009) ihn dann für zulässig, wenn die Interpretation und Nutzung der Testwerte eindeutig festgelegt ist.

Trotz der inzwischen starken Verbreitung des Verständnisses von Validität als Eigenschaft der Testwertinterpretation, gibt es weiterhin Kritik daran (z. B. bewerten Borsboom, Mellenbergh und van Heerden (2004) sie als unpraktikabel; s. auch Kane (2009) und Lissitz und Samuelsen (2007)). Newton (2012) schlägt vor, die konzeptionelle Klärung an zwei zentralen Fragestellungen auszurichten:

1. Wie ist Validität *definiert*? Diese Frage beinhaltet eine Beschreibung des Konzepts und dessen Anwendungsbereiche sowie Grenzen und Verhältnisse zu anderen (angrenzenden) Konzepten.
2. Wie kann Validität *nachgewiesen* werden? Dies beinhaltet insbesondere die Frage nach angemessener (empirischer und/oder analytischer) Evidenz als Argumentationsgrundlage für Validität.

Newton (2012) diskutiert insbesondere die erste Frage vor dem Hintergrund der theoretischen Entwicklung des Validitätsbegriffs ausführlich. Dazu verweist er ebenfalls auf das dargestellte ganzheitliche Verständnis von Validität als Eigenschaft der Testwertinterpretation und schlägt eine Formulierung vor, die auf den Prozess der Interpretation von Testwerten fokussiert und damit die drei Elemente Testinstrument, Testwertinterpretation und -nutzung umfasst:

> „[V]alidity is a property of an assessment-based decision-making procedure. [...] Although assessment ultimately serves the purpose of decision making, the discourse of validity concerns the effectiveness of the mechanism for reaching decisions, not the actual decisions and not the actual interpretations used to make them“ (Newton, 2012, S. 18).

Bezüglich der zweiten von Newton (2012) dargestellten Problematik des Nachweises der Validität kann festgehalten werden, dass weder Art noch Ausmaß der dafür benötigten Evidenz unabhängig vom konkreten Kontext festzulegen sind. Vielmehr müssen die Interpretation der Testwerte und der Verwendungszweck des Tests in die Betrachtung einbezogen werden, denn inwieweit Interpretationen und Schlussfolgerungen unabhängig von der Testkohorte oder vom Kontext sind, kann

nicht pauschal beantwortet werden (Hubley und Zumbo, 2011; Messick, 1995; Kane, 1992). Das Ziel ist daher die Entwicklung einer plausiblen bzw. überzeugenden Argumentation (*Validitätsargumentation*) für die Gültigkeit der Testwertinterpretation im jeweiligen Kontext des Testeinsatzes (Kane, 1992). Hierzu bedarf es einerseits offenbar Evidenz, die die Testwertinterpretation stützt, andererseits kann die Implausibilität anderer Testwertinterpretationen aufgezeigt werden (Kane, 1992). Darüber hinaus sollen auch Faktoren berücksichtigt werden, die die Gültigkeit der Testwertinterpretation möglicherweise gefährden (Kane, 1992). Insgesamt ist also eine holistische, kohärente Argumentation anzustreben, die „ihren Ausdruck aber gerade nicht im additiven Vorliegen einzelner Eigenschaften [findet]“ (Leuders, 2014, S. 12), sondern in der Integration der unterschiedlichen untersuchten Validitätsaspekte (Messick, 1989), was in den Teststandards von AERA, APA und NCME folgendermaßen zusammengefasst ist:

> „A sound validity argument integrates various strands of evidence into a coherent account of the degree to which existing evidence and theory support the intended interpretation of test scores for specific uses.“ (American Educational Research Association, American Psychological Association & National Council on Measurement in Education, 2014, S. 21)

Insbesondere muss der vorläufige Charakter der Validitätsargumentation berücksichtigt werden, denn Validität kann nie abschließend nachgewiesen werden (American Educational Research Association, American Psychological Association & National Council on Measurement in Education, 2014). Einerseits können veränderte Rahmenbedingungen und Bedarfe zu neuen bzw. veränderten Testwertinterpretation und deren Verwendung führen, andererseits kann zusätzliche (empirische) Evidenz zur weiteren Generalisierbarkeit, aber auch Einschränkung der Validität führen (Kane, 2013, S. 37).

Zusammenfassend ist die Evaluation der Validität also als andauernder und dynamischer argumentativer Prozess aufzufassen (Leuders, 2014; Kane, 1992), in welchem dargelegt wird, inwieweit theoretische Begründungen und empirische Evidenz eine bestimmte Interpretation der Testwerte und die zugehörige Nutzung stützen (Messick, 1989).[41]

Da die Entwicklung einer Validitätsargumentation stark von der mit einem Testeinsatz verbundenen Testwertinterpretation und Nutzungsabsicht abhängt, ist die Konzeption von Validität mit großer Offenheit bezüglich des Vorgehens bei der Entwicklung der Validitätsargumentation verbunden. Unklar bleibt bspw., welche Form der Evidenz für eine bestimmte Testwertinterpretation und in wel-

[41] Newton (2012) betont, dass Validität als Kriterium nicht als Maß mit unterschiedlichen Ausprägungen zu verstehen ist (eine Testwertinterpretation ist also nicht *teilweise* oder *bis zu einem bestimmten Grad* valide). Stattdessen kann die Überzeugungskraft der Validitätsargumentation - wenn auch nicht explizit beziffert - graduell diskutiert und eingestuft werden.

chem Umfang notwendig ist (Kane, 2013, S. 8; Stobart, 2009; Crooks, Kane und Cohen, 1996). Nachfolgend werden die von Messick (1989; 1995) ausdifferenzierten Validitätsaspekte und der argumentbasierte Ansatz nach Kane (u. a. 1992; 2012; 2013) sowie Dickmanns (2016, S. 67ff.) Umsetzung davon (entwickelt im Rahmen der Validierung eines computerbasierten Testinstruments zur Messung der auf den Physikunterricht bezogenen Experimentierfähigkeiten von Schüler*innen) vorgestellt. Sie bilden die Grundlage für das Validitätskonzept dieser Arbeit.

Das Validitätskonzept nach Messick

Ausgehend vom ursprünglichen Verständnis der Validität als Eigenschaft des Testinstruments (nicht der Testwertinterpretation) wurden in der Mitte des zwanzigsten Jahrhunderts drei Validitätstypen unterschieden, deren Untersuchung jeweils unterschiedliche Evidenz hervorbringt:

- Die *Kriteriumsvalidität* gibt Auskunft darüber, inwiefern die Testwerte mit weiteren Variablen in Zusammenhang stehen, die Informationen über das gleiche Merkmal liefern (z. B. können Schulnoten und Maße für die Intelligenz von Schüler*innen in Beziehung gesetzt werden (Bühner, 2011, S. 63; Shaw und Crisp, 2011)). Je nach zeitlichem Zusammenhang unterscheidet Bühner (2011, S. 63) z. B. zwischen der Vorhersagevalidität und der retrospektiven Validität.
- Die *Inhaltsvalidität* liefert Aufschluss über die angemessene Abbildung des Merkmals durch die Testitems und wird i. d. R. über Urteile von Expert*innen erfasst (Hartig, Frey und Jude, 2012; Bühner, 2011, S. 61f.; Shaw und Crisp, 2011).
- Die *Konstruktvalidität* umfasst theoriebasierte und empirisch geprüfte Zusammenhänge zwischen den Testergebnissen sowie zugrundeliegenden theoretischen Konstrukten (Schmiemann und Lücken, 2014; Bühner, 2011, S. 63f.; Shaw und Crisp, 2011).

Die Entwicklung hin zu einem holistischen Validitätsbegriff gemäß Messicks (1989; 1989) Verständnis geht mit der Integration von Kriteriums- und Inhaltsvalidität in die als übergeordneter Bezugsrahmen verstandene Konstruktvalidität einher (Messick, 1989; s. auch Schaper, 2014; Hartig, Frey und Jude, 2012). Auch fordert er die Berücksichtigung von Evidenz bezüglich der Relevanz und Nützlichkeit der ermittelten Testwerte sowie die mit der Testwertinterpretation und -nutzung verbundenen Wertvorstellungen und Konsequenzen im Rahmen der Validierung (Messick, 1989). Dabei betont er, dass diese Aspekte nicht als unterscheidbare „type[s] of validity“ (Messick, 1989, S. 6) aufzufassen sind, sondern sich gegenseitig ergänzen, teilweise auch überlappen (er formuliert sogar: „validity judgements are

	Testwertinterpretation	Testnutzung
Basierend auf Evidenz	(Konstrukt-)Validität	(Konstrukt-)Validität + Relevanz/Nützlichkeit des Tests
Basierend auf Konsequenzen	(Konstrukt-)Validität + Wertvorstellungen	(Konstrukt-)Validität + Relevanz/Nützlichkeit des Tests + soziale Konsequenzen

Tabelle 6.1: Differenzierung des holistischen Validitätskonzepts in Argumentationsgrundlagen für den Testwertinterpretationen und deren Nutzung nach Messick (1995, S. 748).

value judgements"; (Messick, 1995, S. 748)) und insofern in der Beurteilung der Validität umfassend berücksichtigt werden (Messick, 1995).[42]

Messick (1989) bezieht sich also auf ein holistisches Verständnis des Validitätsbegriffs, das die Beurteilung der Validität einer Testwertinterpretation und -nutzung unter Einbezug unterschiedlicher Evidenz vorsieht. Auch er verweist demnach auf die Prämisse, dass nicht ein Testinstrument, sondern eine Testwertinterpretation validiert wird und fasst zusammen:

> „[V]alidity is an integrated evaluative judgment of the degree to which empirical evidence and theoretical rationales support the *adequacy* and *appropriateness* of *inferences* and *actions* based on test scores or other modes of assessment." (Messick, 1989, S. 5, Hervorh. i. O.)

Messick (1989) betont außerdem die Bedeutung ethischer Betrachtungen (u. a. bezüglich sozialer Konsequenzen und der Relevanz) eines Testeinsatzes und stellt diese im Rahmen der Validitätsbetrachtung heraus. Dafür differenziert er zwei Facetten von Validität: Er unterscheidet zwischen auf Evidenzen und Konsequenzen basierter Argumentation für die Validität der Testwertintepretation und -nutzung als zu evaluierendem Ergebnis (s. Tabelle 6.1 Messick, 1995, S. 748f.). Er bewertet diese Unterteilung allerdings selbst als unscharf, da die Facetten nicht nur miteinander verbunden, sondern auch nicht trennscharf sind (Messick, 1995). Hubley und Zumbo (2011) führen diese Kritik fort und schlagen vor, in allen Bereichen der Validierung (bei der Definition des Konstrukts, der Konstruktion eines Tests, der Testwertinterpretation und der Evaluation von Konsequenzen) eine Wechselwirkung sowohl zur zugrundeliegenden Theorie als auch zu beeinflussenden Wertvorstellungen anzunehmen und zu berücksichtigen.

[42] Diese Ausdifferenzierung kritisiert Shepard (2016) stark, da die Facetten in der Literatur als trennbar rezipiert wurden und bezeichnet sie daher als einen „Fehler" (Shepard, 2016, S. 296).

Validitätsaspekte

Messick (1995) schlägt als Strukturierungshilfe für die Validierung von Tests die Unterscheidung von sechs unterschiedlichen Validitätsaspekten vor, die bei der Sammlung von Evidenz zur Validierung einer Testwertinterpretation berücksichtigt werden sollen: Der *inhaltliche*, *kognitive*, *strukturelle*, *externe* und *konsequentielle Aspekt der Validität* sowie der Aspekt der *Generalisierbarkeit*. Diese wurden z. B. in die Teststandards von American Educational Research Association, American Psychological Association und National Council on Measurement in Education (2014, S. 13ff.) aufgenommen. Die Orientierung an diesen Aspekten soll eine „systematische, alle Schritte einer Kompetenzmodellierung durchdringende Validitätsanalyse“ (Leuders, 2014, S. 12) ermöglichen, wobei die einzelnen Aspekte aber nicht isoliert voneinander betrachtet werden können, sondern in eine umfassende Beurteilung der Validität einer Testwertinterpretation zusammengefasst werden müssen (Hubley und Zumbo, 2011; Messick, 1995). Sie dienen weiterhin als Kriterien bei der Beurteilung der Validität (Messick, 1995) und werden nachfolgend näher beschrieben.

Der *inhaltliche Aspekt der Konstruktvalidität* (*inhaltliche Validität*[43]) umfasst die Evaluation der Relevanz der Testinhalte sowie deren Repräsentativität (Messick, 1995). Sireci (1998) hält ergänzend eine präzise (theoretische) Beschreibung des Konstrukts für sinnvoll sowie ein angemessenes Vorgehen bei der Aufgabenkonstruktion für notwendig (s. auch American Educational Research Association, American Psychological Association und National Council on Measurement in Education, 2014, S. 14f.). Im Rahmen der theoretischen Erfassung des Konstrukts soll insbesondere verstanden werden, an welchen Charakteristika des Konstrukts sich die Aufgabenschwierigkeit festmachen lässt (Messick, 1995). Zur Test- bzw. Aufgabenkonstruktion bedarf es zunächst der „Setzung eines inhaltlichen Rahmens“ (Leuders, 2014, S. 12), der den inhaltlichen Umfang sowie die angestrebte Auflösung für die Erfassung eines Konstrukts festlegt (Schaper, 2014). Um eine inhaltlich angemessene Aufgabenauswahl zu treffen, muss sichergestellt werden, dass nicht nur alle Aufgaben relevante Inhalte darstellen, sondern die Gesamtheit der Aufgaben das interessierende Konstrukt repräsentativ abbildet. Hierzu können die mit dem Konstrukt verbundenen Anforderungen, z. B. durch systematische Literaturrecherchen zum Konstrukt oder Befragungen von Expert*innen, analysiert werden (Dickmann, 2016, S. 60; Schaper, 2014; Messick, 1995).

Nach Hartig, Frey und Jude (2012) ist zur Evaluation der inhaltlichen Validität bei operational definierten Merkmalen für jedes Item festzustellen, ob „das Item

[43] Hier und im Folgenden werden die verkürzten Übersetzungen der Validitätsaspekte nach Leuders (2014) übernommen, die zwar nicht jeweils das Wort „Aspekt“ beinhalten, aber im Sinne von Messick (1995) nicht als verschiedene Arten von Validität, sondern als zusammenhängende Aspekte des holistischen Validitätsbegriffs zu verstehen sind, die zu einer umfassenden Beurteilung der Validität auf Grundlage unterschiedlicher Evidenz herangezogen werden (Sireci, 1998).

Teil der interessierenden Gesamtheit möglicher Items“ (Hartig, Frey & Jude, 2012, S. 152) ist. Bezogen auf die Gesamtheit der Items ist zu prüfen, ob „die Items eine repräsentative Auswahl der interessierenden Gesamtheit möglicher Items dar[stellen]“ (Hartig, Frey & Jude, 2012, S. 152), wobei diese Frage bezüglich der Repräsentativität nach Hartig, Frey und Jude (2012, S. 152) entlang dreier Kriterien einschätzen lässt:

- Alle relevanten Inhalte sind vorhanden,
- Die Inhalte stehen in einem angemessenen Verhältnis zueinander
- Es sind keine Inhalte enthalten, die sich auf etwas Irrelevantes beziehen

Zur Evaluation des inhaltlichen Validitätsaspekts werden üblicherweise Urteile von Expert*innen herangezogen (Messick, 1995).

Sireci (1998) merkt an, dass die Evaluation der hier diskutierten Aspekte der inhaltlichen Validität zu evaluativen Aussagen bezüglich der Konzeption eines Testinstruments führen und demnach mit dem u. a. Messicks (1995) holistischen Verständnis von Validität in Konflikt stehen. Da die sorgfältige inhaltliche Testentwicklung aber Voraussetzung für eine fundierte Testwertinterpretation ist, hält er die Verwendung des Validitätsbegriffs an dieser Stelle dennoch für praktikabel (Sireci, 1998).

Die *kognitive Validität* bezieht sich auf die Passung zwischen der theoretischen Modellierung des Konstrukts und den tatsächlich ablaufenden kognitiven Prozessen, die während der Bearbeitung des Tests bzw. einzelner Testitems angeregt werden (Leuders, 2014). Damit wird die oben beschriebene Forderung nach Repräsentativität des Testinstruments im Rahmen der inhaltlichen Validität erweitert und gilt auch für „die Prozesse, die zur Lösung kognitiver Testaufgaben benötigt werden“ (Schaper, 2014, S. 29). Um die kognitive Validität eines Tests sicherzustellen, werden qualitative Informationen über die Bearbeitungsprozesse erhoben. Dazu werden beispielsweise Think-Aloud- oder Stimulated-Recall-Interviews geführt, die Einblicke in diese kognitiven Prozesse ermöglichen, indem Teilnehmende ihre Überlegungen verbalisieren (z. B. Sandmann, 2014; Konrad, 2010; Ericsson und Simon, 1993), Eye-Tracking-Analysen genutzt werden (z. B. Frank, 2022; Maddox et al., 2018; Hofmann, 2011) oder das Antwortverhalten über die verschiedenen Teile des Tests auf Kohärenz bzw. auffallende Strukturen analysiert wird (American Educational Research Association, American Psychological Association und National Council on Measurement in Education, 2014, S. 15; Leuders, 2014; Schaper, 2014; Messick, 1995). Weiterhin können Urteile von Expert*innen zu Rate gezogen werden, sofern diese auf „systematische[m] Wissen über kognitive Prozesse bei der Aufgabenbearbeitung“ (Leuders, 2014, S. 14f.) z. B. aus eigener Forschungspraxis aufbauen (Schaper, 2014). Mit der Erhebung der kognitiven Prozesse während der Testbearbeitung können auch Informationen über (konstruktirrelevante) Einflüsse auf den Bearbeitungsprozess erlangt werden, die u. a. abhängig vom Aufgaben-

format auftreten. Wird bspw. ein Test im Multiple-Choice-Format bearbeitet, könnte beobachtet werden, dass die Aufgaben im Ausschlussverfahren bearbeitet werden (d. h. unplausible Antworten schrittweise ausgeschlossen werden) (Leuders, 2014; Schaper, 2014). Eine Beurteilung der kognitiven Validität berücksichtigt dann den „Anteil der intendierten kognitiven Prozesse bei der Bearbeitung der Aufgaben“ (Hadenfeldt & Neumann, 2014, S. 331). Analysen der kognitiven Prozesse bei der Aufgabenbearbeitung können außerdem die Rekonstruktion von Schwierigkeitsmerkmalen ermöglichen (Hartig, Frey & Jude, 2012).

Die *strukturelle Validität* befasst sich mit der Passung des gewählten Messmodells bzw. „scoring models“ (Messick, 1995, S. 745) (der Vorschrift zur Überführung einer Testbearbeitung in Testwerte) zur theoretischen Modellierung des gemessenen Konstrukts. Im Rahmen der klassischen Testtheorie ist hier insbesondere zu prüfen, ob die theoretisch angenommene Dimensionalität eines Konstrukts und die Gewichtung verschiedener Items zueinander auch empirisch bestätigt werden können (Schaper, 2014; Leuders, 2014). Hierzu können „konkurrierende Messmodelle“ (Schaper, 2014, S. 31), die sich u. a. in Bezug auf die angenommene Dimensionalität unterscheiden können, bezüglich ihrer Modellgüte verglichen werden. Wird also bspw. angenommen, dass das Professionswissen sich aus den drei in Kapitel 5 beschriebenen Facetten Fachwissen, fachdidaktisches und pädagogisches Wissen zusammensetzt, sollte sich diese Struktur in Bearbeitungen von Professionswissenstests rekonstruieren lassen. Gleichzeitig sollen auch die theoretischen Annahmen bezüglich der Dimensionalität des Konstrukts nachvollziehbar begründet werden (z. B. anhand notwendiger kognitiver Prozesse für eine Aufgabenbearbeitung) (Leuders, 2014). Unter Umständen ist zu berücksichtigen, dass Items von verschiedenen Stichproben (abhängig von bestimmten Merkmalen, wie z. B. dem Alter oder Geschlecht) systematisch unterschiedlich bearbeitet werden. Im Falle eines solchen *differential item functioning* sind unter Umständen für die verschiedenen Stichproben unterschiedliche Modellierungen adäquat (American Educational Research Association, American Psychological Association & National Council on Measurement in Education, 2014, S. 16).

Ein Aspekt der oben beschriebenen inhaltlichen Validierung ist die Sicherstellung einer repräsentativen Aufgabenauswahl, die eine Verallgemeinerbarkeit der Testergebnisse von den im Instrument enthaltenen Aufgaben auf das abgebildete Konstrukt sicherstellen soll. Der Aspekt der *Generalisierbarkeit* umfasst darüber hinaus die Verallgemeinerbarkeit der Testwertinterpretation über unterschiedliche Messzeitpunkte, „Verfahren der Antwortbewertung“ (Leuders, 2014, S. 18) und teilnehmende Stichproben (Leuders, 2014; Schaper, 2014; Messick, 1995). Anzustreben ist die Entwicklung eines Tests mit Testwertinterpretation, die möglichst umfassend Gültigkeit behält (Leuders, 2014; Schaper, 2014). Um die Generalisierbarkeit der Testwerte in Bezug auf variierende Testaufgaben zum gleichen Merkmal zu erhöhen, können Multimatrixdesigns verwendet werden, die

es (bei einer ausreichend großen Gruppengröße) ermöglichen, eine höhere Zahl von Testaufgaben in gleicher Zeit einzusetzen.

Die *externe Validität* befasst sich mit theoretisch erwarteten und empirisch beobachtbaren Zusammenhängen zwischen Testwerten und verschiedenen externen Kriterien (Leuders, 2014; Messick, 1995).[44] Externe Kriterien können dabei Testwerte anderer Tests sein, die das gleiche Merkmal oder auch abweichende Merkmale messen, aber auch unabhängig von Tests erfasste Merkmale, die die Testwerte vorhersagen sollen (z. B. könnte ein Zusammenhang zwischen Testwerten zum Fachwissen und der Abiturnote hergestellt werden). Zwischen den Testwerten und externen Kriterien wird üblicherweise eine Korrelation berechnet (Schmiemann & Lücken, 2014). In Bezug auf die Art der externen Kriterien wird zwischen der *konvergenten* und *diskriminanten Validität* unterschieden (American Educational Research Association, American Psychological Association und National Council on Measurement in Education, 2014, S. 16f.; Leuders, 2014; Messick, 1995). Konvergente Validität liegt vor, wenn enge Zusammenhänge (d. h. hohe Korrelationen) zwischen den Testwerten und externen Kriterien erwartet und bestätigt werden, da diese ähnliche Konstrukte erfassen; werden niedrige Korrelationen erwartet, wird die diskriminante Validität untersucht (Schaper, 2014; Hartig, Frey und Jude, 2012).[45] Zusätzlich kann die externe Validität in Bezug auf unterschiedliche Messzeitpunkte systematisiert werden: Von *Übereinstimmungsvalidität* (auch *Konkurrenzvalidität*) spricht man, wenn zwei Tests zum (annähernd) selben Zeitpunkt durchgeführt werden, bzw. ein anderes externes Kriterium zeitnah zum Test erhoben wird. Wird das externe Kriterium zeitlich versetzt erhoben, ist also von Interesse, ob Testwerte die Merkmalsausprägung zu einem späteren Zeitpunkt vorhersagen, handelt es sich um *Vorhersagevalidität* (auch *prognostive* oder *prädiktive Validität*) (Schaper, 2014; Schmiemann und Lücken, 2014; Hartig, Frey und Jude, 2012; Cronbach und Meehl, 1955). Schmiemann und Lücken (2014) ergänzen die Systematik um die *retrospektive Validität*, die die Untersuchung von Zusammenhängen zwischen Testwerten und zu einem früheren Zeitpunkt erhobenen Merkmalen beschreibt. Zur Strukturierung der externen Validierung fordern Cronbach und Meehl (1955) die Einbindung des untersuchten Konstrukts bzw. Testinstruments in ein nomologisches Netzwerk. Im Rahmen eines nomologischen Netzwerks werden Zusammenhänge zwischen Testwerten sowie theoretischen Konstrukten und von Testwerten und theoretischen Konstrukten untereinander (theoretisch begründet) postuliert. Sie dienen als Grundlage für den Validierungsprozess unter Nutzung externer Kriterien (Schaper, 2014; Hartig, Frey und Jude, 2012; Cronbach und Meehl, 1955).

[44] Die externe Validität entspricht damit der in Kapitel 6.1.3 beschriebenen *Kriteriumsvalidität* (Schaper, 2014; Hartig, Frey und Jude, 2012; Lienert und Raatz, 1998, S. 11).

[45] Zur gleichzeitigen Erfassung konvergenter und diskriminanter Validität sei auf die Methode der *Multitrait-Multimethod-Analyse* verwiesen, die z. B. von Schermelleh-Engel und Schweizer (2012) beschrieben wird. Dabei werden die jeweiligen Zusammenhänge zwischen verschiedenen, mehrfach auf unterschiedliche Weise erhobenen, Konstrukten betrachtet.

Die *konsequentielle Validität* befasst sich mit der Frage nach den kurz- und langfristig auftretenden Konsequenzen, die sich aus einer Testwertinterpretation und der Nutzung dieser ergeben. Hierbei sollen einerseits intendierte Konsequenzen bzw. Ziele des Testeinsatzes berücksichtigt werden, die als positive Effekte angestrebt werden (z. B. eine Verbesserung der Fördermöglichkeiten für Schüler*innen auf Grundlage einer individuellen Diagnose (Leuders, 2014)). Andererseits soll geprüft werden, ob nicht anvisierte Konsequenzen (i. d. R. negative Auswirkungen bzw. Nebenwirkungen, die z. B. mit einer unfairen Bewertung einer Bearbeitung einhergehen) ausbleiben (Schaper, 2014; Hubley und Zumbo, 2011; Messick, 1995). Dafür ist insbesondere die Passung der theoretischen Modellierung des interessierenden Konstrukts zur intendierten Nutzungsabsicht zu überprüfen (Schaper, 2014). Gefährdend für die Validität ist u. a. eine zu weit gefasste Nutzungsabsicht zum Test („straw man fallacy“ (Kane, 2013, S. 18)). Sollen Ergebnisse eines Tests zu diagnostischen Zwecken herangezogen werden, gilt die *Instruktionssensitivität* eines Tests als Voraussetzung für die konsequentielle Validität (Berger, Kulgemeyer & Lensing, 2019): Wenn angeleitete Lernprozesse auf Grundlage von Testwerten angepasst oder ausgewählt werden, muss sichergestellt sein, dass sich erfolgreich genutzte Lerngelegenheiten in den Testwerten widerspiegeln (Naumann et al., 2019).

Das Konzept der konsequentiellen Validität ist umstritten, wobei die Kritik darin besteht, dass eine Prüfung der konsequentiellen Validität praktisch kaum umsetzbar sei und auch „eine missbräuchliche Verwendung niemals ausgeschlossen werden [könne]“ (Schaper, 2014, S. 41). Hier stellt sich die Frage nach einer (Mit-)Verantwortung der Testentwickler*innen für die angemessene Interpretation und Nutzung von Testwerten. Reckase (1998) beschreibt die Unmöglichkeit, die Konsequenzen eines Tests schon im Rahmen der Entwicklung umfassend zu evaluieren, da man hier auf antizipierte Konsequenzen beschränkt sei. Da die Evaluation der Validität als „fortwährende[r] Prozess“ (Schaper, 2014, S. 26) auch neue Evidenz einbeziehen muss, muss die Untersuchung der konsequentiellen Validität vor dem regulären Testeinsatz als vorläufig gelten. Die Beurteilung der Validität muss unter Berücksichtigung ggf. auch nach der Entwicklungsphase auftretender, nicht antizipierter Konsequenzen überprüft werden. So sind Testentwickler*innen in der Verantwortung, Konsequenzen zu antizipieren, die sich aus der Testwertinterpretation ergeben sowie die der Testentwicklung zugrundeliegenden Theorien transparent darzustellen und z. B. keine irreführenden Titel zu verwenden (Reckase, 1998). Tatsächlich auftretende (unintendierte) Konsequenzen werden vor allem Anwender*innen in der Praxis auffallen (Kane, 2001), sodass die Evaluation eines Tests auch über die Entwicklungsphase hinaus bedeutsam

bleibt. Im Falle einer inadäquaten Anwendung eines Tests sieht Shepard (1997) explizit die Testanwender*innen in der Verantwortung:[46]

> „When users appropriate tests for purposes not sanctioned and studied by the test developers, users become responsible for conducting the needed validity investigation." (Shepard, 1997, S. 8)

Zusammenfassend soll eine Validierung also evidenzbasiert und begründet Antwort auf die Frage geben, inwiefern eine vorgeschlagene Interpretation und Nutzung von Testwerten angebracht ist (Messick, 1989). Dabei ist zwar Evidenz vorzubringen, die die Testwertinterpretation stützt, gleichzeitig soll die Validität gefährdende Evidenz berücksichtigt werden. Als zentrale Bedrohungen für die Validität benennt Messick (1989; 1995) die *Unterrepräsentation des Konstrukts* von Interesse sowie *konstruktirrelevante Varianz*. Konstruktirrelevante Varianz beschreibt die (unintendierte) Beeinflussung der Schwierigkeit des Testinstruments und damit der Testwerte und ihrer Varianz durch externe Faktoren, wie bspw. die Lesefähigkeit bei der Bearbeitung von Aufgaben zum physikalischen Fachwissen (Messick, 1989; 1995). Die so erzeugte Varianz kann „nicht im Hinblick auf das interessierende Konstrukt interpretiert werden" (Dickmann, 2016, S. 58). Bildet ein Test das interessierende Konstrukt nicht umfassend (z. B. bezüglich der Dimensionalität des Konstrukts) ab, liegt eine *Unterrepräsentation des Konstrukts* vor (Messick, 1995). Dies wäre bspw. der Fall, wenn ein Test zum physikdidaktischen Wissen kein Wissen zur Facette Schülervorstellungen einfordert.

Die von Messick (1995) vorgeschlagene Berücksichtigung der sechs Validitätsaspekte für eine umfassende Beurteilung der Validität, erscheint aufgrund des damit verbundenen hohen (Zeit-)Aufwands i. d. R. nicht praktikabel (Wolming & Wikström, 2010). Auch Messick (1995) erkennt diese Schwierigkeit und empfiehlt, einzelne Validitätsaspekte abhängig von der konkreten Testwertinterpretation und -nutzung zu priorisieren. Zur Auswahl der relevanten Aspekte verweist er auf argumentbasierte Ansätze der Validierung, z. B. den nachfolgend vorgestellt Ansatz von Kane (1992; 2001; 2013).

Der Argument-Based-Approach von Kane

Auch Kanes (u. a. 1992; 2013) Verständnis des Validitätsbegriffs ist ein ganzheitliches. Demnach gibt Validität Aufschluss darüber, bis zu welchem Grad „empirische und theoretische Evidenz für Angemessenheit von Interpretationen und Handlungen, die auf diagnostischen Messungen beruhen, vorliegt" (Schaper, 2014, S. 25). Validiert werden also Interpretationen von Testwerten und deren Nutzung (z. B. Kane, 2009; 1992). Dazu schlägt Kane (u. a. 2013; 1992) einen *argumentbasierten*

[46] Treten allerdings systematisch unangemessene Testwertinterpretation auf, deren Ursache z. B. in der Darstellung des Testinstruments für die Anwender*innen auffindbar ist, ist dies in der Beurteilung der Validität zu berücksichtigen (Newton, 2012).

Ansatz der Validierung (*argument-based approach to validatation*) vor, den er als „realistic and pragmatic framework for evaluating the interpretations and uses of test scores“ (Kane, 2009, S. 49) versteht.

Grundlegend für diesen Ansatz ist die Entwicklung zweier Argumente bzw. Argumentationen:

- Eines *Interpretations-Nutzungs-Arguments (INA)*[47], das die intendierte Interpretation der Testwerte sowie darin inkludierte Annahmen und deren Nutzung transparent macht sowie
- eine *Validitätsargumentation*, die mit Hilfe geeigneter Evidenz das INA plausibel machen soll (Kane, 2013, S. 9ff.; Kane, 1992) [48]

Das INA setzt damit - und das ist das zentrale Charakteristikum des argumentbasierten Ansatzes - den Rahmen für die nachfolgende Validitätsargumentation und bildet die Grundlage für die Auswahl der für die Validierung notwendigen Evidenz (Kane, 2012; Kane, 2009). Dabei können verschiedene Aspekte, wie die Generalisierbarkeit der Ergebnisse über die Rahmenbedingungen hinweg oder Entscheidungen, die auf Grundlage der Testwerte gezogen werden, berücksichtigt werden (Kane, 1992). Dass die Validitätsargumentation jeweils abgestimmt auf das INA entwickelt wird, führt einerseits dazu, dass keine „standardisierte[n] Routineverfahren“ (Schmiemann & Lücken, 2014, S. 108) angewendet werden können; andererseits ermöglicht diese Flexibilität die Anwendbarkeit des argumentbasierten Ansatzes der Validierung für beliebige INAs (Kane, 1992). Um explizit die Konsequenzen einer Testwertinterpretation und -nutzung zu berücksichtigen, schlägt Kane (2001) die Unterscheidung zwischen einem deskriptiven Teil des INA vor, das die *Interpretation* der Testwerte z. B. als Maß für ein bestimmtes theoretisches Konstrukt fokussiert und einem präskriptiven Teil des INA, das die Konsequenzen fokussiert, die sich aus der *Nutzung* der Testwerte (z. B. im Rahmen von Zulassungsentscheidungen) ergeben (s. auch Meinhardt, Rabe und Krey (2018)). Das Validitätsargument ermöglicht eine globale Einschätzung der Plausibilität das INA. Dazu wird das INA entlang dreier Kriterien evaluiert, die Dickmann (2016, S. 63) in Form von Leitfragen zusammenfasst.

- *Clarity of the Argument*: „Sind die Aussagen und Annahmen im INA sowie die Maßnahmen zu deren Stützung detailliert genug beschrieben, um die Argumentation für die Testwertinterpretationen nachvollziehen zu können?“ (Dickmann, 2016, S. 63)
- *Coherence of the Argument*: „Ist das INA kohärent, d. h. ermöglicht es eine schlüssige Argumentation?“ (Dickmann, 2016, S. 63). Hierbei wird auch die Passung der Schlussfolgerungen zur Theorie berücksichtigt (Kane, 1992).

[47] In früheren Arbeiten wird es Interpretationsargument („interpretive argument“ (Kane, 1992, S. 527)) genannt, die Nutzung der Testwerte ist aber auch hier schon mit gemeint (Kane, 2013, S. 2).

[48] Übersetzungen in Anlehnung an Dickmann (2016, S. 63)

- *Plausibility of Assumptions*: „Sind die Aussagen und Annahmen im INA a priori plausibel oder durch Evidenz gestützt?“ (Dickmann, 2016, S. 63). Dabei ist sowohl die Annahmen stützende als auch widersprechende Evidenz zu berücksichtigen (Kane, 1992). Es besteht insbesondere die Gefahr, „‚hidden‘ assumptions“ (Kane, 1992, S. 528) zu übersehen und daher nicht mit Evidenz zu stützen.

Kane (1992) befürwortet eine Triangulation der vorgebrachten Evidenz und das Widerlegen antizipierter Gegenargumente zum INA. Dabei gilt: „[t]he most questionable assumption deserve the most attention“ (Kane, 1992, S. 530). Als besonders starke Argumente für die Plausibilität eines INA bewertet Kane (2012) die Evaluation alternativer INAs, um die Überlegenheit des gewählten INA zu stützen. Insgesamt kann ein Testwertinterpretation und -nutzung insoweit als valide angesehen werden, wie das INA klar und kohärent ist und die damit verbundenen Annahmen und Aussagen plausibel sind - ein Beweis für die Validität einer Testwertinterpretation und -nutzung ist nicht möglich (Kane, 2013, S. 13). Vielmehr beschreibt auch Kane (1992) INAs als vorläufig, ggf. mit Bedarf für Anpassungen abhängig von neuer Evidenz bzw. spezifischen Rahmenbedingungen.

Im Rahmen des Validierungsprozesses unterscheidet Kane (z. B. 2013; 2009) zwischen der *Entwicklungsstufe* (*development stage* (Kane, 2013, S. 16)) und der i. d. R. zeitlich nachgelagerten *Bewertungsstufe* (*appraisal stage* (Kane, 2013, S. 17)). In der Entwicklungsstufe steht die Entwicklung eines Tests mit dazugehörigem INA im Fokus, für dessen Gültigkeit Evidenz gesammelt wird (Kane, 2013, S. 17). Da die Sammlung von Evidenz hier vorrangig der Rechtfertigung des INA gilt, ist ein confirmation bias nicht auszuschließen (Kane, 2012). In der Bewertungsstufe wird das INA möglichst objektiv (idealerweise von neutralen Forschenden) kritisch geprüft (Kane, 2013, S. 17).

Kane (2013, S. 18f.) beschreibt die fünf nachfolgenden verbreiteten Fehlschlüsse bei der Validierung gemäß argumentbasiertem Ansatz:

- Der *begging-the-question fallacy* beschreibt, dass Elemente des INA ohne ausreichende Prüfung als abgesichert angenommen werden (z. B. die Generalisierbarkeit der Ergebnisse über einen bestimmten Kontext hinaus). Typischerweise geht dies mit mangelnder Passung zwischen formulierter INA und (intendierter) Testwertinterpretation und -nutzung einher.
- Im Falle eines *straw man fallacy* liegt eine für die Testwertinterpretation und -nutzung unnötig umfassende INA vor.
- Unter *reification fallacy* ist die fehlerhafte Interpretation einer regelmäßig auftretenden Beobachtung als Ausprägung eines bestimmten Merkmals zu verstehen, obwohl für die kausale Beziehung zwischen Merkmal und Beobachtung nicht ausreichend Evidenz vorliegt.

- Unter der *gilding the lily fallacy* versteht er das Überdecken fehlender Evidenz für einen Teil des INA, indem zusätzliche (redundante) Evidenz für einen anderen Teil des INA geliefert wird.
- Der *fallacy of statistical surrogation* beschreibt einen Fehlschluss auf Grundlage fälschlicherweise angewendeter statistischer Konzepte (z. B. darf eine Schlussfolgerung über Kausalität nicht auf Grundlage einer berechneten Korrelation gezogen werden (Dickmann, 2016, S. 65)).

Im Rahmen der Validierung eines computerbasierten Testverfahrens zur Messung der experimentellen Kompetenz von Schüler*innen am Ende der Sekundarstufe I formuliert Dickmann (2016, S. 67) das INA entlang der Übersetzungsschritte, die erforderlich sind, um einer Zieldomäne Testwerte zuweisen zu können (s. Abbildung 6.1; analog auch von Bartels (2018, S. 20) genutzt): Zunächst wird (1) die Zieldomäne (das zu untersuchende Konstrukt) in Aufgaben (Testitems) übersetzt.[49] Im Rahmen des Testeinsatzes wird durch die Bearbeitung des Tests (2) eine Übersetzung der Aufgaben in Verhalten von Teilnehmenden vorgenommen. Abschließend wird (3) das beobachtbare Verhalten der Teilnehmenden in Testwerte übersetzt.

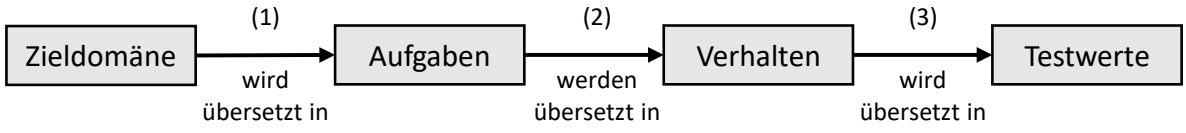


Abbildung 6.1: Darstellung der erforderlichen Übersetzungsschritte, um Ausprägungen einer Zieldomäne Testwerte zuzuweisen nach Dickmann (2016, S. 67)

Diese Übersetzungsschritte sind Grundlage für das INA, das den Bezugsrahmen für die Evaluation bildet, indem zentrale übergeordnete Anforderungen an die einzelnen Schritte formuliert werden:

- Anforderung an Teil 1 des INA: „Die Aufgaben umfassen relevante und repräsentative Inhalte und Anforderungen aus der Zieldomäne.“ (Dickmann, 2016, S. 76)
- Anforderung an Teil 2 des INA: „Die beobachtete Performanz [während der Testbearbeitung] passt zur beabsichtigten Performanz“ (Dickmann, 2016, S. 76)

[49] Der Übergang von theoretischer Beschreibung des Konstrukts und der zugehörigen Operationalisierung ist dabei nicht trennscharf (Dickmann, 2016, S. 64).

- Anforderung an Teil 3 des INA: „Die beobachtete Performanz wird in geeigneter Art und Weise in Testwerte überführt.“ (Dickmann, 2016, S. 76)

Die verschiedenen Teile des INA werden zur Prüfung durch Annahmen präzisiert, sodass es in belegbare Abschnitte untergliedert wird (z. B. wird für den Abgleich von beobachteter und intendierter Performanz angenommen, dass die Testbearbeitung „überwiegend experimentbezogene Überlegungen“ (Dickmann, 2016, S. 71) hervorruft). Kann die Annahme theoretisch begründet oder evidenzbasiert als erfüllt bewertet werden, stützt dies die Anforderung des INA. Im Rahmen der Evaluation des INA können unterschiedliche Aspekte der Validität angesprochen werden (wird also z. B. geprüft, ob während der Testbearbeitung tatsächliche experimentbezogene Überlegungen dominieren, kann dies der kognitiven Validität nach Messick (1995) zugeordnet werden; analog für die Erklärfähigkeit auch Bartels (2018, S. 20)). Liegt die Validitätsargumentation vollständig vor (d. h. mit Einschätzungen für alle Annahmen, die das INA stützen sollen), kann bewertet werden, inwiefern die folgende Testwertinterpretation als gültig angenommen werden kann:

> „[D]ie Testwerte können valide als Ausdruck der Zieldomäne aufgefasst werden“ (Bartels, 2018, S. 20)

Bartels (2018, S. 85ff.) adaptiert dieses Vorgehen zur Ausgestaltung des INA und der Validitätsargumentation für die Evaluation eines performanznahen Tests zur Messung der Erklärfähigkeit von Physik-Lehramtsstudierenden. Crooks, Kane und Cohen (1996) schlagen zur Validierung ein Kettenmodell vor, dessen Kettenglieder z. B. die Darstellung der Aufgaben (*Administration*) oder die Bestimmung von Testwerten (*Scoring*) beinhalten; sie führen aber auch Entscheidungen (*Decision*) und daraus resultierende Folgen als Kettenglieder auf, die sich Schüler*innen nach der Testdurchführung ergeben (*Impact*). Damit tragen sie auch dem Aspekt der konsequentiellen Validität Rechnung (Crooks, Kane & Cohen, 1996). Entscheidend ist die Identifikation des schwächsten Kettenglieds, da dieses der wichtigste limitierende Faktor für die Beurteilung der Validität darstellt (Crooks, Kane & Cohen, 1996).

Zusammenfassend soll eine Validierung also evidenzbasiert und begründet Antwort auf die Frage geben, inwiefern eine vorgeschlagene Interpretation und Nutzung von Testwerten angebracht ist (Messick, 1989). Dabei ist zwar Evidenz vorzubringen, die die Testwertinterpretation stützt, gleichzeitig soll die Validität gefährdende Evidenz berücksichtigt werden. Als Bezugsrahmen empfiehlt Kane (z. B. 1992; 2013) die Formulierung eines INA, das in der Validitätsargumentation mit theoretischer und empirischer Evidenz gestützt wird. Zur Auswahl dieser Evidenz können die Validitätsaspekte nach Messick (1995) herangezogen werden (Kinnear et al., 2021). Die erforderliche Evidenz ist dabei so zu wählen, dass das INA ausreichend gestützt ist, um die mit der Testentwicklung intendierte Testwertinterpretation und -nutzung als angemessen einzustufen.

Zum Begriff der Augenscheinvalidität

Verschiedene Autoren stellen die *Augenscheinvalidität* als eine Form der Validität heraus, die sich daran bemisst, inwiefern es Lai*innen[50] glaubhaft erscheint, dass ein Test einem Validitätsanspruch gerecht wird, bspw. also tatsächlich zur Erfassung eines Merkmals geeignet ist (z. B. Moosbrugger und Kelava, 2012; Pospeschill, 2010, S. 27). Die Augenscheinvalidität gibt damit Aufschluss über die „offensichtliche Akzeptanz eines Tests durch Laien“ (Pospeschill, 2010, S. 27); Kersting (2008) operationalisiert die Augenscheinvalidität als eine Skala zur Erfassung der Akzeptanz von Tests zur Eignungsdiagnostik. Problematisch ist allerdings, dass die Augenscheinvalidität Einfluss auf die Bearbeitung von Tests nehmen kann (z. B. im Rahmen eines veränderten Antwortverhaltens vor dem Hintergrund sozialer Erwünschtheit (Pospeschill, 2010, S. 27)). Auch ist möglich, dass ein Test zwar für Lai*innen angemessen erscheint (also augenscheinvalide ist), die Testwertinterpretation aber trotzdem nicht als valide bewertet werden kann (z. B. wenn durch komplexe Satzkonstruktionen beim Lesen von Aufgabenstellungen konstruktirrelevante Schwierigkeiten vorliegen). Für eine klare begriffliche Trennung zwischen dem holistischen Validitätskonzept (in welchem die Augenscheinvalidität bzw. Akzeptanz eines Tests durchaus von Relevanz sein kann, wenn bspw. die Annahme der mit einem Testeinsatz verbundenen Konsequenzen betrachtet wird) und der Augenscheinvalidität (im Sinne der Akzeptanz eines Tests und damit verbundener Konsequenzen), wird im Folgenden ausschließlich der Begriff der *Akzeptanz* verwendet und in Kapitel 6.1.5 ausführlich beschrieben.

6.1.4 Nebengütekriterien

Zusätzlich zu den drei Hauptgütekriterien *Objektivität*, *Reliabilität* und *Validität* werden bei der Beurteilung der Qualität eines Testinstruments verschiedene Nebengütekriterien angelegt, die nachfolgend kurz beschrieben und bezüglich ihres Verhältnisses zu den Hauptgütekriterien eingeordnet werden sollen. Die Evaluation der Nebengütekriterien erfolgt grundsätzlich argumentativ, für sie sind i. d. R. keine „zahlenmäßigen Kennwert[e]“ (Lienert & Raatz, 1998, S. 12) bestimmbar.

Das Gütekriterium der *Normierung* (auch *Eichung*) fordert die Entwicklung eines Bezugssystems zur eindeutigen Einordnung der individuellen Testergebnisse in eine Referenzstichprobe (Moosbrugger & Kelava, 2012) - sofern die Einordnung nicht in ein vorliegendes Kriterien- bzw. Kategoriensystem erfolgt (Bühner, 2011, S. 72).

[50] Diese Bezeichnung ist ebenso unscharf und abhängig vom Forschungsinteresse wie die der Expert*innen, denen die Lai*innen klassischerweise gegenüber gestellt werden (Bogner, Littig & Menz, 2014). In Bezug auf den Einsatz eines Testinstruments beschreibt der Begriff die Personen, denen das Wissen um die theoretischen Hintergründe der Testkonstruktion fehlt; nach Döring und Bortz (2016, S. 445) ist hier insbesondere die Einschätzung der getesteten Personen gemeint.

Im eindimensionalen Fall entspricht das Bezugssystem der Ergebnisverteilung der Referenzstichprobe (Lienert & Raatz, 1998, S. 11). Zusätzlich können so „Ergebnisse verschiedener Tests vergleichbar gemacht“ (Himme, 2007, S. 375f.) werden. Die Norm kann dabei auf verschiedene Weise dargestellt werden: Unmittelbar durch die erreichten Testwerte der Referenzstichprobe oder durch transformierte Werte, wie z. B. Prozenträngе, die den Prozentsatz der Personen angeben, die besser bzw. schlechter abgeschnitten haben und als besonders anschaulich gelten (Moosbrugger und Kelava, 2012; Fisseni, 2004, S. 74ff.). Pospeschill (2010, S. 28f.) empfiehlt pauschal Stichprobengrößen ab 300 Personen, um eine repräsentative Referenzstichprobe er erhalten.

Nach Lienert und Raatz (1998, S. 12) ist ein Test dann *vergleichbar*, wenn zu diesem entweder Paralleltestformen oder „validitätsähnliche Tests“ (Lienert & Raatz, 1998, S. 12) vorliegen, die ein ähnliches oder sogar dasselbe Merkmal erfassen. Durch den zusätzlichen Einsatz paralleler bzw. validitätsähnlicher Tests werden eine intraindividuelle Reliabilitäts- bzw. Validitätskontrolle ermöglicht (Lienert & Raatz, 1998, S. 12) sowie der Einsatz der Paralleltestformen für wiederholte Testungen zur Reduktion von Erinnerungseffekten bzw. Vermeidung von Abschreiben in Gruppensituationen (Bühner, 2011, S. 72).

Die *Ökonomie* eines Tests bemisst sich an der benötigten Durchführungszeit, dem Materialverbrauch, der Komplexität in der Handhabung sowie der Möglichkeit, den Test als Gruppentest durchzuführen und „schnell und bequem auszuwerten“ (Lienert & Raatz, 1998, S. 12). Ein Test ist also umso ökonomischer, je weniger Ressourcen seine Durchführung und Auswertung beansprucht. Insbesondere der zeitliche Aufwand eines Tests kann durch eine (computergestützte) Automatisierung von Einsatz und Auswertung reduziert werden (Moosbrugger & Kelava, 2012)

Das Kriterium der *Nützlichkeit* soll sicherstellen, dass für einen Test ein „praktisches Bedürfnis“ (Lienert & Raatz, 1998, S. 13) besteht. Dieses liegt bspw. dann vor, wenn mit dem (neuen) Testinstrument eine Weiterentwicklung der Theorie abgebildet wird (Bühner, 2011, S. 73). Die Nützlichkeit ist dann nicht erfüllt, wenn andere Tests vorliegen, die dasselbe Merkmale mit gleicher Güte erfassen (Bühner, 2011, S. 73; Lienert und Raatz, 1998, S. 13).

Die *Skalierung* betrifft die Bildung von Testwerten im Rahmen der Testauswertung. Zu prüfen ist, inwiefern „gültige Verrechnungsvorschrift[en]“ (Bühner, 2011, S. 67) angewendet werden, d. h. inwiefern (inter- und intraindividuell) unterschiedliche Merkmalsausprägungen sowie deren Relation zueinander - und bei vorliegendem Intervallskalenniveau auch die Stärke der Unterschiede - angemessen abgebildet werden (Moosbrugger und Kelava, 2012; Pospeschill, 2010, S. 29).

Die Berücksichtigung des Kriteriums der *Zumutbarkeit* soll sicherstellen, dass die *getestete* Person „in zeitlicher, psychologischer sowie körperlicher Hinsicht nicht über Gebühr belastet“ (Moosbrugger & Kelava, 2012, S. 22) wird, wobei

die Belastung abhängig vom Kontext des Tests durchaus variabel ist (Pospeschill, 2010, S. 30).

Ein Test soll außerdem möglichst *unverfälschbar* sein. Es soll vermieden werden, dass die bearbeitende Person ihr Testergebnis (bewusst) z. B. durch sozial erwünschtes Antwortverhalten verzerren kann, was insbesondere dann möglich wird, wenn das Messprinzip leicht zu durchschauen ist (Moosbrugger und Kelava, 2012; Pospeschill, 2010, S. 31). Da die bewusste Beeinflussung der Testergebnisse nie ausgeschlossen werden kann, sollten Möglichkeiten, diese zu erkennen und ggf. aufzudecken, bei der Interpretation der Ergebnisse berücksichtigt werden (Bühner, 2011, S. 73f.).

Das Gütekriterium der *Fairness* soll sicherstellen, dass die Bestimmung der Testwerte nicht durch irrelevante Kriterien beeinflusst wird und auch im Rahmen der aus den Testwerten gezogenen Schlussfolgerungen Personen(-gruppen) keinerlei Diskriminierung erfahren (American Educational Research Association, American Psychological Association und National Council on Measurement in Education, 2014, S. 49f.; Moosbrugger und Kelava, 2012; Bühner, 2011, S. 73). Zur Prüfung der Testfairness können statistische Analysen bezüglich der (Un-)Gleichbehandlung verschiedener Gruppen durchgeführt werden (Bühner, 2011, S. 73).

Pospeschill (2010, S. 32) beschreibt als wesentliche Merkmale der *Transparenz*, dass die Instruktionen zu einem Test für die teilnehmenden Personen verständlich sind (oder z. B. durch vorgeschaltete Beispielaufgaben gemacht werden) und die Teilnehmenden im Nachhinein über die Intention des Tests sowie die Ergebnisse unterrichtet werden.

Die Frage nach der *Generalisierbarkeit* von Testergebnissen bezieht sich auf die Verallgemeinerbarkeit der Ergebnisse über die konkreten Testaufgaben sowie Erhebungsbedingungen wie teilnehmende Personen oder Zeitpunkt und Ort der Befragung hinaus, um Aussagen über die Domäne oder das Konstrukt zu treffen, das mit dem Test erfasst werden soll (Himme, 2007; Messick, 1995).[51]

6.1.5 Akzeptanz

Die Akzeptanz von Tests wird zuweilen der Augenscheinvalidität zugeordnet (s. Kapitel 6.1.3), teilweise aber auch als eigenständiges Nebengütekriterium erfasst (Kersting, 2008). Insbesondere für den freiwilligen Einsatz eines Tests erscheint es zentral, dass Teilnehmende ihm gegenüber positiv eingestellt sind und ihn annehmen. Für die in dieser Arbeit entwickelte und evaluierte Lernumgebung aus Testinstrument und Assessment-Feedback (s. Teil II) erscheint die Betrachtung der Akzeptanz als Voraussetzung für die ernsthafte Testbearbeitung notwendig. Zu-

[51] Dieses Verständnis ist deckungsgleich mit dem in Kapitel 6.1.3 beschriebenen Validitätsaspekt der Generalisierbarkeit nach Messick (1995).

sätzlich ist davon auszugehen, dass die Wahrnehmung der im Assessment-Feedback dargestellten Testergebnisse abhängig von der Akzeptanz der Lernumgebung ist und diese so Einfluss auf die Konsequenzen nimmt, die sich für Teilnehmende aus der Testbearbeitung ergeben. Insbesondere ist die Lernumgebung so konzipiert, dass sie von (angehenden) Lehrkräften nach Bedarf und wiederholt eingesetzt werden kann, um eine fundierte Selbsteinschätzung der Reflexionsfähigkeit zu ermöglichen. Gerade für die selbstständige Nutzung der Lernumgebung erscheint die positive Einstellung gegenüber der Lernumgebung notwendig. Daher wird nachfolgend das dieser Arbeit zugrundeliegende Verständnis des Akzeptanzbegriffs beschrieben und dieser modelliert.

Der Akzeptanzbegriff

Der Akzeptanzbegriff findet in vielfältigen Kontexten Anwendung im alltäglichen Sprachgebrauch (Lucke, 1995, S. 31ff.), wird dort allerdings ebenso wie im Bereich der Forschung nicht eindeutig verwendet (Schmalz, 2019, S. 21). Verschiedene Definitionen des Akzeptanzbegriffs stimmen darin überein, dass Akzeptanz „eine Form der nicht negativen Bewertung eines Phänomens“ (Schmalz, 2019, S. 22) ist. So definiert z. B. auch Simon (2001) den Akzeptanzbegriff „im Widerspruch zum Begriff der Ablehnung“ (Simon, 2001, S. 87): Akzeptanz „bezeichnet die positive Annahmeentscheidung einer Innovation durch die Anwender“ (Simon, 2001, S. 87). Diese eher allgemeine Definition wird auch in anderen Arbeiten übernommen (Taherdoost, 2018) und hier zugrunde gelegt. Allgemeines Ziel der Untersuchung der Akzeptanz von Produkten (i. d. R. von Innovationen) ist, „Gründe für die Annahme bzw. Ablehnung von Innovationen [auf Seite der Anwender*innen] zu erschließen“ (Klosa, 2016, S. 74).

Verbreitet ist die Berücksichtigung von Einstellungen gegenüber einem Akzeptanzobjekt (z. B. dem in Bezug auf Akzeptanz untersuchten Produkt, z. B. einem Test) und dem tatsächlichen Verhalten (z. B. der Nutzung des Tests) als Formen bzw. Komponenten der Akzeptanz (Olbrecht, 2010, S. 19). Kollmann (1998) legt seinem dynamischen Akzeptanzmodell ein Verständnis der Akzeptanz als „freiwillige Übernahme eines Objekts“ (Kollmann, 1998, S. 64) (z. B. einer innovativen Technologie) zugrunde und unterscheidet zwischen einer *Einstellungsebene* (auch „Einstellungsakzeptanz“; Müller-Böling und Müller, 1986, S. 25) und einer *Verhaltensebene* (auch „Verhaltensakzeptanz“; Müller-Böling und Müller, 1986, S. 26) der Akzeptanz:

1. Die Akzeptanz auf *Einstellungsebene* berücksichtigt eine affektive und kognitive Komponente in Bezug auf die Innovation bzw. das Akzeptanzobjekt und ist als (positive) Erwartungshaltung in Bezug auf die Nutzung der Innovation nicht direkt beobachtbar (Kollmann, 1998, S. 67; Simon, 2001, S. 87).

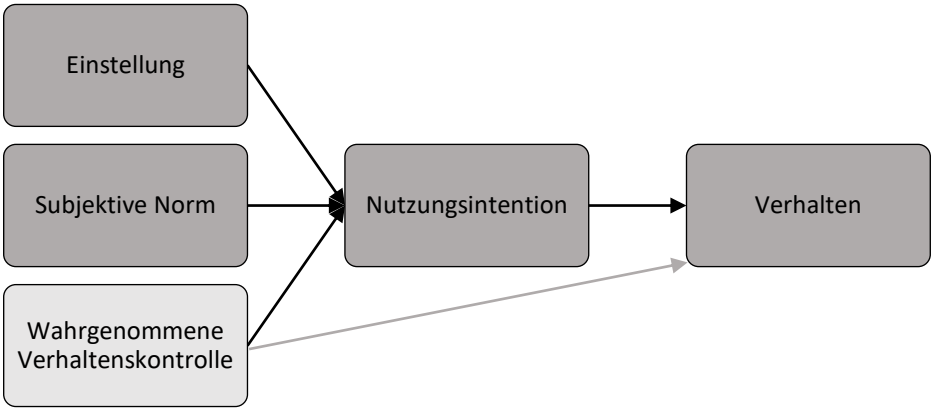


Abbildung 6.2: Darstellung der TRA (nur dunkle Blöcke) sowie leicht vereinfachte Darstellung der TPB nach Ajzen (2002 (überarbeitet 2006), S. 1) unter Auslassung des Faktors der tatsächlichen Verhaltenskontrolle

2. Auf der *Verhaltensebene* erfolgt zunächst die Übernahme einer Innovation (die Nutzung wird bspw. durch den Kauf möglich gemacht), die positive Einstellung wird in eine Handlung umgesetzt (Kollmann, 1998, S. 67; Olbrecht, 2010, S. 20) und kann anschließend und für eine „freiwillige, konkrete und aufgabenbezogene bzw. problemorientierte Nutzung“ (Kollmann, 1998, S. 68) übernommen werden.

Unterschiede sind je nach Fachdisziplin bzw. Anwendungskontext in Bezug auf die Frage auszumachen, ob Akzeptanz graduell oder nur dichotom ausgeprägt sein kann (Simon, 2001, S. 87). Steht die Nutzung einer Innovation (z. B. zur Bereitstellung und zum Austausch von Lehrmaterialien) im Fokus, hält Simon (2001, S. 88) eine graduelle Einschätzung für angemessen, die die Intensität der Auseinandersetzung widerspiegelt.

Akzeptanzmodelle

Die *Theorie des überlegten Handelns* (*Theory of reasoned action, TRA*) von Fishbein und Ajzen (1975) dient dazu, das Verhältnis von Verhalten und Einstellung von Personen zu erklären (Olbrecht, 2010, S. 21). Unmittelbarer Prädiktor für ein Verhalten ist gemäß der TRA die Verhaltensintention, also die „individuelle, subjektive Wahrscheinlichkeit, ein Verhalten in einer mehr oder weniger genau terminierten Zukunft auszuführen“ (Olbrecht, 2010, S. 21). Diese wird wiederum durch zwei Faktoren bestimmt (s. Abbildung 6.2): (1) Die persönliche Einstellung gegenüber einem Verhalten sowie (2) die subjektiven Norm, also der Einfluss der vermuteten Erwartungshaltung relevanter Bezugspersonen (Olbrecht, 2010, S. 21f.; Graf, 2007, S. 34f.).

Die Gültigkeit der TRA ist insofern eingeschränkt als nur solches Verhalten vorhergesagt werden kann, das „vollständig der willentlichen Kontrolle“ (Graf, 2007, S. 35) unterliegt. Um dieser Einschränkung zu begegnen, erweitern Ajzen (1985) und Ajzen und Madden (1986) die TRA zur *Theorie des geplanten Verhaltens* (*Theory of Planned Behaviour, TPB*). Die TPB umfasst die TRA sowie eine Erweiterung des Modells um die Komponente der *wahrgenommenen Verhaltenskontrolle*, welche sowohl Einfluss auf die Verhaltensintention nimmt, als auch – im Gegensatz zu den anderen Variablen – das Verhalten direkt beeinflusst (s. Abbildung 6.2) (Taherdoost, 2018; Olbrecht, 2010, S. 23; Graf, 2007). Je positiver die Einstellung gegenüber einem Verhalten und die subjektive Norm ist bzw. je größer die wahrgenommene Verhaltenskontrolle ist, desto stärker ist die Verhaltensintention (Ajzen, 2002 (überarbeitet 2006)). Im Vergleich zur TRA ermöglicht die TPB empirisch eine deutlich bessere Vorhersage von Verhalten (Graf, 2007).[52]

Die TRA ist auch Grundlage für weitere Akzeptanzmodelle (Taherdoost, 2018), von denen hier noch auf das *Technologieakzeptanzmodell* (*Technology Acceptance Model, TAM*) eingegangen werden soll (s. Abbildung 6.3). Dieses wurde von Davis (1989) für die Vorhersage der Akzeptanz von Technologien bzw. technischen Systemen entwickelt. Ebenso wie in der TRA und TPB ist die Verhaltensintention unmittelbare Ursache für das tatsächliche Verhalten, also die Nutzung des Systems. Die Intention wird durch die beiden Faktoren (1) wahrgenommener Nutzen (die Akzeptanz von Anwendungen steigt, wenn Anwendende einen Nutzen für ihren Beruf sehen (Davis, 1989)) und (2) wahrgenommene Bedienbarkeit (die Akzeptanz von Anwendungen steigt, wenn die Nutzung möglichst wenig Aufwand auf Seiten der Anwendenden erfordert (Davis, 1989)) bestimmt (Übersetzung nach Olbrecht (2010, S. 24)). Die Gültigkeit des TAM ist in verschiedenen Bereichen empirisch abgesichert (z. B. zur Nutzung von E-Learning von Olbrecht (2010), zur Nutzung eines Wikis für Gruppenarbeiten von Cheng (2019) und zur Nutzung von Kalkulationstabellen von Mathieson (1991)), wobei eine Literaturanalyse zeigt, dass die Vorhersagegültigkeit bei einer objektiven Messung des tatsächlichen Verhaltens im Gegensatz zu einer Selbstauskunft insgesamt abnimmt (Turner et al., 2010).

In zahlreichen Studien wurden verschiedene Erweiterungen des TAM vorgeschlagen und evaluiert (Marangunić & Granić, 2015). Venkatesh und Davis (2000) schlagen das TAM2, eine Erweiterung des TAM vor, in welchem auch soziale Einflüsse und kognitive Einflüsse auf den wahrgenommenen Nutzen eines Systems berücksichtigt werden.[53] Die sozialen Einflussfaktoren setzen sich zusammen aus der *subjektiven Norm*, der *Freiwilligkeit* der Übernahme eines Systems und dem *Ansehen*, was mit einer Systemnutzung verbunden ist (Venkatesh & Davis, 2000).

[52] Hinweise und Beispiele zur Konstruktion von Fragebögen, um Verhalten nach der TPB zu untersuchen, stellen Ajzen (2002 (überarbeitet 2006)) und Graf (2007) zur Verfügung.

[53] Darüber hinaus gehende Erweiterungen (z. B. das TAM3) werden in dieser Arbeit aufgrund der stetig zunehmenden Ausdifferenzierung und dem entsprechend zunehmenden Erhebungsaufwand für einen Teilaspekt dieser Arbeit nicht berücksichtigt.

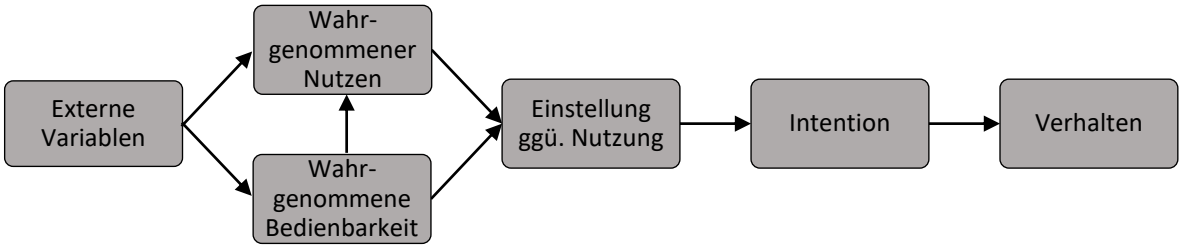


Abbildung 6.3: Darstellung des TAM (Davis, 1989) nach Turner et al. (2010, S. 472).

Die *subjektive Norm* beschreibt – analog zur TRA - den Einfluss der vermuteten Erwartungshaltung relevanter Bezugspersonen bezüglich eines Verhaltens auf den wahrgenommenen Nutzen eines Systems (Venkatesh & Davis, 2000). Damit verbunden ist der Faktor der *Freiwilligkeit*: Bei der verpflichtenden Nutzung eines Systems ist der Einfluss der subjektiven Norm (im Gegensatz zur freiwilligen Nutzung) von signifikanter Bedeutung (Venkatesh & Davis, 2000). Der Faktor *Ansehen* verweist auf den Zusammenhang zwischen der Nutzung des Systems und dem Status im Sozialgefüge (z. B. am Arbeitsplatz (Venkatesh & Davis, 2000)). Auf kognitiver Ebene wird die Beurteilung der *Relevanz* bzw. *Anwendbarkeit* eines Systems für den eigenen Beruf berücksichtigt (Venkatesh & Davis, 2000) sowie die Frage danach, wie gut das System die Aufgaben bzw. Anforderungen erfüllt, für die es eingesetzt wird (*Ergebnisqualität* (Venkatesh & Davis, 2000)). Der Faktor *Ergebnisnachweisbarkeit* beschreibt, inwiefern die Ergebnisse einer Systemnutzung greifbar sind bzw. von Anwendenden beschrieben werden können (Venkatesh & Davis, 2000). Die Gültigkeit des TAM2 wurde von Venkatesh und Davis (2000) für vier Feldstudien mit unterschiedlichen Rahmenbedingungen nachgewiesen.

6.2 Klassifikation unterschiedlicher Testtypen nach Miller (1990)

Für den Kontext der medizinischen Ausbildung schlägt Miller (1990) ein Rahmenmodell zur Ausdifferenzierung von Tests zur Messung des pädagogischen Wissens und Handelns von angehenden Ärzt*innen vor und unterscheidet zwischen vier Typen von Tests: *Wissenstests*, *Kompetenztests*, *Performanztests* und *Handlungstests*. Diese Klassifizierung wurde für die Testentwicklung in der Domäne der Lehrkräftebildung z. B. von Schröder et al. (2020) (zur Messung der Fähigkeit zur Unterrichtsplanung), Kempin, Kulgemeyer und Schecker (2018) (zur Messung der Fähigkeit, Physikunterricht zu reflektieren), Bartels (2018) und Kulgemeyer (2016) (im Rahmen der Entwicklung von Testinstrumenten zur Messung der

Fähigkeit, physikalische Sachverhalte zu erklären) und Rath (2017) (im Bereich der Diagnose von Schülervorstellungen) adaptiert. Im Sinne von Millers (1990) Klassifikation können die vier Testtypen folgendermaßen charakterisiert werden (s. auch Kulgemeyer et al., 2023; Bartels, 2018, S. 23):

1. Mit einem *Wissenstest* (test for knowledge) wird deklaratives Wissen geprüft. Wissenstests liegen üblicherweise in Form von schriftlichen Tests vor.
2. Mit *Kompetenztests* (tests for competence) wird geprüft, ob eine Person ihr (deklaratives) Wissen zur Lösung professioneller Anforderungssituationen anwenden kann. Auch dafür werden üblicherweise schriftliche Tests eingesetzt, in denen die Lösung beschrieben wird.
3. In einem *Performanztest* (test for performance) wird das professionelle Handeln einer getesteten Person in standardisierten Situationen erfasst. Die Bewertung erfolgt entweder im Rahmen der unmittelbaren Beobachtung des Handelns oder im Nachhinein auf Grundlage einer Audio- oder Videoaufnahme. Bartels (2018, S. 30) identifiziert drei Charakteristika von Performanztests:
   - Relevanz: Das Setting eines Performanztests bildet Situationen von zentraler Bedeutung für den Beruf nach, in denen professionelles Handeln erforderlich ist.
   - Authentizität: Performanztests zeichnen sich durch ein hohes Maß an Authentizität bzw. Realitätsbezug zur tatsächlichen beruflichen Situation aus.
   - Interaktivität: Im Rahmen von Performanztests findet „reale Interaktion zwischen den Proband[*inn]en und der Testumgebung statt“ (Bartels, 2018, S. 30).
4. Ein *Handlungstest* (test for action) dient ebenfalls der Erfassung des professionellen Handelns, allerdings in realen beruflichen Anforderungssituationen, z. B. während des Unterrichtens. Auch hier erfolgt die Bewertung auf Grundlage von Beobachtungen in der Situation bzw. Audio- oder Videoaufnahmen.

Während die ersten beiden Typen sich also explizit auf das Wiedergeben von Wissen und dessen Anwendung auf konkrete Situationen beziehen, gehen die unteren beiden darüber hinaus. Performanz- und Handlungstests erfordern tatsächliches (professionelles) Handeln in „Laborsituationen“ (Vogelsang, 2014, S. 91) bzw. realen Anforderungssituationen.

Für die Übersetzung und Übertragung von Millers (1990) Klassifikation auf den Kontext der Lehrkräftebildungsforschung bedarf es einer begrifflichen Klärung des verwendeten Kompetenz- und Performanzbegriffs (Bartels, 2018, S. 23f.). Während der Begriff der Performanz in der Lehrkräftebildungsforschung häufig für die in einem Test erbrachte Leistung bzw. das in einem Test beobachtbare Verhalten

von Teilnehmenden gebraucht wird, meint Miller (1990) damit die Ausführung adäquater Handlungen in einer Anforderungssituation („*show how* [to] do it“ (Miller, 1990, S. 63, Hervorh. i. O.)). Andernfalls müsste jeder Test als Performanztest klassifiziert werden (Bartels, 2018, S. 23). Das Verständnis des Kompetenzbegriffs als die „Fähigkeit, Lösungen für professionelle Anforderungen *beschreiben* zu können“ (Bartels, 2018, S. 23, Hervorh. i. O.), steht in Konflikt mit dem Kompetenzbegriff nach Weinert (2002b) (s. Kapitel 5.3), nach welchem Kompetenz als Zusammenspiel von Wissen und motivationalen, volitionalen und affektiven Komponenten verstanden wird. Kompetenzen in diesem Sinne sind zentraler Bestandteil der professionellen Handlungskompetenz und insofern auch für Handlungen relevant, wie sie in Performanz- und Handlungstest eingefordert werden. Diese begrifflichen Konflikte lassen sich auflösen, indem die Unterscheidung der Testtypen vorrangig auf die Nähe der Testsituation und Testanforderungen zu realem beruflichen Handeln bzw. der zu messenden Kompetenz bezogen wird. Ein Handlungstest ist demnach maximal authentisch, da reales Handeln erfasst wird, dieses kann allerdings aufgrund der Vielzahl möglicher Einflussfaktoren nur schwer zuverlässig getestet werden (Miller, 1990). Ein Wissenstest ermöglicht hingegen eine besonders zuverlässige Messung bei wenig Authentizität in Bezug auf professionelles Handeln. Er weist den niedrigsten Grad an Realitätsbezug auf, da lediglich deklaratives Wissen abgefragt wird, dieses aber nicht angewendet werden muss. Analog nimmt die Komplexität vom Wissenstest bis hin zum Handlungstest zu (Kulgemeyer et al., 2023). Bartels (2018, S. 24) folgert, dass mindestens Kompetenztests nach Miller (1990) sowie Tests mit höherer Authentizität zur Messung der professionellen Handlungskompetenz (im Sinne von Kapitel 5.3) von (angehenden) Lehrkräften geeignet sind. Wissenstests bilden die Komplexität realer Handlungssituationen, die ggf. die Anwendung von Wissen erfordern, nicht ab und ermöglichen demnach nur die Erfassung von (deklarativem) Professionswissen, welches eine Disposition für die professionelle Handlungskompetenz und damit professionelles Handeln darstellt (s. Kapitel 5.3 und 5.4). Ebenfalls kann ohne bekannte Wirkmechanismen, die professionelles Wissen, professionelle Handlungskompetenz und beobachtetes Handeln zueinander in Beziehung setzen, aus beobachtetem (adäquatem) Handeln in Performanz- oder Handlungstests nicht zwangsläufig auf eine zugrundeliegende Kompetenz- oder Wissensbasis geschlossen werden (Vogelsang, 2014, S. 98f.); die Überschätzung der individuellen Merkmale von Lehrkräften bei Vernachlässigung der Kontextfaktoren wird auch als *Attributionsfehler* bezeichnet (Kennedy, 2010).[54]

[54]Legt man das Modell von Kompetenz als Kontinuum nach Blömeke, Gustafsson und Shavelson (2015) zugrunde, ergibt sich ein ähnliches Bild. Zwar kann auch auf Grundlage dieses Modells argumentiert werden, dass alle Testtypen professionelle Handlungskompetenz messen, da auch das Wissen ein Bestandteil der Kompetenz als Kontinuum ist. Es bleibt aber die Unterscheidung bestehen, dass Wissenstests kognitive Fähigkeiten messen, die als Dispositionen für professionelles Handeln dienen und authentischere Tests die situationsspezifischen Fähig-

Bartels (2018, S. 58ff.) ergänzt Millers (1990) Klassifikation im Kontext der Lehkräftebildungsforschung um sogenannte *performanznahe Tests* (Bartels, 2018, S. 58ff.). Ein performanznaher Test erfasst ähnlich einem Kompetenztest, inwieweit eine Person „Lösungen für professionelle Anforderungen“ (Bartels, 2018, S. 23) finden kann, wobei diese Anforderungen - in Annäherung an Performanztests - berufsrelevante und möglichst authentische Situationen darstellen. Durch eine Reduktion insbesondere der Interaktivität (und damit unter Umständen auch der Authentizität) wird ein hohes Maß an Standardisierung der Situationen erreicht. Realisiert wird die standardisierte Darstellung der Situation i. d. R. durch (Unterrichts-)Vignetten, die in unterschiedlichen Formaten vorliegen können (z. B. in Textform oder als Videoaufnahmen (Bartels, 2018, S. 33ff.; Rehm und Bölsterli, 2014)).[55] Performanznahe Tests seien in Bezug auf ihren Grad an Realitätsbezug zwischen Kompetenz- und Performanztests anzusiedeln. In Abgrenzung zu realitätsferneren Kompetenztests sollen Antworten in einem performanznahen Test als „Indikator[en] für das tatsächliche Handeln in einer realen Situation“ (Bartels, 2018, S. 58) gewertet werden können (indem Handlungen bspw. beschrieben oder Handlungsoptionen aus einer vorgegebenen Auswahl ausgewählt werden, also jeweils in einer nicht ausführenden Darstellungsform erfasst werden). Dieses Kriterium erscheint unscharf, da unklar bleibt, welche Eigenschaften performanznaher Tests sie hierzu geeigneter machen, als Kompetenztests es wären. Von Performanztests lassen sie sich nach Bartels (2018) unter Betrachtung der für diese charakteristischen Eigenschaften unterscheiden: Während die Relevanz eines Settings (z. B. einer Erklärsituation im Instrument von Bartels (2018) oder der Wahrnehmung und Analyse von Unterricht im Projekt LUV (Seidel & Prenzel, 2007)) durchaus auch in performanznahen Instrumenten nachvollziehbar gemacht werden kann, ist die Herstellung einer authentischen Handlungssituation nicht immer möglich - insbesondere da kaum echte Interaktion in der standardisierten Testumgebung hergestellt werden kann (Bartels, 2018, S. 59ff.). Die Beschreibung der Kategorie performanznaher Tests deutet grundsätzlich an, dass zu den vier von Miller (1990) unterschiedenen Typen feinere Abstufungen definiert werden könnten. Gleichzeitig sind die einzelnen Typen nicht unbedingt trennscharf definiert: Offen ist bspw. wie stark ein Handlungstest standardisiert werden könnte, bis dieser als Performanztest einzustufen ist. Werden in einem Handlungstest zur Planung von Unterricht bspw. lediglich Planungen für 45-minütige Stunden zum Themenbereich der Optik berücksichtigt, können beide Kriterien als Maßnahmen zur Standardisierung der Situation verstanden werden - auch wenn die Planungen

keiten oder das beobachtbare Verhalten in unterschiedlich stark standardisierten Situationen erfassen.

[55] Die Verwendung von Vignetten ist allerdings keinesfalls hinreichendes Kriterium für performanznahe Tests. Auch Kompetenztests können auf Vignetten zurückgreifen, wenn Wissen auf die in der Vignette dargestellte Situation angewendet werden soll, aber keine unmittelbar berufsrelevante bzw. authentische Situation vorliegt (z. B. zu Schülervorstellungen im Test von Gramzow, Riese und Reinhold (2013)).

im Sinne des Handlungstests real sind. Insbesondere grenzt Bartels (2018) performanznahe Tests vor allem gegenüber den authentischeren Performanztests ab; es bleibt allerdings unklar, inwiefern ein performanznaher Test von Kompetenztests abzugrenzen ist. Die Einordnung bleibt unter Berücksichtigung der adressierten Zieldomäne zu begründen. Aus diesem Grund werden auch nachfolgend vorrangig nur die vier Testtypen nach Miller (1990) unterschieden; performanznahe Tests werden unter Kompetenztests subsumiert.

6.3 Methoden zur Erfassung des professionellen Wissens und Handelns entlang Millers Klassifikation

Miller (1990) legt mit seiner Klassifizierung der verschiedenen Testtypen exemplarische Zuordnungen unterschiedlicher Methoden zu den Typen nahe, z. B. schriftliche Tests für die Erfassung deklarativen Wissens oder die medizinische Diagnose von Krankheitsbildern, die von Schauspieler*innen („standardisierten Patienten“ (Miller, 1990, S. 65)) verkörpert werden, um professionelles Handeln in einer standardisierten Situation (nach Miller (1990): Performanz bzw. *shows how*) zu erfassen. Systematisch ergänzen Kulgemeyer et al. (2023) die Klassifizierung um diese zweite Dimension: die verwendete Methode. Sie widmen sich in ihrem Artikel der Erfassung von „everyday classroom experiences“ (Kulgemeyer et al., 2023, S. 609) und unterscheiden exemplarisch schriftliche Tests, Videovignettentests (s. Kapitel 6.4), Performanztests im Sinne von Simulationen oder isolierten unterrichtlichen Handlungen (Microteaching) und beobachteten Unterricht, wobei ein schriftlicher Test vorrangig einen Wissenstest und unter Umständen noch einen Kompetenztest darstellen kann. Weniger standardisierte Methoden erfassen handlungsnahe Fähigkeiten authentischer und sind als Kompetenztests (z. B. durch die Herstellung unterrichtlicher Bezüge mit Hilfe von Videovignetten), Performanztests (z. B. in Form von Simulationen) oder Handlungstests (in Form von Unterrichtsbeobachtungen) im Sinne von Millers (1990) Zuordnung einzustufen. Mit zunehmender Authentizität (z. B. durch offene Antwortformate oder die Interaktion mit Schüler*innen im Rahmen von nicht standardisierten Unterrichtsbeobachtungen) geht also einerseits eine Reduktion der Standardisierung der Testsituation einher. Die Herstellung objektiver Bedingungen wird dadurch erschwert; es müssen bspw. Schauspieler*innen trainiert werden, die sich wiederholt gleich verhalten und präzise Bewertungsvorlagen entwickelt werden (Miller, 1990). Andererseits sind solche Tests mit einem erhöhten u. a. zeitlichen und personellen Aufwand in Bezug auf die Durchführung und Auswertung verbunden. Stärker standardisierte und somit weniger aufwendig auszuwertende Tests sind also grundsätzlich ökonomischer.

Diese Zuordnung der Methoden zu Millers (1990) verschiedenen Testtypen darf nicht als allgemeingültig missverstanden werden, sondern ist abhängig von der untersuchten Zieldomäne. Soll beispielsweise die Fähigkeit erfasst werden, im Rahmen der Unterrichtsplanung ein adäquates Tafelbild zu entwickeln oder einen diagnostischen Bericht auf Grundlage vorliegender Beobachtungen zu verfassen, sind schriftliche Testformate durchaus authentisch und können als Performanztests dienen - selbst Handlungstests basieren dann auf schriftlichen Dokumenten. Das Spektrum von hoher Authentizität und Komplexität und niedriger Standardisierung hin zu niedriger Authentizität und Komplexität und hoher Standardisierung wird dann nicht grob durch Variation der Methode verändert, sondern mit Hilfe der Aufgabenstellung, die bspw. Rahmenbedingungen bezüglich des zu planenden Tafelbilds vorgeben kann. Eine präzise Formulierung der Zieldomäne ist also Voraussetzung für die Bewertung einer Methode in Bezug auf die Authentizität des Tests bzw. dessen Nähe zur realen beruflichen Situation.

Im Rahmen der Erforschung und Evaluation der Lehrkräftebildung sind häufig das professionelle Wissen und Handeln bzw. deren Bestandteile Untersuchungsgegenstände. Während Large-Scale-Assessments aus testökonomischen Gründen vor allem auf schriftliche Testinstrumente zur Erfassung von deklarativem und angewendetem Wissen setzen, dient auch seit mehreren Jahrzehnten - allerdings üblicherweise in kleinen Stichprobenumfängen - videografierter (realer) Unterricht als Untersuchungsgegenstand zur Erfassung des professionellen Handelns. In jüngeren Forschungsprojekten werden zunehmend auch Performanztests und performanznahe Tests eingesetzt, um professionelles Handeln zu untersuchen. Abhängig vom jeweiligen Untersuchungsziel bzw. der angestrebten Authentizität des Testformats bieten sich unterschiedliche Methoden an. Nachfolgend werden einige Beispiele für die unterschiedlichen Testtypen und ihre Anwendungsbereiche vorgestellt, wobei auf Zieldomänen fokussiert wird, die sich u. a. durch die Interaktion z. B. mit Schüler*innen im Unterricht auszeichnen. Selbsteinschätzungstests werden dabei nicht berücksichtigt. Sie sollten nicht als objektives Verfahren zur Messung von Fähigkeiten (miss-)verstanden und eingesetzt werden, da fraglich ist, inwiefern sie tatsächlich die interessierende Fähigkeit erfassen und nicht z. B. durch Selbstkonzepte beeinflusst werden (Kadluba und Obersteiner, 2024; Holtz und Gnambs, 2017; König, Kaiser und Felbrich, 2012).

6.3.1 Wissens- und Kompetenztests

Zur Untersuchung des Professionswissens (angehender) Lehrkräfte sind schriftliche Testformate im Paper-Pencil-Format oder in Form eines digitalen Instruments weit verbreitet (z. B. Riese et al. (2015) im Fach Physik und Tepner et al. (2012) für die Naturwissenschaften allgemein). Im Rahmen von Wissenstests können gegebene Antworten (sowohl frei formulierte Antworten, als auch z. B. die Bearbeitung von Multiple-Choice-Items) eindeutig als richtig oder falsch identifiziert werden

(Vogelsang, 2014, S. 95). Schriftliche Testinstrumente ermöglichen dabei eine vergleichsweise ökonomische Durchführung und Auswertung; sie werden daher auch in großen Stichproben (in Large-Scale-Assessments) vorrangig eingesetzt. Analog können auch Kompetenztests (die die Anwendung von Wissen in bestimmten Situationen einfordern (Miller, 1990, S. 63)) schriftlich umgesetzt werden. Sollen im Rahmen eines Tests beispielsweise Schülervorstellungen zum Kraftbegriff benannt werden, ist dies eine reine Abfrage des Wissens, also eine Aufgabe für einen Wissenstest. Soll aber aus einer abgedruckten Aussage die zugrundeliegende Schülervorstellung erkannt werden, muss das Wissen um die Schülervorstellung vorhanden sein und von den Proband*innen auf die Aussage *angewendet* werden können, es handelt sich dann um eine Aufgabe für einen Kompetenztest.

Auch Interviews können als Wissens- und Kompetenztests umgesetzt werden. Sie sind vorwiegend im englischsprachigen Forschungsraum verbreitet (Vogelsang, 2014, S. 94). Analog zu schriftlichen Tests können unterschiedliche Fragen dazu genutzt werden, um einerseits unmittelbar das Wissen der interviewten Person zu erfragen, indem z. B. fachliche Zusammenhänge wiedergegeben werden sollen (z. B. zum fachdidaktischen Wissen im Bereich der Elektrostatik (Melo, Cañada-Cañada, González-Gómez & Jeong, 2020)). Andererseits können Anwendungskontexte bzw. Beispiele eingesetzt werden, die auf die Erfassung des angewendeten Wissens (im Sinne eines Kompetenztests) abzielen und so Rückschlüsse auf das vorhandene Wissen ermöglichen sollen (z. B. für mathematikdidaktisches Wissen (McCray und Chen, 2012; Hillje, 2012)). Dabei werden entweder Verhaltensintentionen zu beschriebenen (fiktiven) Situationen erfasst oder retrospektiv Begründungen für beobachtetes Verhalten von den Lehrkräften erfragt. Während die pure Abfrage von Wissen im Rahmen von schriftlichen Tests ökonomischer durchgeführt werden kann, da weniger (personelle) Ressourcen benötigt werden, als bei der Durchführung von Interviews, ermöglichen diese ggf. differenziertere Einblicke durch Rückfragen sowie die Erfassung von Lösungsprozessen (Vogelsang, 2014, S. 94f.).

In Bezug auf die Erfassung handlungsnaher Fähigkeiten, die in komplexen, interaktiven Situationen zum Ausdruck kommen (z. B. während des situativ angepassten Handelns während des Unterrichtens einer Klasse), können allerdings sowohl schriftliche Tests als auch Interviews nicht mit Sicherheit die reale Handlung erfassen, sondern maximal Handlungsintentionen (Cauet, 2016, S. 173ff.; Vogelsang und Reinhold, 2013).

6.3.2 Handlungstests

Handlungstests kommen zum Einsatz, wenn reales Unterrichtshandeln erfasst wird und werden meist in Form von Unterrichtsbeobachtungen umgesetzt. Sie bilden die höchste Stufe in Millers (1990) Klassifikation. Im Rahmen der Lehramtsausbildung sind Handlungstests z. B. in Form von unterrichtspraktischen Prüfungen

im Rahmen des Referendariats etabliert. In der Professionalisierungsforschung kommen hingegen kaum (reine) Handlungstests zum Einsatz, die der Erfassung der unterrichtlichen Fähigkeiten von Lehrkräften dienen. Untersuchungen des Unterrichts werden üblicherweise genutzt, um Erkenntnisse über Zusammenhänge zwischen (mit Wissens- oder Kompetenztests) erfassten Fähigkeiten und dem unterrichtlichen Handeln zu gewinnen. Zu Forschungszwecken werden zur Analyse des unterrichtlichen Handelns üblicherweise Videoaufnahmen des Unterrichts herangezogen. Die Videoaufnahmen werden anschließend mit Hilfe von Kodierleitfäden unter bestimmten Gesichtspunkten der Unterrichtsqualität (z. B. hinsichtlich der kognitiven Aktivierung, Strukturierung oder Klassenführung (Cauet, 2016, S. 134ff.; Vogelsang, 2014, S. 296ff.)) ausgewertet. Groß angelegte Studien, die sich explizit dem Physikunterricht widmen, sind selten. Hier ist vorrangig das Projekt *QuIP*[56] zu nennen, in dessen Rahmen insgesamt 103 Stunden Physikunterricht videografiert wurden (Neumann, Fischer, Labudde & Viiri, 2014), um typische Muster im Unterricht und jene Bedingungen aufzudecken, unter denen sie erfolgreich sind (Fischer, Neumann, Labudde & Viiri, 2014). Weitere groß angelegte Videostudien fokussieren meist auf den Bereich der Mathematik (Seidel et al., 2006)[57] Selbst diese groß angelegten Forschungsprojekte sind allerdings kaum in der Lage, Effekte in der für Interventionen typischerweise geringen Effektstärke abzubilden (Bartels, 2018, S. 26). Häufiger lassen sich Projekte mit nur kleinen Stichproben finden (z. B. im Projekt *ProWiN* mit 23 videografierten Unterrichtsstunden (Cauet, Liepertz, Borowski & Fischer, 2015)). Millers (1990) Klassifikation folgend bilden solche Videostudien, die realen Unterricht auswerten, das Handeln von Lehrkräften als eine Umsetzung von Handlungstests am authentischsten ab. Dies macht die weitreichende Standardisierung der Unterrichtssituation unmöglich. Auch wenn die Erhebung z. B. auf bestimmte Schulformen, Jahrgangsstufen und fachliche Inhalte beschränkt wird (z. B. Seidel et al., 2005), können die vielfältigen Kontextbedingungen und Einflussfaktoren (wie z. B. die räumlichen Gegebenheiten, das an der Schule vorhandene (Experimentier-)Material oder das Verhalten der Schüler*innen) nicht standardisiert werden.[58] Darüber hinaus wird die Umsetzung solcher Studien praktisch dadurch erschwert, dass einerseits im Vorfeld viele Akteure der Umsetzung zustimmen müssen (u. a. die Lehrkräfte, Schüler*innen und ggf. deren Erziehungsberechtigte (Seidel et al., 2005)) und andererseits die Aufnahme und Auswertung der videografierten Stunden mit hohem insbesondere personellen Ressourcenbedarf einhergeht.

[56] *Quality of Instruction in Physics*

[57] Eine Übersicht über groß angelegte Videostudien findet sich in Janík, Seidel und Najvar (2009).

[58] Kulgemeyer und Riese (2018) beschreiben zwar, dass diese Einflussfaktoren in großen Stichproben ihre Bedeutung verlören, es aber selbst in großen Forschungsprojekten nicht realistisch ist, diese Stichprobengrößen zu erreichen.

6.3.3 Performanztests

Während also (schriftliche) Wissens- und Kompetenztests zwar eine hohe Standardisierung aufweisen und auch in großen Stichproben mit vergleichsweise geringem Ressourcenbedarf eingesetzt werden können, können sie professionelles Handeln kaum authentisch erheben. Andersherum ermöglichen Handlungstests die Erfassung realen professionellen Handelns bei hohem Ressourcenbedarf. In jüngerer Vergangenheit werden vermehrt Performanztests eingesetzt, die ein Mittelmaß in Bezug auf die beiden Kriterien der Authentizität und Standardisierung aufweisen, indem die Komplexität der untersuchten Unterrichtssituationen so reduziert wird, dass einzelne, abgrenzbare unterrichtliche Anforderungen gezielt untersucht werden können (Albu und Lindmeier, 2023; Vogelsang und Cauet, 2017, S. 93).

Sie ermöglichen die Erhebung bedeutsamer beruflicher Anforderungen unter standardisierten Bedingungen. Charakteristisch für Performanztests sind, wie oben bereits beschrieben, (a) die Relevanz der Anforderungssituation für das berufliche Handeln, (b) ein hohes Maß an Authentizität der Testsituation und (c) die Möglichkeit zur Interaktion zwischen Proband*innen und Testsituation (Bartels, 2018, S. 30). Sie lassen sich in verschiedenen beruflichen Kontexten finden, von denen z. B. Bartels (2018, S. 27ff.) einige darstellt. Ein prominentes und in der Ausbildung von Ärzt*innen etabliertes Beispiel für Performanztests sind *Objective Structured Clinical Examinations* (*OSCEs*), die als praktisches Prüfungsformat in der Ausbildung von Ärzt*innen dienen. Im Rahmen von OSCEs werden mit Schauspieler*innen („standardisierten Patienten"; Miller, 1990, S. 65) z. B. Anamnesegespräche oder Untersuchungen einzelner Körperteile simuliert (Harden, 1988). Zunehmend werden Performanztests nicht nur als Prüfungsformat eingesetzt, sondern dienen - in Kombination mit Feedback zur Performanz der Studierenden - auch als Lerngelegenheit (Brazeau, Boyd & Crosson, 2002).

Auch in der Ausbildung von Lehrkräften kommen Performantests als Prüfungsformate zum Einsatz. In den USA z. B. das Performance Assessment for California Teachers (PACT), das die Fähigkeiten, Unterricht zu planen, durchzuführen, zu überprüfen und zu reflektieren in einem Portfolio erfasst oder das darauf aufbauende educative Teacher Performance Assessment (edTPA), das in diversen Staaten der USA eingesetzt wird, um die Fähigkeiten im Unterrichten von (angehenden) Lehrkräften zu erfassen (Reagan et al., 2016). Das edTPA beinhaltet die Planung einer Unterrichtseinheit von drei bis fünf aufeinanderfolgenden Stunden sowie deren Durchführung und die Überprüfung des Lernfortschritts der Schüler*innen (Reagan et al., 2016). Nachfolgend werden einige vorrangig zu Forschungszwecken entwickelte Instrumente, die sich auf die Erfassung spezifischer Fähigkeiten beziehen, exemplarisch vorgestellt. Allen Instrumenten ist gemein, dass sie die von Bartels (2018, S. 30) beschriebenen Eigenschaften Relevanz, Authentizität und Interaktion aufweisen und sich auf Physikunterricht konzentrieren, also in der Ausbildung von Physiklehrkräften eingesetzt werden (können). Die methodische

Umsetzung variiert dennoch deutlich, z. B. in Bezug darauf, ob mit realen Personen interagiert wird oder mit einer digitalen Testumgebung.

Dialogischer Erklärtest (DET)

Kulgemeyer und Tomczyszyn (2015) stellen den DET zur Messung der Erklärfähigkeit Physiklehramtsstudierender im Rahmen einer standardisierten Erklärsituation vor. Nach einer zehnminütigen Vorbereitungszeit erklären Proband*innen wiederum innerhalb von zehn Minuten einen physikalischen Sachverhalt (z. B. warum man auf nasser Fahrbahn leichter aus der Kurve gleitet); Adressat*in ist ein*e Schüler*in der zehnten Klasse ohne nennenswerte Vorkenntnisse zum Thema (Kulgemeyer & Tomczyszyn, 2015). Um eine standardisierte, vergleichbare Gesprächssituation herzustellen, sind die Adressat*innen trainiert, gewisse Rückfragen zu stellen, um jeweils „vergleichbare Anforderungen an die Variation der Variablen Mathematisierungsgrad, Beispiele, Sprachniveau und Darstellungsformen“ (Riese et al., 2015, S. 66) herzustellen. Die Erklärsituation wird videografiert und mit Hilfe eines Kategoriensystems hinsichtlich der Erklärqualität (Sachgerechtheit und Adressatengemäßheit) ausgewertet (Kulgemeyer & Tomczyszyn, 2015). In Bezug auf die Güte des DET berichten Kulgemeyer und Tomczyszyn (2015) für das ermittelte Maß eine zufriedenstellende Reliabilität von $\alpha_{\text{Cronbach}} = 0,77$ sowie eine hohe Übereinstimmung von 90 % ($\kappa_{\text{Cohen}} = 0,75$) mit Ratings von Expert*innen bezüglich der Auswahl einer qualitativ höherwertigen Erklärung aus zwei vorliegenden Erklärungen.

Φactio - Unterrichtsminiaturen

Ziel der Studie Φactio ist die Untersuchung von Zusammenhängen zwischen dem Professionswissen und den Überzeugungen angehender Lehrkräfte und deren Unterrichtsqualität. Dazu führen Proband*innen ca. zwölfminütige Unterrichtsminiaturen mit einer Klassenhälfte zu je einem Freihandexperiment aus der Mechanik durch (Korneck, Krüger und Szogs, 2017; Korneck, Oettinghaus, Kunter und Redinger, 2016). Diese Einschränkungen dienen der Komplexitätsreduktion der Testsituation (Korneck, Oettinghaus, Kunter & Redinger, 2016). Die Unterrichtsminiaturen werden videografiert und anschließend reflektiert, bevor die Proband*innen die Unterrichtsminiaturen (in überarbeiteter Form) mit der zweiten Klassenhälfte durchführen; auch diese Miniaturen werden zur Auswertung videografiert (Korneck, Oettinghaus, Kunter & Redinger, 2016). Die Auswertung der Unterrichtsminiaturen in Bezug auf die Unterrichtsqualität erfolgt mit Hilfe eines Kodiermanuals mit 17 Subskalen (u. a. der kognitiven Aktivierung der Schüler*innen sowie der Klassenführung) und insgesamt 134 Items; für die Analyse eines Videos wird etwa eine halbe Stunde veranschlagt (Korneck, Krüger & Szogs, 2017). Korneck, Krüger und Szogs (2017) berichten zufriedenstellende

interne Konsistenzen und Interraterübereinstimmungen für die unterschiedlichen Subkategorien.

Performanztests zur Unterrichtsplanung

Schröder et al. (2020) stellen den im Rahmen des Projekts *ProfiLe-P+* entwickelten Performanztest zur Messung der Fähigkeit zur Unterrichtsplanung vor. Im Test sollen Proband*innen eine 45-minütige Unterrichtsstunde zum dritten Newtonschen Axiom planen, wobei zur Standardisierung der Planungssituation zusätzlich zum Thema der Unterrichtsstunde die Unterrichtsziele vorgegeben sowie Rahmenbedingungen der Lerngruppe beschrieben werden (Schröder et al., 2020). Zur Auswertung der Unterrichtsplanungen wurde ein Kategoriensystem mit insgesamt 45 in der Bewertung berücksichtigten Subkategorien zu sieben Oberkategorien (z. B. „Fachlicher Inhalt und fachliche Korrektheit“ oder „Darstellung und Einsatz von Experimenten“) entwickelt (Schröder et al., 2020, S. 111ff.). Basierend auf dem Einsatz des Instruments vor und nach dem Praxissemester berichten Schröder et al. (2020) positive Ergebnisse hinsichtlich unterschiedlicher Gütemaße, wie eine hohe Interraterübereinstimmung ($\kappa_{\text{Cohen}} = 0,75$) sowie die Übereinstimmung der Rangfolge dreier exemplarisch ausgewählter Planungsprodukte bei einer Bewertung entlang des Kodiermanuals sowie durch Expert*innen (Fachleitende). Weiterhin werden signifikante Zusammenhänge zwischen den Testwerten zur Planungsfähigkeit und Testwerten zum fachdidaktischen und pädagogischen Wissen berichtet; zum Fachwissen konnte keine Korrelation beobachtet werden. Aufbauend auf diesen Test adaptierten Jordans und Riese (2023) im Rahmen einer Replikationsstudie den Performanztest von Schröder et al. (2020), um explizit den Einbezug digitaler Medien bei der Planung zu berücksichtigen. Ergebnisse hierzu sind bislang nicht publiziert.

Reflexionsperformanztest

Kempin, Kulgemeyer und Schecker (2018) entwickelten (ebenfalls im Rahmen des Projekts *ProfiLeP+*) einen Performanztest zur Messung der Reflexionsfähigkeit von Physik-Lehramtsstudierenden. In diesem Instrument werden Proband*innen in eine fiktive kollegiale Beratungs- und Reflexionssituation versetzt und geben einem fiktiven Mitpraktikanten Rückmeldungen zu einer von ihm durchgeführten Doppelstunde Physikunterricht. Der Unterricht wird mit Hilfe von geskripteten Unterrichtsausschnitten (Videovignetten) dargestellt. Die Rückmeldungen werden direkt in ein Mikrofon gesprochen, audiografiert und nach einem Kodiermanual ausgewertet. Für die Auswertungen werden gute Interraterübereinstimmungen berichtet (Vogelsang et al., 2019) sowie eine zufriedenstellende Reliabilität von $\alpha_{\text{Cronbach}} = 0,74$ (Kulgemeyer et al., 2021, S. 3047). Eine ausführlichere Vorstellung des Instruments erfolgt in Kapitel 7.7.3.

6.4 (Video-)Vignetten-Tests

Vignetten sind kurze Darstellungen einzelner Szenen, die es Testpersonen ermöglichen sollen, sich in eine bestimmte Situation hineinzuversetzen bzw. diese zu beurteilen (Jeffries und Maeder, 2011; Steiner und Atzmüller, 2006). Sie sind bereits seit den 1950er Jahren als Methode der Sozialwissenschaften bekannt (Steiner und Atzmüller, 2006; Jeffries und Maeder, 2005) und werdend zunehmend auch im Bereich der Lehrkräftebildungsforschung eingesetzt, um situationsspezifische Fähigkeiten (insb. in Performanztests) zu erfassen (Kaiser et al., 2015; Brovelli, Bölsterli, Rehm und Wilhelm, 2013). Im Kontext der Lehrkräftebildungsforschung stellen Vignetten dabei üblicherweise problembehaftete oder zumindest in gewisser Weise offene, also in Bezug auf ihre Auflösung bzw. den Fortgang nicht eindeutige, Situationen aus dem Unterricht oder erweitertem schulischen Kontext dar, „zu deren erfolgreicher Bewältigung bestimmte Kompetenzen notwendig sind“ (Rehm & Bölsterli, 2014, S. 215), bzw. die anregend für Diskussion oder Reflexion der Situation sein sollen (Skilling und Stylianides, 2020; Jeffries und Maeder, 2005; 2011; Veal, 2002).[59] Sie werden daher nicht nur im Kontext von Tests, sondern auch als Lernmaterial eingesetzt (z. B. Krumphals und Haagen-Schützenhöfer, 2020; Jeffries und Maeder, 2011; Veal, 2002; Goffree und Oonk, 1999). Verbreitet ist der sogenannte „advokarische Ansatz“ (Oser, Heinzer & Salzmann, 2010, S. 8ff.) bei dem das in der Vignette dargestellte Handeln der Lehrkräfte beurteilt wird. Aus den Aussagen werden dann Rückschlüsse auf die Kompetenzen der Teilnehmenden gezogen (Rehm und Bölsterli, 2014; Bischoff, Brühwiler und Baer, 2005).

Vignetten liegen üblicherweise in schriftlicher (*Textvignette*) oder videografierter Form (*Videovignette*[60]) vor und weisen eine Länge von etwa 50-200 Wörtern bzw. zwei bis vier Minuten auf (Skilling und Stylianides, 2020; Rehm und Bölsterli, 2014).[61] Grundlage für (die Entwicklung von) Vignetten sind häufig reale Situationen, die (unmittelbar in der Situation oder im Nachhinein reduziert) gefilmt oder beschrieben werden, wobei die Entwicklung von Vignetten basierend auf realen Situationen (also nicht die Nutzung von Mitschnitten aus realem Unterricht) eine Fokussierung auf bestimmte Problemstellungen, die „essence of a classroom situation“ (Veal, 2002), ermöglicht. Rehm und Bölsterli (2014) empfehlen, Vignetten basierend auf solchen realen Unterrichtssituationen zu entwickeln, deren Verbesserungspotenzial auf Grundlage professioneller Kompetenzen identifiziert werden kann. Auch Jeffries und Maeder (2005) und Skilling und Stylianides (2020)

[59] In dieser Arbeit wird die Bezeichnung *Vignette* ausschließlich für die Darstellung einer Situation verwendet. In der Literatur werden teilweise darüber hinaus auch zugehörige Aufgaben inkludiert; die Situationsbeschreibung wird dann bspw. als *Vignettenstamm* bezeichnet, zu dem unterschiedliche Fragestellungen diskutiert werden könnten (Ollesch, Dörfler & Vogel, 2018).

[60] Oser, Heinzer und Salzmann (2010) verwenden hier den Begriff der *Filmvignette*.

[61] Seltenere Darbietungsformen sind Bildvignetten wie Comics oder mündliche Darstellungen (Skilling und Stylianides, 2020; Jeffries und Maeder, 2005).

definieren Vignetten als „reallife situations“ (Jeffries & Maeder, 2005, S. 20) bzw. „simulating real events“ (Skilling & Stylianides, 2020, S. 541); damit ist aber nicht zwangsläufig die Darstellung tatsächlicher Geschehnisse verbunden, sondern die Forderung nach hoher Authentizität der Vignetten (Jeffries & Maeder, 2005). Der Einsatz von Vignetten stellt für interaktive unterrichtliche Situationen grundsätzlich eine Komplexitätsreduktion der dargestellten Situation dar. Im Vergleich zu realen Situationen können Videovignetten zwar komplexe Situationen mit realen zeitlichen Abläufen darstellen (Friesen, Kuntze und Vogel, 2018; Blömeke, Gustafsson und Shavelson, 2015), aber schon aufgrund der vorgegebenen Kameraperspektive nur einen Teil der Situation zeigen (Kaiser et al., 2015; Oser, Heinzer und Salzmann, 2010), was eine „Aufmerksamkeitssteuerung“ (Neuweg, 2015b, S. 381) bedingt. Analog gehen schriftliche Vignetten mit einer Vorauswahl der beschriebenen Details und damit Reduktion der „Kontextinformationen“ (Friesen, Kuntze & Vogel, 2018, S. 161) zur Situation einher. Insbesondere „simultan ablaufende Prozesse“ (Brovelli, Bölsterli, Rehm & Wilhelm, 2013, S. 311) können nicht abgebildet werden.[62] Im Rahmen von Vignettentests können Rahmenbedingungen der dargestellten Szene beschrieben werden, um die Diskussion größerer Unterrichtszusammenhänge zu ermöglichen (Brovelli, Bölsterli, Rehm & Wilhelm, 2013).[63] Trotz der beschriebenen Einschränkungen in Bezug auf die realitätsgetreue Darstellung von Unterricht in Vignetten ermöglicht der Einsatz von Vignettentests die proximale Erfassung handlungsnaher Fähigkeiten von (angehenden) Lehrkräften (Brovelli, Bölsterli, Rehm & Wilhelm, 2013). Vorhersagen über das tatsächliche Verhalten der (angehenden) Lehrkräfte auf Grundlage von Vignettentests basieren im Gegensatz zu nicht kontextualisierten Tests auf „anforderungsnähere[n] Variablen“ (Neuweg, 2015b, S. 377), da Vignetten die Situativität und Kontextgebundenheit und damit Komplexität von Unterricht repräsentativ abbilden können (Blömeke, Gustafsson und Shavelson, 2015; Oser, Heinzer und Salzmann, 2010). Auch können Vignetten unter verschiedenen Gesichtspunkten diskutiert werden (z. B. einem fachdidaktischen und pädagogischen Fokus), was die Erfassung einer „relativ große[n] Bandbreite professioneller Kompetenz“ (von Aufschnaiter & Blomberg, 2010, S. 365) ermöglicht.

Im Rahmen des Forschungsprojekts *LUV - Lernen aus Unterrichtsvideos* analysierten Lehrkräfte Unterrichtsvideos im Rahmen standardisierter Ratingverfahren in Bezug auf das Vorliegen lernrelevanter Situationen (Seidel & Prenzel, 2007). Die Lehrkräfte akzeptierten das videogestützte Verfahren und beurteilten (insbesondere die verwendeten Videoaufnahmen) positiv in Bezug auf die Authentizität,

[62] Wird die Komplexität der Vignetten besonders stark reduziert, wird die Vielschichtigkeit der Situation nicht mehr abgebildet; solche Vignetten eignen sich zur Erfassung des professionellen Wissens, aber nicht der professionellen Handlungskompetenz (z. B. in Performanztests (Brovelli, Bölsterli, Rehm & Wilhelm, 2013)).

[63] Brovelli, Bölsterli, Rehm und Wilhelm (2013) beschreiben dies als Möglichkeit der Komplexitätserhöhung in schriftlichen Vignetten. Ein analoges Vorgehen ist aber auch durch die Ergänzung von (schriftlichen) Informationen in Videovignetten denkbar.

die kognitive Belastung und die Stimulation (Seidel & Prenzel, 2007). In einem vergleichenden Einsatz von Text-, Comic- und Videovignetten zur Erfassung der fachdidaktischen Analysekompetenz zum Umgang mit Darstellungen Mathematiklehramtsstudierender wurden alle drei Vignettentypen positiv in Bezug auf die wahrgenommene Authentizität, Motivation, Immersion und Resonanz bewertet (Friesen, Kuntze & Vogel, 2018). Gegenüber Text- und Comicvignetten wurden die verwendeten Videovignetten allerdings als signifikant weniger authentisch eingeschätzt, was Friesen, Kuntze und Vogel (2018) darauf zurückführen, dass die videografierten Unterrichtsausschnitte nachgestellt waren bzw. die Videovignetten eine vergleichsweise hohe Individualität in Bezug auf „spezifische, individuelle Merkmale der gezeigten Personen oder Klassenräume“ (Friesen, Kuntze & Vogel, 2018, S. 172) aufweisen, die die Einschätzung der Studierenden individuell beeinflusst haben könnte.

Zusätzlich zur Unterscheidung der Darbietungsformen der Vignetten können in Vignettentests die Antwortformate unterschieden werden. Typischerweise werden schriftliche Antwortformate in einem offenen Format (mit Freitextantworten; z. B. von Rehm und Bölsterli (2014), Brovelli, Bölsterli, Rehm und Wilhelm (2013), Riese und Reinhold (2010) und Baer et al. (2007)) oder geschlossenen Format (z. B. mit Multiple-Choice-Items oder Likert-Skalen u. a. von Tepner und Dollny (2014), Seidel, Blomberg und Stürmer (2010) und Oser, Heinzer und Salzmann (2010)) verwendet, die im Nachhinein bewertet werden können. Seltener werden auch verbale Antwortformate genutzt; die geäußerten Antworten werden hier zur Bewertung üblicherweise aufgezeichnet (z. B. von Kempin, Kulgemeyer und Schecker (2018) oder Knievel, Lindmeier und Heinze (2015)). Auch können die unterschiedlichen Antwortformate kombiniert werden. Manuale für die Auswertung der gegebenen Antworten sind häufig auf Grundlage von Expert*innenurteilen entwickelt (z. B. von Bartels, 2018, S. 94f., Tepner und Dollny (2014) oder Oser, Heinzer und Salzmann (2010)). Bei der Auswahl des Antwortformats sind Vor- und Nachteile der unterschiedlichen Formate abzuwägen. Die Verwendung offener Antwortformate (schriftlich oder verbal) weist gegenüber geschlossenen Formaten eine höhere Authentizität der Testsituation in Bezug auf reales professionelles Handeln auf. Von Aufschnaiter und Blomberg (2010) sehen eine Notwendigkeit für offene Aufgabenformate, um Handlungsnähe in Testsituationen zu realisieren. Gleichzeitig gehen offene Antwortformate mit einer komplexeren und häufig weniger reliablen Auswertung einher; hier ist der Einsatz geschlossener Antwortformate ökonomischer (Brovelli, Bölsterli, Rehm und Wilhelm, 2013; von Aufschnaiter und Blomberg, 2010). Diese ermöglichen zwar eine ökonomische Erfassung und vollständige Standardisierung und damit vermutlich bessere Vergleichbarkeit der Antworten mit Expert*innenurteilen, können aufgrund der vorgegebenen Antwortoptionen allerdings das Antwortverhalten beeinflussen, da sie Proband*innen Aufschluss über den Erwartungshorizont geben können (von Aufschnaiter & Blomberg, 2010), die Antwortoptionen unterschiedlich interpretiert werden (Stecher

et al., 2006) und möglicherweise auf Aspekte aufmerksam machen können, die Proband*innen in einem offenen Antwortformat nicht adressiert hätten (Kaiser et al., 2015). Tepner und Dollny (2014) beschreiben darüber hinaus eine Schwierigkeit, kontextualisierte Antwortalternativen in Multiple-Choice-Aufgaben als eindeutig richtig oder falsch einzustufen.

Bartels (2018, S. 39) stellt fest, dass Vignettentests kaum in der Lage sind, Performanz in Millers (1990) Sinne zu erfassen, insbesondere da keine Interaktion zwischen Testperson und z. B. Lerngruppe in der Vignette stattfinden kann. Stattdessen ermöglichen Vignettentests die Erfassung des hypothetischen Verhaltens bzw. von Verhaltensintentionen der Testpersonen in der Situation, die im Gegensatz zu realem Handeln im Unterricht weniger spontan ablaufen (können) (Kaiser et al., 2015). Inwiefern diese Antworten Rückschlüsse auf das Handeln der Lehrkräfte in vergleichbaren Situationen ermöglichen, wird teilweise als offene Frage der Validität gesehen (Stecher et al., 2006; Kennedy, 1999); gemäß der Theorie des geplanten Verhaltens werden solche Handlungsintentionen (die z. B. als Reaktionen auf Vignetten erfragt werden können) allerdings als geeignete Prädiktoren für tatsächliches Verhalten angenommen (Ajzen, 2002 (überarbeitet 2006)), was umso wahrscheinlicher erscheint, je stärker und stabiler die Intention ist (Conner & Norman, 2022).

Skilling und Stylianides (2020) stellen ein Rahmenmodell zur Entwicklung von Vignettentests vor, das über die Unterscheidung der Darbietungsformen der Vignette und Antwortformate im Test hinausgeht. Sie charakterisieren die Eigenschaften von Vignettentests entlang der drei Bereiche *inhaltliche Konzeption* (der mit der Vignette erfasste Inhalt, die realistische Darstellung und Funktion der Vignette), *Design* (u. a. Länge der Vignette, Antwortformat) und *Administration* (die Formulierung klarer Instruktionen, der zeitliche Rahmen und Kontext des Einsatzes und die Darbietungsform bzw. das Darbietungsmedium).[64]

Insgesamt verspricht die Verwendung von schriftlichen (Video-)Vignetten-Tests die „standardisierte, zeitökonomische Erhebung valider Daten“ (Rehm & Bölsterli, 2014, S. 213) und somit, die Vorteile zweier Methoden zu vereinen: die ökonomische Erhebung von professionellem Wissen und Handeln (bzw. Handlungsintentionen oder „Reaktionstendenzen“ (Seidel & Prenzel, 2007, S. 202)) durch schriftliche Testformate sowie das Erreichen hoher Authentizität durch „kontextualisierte und situationsspezifische Testaufgaben“ (Rutsch, Vogel, Rehm und Dörfler, 2018, S. 17; vergleichbar auch Seidel und Prenzel (2007)).

[64] Skilling und Stylianides (2020) sprechen in diesem Rahmenmodell abstrakt von *der Vignette*. Die Eigenschaften lassen sich aber durchaus auf Tests übertragen, die sich aus mehreren Vignetten und dazugehörigen Aufgaben zusammensetzen. Dazu kann das Rahmenmodell für jede Vignette angewendet werden und Eigenschaften, die (auch) das Instrument als Ganzes betreffen (z. B. die Setzung eines zeitlichen Rahmens), können zusätzlich umfassend berücksichtigt werden.

6.5 Zwischenfazit: Messung von professionellem Wissen und Handeln

Die Erfassung von Professionswisssen und insbesondere der professionellen Handlungskompetenzen (angehender) Lehrkräfte, zu der auch die später näher beschriebene Reflexionsfähigkeit zu zählen ist, ist aufgrund der Komplexität unterrichtlicher Situationen mit gewissen Schwierigkeiten verbunden. Selbstverständlich sind - wie für sonstige Testverfahren - Gütekriterien der empirischen Forschung zu berücksichtigen, die in Kapitel 6.1 dargestellt wurden.

In dieser Arbeit wird zur Evaluation der entwickelten Lernumgebung eine Validierung gemäß dem argumentbasierten Ansatz von Kane (z. B. 1992; 2013) angestrebt, wobei die Validitätsaspekte nach Messick (1995) berücksichtigt werden, um breite Evidenz einzubeziehen. Dieser Ansatz ermöglicht eine Fokussierung und Strukturierung der Validitätsargumentation abhängig von der angestrebten Testwertinterpretation und -nutzung. In diesem Rahmen werden Argumente für die Validität gesammelt, die für eine stringente Argumentation auch andere Gütekriterien voraussetzen bzw. umfassen. So sind bspw. die objektive Erfassung und Auswertung der Bearbeitungen und das Vorliegen einer reliablen Skala Voraussetzungen dafür, dass die Testwerte für statistische Analysen genutzt werden können. Die Nützlichkeit des Testinstruments wird als Vorbedingung für dessen Entwicklung verstanden und daher nicht evaluiert, sondern vor Entwicklung des Tests begründet (s. Kapitel 9). Im Folgenden werden die unterschiedlichen Gütekriterien also nicht als unabhängig voneinander aufgefasst, sondern werden - soweit sie für die Validierung der Testwertinterpretation bedeutsam sind - unter dem Validitätsbegriff subsumiert bzw. ergänzend diskutiert.

Zur Klassifizierung der verschiedenen Testformate, die zur Erfassung von professionellem Wissen und Handeln eingesetzt werden, wird auf die Unterscheidung von Wissens-, Kompetenz-, Performanz- und Handlungstests nach Miller (1990) zurückgegriffen. Die Testformate gehen (in der genannten Reihenfolge) mit zunehmender Authentizität der Handlungssituation einher, büßen gleichzeitig aber an Möglichkeiten zur Standardisierung und damit auch Effizienz der Testsituation (durch höheren personellen und zeitlichen Aufwand in der Erhebung und Auswertung des Tests) ein (s. Kapitel 6.2). Für die Erfassung von professionellem Wissen und Handeln von Lehrkräften ist das Testformat (insbesondere die Darstellung der Handlungssituationen bei der Erfassung professionellen Handelns sowie das Antwortformat) in diesem Spannungsfeld passend zum Ziel des Testeinsatzes zu wählen (s. Kapitel 6.3). Zur Herstellung einer möglichst authentischen Handlungssituation bei gleichzeitig hoher Standardisierung wird für das entwickelte Testinstrument auf Videovignetten in Kombination mit geschlossenen Antwortformaten zurückgegriffen (s. Kapitel 6.4).

7 Reflexion von Physikunterricht

7.1 Vorläufige Definition: Der gemeinsame Kern des Reflexionsbegriffs

Es bestehen bedeutsame Unterschiede im Verständnis des Reflexionsbegriffs bei verschiedenen Autor*innen (von Aufschnaiter, Fraij und Kost, 2019; Clarà, 2015; Beauchamp, 2006, S. 48). Um dennoch einen Überblick über in der Literatur vorliegende Begriffsverständnisse und empirische Erkenntnisse zur Reflexion für (angehende) Lehrkräfte zu ermöglichen, wird nachfolgend zunächst der gemeinsame Kern der meisten Definitionen wiedergegeben:[65]

Weitgehend Einigkeit besteht bezüglich der Charakterisierung von Reflexion (bzw. dem Reflexions*prozess*) als eine spezielle Form des (Nach-)Denkens (z. B. Häcker, 2017; Roters, 2011, S. 112; Hatton und Smith, 1995); teilweise gilt sie als höherwertig gegenüber anderen Formen des Nachdenkens (von Aufschnaiter, Fraij und Kost, 2019; Aeppli und Lötscher, 2016). Üblicherweise wird Reflexion im Sinne eines Denkprozesses als mit einem bestimmten Zweck - wie der Lösung einer als problematisch empfundenen Situation - verbunden verstanden (von Aufschnaiter, Fraij und Kost, 2019; Aeppli und Lötscher, 2016) und umfasst verschiedene Denkaktivitäten wie u. a. das Beschreiben und Beurteilen von Situationen oder das Problemlösen (Meier, Vogelsang, Watson und Schaper, 2022; Beauchamp, 2006, S. 68ff.).

7.2 Zur Bedeutung von Reflexion für (angehende) Lehrkräfte

Es besteht breiter Konsens darüber, dass der Reflexion für Lehrkräfte eine zentrale Bedeutung zukommt, sie wird bspw. als „Schlüsselkompetenz“ (Combe & Kolbe, 2008, S. 859) für Lehrkräfte beschrieben. Bezogen auf die Kerntätigkeit von Lehrkräften, das Unterrichten, umfasst die Reflexion den Abgleich von geplanten und tatsächlichen Unterrichtsabläufen. Herzog (1995) beschreibt sie als „dritte Phase der Handlung“ (Herzog, 1995, S. 256). Demnach dient Reflexion dazu, „[n]ach

[65] Eine Konkretisierung des Begriffsverständnisses für die in dieser Arbeit vorgestellte Lernumgebung und deren Evaluation erfolgt in Kapitel 9.

dem Unterricht [...] klüger zu werden“ (Schulz, 1966, S. 22), indem Erkenntnisse für die Gestaltung zukünftigen Unterrichts gewonnen werden. Außerdem wird vor dem Unterricht reflektiert, um „dessen Struktur in der Planung [zu] antizipieren, d. h. [...] auf die gewünschten Folgen hin [zu] konstruieren“ (Schulz, 1966, S. 22). Diese Planung wiederum wird im Anschluss an den Unterricht reflektiert und für die nachfolgende Planung berücksichtigt. Reflexion soll also im Rahmen der Evaluation und Planung von durchgeführtem bzw. zukünftigem Unterricht und der Entwicklung alternativer Handlungsoptionen für zukünftigen Unterricht unmittelbar zur Unterrichtsqualität beitragen (z. B. von Aufschnaiter, Fraij und Kost, 2019; Hilzensauer, 2017, S. 27; Helmke, 2015, S. 116).

Auf einer dem konkreten Unterricht übergeordneten Ebene wird Reflexion als Instrument zur Relationierung von Theorie und Praxis verstanden (von Aufschnaiter, Fraij und Kost, 2019; 2019; Leonhard und Rihm, 2011; Gruber und Rehrl, 2005, S. 3; Herzog, 1995). Ihr wird eine Vermittlungsfunktion zwischen Erlebnissen und Erfahrungen und dem professionellen Wissen zugesprochen, die zur Professionalisierung beiträgt (z. B. von Aufschnaiter, Hofmann, Geisler und Kirschner, 2019; Hilzensauer, 2017, S. 27; Baumert und Kunter, 2006; Leonhard und Rihm, 2011; Neuweg, 2005; 2020). So nimmt Reflexion auch mittelbar Einfluss auf den Unterricht und dessen Qualität (von Aufschnaiter, Hofmann, Geisler & Kirschner, 2019). McAlpine et al. (1999) beschreiben Reflexion als „mechanism for turning experience into knowledge“ (McAlpine et al., 1999, S. 116). Dieser Mechanismus wird üblicherweise als zyklische Abfolge von Handlung und Reflexion (z. B. Korthagen und Vasalos, 2005) bzw. Planung der Handlung, Ausführung und Reflexion (Herzog, 1995) modelliert und ermöglicht so die „stetige Weiterentwicklung der [p]rofessionellen Kompetenz“ (Szogs, Kobl, Volmer & Korneck, 2019) - teilweise wird Professionalisierung sogar mit der „Steigerung von Reflexivität“ (Häcker, 2022, S. 94; auch Reh, 2004, S. 363) gleichgesetzt.[66] Reflexion kann dann dazu dienen, das eigene Handeln (theoriebasiert) zu begründen (Häcker, 2019). Auch (unbewusste) subjektive Theorien können aufgebrochen und „unter Bezugnahme auf wissenschaftliche Theorien begründbar“ (Gruber & Rehrl, 2005, S. 13f.) gemacht werden. Postholm (2008) fasst zusammen:

> „Teachers can reflect independently on their own by using theory as an analytical tool in this activity. In this way teachers can use theory both to analyse and understand their practice, to see possibilities in the practice, and furthermore, to improve their practice.“ (Postholm, 2008, S. 1721)

Aufgrund der hohen Situativität unterrichtlicher Situationen ist „die Theorie gegenüber der Praxis notwendigerweise defizitär“ (Herzog, 1995, S. 256). Wissensbestände bzw. Theorien können unterrichtliche Situationen nicht vollständig

[66] Zum Verhältnis der Begriffe Reflexion und Reflexivität siehe Kapitel 7.4; Häcker (2022) scheint den Begriff der Reflexivität als „Bewusstsein über das eigene Tun“ (Häcker, 2022, S. 96) nach Combe und Kolbe (2008, S. 859) zu verwenden.

erfassen (Gruber & Rehrl, 2005, S. 4). Hinzu kommt, dass (angehende) Lehrkräfte Schwierigkeiten damit haben, ihr Wissen mit unterrichtlichen Situationen in Verbindung zu bringen, was zu einem wahrgenommenen „Bruch“ zwischen Theorie und Praxis führt (König & Rothland, 2018, S. 5). Reflexion dient - auch als Teil von Praxisphasen und insbesondere des Praxissemesters (König & Rothland, 2018)[67] - dazu, diesen Bruch zu überwinden (König, Darge, Klemenz und Seifert, 2018; Herzig, Grafe und Reinhold, 2005[68]; Herzog, 1995).

In Anbetracht der Komplexität und Ungewissheit, die Lehrkräfte bei der Ausübung ihrer Profession immer wieder vor neue Herausforderungen stellt und Unterricht kaum beherrschbar erscheinen lässt (z. B. Cramer et al., 2019; Herzog, 1995), kann die (universitäre) Ausbildung Lehrkräften nicht für alle Situationen anwendbares Wissen und angemessene Handlungsstrategien bereitstellen. Vielmehr müssen Lehrkräfte in der Lage sein, aus ihren Erfahrungen zu lernen und selbst adäquate Handlungsstrategien entwickeln, um unbekannte Situationen bewältigen zu können. Hierfür gilt Reflexion als bedeutsamer Faktor (Hiebert, Morris, Berk und Jansen, 2007; Gruber, Heid und Bauer, 2005, S. 16). Damit verbunden werden auch die adäquate Ausbildung von und der flexible Umgang mit beruflichen Routinen durch Reflexion begünstigt (Kobl, 2021, S. 32; Roters, 2011, S. 104f.; Schmelzing, Wüsten, Sandmann und Neuhaus, 2010; Herzig, Grafe und Reinhold, 2005).

Darüber hinaus weist Reflexion Relevanz in Bezug auf das Wohlbefinden und die psychische Gesundheit von Lehrkräften auf (Wyss, 2013, S. 63). Beispielsweise wird vermutet, dass durch Reflexion das Burn-out Risiko durch einen Abgleich von „eigenen Ansprüchen und realen Handlungsmöglichkeiten“ (Dauber, 2006, S. 31) reduziert werden könnte. Wubbels und Korthagen (1990) beobachten einen Zusammenhang zwischen der Förderung reflexiven Unterrichtens und der Berufszufriedenheit von Lehrkräften, der auch für Englischlehrkräfte im Iran bestätigt wurde (Aliakbari, Khany & Adibpour, 2020).

Die hohe Bedeutung von Reflexion für Lehrkräfte schlägt sich auch in Standards für die Lehrerbildung nieder. Sie wird neben der Planung und Organisation von Lehr- und Lernprozessen als „Kernaufgabe“ (Sekretariat der Kultusministerkonferenz, 2004, S. 3) verstanden. Auch wenn die Reflexion von Unterricht vor allem in der zweiten Ausbildungsphase im Fokus steht, ist bspw. die Reflexion von „Grundlagen und Ergebnisse[n] der Schul- und Unterrichtsforschung“ (Sekretariat der Kultusministerkonferenz, 2004, S. 8) oder von „persönlichen berufsbezogenen Wertvorstellungen und Einstellungen“ (Sekretariat der Kultusministerkonferenz, 2004, S. 14) schon für die erste Ausbildungsphase vorgesehen. Insofern und vor dem oben beschriebenen Hintergrund, dass Reflexion entscheidend für professionelles Handeln in vielfältigen, immer wieder neuen schulischen Herausforderungen

[67] Die Etablierung des Praxissemesters in der universitären Ausbildung von Lehrkräften war mit dem Ziel verbunden, eine engere Verzahnung von Theorie und Praxis zu erreichen (König & Rothland, 2018).

[68] Sie sprechen dabei auch vom „reflexiven Lernen“ (Herzig, Grafe & Reinhold, 2005, S. 51).

ist, überrascht die Forderung nicht, dass bereits Studierende Reflexion erlernen sollen bzw. die Fähigkeit zur Reflexion bereits im Studium vorausgesetzt wird (Hilzensauer, 2017, S. 14).

7.3 Zum vielfältigen Verständnis des Reflexionsbegriffs

Im Rahmen der vorläufigen Definition (s. Kapitel 7.1 auf Seite 79) wurde die Reflexion mit dem Denkprozess (Reflexionsprozess) gleichgesetzt. Ebenfalls häufig als Reflexion(en) bezeichnet werden die Darstellungen bzw. Produkte einer Reflexion (z. B. Reflexionstexte oder verbalisierte Reflexionen), die als Ausdruck des Denkprozesses verstanden werden. Nowak (2023, S. 3) fordert daher eine begriffliche Unterscheidung zwischen dem Reflexions*prozess* und Reflexions*produkt*, die zwar vor dem Hintergrund sinnvoll erscheint, dass Reflexionsprodukte (z. B. schriftliche Texte) eine Aufbereitung des Denkprozesses in einem spezifischen Kontext (z. B. einer Prüfungssituation) darstellen. Indem die Reflexionsprodukte aber den Zugang zum Reflexionsprozess darstellen, kommt ihnen eine zentrale Bedeutung für die Untersuchung des Reflexionsprozesses zu. Eine präzise begriffliche Trennung erscheint daher auch für diese Arbeit kaum möglich.

Obwohl viele Arbeiten als geteilte Grundlage für das Verständnis des Reflexionsbegriffs die Arbeiten von Dewey (1910/2002; 1933) und Schön (1983; 1987) benennen (Clarà, 2015), wurden im Verlauf der letzten Jahrzehnte immer wieder zentrale Unklarheiten im Begriffsverständnis bzw. Unterscheidungsmerkmale verschiedener Definitionen herausgearbeitet. So verweisen z. B. Hatton und Smith (1995, S. 34f.) auf die vier Fragen, (i) ob Reflexion als Denkprozess *über* Handlungen verstanden werden soll, oder diese inkludiert, (ii) in welchem zeitlichen Rahmen bzw. zeitlichen Bezug zur reflektierten Handlung eine Reflexion stattfindet, (iii) ob diese grundsätzlich mit der Lösung eines Problems verknüpft ist und (iv) inwiefern soziokulturelle und moralische/ethische Kontexte im Rahmen einer Reflexion berücksichtigt werden. Szogs, Kobl, Volmer und Korneck (2019) benennen weitere Unterscheidungsmerkmale wie u. a. die an einer Reflexion teilnehmenden Personen, verschiedene inhaltliche Ebenen von Reflexion (wenn z. B. um das Handeln einer Lehrkraft in einer bestimmten Situation oder abstrakter um das Einfinden in die Rolle der Lehrkraft in den Blick genommen wird). Insgesamt lassen sich also eine Vielzahl von Unterscheidungsmerkmalen in Konzeptualisierungen des Reflexionsbegriffs ausmachen, die im Folgenden überblicksartig dargestellt werden.[69]

[69] Zusätzlich zu den folgenden Aspekten von (Definitionen von) Reflexion wird teilweise unterschieden, ob eine Reflexion lediglich den Denkprozess über eine Handlung umfasst oder diese Teil der Reflexion ist (z. B. Aeppli und Lötscher, 2016; Herzog, 1995) Näheres dazu befindet sich im Rahmen der Vorstellung verschiedener Prozessmodelle von Reflexion in

Auch wenn verschiedene Definitionen darin übereinstimmen, dass Reflexionen mit einem bestimmten Ziel durchgeführt werden (z. B. Meier, Vogelsang, Watson und Schaper, 2022; Aeppli und Lötscher, 2016), lassen sich Unterschiede in der Art der *Zielstellung* einer Reflexion ausmachen, die nach von Aufschnaiter, Fraij und Kost (2019) dichotom als *internal* oder *external* eingeordnet werden können. Definition von Reflexionen mit (rein) externaler Zielsetzung fokussieren auf die Lösung eines Problems, das die Ausgangslage der Reflexion bildet (diesem Verständnis folgt z. B. Clarà (2015)). Dewey beschreibt zwei aufeinanderfolgende Phasen der Reflexion: Auf „einen Zustand der Beunruhigung, des Zögerns [oder] Zweifelns“ (Dewey, 1910/2002, S. 13) folgt ein „Akt des Forschens oder Suchens“ (Dewey, 1910/2002, S. 13), dessen Ergebnis die Auflösung des unklaren Zustands ermöglichen soll.[70] Im schulischen Kontext zielt eine solche externale Zielstellung insbesondere auf eine Qualitätssteigerung des Unterrichts ab. Internale Zielstellungen von Reflexionen dienen hingegen der individuellen Weiterentwicklung und Professionalisierung (von Aufschnaiter, Fraij & Kost, 2019). Während einige Autor*innen den Reflexionsbegriff explizit auf solche internalen Zielstellungen eingrenzen (z. B. Lohse-Bossenz, Schmitt, Lenske und Gold, 2023; von Aufschnaiter, Fraij und Kost, 2019; Leonhard und Rihm, 2011), lassen sich in der Literatur durchaus Definitionen finden, die beide Zielstellungen unter dem Reflexionsbegriff subsumieren (z. B. Nowak, 2023, S. 159; Kempin, Kulgemeyer und Schecker, 2020; Aeppli und Lötscher, 2016[71]).[72]

Anschließend an die unterschiedlichen Zielstellungen, die mit dem Reflexionsbegriff verbunden werden, werden Unterschiede in Bezug auf die einbezogenen Denkgegenstände einer Reflexion deutlich.[73] Einerseits können Denkgegenstände auf unterschiedlichen Abstraktionsniveaus betrachtet werden, im Kontext des

Kapitel 7.5.1). Auf eine Differenzierung des Reflexionsbegriffs nach Reflexions*tiefe* (Szogs, Kobl, Volmer & Korneck, 2019) und ihrem Ertrag (von Aufschnaiter, Fraij & Kost, 2019) wird zusammenhängend mit der Frage nach der Qualitätsbewertung von Reflexionen in Kapitel 7.5.2 eingegangen.

[70] Werden nicht ausschließlich problembehaftete Situationen reflektiert, widerspricht das zwar Deweys (1910/2002) ursprünglichem, auf Problemlösung ausgerichteten Reflexionsverständnis, wird aber z. B. von von Aufschnaiter, Fraij und Kost (2019) explizit als gewinnbringend für die Festigung von gelungenen Verhaltensweisen erachtet.

[71] Aeppli und Lötscher (2016) benennen die „Blickrichtungen“ „Au[ß]enwelt“ (Aeppli & Lötscher, 2016, S. 84) die u. a. das eigene oder beobachtetes Verhalten umfasst und eine „Introspektion“ (Aeppli & Lötscher, 2016, S. 84), zur Herstellung des Selbstbezugs und der Berücksichtigung von u. a. Gefühlen und Überzeugungen.

[72] Von Aufschnaiter, Fraij und Kost (2019) sehen in der Reflexion mit externaler Zielstellung Parallelen zu anderen analytischen Tätigkeiten, u. a. zur Diagnose. Diese dient in einem (hierarchischen) Lehr-Lern-Kontext üblicherweise der Feststellung von Personenmerkmalen der Lernenden und darauf aufbauend der Ableitung von Maßnahmen zur Optimierung des Lernprozesses für diese Personen (Rath, 2017, S. 19; von Aufschnaiter et al., 2015).

[73] Sie werden auch als „Reflexionsinhalte“ (Szogs, Kobl, Volmer & Korneck, 2019, S. 318) bezeichnet und sind Teil der von von Aufschnaiter, Fraij und Kost (2019) beschriebenen „Charakteristika“, welche darüber hinaus den Ertrag einer Reflexion umfassen (dieser wiederum ist beobachtbar in der „Veränderung [oder] Stabilisierung von Bestehendem“ (von Aufschnaiter, Fraij & Kost, 2019, S. 147))

Unterrichts z. B. konkrete Unterrichtsplanungen oder die professionelle Handlungskompetenz im Allgemeinen (Szogs, Kobl, Volmer und Korneck, 2019; ähnlich auch im Zwiebelmodell von Korthagen und Vasalos (2005) oder in der Unterscheidung zwischen Mikro-, Meso- und Makrobenene von Wyss (2013, S. 49f.)). Nach einem Reflexionverständnis mit internaler Zielstellung muss außerdem stets ein persönlicher Bezug hergestellt werden. Herzog (1995) fasst für die Lehramtsausbildung zusammen:

> „Einige meinen [mit Reflexion] das schlichte Nachdenken über irgend etwas, wozu auch die eigenen Erziehungsvorstellungen gehören können, andere grenzen die Reflexion auf ein Nachdenken über Verhaltensweisen ein. Einige sehen als Themen von Reflexion zeitlich enge Einheiten (wie Handlungen), andere denken an zeitlich ausgedehnte Einheiten (wie biographische Abläufe, Lebensphasen etc.). Einige konzentrieren die Reflexion auf (akute) Probleme, andere vertreten einen problemunabhängigen Ansatz, bei dem auch über unproblematische Verhaltensweisen nachgedacht wird.[...] Einige grenzen die Reflexion auf das Feld des Bewusstseins ein, andere denken auch an die Aufdeckung unbewusster oder (bisher) unartikulierter Wissensbestände oder Handlungsmotive.“ (Herzog, 1995, S. 253)

Die Ausdifferenzierung des *zeitlichen Zusammenhangs* zwischen der Situation und der Reflexion[74] geht auf Schöns (1983, S. 278ff.) Unterscheidung zwischen der *reflection-in-action* (der Reflexion *während* einer Handlung) und *reflection-on-action* (der Reflexion *nach* einer Handlung) zurück, die breite Akzeptanz gefunden hat (z. B. Aeppli und Lötscher, 2016; Wyss, 2008). Darüber hinaus wird von einigen Autor*innen eine Reflexion im Vorfeld der Handlung berücksichtigt, die bspw. die Antizipation möglicherweise auftretender Probleme umfasst (*reflection-for-action* (Roters, 2011, S. 118) oder *reflection-for-practice* (Thompson & Pascal, 2012, S. 322)) und als Teil der Unterrichtsplanung einzuordnen ist (Nowak, 2023, S. 4).

Weiterhin können Definitionen des Reflexionsbegriffs in Bezug auf die an einer Reflexion *(aktiv) teilnehmenden Personen* unterschieden werden. Zweifelsohne kann eine Person über eigene Handlungen bzw. eigenen Unterricht reflektieren (*Selbstreflexion*). Strittig ist hingegen, ob auch über fremden bzw. beobachteten Unterricht reflektiert werden kann (*Fremdreflexion*). Augenscheinlich passend zur Forderung des Selbstbezugs bzw. einer internalen Zielstellung von Reflexion schließen dies u. a. von Aufschnaiter, Fraij und Kost (2019) aus, während andere Autor*innen beide Formen unter dem Begriff der Reflexion subsumieren (teilweise in Form von gemeinsamen Reflexionen, sodass in einer Gruppenkonstel-

[74] Da einige Modellierungen des Reflexionsprozesses die reflektierte Handlung bzw. das Erproben von Handlungsalternativen als Teil der Reflexion verstehen (s. Kapitel 7.5.1), sei an dieser Stelle explizit darauf verwiesen, dass mit Reflexion hier lediglich der *Denk*prozess gemeint ist.

lation Selbst- und Fremdreflexion gleichzeitig auftreten; z. B. von Wyss (2013, S. 48f., 55)). Grundsätzlich kann allerdings auch auf Grundlage von beobachtetem Unterricht ein Selbstbezug hergestellt werden, indem z. B. eigene Erfahrungen in die Reflexion einbezogen werden oder ein persönlicher Mehrwert abgeleitet werden kann (Merkert, Lohse-Bossenz, Neuber & Lenske, 2023). Sind mehrere Personen anwesend, kann zusätzlich unterschieden werden, welche Rollen (z. B. Mentor*in/Coach*in oder Kolleg*in/Peer) diese innehaben (Wyss, 2013, S. 48; Roters, 2011, S. 120f.). Gemeinsame Reflexionsprozesse werden dabei - sofern als Reflexion verstanden - als bereichernd gegenüber der reinen Selbstreflexion erachtet, da neue Perspektiven aufgezeigt und unbewusste Aspekte aufgedeckt werden können (Wyss, 2013, S. 48f.; Roters, 2011, S. 121; Herzog, 1995).

Auch die *Offenheit* einer Reflexion kann „in einem Spektrum von stark geleitet (z. B. durch maßgebende Leitfragen) bis sehr offen (keine Vorgaben)“ (Szogs, Kobl, Volmer & Korneck, 2019, S. 318) verortet werden.[75] Geschlossene Reflexionen werden dabei üblicherweise *von außen initiiert* bzw. eingefordert (von Aufschnaiter, Fraij und Kost, 2019; Wyss, 2013, S. 49).[76] Dies bedeutet allerdings im Umkehrschluss nicht, dass von außen initiierte Reflexionen (durch Arbeitsaufträge, irritierende Momente in beobachtetem Unterricht ö. ä.) stets geschlossen sind, d. h. auf vordefinierte Kriterien zurückgreifen.

Auf einer formalen Ebene kann darüber hinaus die *mediale Umsetzung*[77] einer Reflexion ausdifferenziert werden, wobei typischerweise mündliche (häufig dialogische) und schriftliche Reflexionen unterschieden werden (Szogs, Kobl, Volmer und Korneck, 2019; Wyss, 2013, S. 50). Einige Autor*innen weisen auf die Notwendigkeit der Verbalisierung von Reflexionen hin (Häcker, 2022) - zumindest für die Erfassung im Forschungskontext (Leonhard & Rihm, 2011).

Auch die (explizite) Berücksichtigung affektiver bzw. emotionaler Aspekte wird gefordert (z. B. Aeppli und Lötscher, 2016; Poldner et al., 2014; Korthagen und Vasalos, 2005), da diese stets mit Reflexion einhergehen (Herzog, 1995).

Wie an verschiedenen Stellen bereits angemerkt, sind die Unterscheidungskriterien nicht pauschal als unabhängig zu verstehen. So nimmt bspw. die Annahme einer internalen Zielstellung von Reflexion Einfluss auf die Denkgegenstände, indem stets (zusätzlich) ein Selbstbezug hergestellt werden muss. Reine Fremdreflexionen sind demnach nicht unbedingt mit einer internalen Zielstellung vereinbar. Auch Vorgaben, die die Offenheit einer Reflexion einschränken, können gleichsam Denkgegenstände einschränken. Ebenso führt der zeitliche Zusammenhang zwischen Handlung und Reflexion zu möglichen Einschränkungen der Denkgegenstände. So ist zu hinterfragen, ob bspw. eher abstrakte Aspekte (z. B. biografischer Natur)

[75] Wyss (2013, S. 46ff.) unterscheidet begrifflich zwischen *geschlossenen* und *offenen* Reflexionen, Aeppli und Lötscher (2016) verweisen auf unterschiedlich starke „Steuerung“ (Aeppli & Lötscher, 2016, S. 80), ohne diese näher zu beschreiben.

[76] Häcker (2022, S. 104) spricht hier von *veranlasster* Reflexion.

[77] Auch *Modalität* (Szogs, Kobl, Volmer & Korneck, 2019, S. 319)

in-action reflektiert werden können oder in einer reflection-in-action auf vorgegebene Kriterien Rücksicht (im Sinne einer geschlossenen Reflexion) genommen werden kann.

7.4 Reflexionskompetenz, Reflexionsfähigkeit und Reflexionsbereitschaft als Personenmerkmale

Während bisher das Verständnis des Begriffs Reflexion (vorrangig im Sinne des Denkprozesses) im Fokus stand, dienen die Begriffe der Reflexions*fähigkeit*, Reflexions*bereitschaft*, Reflexions*kompetenz* und teilweise auch *Reflexivität* dazu, den Blick weg vom Prozess hin zu der eine Reflexion ausführenden Person bzw. hin zu Personenmerkmalen zu lenken, die mit Reflexion verbunden sind.

7.4.1 Reflexionskompetenz, Reflexionsfähigkeit und Reflexionsbereitschaft

Sowohl die Qualität einer Reflexion als auch die Frage, ob sie überhaupt durchgeführt wird, hängen von den personalen Bedingungen der reflektierenden Person ab. Im Sinne des Kompetenzbegriffs nach Weinert (2002b) können diese Bedingungen wie folgt als Reflexionskompetenz zusammengefasst werden:

> Reflexionskompetenz umfasst die bei Individuen verfügbaren oder durch sie erlernbaren kognitiven Fähigkeiten und Fertigkeiten sowie die damit verbundenen affektiv-motivationalen Bereitschaften und Fähigkeiten, um Situationen erfolgreich und verantwortungsvoll zu reflektieren.

Die *Reflexionskompetenz* ist damit als Teil professioneller Handlungskompetenz einzuordnen. Von Aufschnaiter, Fraij und Kost (2019) greifen für die Modellierung der Reflexionskompetenz[78] auf das Modell von Kompetenz als Kontinuum (Blömeke, Gustafsson & Shavelson, 2015) zurück und unterscheiden demgemäß zwischen unterschiedlichen kognitiven und affektiv-motivationalen Dispositionen. Angelehnt an diese Unterscheidung kann das Konzept der Reflexionskompetenz weiter ausdifferenziert werden in Reflexionsfähigkeit und -bereitschaft.

Die *Reflexionsbereitschaft* umfasst dabei die motivationalen, volitionalen und sozialen Bereitschaften für eine Reflexion, die schließlich entscheidend dafür sind, ob bzw. in welchem Ausmaß Reflexionen durchgeführt werden (Szogs, Kobl,

[78] Sie benennen die Reflexionskompetenz als *Reflexivität* (von Aufschnaiter, Fraij & Kost, 2019, S. 152).

Volmer & Korneck, 2019). Die Reflexionsbereitschaft hängt dabei eng zusammen mit den persönlichen Einstellungen und Überzeugungen gegenüber Reflexion (z. B. bezüglich deren Relevanz (von Aufschnaiter, Fraij & Kost, 2019)).

Die *Reflexionsfähigkeit* beinhaltet komplementär die kognitiven Wissensbestände und Fähigkeiten, die die Qualität einer Reflexion bestimmen (Szogs, Kobl, Volmer & Korneck, 2019). Aus der Qualität einer Reflexion kann demnach auf die Reflexionsfähigkeit geschlossen werden. Für (gewinnbringende) Reflexion muss also nicht nur Bereitschaft zur Durchführung dieser vorhanden sein, sondern Individuen müssen darüber hinaus in der Lage dazu sein, zu reflektieren. Basierend auf dem psychologischen Verständnis des Begriffs Fähigkeit als „die Gesamtheit der zur Ausführung einer bestimmten Leistung erforderlichen personalen Bedingungen“ (Wirtz, 2017, S. 556) kann die Reflexionsfähigkeit wie folgt definiert werden:

> Reflexionsfähigkeit ist die Gesamtheit der zur Ausführung einer Reflexion erforderlichen kognitiven personalen Bedingungen.

7.4.2 Zum Begriff der Reflexivität

Die Begriffe Reflexion und Reflexivität werden teilweise synonym oder zumindest ohne präzise Abgrenzung verwendet (Forster, 2014). Häufig wird Reflexivität aber auch als (Reflexions-)Kompetenz verstanden, die „zu gelingenden Reflexionen“ (von Aufschnaiter, Fraij & Kost, 2019, S. 152)[79] führen kann (Cramer et al., 2019 ; Leonhard und Rihm, 2011). Häcker (2017) beschreibt Reflexivität auch auf individueller Ebene als „habitualisierte bzw. institutionalisierte Form“ (Häcker, 2017, S. 23) der Reflexion. Dieses Verständnis wird von anderen Autor*innen übernommen (z. B. Adl-Amini et al., 2022). Passend zum Verständnis von Reflexivität als (Handlungs-)Kompetenz wird Reflexivität als „Professionalisierungsziel“ (Pollmanns et al., 2022, S. 73) ausgerufen und, wie bereits einleitend zur Begründung der Relevanz von Reflexion dargelegt, Professionalisierung als „Steigerung von Reflexivität“ (Häcker, 2022, S. 94) übersetzt.

Gleichzeitig lassen sich andere Differenzierungen des Begriffspaars Reflexion und Reflexivität ausmachen, von denen drei exemplarisch dargestellt werden:

- Hilzensauer (2017) setzt die Begriffe Reflexivität und Reflexions*fähigkeit* gleich und versteht sie als „Voraussetzung für Reflexion“ (Hilzensauer, 2017, S. 13).[80]

[79] Ohne weitere Erläuterungen verwenden sie außerdem - augenscheinlich synonym - den Begriff der „reflexive[n] Kompetenz“ (von Aufschnaiter, Fraij & Kost, 2019, S. 153).

[80] Dass Reflexionsfähigkeit eine Voraussetzung für die Umsetzung von Reflexion ist, stimmt grundsätzlich mit dem Verständnis des Begriffs in dieser Arbeit überein; das Verständnis wird von Hilzensauer (2017) allerdings nicht weiter ausgeführt, sodass unklar bleibt, inwiefern die zugrundeliegenden Konzeptualisierungen übereinstimmen.

- Brendel (2017) versteht Reflexion als „ein Nachdenken über Unterrichtsinhalte, Überzeugungen oder Vorstellungen [...]. Reflexivität bezeichnet ein Nachdenken über eigene Handlungen oder Handlungsalternativen“ (Brendel, 2017, S. 19). Die Trennung entspricht damit eher der Unterscheidung zwischen externaler Zielstellung (nach Brendel (2017): Reflexion) und internaler Zielstellung (nach Brendel (2017): Reflexivität).
- Bohnsack (2020) zieht als Differenzierungsmerkmal zwischen Reflexion und Reflexivität das Ausbrechen aus Handlungsroutinen zurate (s. auch Cramer et al. (2019)). Reflexion beinhalte demnach im Gegensatz zur Reflexivität „[d]en Ausstieg aus dem bisherigen Zirkel und der bisherigen Rahmung bei gleichzeitigem Einstieg in eine/n neue/n“ (Bohnsack, 2020, S. 57).

Aufgrund des stark divergierenden Begriffsverständnisses wird im Rahmen dieser Arbeit auf eine weitere Einordnung und Verwendung des Begriffs der Reflexivität verzichtet. Stattdessen wird der Begriff der Reflexionsfähigkeit als Disposition für die Durchführung von Reflexionen verstanden.

7.5 Modellierung des Reflexionsbegriffs

Auch wenn die Reflexion überwiegend übereinstimmend als Denkprozess verstanden wird, sind verschiedene Ansätze bezüglich dessen konkreter Modellierung auszumachen. Unterschieden werden kann hier zwischen *Prozessmodellen*, die explizit die (nacheinander) ablaufenden Schritte einer Reflexion umfassen, und *Stufenmodellen*, die üblicherweise zur Qualitätsbewertung herangezogen werden.[81] Diverse Modelle stellen Kombinationen aus Prozess- und Stufenmodellen dar; sie können als *integrative Modelle* bezeichnet werden (Lenske & Lohse-Bossenz, 2023, S. 1141).

7.5.1 Prozessmodelle

Aus einer qualitativen Analyse von Beschreibungen der Prozesse in Definitionen des Begriffs Reflexion entwickelt Beauchamp (2006, S. 68ff.) sechs Kategorien von benannten (teilweise mehrstufigen) Denkprozessen. Sie beinhalten das...

- Betrachten einer Situation (*Examine*),
- schrittweise (1) Nachdenken über und (2) Verstehen einer Situation (*Step 1: Think, Step 2: Understand*),

[81] Teilweise wird die Entwicklung von Kompetenz- oder Kompetenzentwicklungsmodellen angestrebt, die sich durch ein empirisches Fundament auszeichnen, grundsätzlich aber Überschneidungen insb. zu Stufenmodellen aufweisen können (vgl. Hilzensauer, 2017, S. 49; Leonhard und Rihm, 2011; Abels, 2011, S. 96f.).

- Problemlösen (*Problem-solve*),
- kritische Analysieren (*Analyze*),
- Evaluieren (*Evaluate*) und
- die aufeinander aufbauenden Schritte (1) Strukturieren des Denkgegenstands, (2) Entwickeln neuer Perspektiven und (3) Transformieren des Denkgegenstands (*Step 1: Construct, Step 2: Develop, Step 3: Transform*).

Wie die mehrstufigen Kategorien bereits andeuten, wird Reflexion häufig als mehrschrittiger Prozess verstanden, in dem unterschiedliche Prozesse aufeinander folgen. Eine exemplarische Übersicht über einige verschiedene Modellierungen der letzten Jahrzehnte ist in Tabelle 7.1 dargestellt. Die Auswahl gibt einen Überblick über Modelle unterschiedlichen Umfangs und unterschiedlicher Feingliedrigkeit.[82] Die vorkommenden Bestandteile der Modelle sind in fünf Phasen gruppiert, wobei nicht alle Modelle alle Phasen beinhalten. Auslöser bzw. Grundlage für eine Reflexion ist eine erlebte Situation bzw. eine gemachte Erfahrung (Phase I). Diese Erfahrung wird nicht von allen Autor*innen in das Modell inkludiert, sie wird also nicht immer als Bestandteil von Reflexion, sondern ggf. ihr vorausgehend verstanden. In den Phasen II bis IV lassen sich die der Reflexion zugehörigen Denkaktivitäten finden. Zunächst findet ein Zurückblicken auf die Situation statt (Phase II), das deren deskriptive Darstellung umfasst (ggf. unter Berücksichtigung der durchlebten Emotionen (z. B. Gibbs, 1988, S. 49; Boud, Keogh und Walker, 1985)). Darauf folgt in Phase III die nähere Analyse[83] dieser Situation. Sie dient dazu, die zentralen Zusammenhänge in der Situation zu erkennen, zu verstehen und zu bewerten. Auf Grundlage der in der vertieften Analyse gewonnenen Erkenntnisse werden Handlungsoptionen zum Umgang mit der erlebten Situation entwickelt (Phase IV) und ggf. umgesetzt (Phase V). Insgesamt wird die Modellierung des Denkprozesses also in einen Handlungsprozess bzw. -zyklus (Erleben - Reflektieren (im Sinne des Denkprozesses) - Umsetzen (Erleben) der Handlungsoptionen) eingebettet (z. B. Aeppli und Lötscher, 2016; Korthagen, 1985). Die Prozessmodelle dürfen allerdings nicht als starr angesehen werden. Dewey (1910/2002) beschreibt die Schwierigkeit der Trennung insbesondere für das Erleben einer (problematischen) Situation (I) und einer ersten Interpretation (III). Auch kann nicht davon ausgegangen werden, dass in Reflexionen stets die modellierte Reihenfolge durchlaufen wird (Riel, 2022, S. 28; von Aufschnaiter, Fraij und Kost, 2019).

[82] Einige weitere Prozessmodelle listet z. B. Lee (2005) auf.

[83] Von Aufschnaiter, Fraij und Kost (2019) kritisieren den Begriff der Analyse an dieser Stelle, da „der gesamte Prozess einen analytischen Zugang abbildet“ (von Aufschnaiter, Fraij & Kost, 2019, S. 151).

	Dewey, 1910/2002, S. 56	Hullfish und Philip, 1963, S. 43f.	Gibbs (1988, S. 49f.)[84]	Copeland, Birmingham, de La Cruz und Lewin (1993, S. 350ff.)	Korthagen (1985, S. 12f.)	Aeppli und Lötscher (2016, S. 88ff.)	Kost (2019, S. 14f.)
			↻		↻		
I	Erfahrung	Presence (and recognition) of a problem situation	Experience		Action	Erleben	
II			Description		Looking Back	Darstellen	Beschreibung
			Feelings				
	spontane Interpretation	Clarification of the problem	Evaluation	Problem Identification			Deutung
III	Benennung der Probleme		Analysis		Awareness	Analysieren	Ursachenforschung
	Entwickeln möglicher Erklärungen		Conclusion				
IV		Hypotheses formed, tested and modified	Action plan	Generating Solutions	Creating alternative Actions	Maßnahmen entwickeln, planen	Konsequenzen ableiten
V		Action taken	Experimentation	Testing Solutions	Trial	Anwenden	(Konsequenzen Umsetzen[85])

Tabelle 7.1: Übersicht über Prozessmodelle zur (Unterrichts-)Reflexion (mit ↻ markierte Modelle sind explizit zyklisch modelliert, die Umsetzung einer Handlung löst demnach eine neue Reflexion aus)

[84] Die Bezeichnung als „experiential learning cycle“ (Gibbs, 1988, S. 49) lässt nicht automatisch auf eine Modellierung des Reflexionsbegriffs schließen. Die Entsprechung wird aber vor dem Hintergrund des Verständnisses von Reflexion als Lernen aus Erfahrungen (z. B. McAlpine et al., 1999, S. 116) nachvollziehbar.

[85] Ergänzung nach von Aufschnaiter, Fraij und Kost (2019, S. 151)

7.5.2 Stufenmodelle

Die Verwendung von Stufenmodellen dient üblicherweise zur Einschätzung der Qualität von (meist schriftlich verfassten) Reflexionen, indem „levels of reflection“ (Poldner et al., 2014, S. 354) definiert werden.[86] Ein viel rezipiertes Modell stammt von van Manen (1977). Er unterscheidet in einem theoretischen Modell drei hierarchisch gestufte „levels of reflexivity“ (van Manen, 1977, S. 226f.); diese werden von diversen Autoren verwendet, um die Qualität von Reflexionen zu beurteilen (Poldner et al., 2014):

- Stufe 1, die „[t]echnical rationality“ (Poldner et al., 2014, S. 351; Ballard, 2006, S. 7; auch „technical reflection“ (Hatton & Smith, 1995, S. 35)), umfasst die Anwendung von Wissen, um ein feststehendes Ziel zu erreichen (van Manen, 1977).
- In Stufe 2, die „practical action“ (van Manen, 1977, S. 226; auch „practical reflection“ (Leijen, Valtna, Leijen und Pedaste, 2012, S. 205; Hatton und Smith, 1995, S. 35)), wird unterrichtliches Handeln auf Grundlage einer Analyse der gegebenen Rahmenbedingungen und Voraussetzungen ausgewählt und umgesetzt (van Manen, 1977). Auch die Ziele, die erreicht werden sollen, werden hinterfragt (Hatton & Smith, 1995).
- Stufe 3, die „critical reflection“ (Ballard, 2006, S. 7; Hatton und Smith, 1995, S. 35; auch „[c]ritical rationality“ (Poldner et al., 2014, S. 351)) geht über die unteren Phasen hinaus, indem zusätzlich moralische und ethische Überlegungen in der Einschätzung berücksichtigt werden (van Manen, 1977).

Kritik an van Manens (1977) Modell wird aufgrund des hierarchischen Verständnisses der Stufen geübt, die eine lineare Entwicklung der Reflexionsfähigkeit suggerieren (Poldner et al., 2014). In Abgrenzung dazu unterscheiden z. B. Hatton und Smith (1995) drei[87] und Valli (1997) fünf Stufen[88], die sie als Typen bzw. Orientierungen von Reflexion bezeichnen. Sie halten es für wichtig, dass diese nicht als „increasingly desirable hierarchy“ (Hatton & Smith, 1995, S. 35) verstanden werden, obwohl sie teilweise aufeinander aufbauen können (Valli, 1997), sondern

[86] Poldner et al. (2014) geben eine Übersicht über die in 40 Publikationen eingesetzten Modelle zur Einstufung der Stufe bzw. Kategorisierung von Reflexionen.

[87] Hatton und Smith (1995) unterscheiden die Formen *descriptive writing* (die reine Wiedergabe von Situationen, die sie nicht als Reflexion einordnen), *descriptive reflection* (die Schilderung von Situationen mit persönlichen Begründungen), *dialogic reflection* (die Diskussion und Abwägung verschiedener Argumente zur geschilderten Situation) und *critical reflection* (die Argumentation unter Einbezug größerer z. B. politischer oder sozialer Kontexte).

[88] Sie unterscheidet *technical reflection* (bezüglich der Umsetzung gesicherter empirischer Erkenntnisse), *reflection-in- and -on-action* (bezüglich des eigenen Unterrichtens), *deliberative reflection* (bezüglich verschiedenster unterrichtlicher Belange wie z. B. Instruktionsstrategien), *personalistic reflection* (bezüglich der individuellen Weiterentwicklung) und *critical reflection* (unter Einbezug sozialer, moralischer und ethischer Dimensionen des Unterrichtens (Valli, 1997, S. 75)).

situationsabhängig unterschiedliche Typen oder Kombinationen verschiedener Typen adäquat sind (Hatton und Smith, 1995; Valli, 1997).

Das Erreichen höherer Stufen geht mit zunehmender Komplexität oder Abstraktion der Reflexion einher. Diese kann aufbauend auf vorliegende Modellierungen des Reflexionsprozesses einerseits durch eine Zunahme an notwendigen Prozessschritten erreicht werden. So formulieren bspw. Poldner et al. (2014) (aufbauend auf Leijen, Valtna, Leijen und Pedaste (2012)) fünf Stufen von *1. Description* hin zu *5. Description, Evaluation, Justification, Dialogue, Transfer* (je Stufe kommt ein Prozessschritt in dieser Reihenfolge hinzu). Windt und Lenske (2016) unterscheiden in der Kodierung zwischen den Elementen *Beschreibungen*, *Bewertungen*, *Begründungen*, *Alternativen* und *Konsequenzen*; eine höhere Stufe wird auch hier mit zunehmender Anzahl an verschiedenen auftretenden Elementen erreicht. Im Gegensatz zum Modell von Poldner et al. (2014) ist die Auswahl der tatsächlich auftretenden Prozessschritte nicht bedeutsam, es wird lediglich die Anzahl unterschiedlicher Elemente bzw. Prozessschritte gezählt.[89] Andererseits kann die Stufung unmittelbar entlang von (distinkten) Prozessschritten vorgenommen werden, die in ihrer Komplexität zunehmen. Hierunter fallen z. B. die Unterscheidung zwischen *descriptive*, *dialogical* und *critical reflection* nach Hatton und Smith (1995, S. 40f.) sowie die Unterscheidung von *Description (Beschreibung)*, *Justification (Erläuterung mit Begründung)*, *Critique (Bewertung mit Begründung)* und *Discussion (Abwägung unter Berücksichtigung von Alternativen bzw. einer Zukunftsperspektive)* nach Leijen, Valtna, Leijen und Pedaste (2012) oder Lees (2005, S. 703) Ausdifferenzierung eines rein deskriptiven *Recall level*, eines *Rationalization level*, das die Erklärung und Generalisierung der Situation anstrebt und des *Reflectivity level*, das insbesondere eine zukunftsgerichtete Perspektive aufweist. Adl-Amini et al., 2022 nehmen abweichend eine Stufung der Reflexionsqualität quer zum Reflexionsprozess vor: Alle Prozessschritte (*Beschreibung*, *Deutung*, Entwicklung einer *Fragestellung*, Einbezug *wissenschaftlichen Wissens* sowie Ziehen von *Schlussfolgerungen*) können in vier unterschiedlich komplexen Stufen vorliegen, die sich durch den Einbezug eines zunehmend größeren Kontexts auszeichnen (von *subjektiv* bis *gesellschaftskritisch* (Adl-Amini et al., 2022, S. 26)).

Wie dieser exemplarische Überblick über verschiedene Stufenmodelle verdeutlicht, weisen sie Unterschiede in Bezug darauf auf, ob sie als hierarchisch zu verstehen sind und ob (ggf. hierarchische) Stufen sich durch eine Zunahme an Prozessschritten auszeichnen oder z. B. distinkte Prozesse darstellen. Darüber hinaus erfolgt auch der Einsatz von Stufenmodellen zur Bewertung von (z. B. schriftlich oder mündlich verbalisierten) Reflexionen nicht konsistent. Üblicherweise wird die Qualität einzelner Textsegmente bewertet (z. B. Leijen, Valtna, Leijen und Pedaste, 2012; Hatton und Smith, 1995). Anschließend kann daraus ein holistisches Urteil gebildet werden. So kategorisieren bspw. Poldner et al. (2014) zunächst einzelne Segmente (in Beschreibung, Bewertung, etc.) und stufen den Text anschließend

[89] Sie sprechen hier vom Kriterium der „Vollständigkeit“ (Windt & Lenske, 2016, S. 285).

abhängig von den aufgetretenen Kategorien ein (ähnlich gehen auch z. B. Kempin, Kulgemeyer und Schecker (2020) bei der Bewertung mündlicher Reflexionen vor).

7.5.3 Weitere Kriterien zur Bewertung der Qualität von Reflexionen

Wie bereits beschrieben, dienen häufig Stufenmodelle zur Beurteilung der Qualität von Reflexionen. Sie ermöglichen eine Einschätzung der *Reflexionstiefe*, d. h. die Intensität der Auseinandersetzung mit dem Inhalt der Reflexion. Eine umfassende Beurteilung der Qualität ist allerdings komplexer und erfolgt häufig unter Berücksichtigung zusätzlicher Merkmale. Üblich ist die zusätzliche Einschätzung der Reflexionsbreite (z. B. Leijen, Valtna, Leijen und Pedaste, 2012; Leonhard und Rihm, 2011; Riel, 2022, S. 71ff.), die sich auf die Anzahl der reflektierten Aspekte (Denkgegenstände) bezieht (Mientus, 2023, S. 18).

Darüber hinaus werden diverse weitere Kriterien für die Bewertung von Reflexionen vorgeschlagen, u. a. (ohne Anspruch auf Vollständigkeit) die folgenden:

- Die *Vollständigkeit* einer Reflexion beschreibt das Vorkommen aller Prozessschritte bzw. Denkaktivitäten (Windt und Lenske, 2016).
- Das Kriterium der *Strukturiertheit* fordert, dass die einzelnen Denkaktivitäten in modellierter Reihenfolge auftreten (Mientus, Wulff, Nowak und Borowski, 2023; Nowak, Kempin, Kulgemeyer und Borowski, 2019; Windt und Lenske, 2016)
- Für Reflexionstexte schlägt Mientus (2023, S. 18) die Berücksichtigung von *Kohärenz* (der logischen Argumentationsstruktur; auch von Aufschnaiter, Fraij und Kost (2019)) vor sowie der *Spezifität* (der Bezugnahme auf die reflektierte Situation bzw. den konkreten Denkgegenstand; auch von Aufschnaiter, Fraij und Kost (2019)) als Qualitätsmerkmale. Inhaltlich ähnlich betrachten Moore-Russo und Wilsey (2014) die Komplexität und inhaltliche Verbundenheit der Bestandteile der Reflexion als unterschiedliche (Qualitäts-)Dimensionen einer Reflexion.
- Die *Mehrperspektivität* umfasst den Einbezug unterschiedlicher „(gleichwertiger) alternativer Deutungen, Ursachen und Konsequenzen“ (von Aufschnaiter, Fraij und Kost, 2019, S. 154; zur Mehrperspektivität auch z. B. Abels (2011, S. 148) und Adl-Amini et al. (2022)) und setzt die Fähigkeit zum Perspektivwechsel voraus (Cramer et al., 2019; Hatton und Smith, 1995).
- Weiterhin kann das Vorliegen bzw. der Grad eines *Theorie- bzw. Empiriebezugs* bewertet werden (von Aufschnaiter, Fraij und Kost, 2019; Adl-Amini et al., 2022).

- Von Aufschnaiter, Fraij und Kost (2019) schlagen eine Berücksichtigung des *Grads des hergestellten Selbstbezugs* vor, was insbesondere dann bedeutsam erscheint, wenn der Reflexionsbegriff mit einer internalen Zielstellung definiert wird.

Windt und Lenske (2016) erfassen darüber hinaus Umfang[90] und Inhalt einer Reflexion sowie deren Realitätstreue als weitere mögliche Einflussfaktoren auf die Qualität von Reflexionen.

Die Kombination verschiedener Qualitätskriterien wird teilweise in den oben beschriebenen Stufenmodellen integriert. So gehen z. B. die Reflexionsniveaus nach Hatton und Smith (1995) mit dem Einbezug einer zusätzlichen (ethischen) Perspektive einher und im Modell von Jahncke, Berding, Porath und Magh (2018) werden Bewertungen in höheren Niveaustufen um (theoretische) Begründungen ergänzt. Alternativ können verschiedene Kriterien in mehrdimensionalen Modellen umgesetzt werden (z. B. unterscheiden Poldner et al. (2014) die Dimensionen Inhalt und Komplexität einer Reflexion). In diesen Fällen werden die Einschätzungen zu unterschiedlichen Qualitätskriterien nicht immer zu einem Gesamturteil kombiniert, sodass keine globale Einschätzung der Qualität vorgenommen wird (z. B. Meier, Vogelsang, Watson und Schaper, 2022; Wyss, 2013, S. 124ff.).

7.5.4 Reflexionsmodell nach Nowak, Kempin, Kulgemeyer und Borowski (2019)

Auf Grundlage verschiedener anderer Stufenmodelle entwickelten Nowak, Kempin, Kulgemeyer und Borowski (2019) ein Modell zur Bewertung der Güte von Reflexionen über Physikunterricht, das Elemente diverser anderer Modellierungen bzw. Operationalisierungen vereint (Nowak, 2023, S. 27).[91] Im Rahmen des Modells werden die nachfolgend näher charakterisierten drei Kriterien *Elemente einer Reflexion*, das Vorliegen einer *Begründung* und die einer Aussage zugrundeliegende *Wissensbasis* berücksichtigt und als dreidimensionales Modell zusammengefasst (s. Abbildung 7.1). Mit der Benennung des Modells ist von einer Einschränkung der Denkgegenstände von Reflexionen auf solche mit Bezug zum Unterricht auszugehen, die im Rahmen des bisher gegebenen Überblicks über Merkmale und Modellierungen von Reflexionen nicht vorlag.

[90] Einen Zusammenhang zwischen Umfang von schriftlichen Reflexionen und deren Qualität stellen Leonhard und Rihm (2011) für das Kriterium der Reflexionsbreite fest.

[91] Nowak, Kempin, Kulgemeyer und Borowski (2019) benennen als Grundlage explizit die Stufenmodelle von Windt und Lenske (2016) und Plöger, Scholl und Seifert (2015) und das zyklische *ALACT*-Modell (z. B. Korthagen und Vasalos, 2005). Vorgängerversionen des Modells finden sich in Kempin, Kulgemeyer und Schecker (2018), Nowak, Liepertz und Borowski (2018) und Nowak, Ackermann und Borowski (2018).

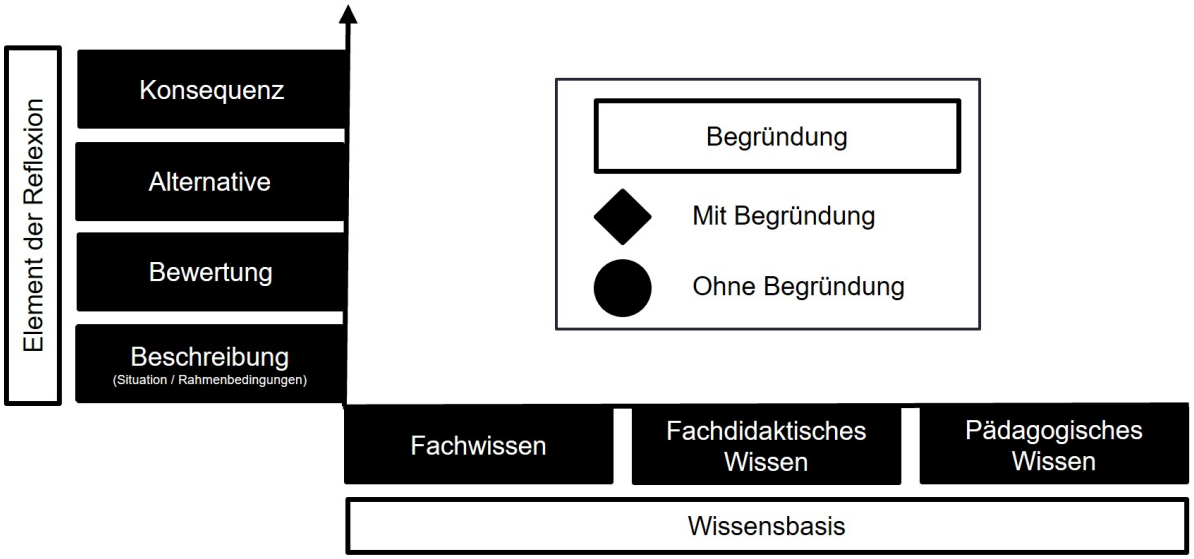


Abbildung 7.1: Modell zur Bewertung der Qualität von Reflexionen von Physikunterricht in drei Dimensionen nach Nowak, Kempin, Kulgemeyer und Borowski (2019, S. 839)

Dimension 1: Element der Reflexion. Unterschieden wird zwischen vier Elementen einer Reflexion, die in hierarchisch in Stufen angeordnet sind.

- Stufe 1: Die (möglichst objektive) *Beschreibung* der erlebten oder beobachteten Unterrichtssituation bzw. zugrundeliegender Rahmenbedingungen (z. B. Ziele des Unterrichts oder Eigenschaften der Lerngruppe)
- Stufe 2: Die *Bewertung* der (beschriebenen) Unterrichtssituation
- Stufe 3: Die Entwicklung von *Alternativen* für die (beschriebene und bzw. oder bewertete) Situation; für positiv bewertete Situationen kann das Beibehalten der durchgeführten bzw. beobachteten Vorgehensweise empfohlen werden
- Stufe 4: Das Ableiten von *Konsequenzen* für (a) den Folgeunterricht, (b) die Weiterentwicklung des professionellen Handelns der unterrichtenden Person und bzw. oder - im Falle einer Fremdreflexion - (c) des professionellen Handelns der beobachtenden Person

Dimension 2: Begründung. In der Dimension Begründung wird unterschieden, ob eine Aussage der Stufen 2 bis 4 (Bewertung, Alternative oder Konsequenz) begründet wird. Begründungen können auf unterschiedlicher Grundlage erfolgen und bspw. erfahrungs- oder theoriebasiert sein. Dies wird bei der Beurteilung dieser Dimension allerdings nicht berücksichtigt; es wird dichotom eingestuft, ob eine Begründung vorliegt.

Dimension 3: Wissensbasis. In der dritten Dimension wird einer Reflexion die zugrundeliegende Wissensbasis zugeordnet. Hierbei wird zwischen den drei

Wissensbasen *Fachwissen*, *fachdidaktisches Wissen* und *pädagogisches Wissen* unterschieden. Die Fachspezifität, die in der Benennung des Modells zur Reflexion von *Physik*unterricht angelegt wird, lässt sich mit der Festlegung der fachspezifischen Wissensbasen auf physikalisches Fachwissen und physikdidaktisches Wissen begründen.

Die Entwicklung des Modells war mit zwei Zielen verbunden: Einerseits dient es der Unterstützung Studierender bei der Reflexion im Rahmen des Projekts *PSI-Potsdam*[92]. Studierenden wird das entwickelte Reflexionsmodell vorgestellt und gemeinsam mit Leitfragen für die Selbstreflexion (z. B. bezüglich der Rahmenbedingungen „Warum habe ich die Methode/das Experiment/... ausgewählt? Was war mein Ziel?“ oder „Was hätte ich anders machen können und warum wäre dieses Handeln erfolgreicher gewesen?“ zur Findung von Alternativen (Nowak, Liepertz & Borowski, 2018, S. 223)) als Hilfestellung für das Verfassen eigener Reflexionen zur Verfügung gestellt (Nowak, Liepertz & Borowski, 2018). Andererseits bildet das Modell die Grundlage für die Einschätzung der Qualität von mündlichen oder schriftlichen Reflexionen in den Projekten *PSI-Potsdam* und *ProfiLeP+* (Nowak, Kempin, Kulgemeyer & Borowski, 2019). Die Umsetzung der Qualitätsbeurteilung erfolgt in den Projekten allerdings unterschiedlich. Im Rahmen der Begleitveranstaltung zum Praxissemester Physik bewerteten Nowak, Liepertz und Borowski (2018) schriftliche Selbstreflexionen der Studierenden in Bezug auf das Vorliegen der im Reflexionsmodell enthaltenen Elemente einer Reflexion, der Strukturiertheit der Reflexion (d. h. das Auftreten der Elemente mit Begründung in der modellierten Reihenfolge), deren Vollständigkeit (d. h. das Auftreten aller Elemente) sowie die innere Logik der Reflexion. Sie ziehen also zusätzliche Kriterien für die Bewertung der Reflexionen hinzu.[93] Kempin, Kulgemeyer und Schecker (2019) nutzen das Modell als Grundlage für die Bewertung von mündlichen Reflexionen in einem Reflexionsperformanztest (s. Kapitel 7.7.3). Für unterschiedliche Aspekte, die die Reflexionsbreite darstellen und unterschiedlichen Wissensbasen zugeordnet werden können, werden die Äußerungen der Studierenden basierend auf der höchsten angesprochenen Stufe bewertet; die Vollständigkeit oder die Reihenfolge der Reflexionselemente werden hier nicht berücksichtigt.

In Bezug auf die in Kapitel 7.5.2 beschriebenen Unterscheidungsmerkmale von Stufenmodellen lässt sich festhalten, dass dieses Reflexionsmodell distinkte Denkprozesse hierarchisch anordnet (eine Alternative ist höherwertiger als eine Bewertung und diese wiederum einer Beschreibung überlegen). Zudem dient das Modell dazu,

[92] *Professionalisierung - Schulpraktische Studien - Inklusion: Potsdamer Modell der Lehrerbildung*

[93] Auf Grundlage dieser Untersuchung schlug Nowak (2023, S. 157ff.) kürzlich eine überarbeitete Version des Reflexionsmodells von Nowak, Kempin, Kulgemeyer und Borowski, 2019 vor, in welcher insbesondere der Bedarf nach Handlungsalternativen bzw. Konsequenzen zu positiven Reflexionsauslösern infrage gestellt wird: Ziel ist in diesem Falle nicht die Weiterentwicklung von Unterricht bzw. der professionellen Kompetenzen, sondern die Zufriedenheit bezüglich des Unterrichts bzw. der eigenen professionellen Kompetenzen (Nowak, 2023, S. 160f.).

einzelne Aussagen bzw. Textsegmente einzustufen. Eine umfassende Bewertung einer Reflexion erfolgt dann auf Grundlage der feineren Einstufungen und ggf. unter Berücksichtigung weiterer Kriterien (u. a. der Vollständigkeit (Nowak, Liepertz & Borowski, 2018)). Die hierarchisch geordneten Elemente einer Reflexion (Beschreibung, Bewertung, Alternative und Konsequenz) stellen als prototypische Reihenfolge für eine Reflexion in gewisser Weise auch ein Prozessmodell dar, das allerdings (insbesondere in Kombination mit dem von Nowak, Liepertz und Borowski (2018) herangezogenen Bewertungskriterium der Strukturiertheit) als rein normativ zu verstehen ist.[94]

7.6 Zum Verhältnis von Reflexion, Noticing und professioneller Unterrichtswahrnehmung

Allgemein ist das Erleben einer (problembehafteten) Situation bzw. eine gemachte Erfahrung Auslöser für die Reflexionen in den verschiedenen (Prozess-)Modellen. Zentral für die Reflexion bedeutsamer Aspekte von Unterricht ist es, diese im Unterrichtsgeschehen zu identifizieren (Göbel & Gösch, 2019). Gleichzeitig wird dieser Schritt nur von einigen Autor*innen als Bestandteil der Reflexion definiert (z. B. von Weber, Neuber und Prilop (2023) und Kost, Kirschner und von Aufschnaiter (2016)) und ist häufig kein expliziter Bestandteil von (Prozess-)Modellen zur Reflexion.[95] Gleiches gilt für das zuletzt ausführlicher vorgestellte Modell von Nowak, Kempin, Kulgemeyer und Borowski (2019). Der Reflexionsprozess beginnt demnach mit der Beschreibung des spezifischen Aspekts bzw. einer Situation. Die Auswahl dieser Situation ist dann kein inhärenter Teil der Reflexion, sondern muss als der Reflexion vorgelagert verstanden werden (Lam und Chan, 2020; Wöhlke, 2020, S. 31).

Zur theoretischen Rahmung dieses Analyseschritts wird auf das Konzept des *Noticing* (auch „Selective Attention“ (Wöhlke, 2020, S. 12) bzw. „selektive Wahrnehmung“ (Weber, Neuber & Prilop, 2023, S. 1237); „call-out“ oder „highlighting“; (van Es & Sherin, 2008, S. 245f.)) zurückgegriffen. Der Begriff des Noticing wird in der Literatur uneindeutig verstanden und genutzt (z. B. Weyers et al., 2023; Weston und Amador, 2023) und teilweise sogar mit professioneller Unterrichtswahrnehmung gleichgesetzt (Chan et al., 2021)). Für diese Arbeit wird dem Verständnis nach Seidel, Blomberg und Stürmer (2010) gefolgt:

[94] Da der tatsächliche Ablauf von Reflexionsprozessen bislang nicht empirisch belegt ist, gilt dies allerdings auch für übrige Prozessmodelle (auch wenn sie nicht zur Bewertung von Reflexionen herangezogen werden; Riel, 2022, S. 29).

[95] Für Modellierungen des Reflexionsprozesses, die das Erleben der Situation bzw. des „Auslöser[s] für den Reflexionsanlass“ (Aeppli & Lötscher, 2016, S. 88) als Teil des Reflexionsprozesses verstehen (vgl. Tabelle 7.1), ist allerdings auch ohne explizite Benennung davon auszugehen, dass auch die Auswahl der bedeutsamen Aspekte unter den Reflexionsbegriff subsumiert wird, da die Auswahl zwischen dem Erleben und der spezifischen Analyse anzusiedeln ist.

> Noticing ist die „wissensgesteuerte Identifikation relevanter Situationen und Ereignisse im Unterrichtsgeschehen“ (Seidel, Blomberg & Stürmer, 2010, S. 297)

Mit Noticing einher geht eine Aufmerksamkeitsfokussierung auf die relevanten Ereignisse im Unterrichtsgeschehen (Wöhlke, 2020, S. 24). In diesem Sinne kann Noticing als Teil der professionellen Unterrichtswahrnehmung verstanden werden. Diese wird üblicherweise weiter ausdifferenziert in die beiden Bestandteile *Noticing* und *Knowledge-based Reasoning*, der „[w]issensgesteuerten Verarbeitung von Unterricht“ (Seidel, Blomberg & Stürmer, 2010, S. 297).[96] Knowledge-based Reasoning umfasst wiederum die wissensbasierte Beschreibung und Erklärung der Situation sowie die Vorhersage von Wirkungen der Situation. Teilweise wird darüber hinaus die Entscheidung über die Umsetzung einer Handlung als Komponente des Knowledge-based Reasoning beschrieben (Wöhlke, 2020, S. 14f.).[97]

Während Noticing als Voraussetzung für die Reflexion bzw. für Knowledge-based Reasoning relativ klar von diesen beiden Prozessen abgegrenzt werden kann, sind die Reflexion und Knowledge-based Reasoning weniger trennscharf. Van Es und Sherin (2008) bezeichnen die professionelle Unterrichtswahrnehmung sogar als „one important dimension within the process of reflection“ (van Es & Sherin, 2008, S. 247). Auch für Kost, Kirschner und von Aufschnaiter (2016) beinhaltet die Reflexion die professionelle Unterrichtswahrnehmung und zusätzlich das Ableiten von Konsequenzen, das aus ihrer Perspektive über das Knowledge-based Reasoning hinausgeht. Wöhlke (2020, S. 31) deutet ein mögliches Unterscheidungsmerkmal im zeitlichen Bezug zur Unterrichtssituation an: Während die professionelle Unterrichtswahrnehmung (insbesondere auch das Knowledge-based Reasoning) sich unmittelbar an die Situation anschließt, kann eine Reflexion auch mit zeitlichem Verzug (als *reflection-on-action* erfolgen). Dieses Kriterium ermöglicht allerdings keine klare Abgrenzung zur *reflection-in-action*. Reflexion und Knowledge-based Reasoning sind theorie- bzw. wissensbasierte Analyseprozesse einer Situation, die zu Schlussfolgerungen z. B. zum weiteren Handeln in der Situation führen können. Dabei legt die professionelle Unterrichtswahrnehmung „einen stärkeren Fokus auf Tätigkeiten der Lernenden, die orchestriert und unterstützt werden müssen, statt auf die eigene Person.“ (Wöhlke, 2020, S. 31). Daraus folgt eine Überschneidung von Reflexion und Knowledge-based Reasoning in Bezug auf „reflexive Prozesse, die zu einer Regulierung des eigenen Handelns zur Förderung von Lernen der Lernenden führen.“ (Wöhlke, 2020, S. 33).

[96] Die begriffliche Unterscheidung von Noticing und *Knowledge-Based*-Reasoning ist irreführend, da beide Prozesse als wissensbasiert verstanden werden.

[97] Van Es und Sherin (2008) sowie darauf aufbauend einige andere Autor*innen verwenden den Begriff *Noticing* vergleichbar zur professionellen Unterrichtswahrnehmung insgesamt (Weyers et al., 2023); sie unterscheiden trotzdem zwischen der Identifikation bedeutsamer Elemente eines Unterrichts und der theoriebasierten Verarbeitung dessen (hier *Knowledge-based Reasoning*) sowie dem Ziehen von Schlussfolgerungen für folgende Handlungen.

Die Frage, wie die Fähigkeit zum Noticing bzw. zur professionellen Unterrichtswahrnehmung mit anderen Aspekten der professionellen Handlungskompetenz zusammenhängt bzw. inwieweit Bestandteile der Handlungskompetenz eine Voraussetzung für sie darstellen, stellt ein Forschungsdesiderat dar (Amador & Weston, 2024). Bekannt ist, dass erfahrenen Lehrkräften Noticing schneller, umfassender und präziser gelingt (Kaiser et al., 2015); Studierende erkennen im Gegensatz zu erfahrenen Lehrkräften nur einen Teil der relevanten Aspekte (Weber, Neuber & Prilop, 2023). Auch das (fachdidaktische) Professionswissen der Lehrkräfte erscheint bedeutsam für die professionelle Unterrichtswahrnehmung (Luft et al., 2024). Darüber hinaus stellen Weber, Neuber und Prilop (2023) fest, dass das deklarative Wissen einen Prädiktor für das Noticing, das konzeptuelle Wissen für die weitere Reflexion (bezeichnet als „wissensgesteuerte Verarbeitung“ (Weber, Neuber & Prilop, 2023, S. 1251)) darstellt.

7.7 Empirische Ergebnisse der Erfassung der Reflexionskompetenz

Wie schon allgemeiner in Bezug auf die Erfassung der professionellen Handlungskompetenz beschrieben, kommen auch bei der Erfassung der Reflexionsfähigkeit und verwandter Konstrukte unterschiedliche Methoden zum Einsatz. Verbreitet ist der Einsatz von Fragebögen zur Selbsteinschätzung sowie die Untersuchung schriftlicher oder mündlicher Reflexionen. In diesem Kapitel soll daher ein Überblick über aktuelle Instrumente und Forschungsstände mit einem Fokus auf die Erfassung der Reflexionsfähigkeit in der ersten (universitären) Ausbildungsphase gegeben werden.

7.7.1 Selbstauskünfte zur Reflexionsbereitschaft und -fähigkeit

Selbsteinschätzungen, häufig in Form von Fragebögen, werden primär zur Erfassung affektiv-motivationaler Aspekte (wie z. B. der Einstellung gegenüber Reflexion) oder der Auskunft über gelebte Reflexionspraxis eingesetzt. Als affektiv-motivationale Komponente der Reflexionskompetenz wird die Reflexionsbereitschaft als Disposition für Reflexion verstanden (Stender, Watson, Vogelsang & Schaper, 2021). Die Bedeutung affektiv-motivationaler Aspekte wird darin sichtbar, dass sie Anteile der Varianz in der Qualität studentischer Reflexionen aufklären kann (Stender, Watson, Vogelsang & Schaper, 2021). Die Bereitschaft zur Ausführung und Darlegung einer Reflexion ist Voraussetzung für die Erfassung der Reflexionsfähigkeit (Leonhard & Rihm, 2011).[98]

[98] Leonhard und Rihm (2011) benennen die Bereitschaft zur Artikulation allgemeiner als Voraussetzung für die Erfassung der Reflexions*kompetenz*. Gleiches gilt aber auch für die

Im Projekt *ScRiPS*[99] wurden in relativ großem Umfang Fragebögen zur Erfassung von Reflexionsbereitschaft sowie der Reflexionspraxis in einem quasiexperimentellen Design eingesetzt, in dem die Interventionsgruppen Reflexionen in verschiedenen Settings durchführen (schriftlich bzw. als kollegiale Reflexion) (Neuber & Göbel, 2018). Für die Reflexionsbereitschaft extrahieren sie auf Grundlage einer Befragung von $N = 245$ Lehramtsstudierenden und $N = 218$ Lehrpersonen im Schuldienst die Skalen zu (1) Einstellungen zur systematischen Reflexion (z. B. „Ich halte es für sinnvoll, Strukturierungshilfen [...] einzusetzen, um den individuellen Reflexionsprozess über den Unterricht zu unterstützen“ (Neuber & Göbel, 2018, S. 12)), (2) Einstellungen zur kollegialen Reflexion (z. B. „Ich halte es für sinnvoll, sich von seinen Kolleginnen und Kollegen eine Rückmeldung zum eigenen Unterricht einzuholen.“ (Neuber & Göbel, 2018, S. 13)), (3) Einstellungen zu Schülerrückmeldungen (z. B. „Es ist mir wichtig zu erfahren, was Schülerinnen und Schüler über meinen Unterricht denken.“ (Neuber & Göbel, 2018, S. 14)), (4) Präferenz der individuellen, selbstbezogenen Reflexion (z. B. „Alleine kann man am besten über seinen Unterricht nachdenken.“ (Neuber & Göbel, 2018, S. 15)) und (5) Relevanz von Unterrichtsreflexion im Lehrerberuf (z. B. „Ich denke, eine Lehrkraft hat wichtigere Aufgaben, als ständig zu reflektieren.“ (Neuber & Göbel, 2018, S. 16)). Für beide Gruppen ergeben sich für jede Skala zufriedenstellende Reliabilitäten von mindestens $\alpha_{\text{Cronbach}} = 0,71$ (Neuber & Göbel, 2018, S. 9ff.).

Für den Bereich der Reflexionspraxis ergeben sich aus der Befragung ausschließlich der im Schuldienst aktiven Lehrkräfte die drei Skalen (1) mangelnde Zeit für Reflexion (z. B. „Ich habe zu wenig Zeit, um meinen Unterricht regelmäßig zu reflektieren“(Neuber & Göbel, 2018, S. 21)), (2) kollegiale Reflexionspraxis (z. B. „Ich reflektiere meinen Unterricht regelmäßig gemeinsam mit meinen Kolleginnen und Kollegen.“ (Neuber & Göbel, 2018, S. 22)) und (3) Nutzung von Schülerrückmeldungen („Wenn ich systematisch über meinen Unterricht nachdenke, hole ich mir vorab eine mündliche Rückmeldung von meinen Schülerinnen und Schülern ein.“(Neuber & Göbel, 2018, S. 22)). Auch hier ergeben sich - mit Ausnahme der Skala zur Nutzung von Schülerrückmeldungen - gute Reliabilitäten von $\alpha_{\text{Cronbach}} \geq 0,81$ (Neuber & Göbel, 2018, S. 19ff.).

Insgesamt können im Projekt *ScRiPS* überwiegend positive Einstellungen gegenüber Unterrichtsreflexion festgestellt werden (Neuber & Göbel, 2019). Dies steht in Einklang mit Ergebnissen anderer Studien, die die Reflexionsbereitschaft von Studierenden und Lehrkräften sowie Selbstauskünfte erheben (z. B. Stender, Watson, Vogelsang und Schaper, 2021 und Fraij (2018) (mit einer Adaption des Fragebogens von Neuber und Göbel (2018))[100]; Wyss, 2013, S. 154f.) und passt

Reflexionsfähigkeit, die einen Teil der Reflexionskompetenz darstellt (vgl. Leonhard und Rihm, 2011).

[99] *Schülerrückmeldungen zum Unterricht und ihr Beitrag zur Unterrichtsreflexion im Praxissemester*

[100] Fraij (2018) bewertet die Ergebnisse zwar nicht explizit, berichtet aber überwiegend Zustimmungswerte um den bzw. über dem theoretischen Mittelwert der Skala.

zu den Selbstauskünften von Lehrkräften, regelmäßig zu reflektieren (Wyss, 2013, S. 249f.).

Kost (2019, S. 34ff.) entwickelte einen Fragebogen u. a. zur Selbsteinschätzung der Reflexionsfähigkeit. Die Items zur Selbsteinschätzung der Reflexionsfähigkeit beziehen sich auf die Aspekte (1) Analyse von Unterricht (z. B. „Nach dem Unterricht kann ich analysieren, ob ich fachliche Fehler gemacht habe“ (Kost, 2019, S. 152)), (2) die Ursachenforschung nach der Analyse (z. B. „Ich kann anhand meiner Unterrichtsanalyse ableiten, ob die Experimente hinsichtlich Schülervorstellungen angemessen geplant waren.“ (Kost, 2019, S. 153)) und (3) Konsequenzen für die Unterrichtsplanung (z. B. „Ich bin aufgrund der Analyse meines Unterrichts in der Lage, mein Lehrerverhalten besser zu planen.“ (Kost, 2019, S. 154)). Im Gegensatz zu den bislang berichteten Einstellungen zur Reflexion und Auskünften über die Häufigkeit der Reflexionen könnten hier verstärkt Aussagen über die Qualität der Reflexionen bzw. ihrer Bestandteile (wie die Benennung von Ursachen oder das Ableiten von Konsequenzen) abgeleitet werden. Er nutzt diese Selbsteinschätzung allerdings lediglich zur näheren Charakterisierung der Stichprobe als heterogen (Kost, 2019, S. 56f.).

Während Selbsteinschätzungen zwar Auskunft über affektiv-motivationale Aspekte liefern könnten, sollten sie (wie in Kapitel 6.2 für Fähigkeiten im Allgemeinen beschrieben) nicht als objektives Maß der Reflexionsfähigkeit interpretiert werden, da fraglich ist, inwiefern sie tatsächlich die interessierende Fähigkeit erfassen (z. B. König, Kaiser und Felbrich, 2012).

7.7.2 Verfahren zur Messung der Reflexionsfähigkeit

Auch zur Messung der Reflexionsfähigkeit kommen Fragebögen zum Einsatz (diverse Beispiele auch über die Domäne der Lehrerbildung hinaus beschreibt Riel (2022, S. 54ff.)). Kobl (2021) entwickelte einen Multiple-Choice Fragebogen zur Erfassung der Reflexionskompetenz (das Instrument ist in Kobl (2021, S. 261ff.) abgedruckt)[101], der einerseits Wissen über Reflexion und Modellierungen von Reflexion abfragt (z. B. „Die von der Kultusministerkonferenz der Länder [KMK] formulierten Lehrerbildungsstandards (2014) unterscheiden vier Kompetenzbereiche. Welcher aufgeführte Punkt gehört nicht dazu?“ (Kobl, 2021, S. 261)) und andererseits die Anwendung dieses Wissens einfordert, indem gegebene Aussagen in bestimmten Stufenmodelle eingeordnet werden sollen (z. B. „Ordnen Sie folgendes Beispiel der Reflexionstiefe nach Hatton und Smith (1995) zu: ‚Dieser Versuch ist als Schülerexperiment im Microscale-Maßstab geeignet.‘“ (Kobl, 2021, S. 262)).

[101] Das Instrument ist in ihrer Arbeit nur ein Bestandteil zur Erfassung der Reflexionskompetenz; zusätzlich nutzt sie ein Manual zur Kodierung schriftlicher Reflexionen in Bezug auf die Reflexionsbreite und -tiefe (Kobl, 2021, S. 109ff.); Ergebnisse ihrer Studie werden daher gesammelt in Kapitel 7.8 berichtet.

Ähnlich berichten Fraij und Kirschner (2017) in ihrem Beitrag zur Tagung der Arbeitsgruppe für Empirische Pädagogische Forschung (AEPF) von einem Test, der Wissen über Reflexion sowie die Bewertung der Qualität von Beispielen in die Reflexionsstufen nach Larrivee (2008) einfordert.[102] Beide Testinstrumente weisen damit Elemente aus Millers (1990) Wissens- und Kompetenztests auf und enthalten zu einem relevanten Teil Wissen *über* den Reflexionsbegriff bzw. dessen Modellierung und geben ebenfalls keinen direkten Aufschluss darüber, in welchem Ausmaß die befragten Personen zu qualitativ hochwertigen Reflexionen in der Lage sind.

Weit verbreitet ist die Erfassung der Reflexionsfähigkeit mit Hilfe von Stufenmodellen (s. Kapitel 7.5.2) auf Grundlage von inhaltsanalytischen Auswertungen von schriftlichen oder mündlichen Reflexionen. Ein frühes Beispiel für die Bewertung von Reflexionstexten stellt die Untersuchung von Hatton und Smith (1995) dar. Verbunden mit dem Ziel, „to investigate the nature of reflection in teaching“ (Hatton & Smith, 1995, S. 40), erfassen sie u. a. schriftliche Reflexionen von $N = 50$ Studierenden in einer Veranstaltung zur Vorbereitung auf ein zehntägiges Praktikum. Sie stellen fest, dass diese Reflexionen überwiegend (60 % - 70 % der Kodierungen) beschreibender Natur sind, was der niedrigsten Reflexionsstufe entspricht. Nur die Reflexionen von acht Studierenden weisen Ansätze von kritischer Reflexion (der höchsten Stufe: Reflexionen unter Einbezug eines größeren Kontextes) auf. Dialogische Reflexionen, die die Diskussion und Abwägung unterschiedlicher Argumente zur reflektierten Situation beinhalteten, werden überwiegend in Reflexionen beobachtet, die auf Grundlage von Diskussionen mit einem „critical friend“ (Hatton & Smith, 1995, S. 41) entstanden sind. Passend dazu beobachten sie in etwa der Hälfte der Reflexionen den Einbezug verschiedener Perspektiven (Hatton & Smith, 1995). Auch Riel (2022, S. 113ff.) greift für die inhaltsanalytische Auswertung studentischer Reflexionstexte auf das Stufenmodell von Hatton und Smith (1995) zurück und untersucht den Einfluss weiterer Faktoren wie der Aufgabenstellung oder des persönlichen Bezugs zur reflektierten Situation auf die Einordnung der schriftlichen Reflexionen. Sie erfasst Reflexionstexte von insgesamt $N = 95$ Studierenden, die aktuell ein Praktikum absolvieren und jeweils eine von drei Reflexionsaufgaben mit unterschiedlich starker persönlicher Involviertheit bearbeiten (Riel, 2022, S. 119ff.). Etwa die Hälfte der Äußerungen der Studierenden stellt nicht als Reflexion einzustufende unbegründete Einschätzungen oder Meinungen dar (Riel, 2022, S. 131). Die überwiegende Anzahl der Studierenden (etwa 47 %) formuliert maximal beschreibende Reflexionen, die höchste Stufe der kritischen Reflexion erreichen lediglich zwei Studierende (Riel, 2022, S. 140). Außerdem stellt sie fest, dass die Studierenden mit zunehmender persönlicher Involviertheit der Aufgabenstellung eine höhere durchschnittliche

[102] Sie berichten eine Korrelation zwischen den Ergebnissen des Wissenstests und der „Selbstwirksamkeitserwartung hinsichtlich der Fähigkeit in problematischen Situationen reflektieren zu können“ (Fraij & Kirschner, 2017); weitere Ergebnisse sind nach Kenntnis der Autorin bislang nicht publiziert.

Reflexionsstufe erreichen (Riel, 2022, S. 142f.) und häufiger eigene Gefühle in die Reflexion einbeziehen, gleichzeitig allerdings weniger Bezüge zu fachwissenschaftlichen Theorien hergestellt werden (Riel, 2022, S. 144f.). Einen Einfluss auf die Motivation ist abhängig von der Aufgabenstellung nicht festzustellen (Riel, 2022, S. 143). Auch scheinen - mit Ausnahme der eingeschätzten Nützlichkeit der Reflexionsaufgabe - verschiedene andere Variablen und Persönlichkeitsmerkmale wie die Motivation, der empfundene Reflexionsgehalt des Reflexionsauftrags, die Offenheit, Ungewissheitstoleranz und Kreativität keinen Einfluss auf die Qualität der Reflexionen zu nehmen (Riel, 2022, S. 163ff.).

Wyss (2013) untersucht vergleichend die Reflexionskompetenz von jungen Lehrkräften im ersten Berufsjahr und erfahrenen Lehrkräften (mit mindestens sechsjähriger Berufserfahrung (Wyss, 2013, S. 114f.)). Sie erfasst mündliche Reflexionen im Rahmen von Stimulated-Recall-Interviews: Die Lehrkräfte reflektieren selbst durchgeführten Unterricht auf Grundlage einer 20-minütigen Videoaufnahme, die als Stimulus für das Interview diente und bei Bedarf für die Reflexion pausiert wurde (Wyss, 2013, S. 120). Die Reflexionen wurden inhaltlich in Bezug auf die Beurteilung der Unterrichtsqualität, die in den Fokus gerückten Personen, die Nennung und Diskussion von Handlungsalternativen sowie die Qualität der Reflexion kodiert (Wyss, 2013, S. 123ff.). Übereinstimmend mit den Ergebnissen von Hatton und Smith (1995) zeigt sich, dass sich die Lehrkräfte zum größten Teil (in etwa der Hälfte ihrer Aussagen) deskriptiv äußern (Wyss, 2013, S. 164) und nur in rund 20 % bei jungen Lehrkräften bzw. 15 % bei erfahrenen Lehrkräften der Aussagen alternative Handlungsoptionen thematisiert werden (Wyss, 2013, S. 167). Dementsprechend erreichen die Reflexionen insgesamt nur niedrige Reflexionsniveaus: die Beschreibung des Unterrichts (Stufe 1) und ggf. dessen Erläuterung (Stufe 2), allerdings ohne eine theoretische (Stufe 3) oder kritische (Stufe 4) bzw. ethisch moralische Erläuterung (Stufe 5; Wyss, 2013, S. 168).

Als neuere Methode werden Machine-Learning-Algorithmen verwendet, um Reflexionstexte zu klassifizieren. So lässt Ullmann (2019) Texte in reflexiv bzw. nicht-reflexiv einteilen sowie in sieben inhaltliche Kategorien (u. a. Beschreibungen, Gefühle und zukünftige (Handlungs-)Absichten) einordnen. Dabei berichtet er eine um durchschnittlich etwa zehn Prozent geringere Genauigkeit als bei einer manuellen Auswertung (Ullmann, 2019). Wulff, Mientus, Nowak und Borowski (2023b) und Mientus, Wulff, Nowak und Borowski (2023) nutzen Machine-Learning-Algorithmen zur Klassifikation von Bestandteilen von Reflexionen nach dem Reflexionsmodell von Nowak, Kempin, Kulgemeyer und Borowski (2019). Die $N = 110$ untersuchten studentischen Reflexionen einer fünfzehnminütigen Videovignette enthalten mit einem Anteil von über 50 % der Segmente überwiegend Beschreibungen und in abnehmender Häufigkeit Bewertungen, Alternativen und Konsequenzen (Mientus, Wulff, Nowak & Borowski, 2023). Zusätzlich wurde der Grad der Strukturierung (*level of structure*) bestimmt, der für die Texte bei durchschnittlich $0,289$ ($SD \approx 0,161$) liegt und sich im Vergleich mit einer manuellen

Bewertung der Reflexionsqualität mit $\kappa_{\text{Cohen}} = 0,64$ ($p < 0,001$) als Indikator für diese herausstellt (Mientus, Wulff, Nowak & Borowski, 2023, S. 697ff.).

Einen anderen Zugang wählen Szogs, Große und Korneck (2020): Angegliedert an die Durchführung und Reflexion von Unterrichtsminiaturen im Projekt Φactio (s. Kapitel 6.3.3) werden Zusammenhänge von Reflexion und Reflexionskompetenz[103] und Unterrichtsqualität untersucht. Dazu wird zwischen zwei Wiederholungen der Unterrichtsminiatur eine mündliche kollegiale Reflexion des selbst durchgeführten bzw. von Kommiliton*innen beobachteten Unterrichts durchgeführt und zur Auswertung videografiert. In Bezug auf die Reflexionstiefe stellen sie fest, dass die Hälfte der Aussagen Rückfragen (5 %), Beschreibungen (19 %) oder Interpretationen (26 %) darstellen, während nur 20 % der Aussagen Handlungsalternativen für die erneute Durchführung der Unterrichtsminiatur darstellen und deren Entwicklung das Ziel der kollegialen Reflexionen war und lediglich rund 5 % der Aussagen begründet wurden (Szogs, Große & Korneck, 2020). Unter Berücksichtigung der Zusammenhänge zur Unterrichtsqualität zeigt sich ein positiver Effekt, „wenn die Unterrichtsanalysen diskutiert und abgewogen werden. Dahingegen können Reflexionen dann hinderlich sein, wenn der Unterricht lediglich nacherzählt wird oder der*die Unterrichtende mit Ideen ‚bombardiert' wird" (Szogs, Große & Korneck, 2020, S. 501). Bedeutsam für die Entwicklung der Unterrichtsqualität ist außerdem der Anteil an begründeten Aussagen in der Reflexion. Die Autor*innen schließen, dass die Unterrichtsqualität nicht unbedingt von zunehmender Reflexionstiefe profitiert, „sondern ein adäquates Zusammenspiel der Prozesselemente notwendig zu sein [scheint]" (Szogs, Große & Korneck, 2020, S. 501). Gleichzeitig stellen sie fest, dass die Studierenden, deren Unterricht eine höhere Qualität aufweist, tendenziell (nicht signifikant) weniger beschreiben als andere Studierende, allerdings signifikant häufiger interpretieren, positive Aspekte rückmelden und Aussagen begründen (Szogs, Große und Korneck, 2021; Szogs, Große und Korneck, 2020). Aufbauend auf diese Ergebnisse soll ein flexibel einsetzbares Ratingmanual zur Bewertung von Reflexionen entwickelt und validiert werden (Große, Szogs, Lamprecht & Korneck, 2021).

Wie die hier exemplarisch vorgestellten Studien bereits andeuten, liegen diverse Befunde dazu vor, dass Reflexionen (angehender) Lehrkräfte überwiegend niedrige Niveaus aufweisen, was sich u. a. in hohen deskriptiven Anteilen, wenigen abgeleiteten Handlungsalternativen und Konsequenzen oder fehlender Diskussion bzw. Multiperspektivität in der Reflexion niederschlägt (z. B. Mientus, Wulff, Nowak und Borowski, 2023; Riel, 2022, S. 131, 140ff.; Weber, Elstrodt-Wefing und Hoge, 2022; Körkkö, Kyrö-Ämmälä und Turunen, 2016; Wyss, 2013, S. 164ff.; Abou Baker El-Dib, 2007) und mit mangelnder Strukturiertheit einhergeht (Mientus, Wulff, Nowak & Borowski, 2023). Vor dem Hintergrund des großen Mehrwerts, der der Reflexion für die qualitativ hochwertige Gestaltung von Unterricht und der individuellen Professionalisierung zugesprochen wird (s. Kapitel 7.2) und dem

[103] Sie benennen sie als „Reflexivität" (Szogs, Große & Korneck, 2021, S. 234)

verbreiteten Einsatz von Aufträgen zur Reflexion schon im Studium (Hilzensauer, 2017, S. 14), kann ein Bedarf zur Förderung der Reflexionsfähigkeit (angehender) Lehrkräfte konstatiert werden. Auf diesbezügliche Ansätze wird in Kapitel 7.8 eingegangen.

Die zuletzt beschriebenen Studien greifen in der Beurteilung der Reflexionsfähigkeit (bzw. -kompetenz) zwar auf reale Reflexionen zurück.[104] Insbesondere die häufig eingeforderten schriftlichen Reflexionen stellen allerdings keine für Lehrkräfte und ihren schulischen Alltag authentischen Reflexionsprodukte dar. Diesem Problem nehmen sich Kempin, Kulgemeyer und Schecker (2018) mit dem nachfolgend vorgestellten Reflexionsperformanztest an.

7.7.3 Reflexionsperformanztest von Kempin, Kulgemeyer und Schecker (2018)

Im Rahmen des Projekts *ProfiLe-P+* entwickelten Kempin, Kulgemeyer und Schecker (2018) einen Performanztest zur Erfassung der Reflexionsfähigkeit von (angehenden) Physiklehrkräften. Dieser Performanztest bildet die Grundlage für das in dieser Arbeit entwickelte Testinstrument (s. Kapitel 10). Daher werden nachfolgend das Testinstrument selbst sowie die zentralen im Projekt *ProfiLe-P+* gewonnenen Erkenntnisse ausführlicher vorgestellt.

Aufbau des Reflexionsperformanztests

Im Rahmen des Reflexionsperformanztests[105] werden die (angehenden) Lehrkräfte in eine kollegiale Beratungs- und Reflexionssituation versetzt und dazu aufgefordert, der Persona Robert, einem fiktiven Mitpraktikanten im Praxissemester, Feedback zu einer Doppelstunde seines Physikunterrichts zu geben. Thema der Unterrichtsstunde ist - nach einer Wiederholung der Newtonschen Axiome zu Beginn - die Einführung des Impulsbegriffs und des Konzepts der Impulserhaltung. Die Struktur des Instruments ist in Abbildung 7.2 dargestellt.

In der Bearbeitung des Performanztests erhalten die Lehrkräfte zunächst Informationen bezüglich der Rahmenbedingungen der Unterrichtsstunde: eine Beschreibung der Lerngruppe und deren Vorwissen aus vorherigen Unterrichtsstunden sowie eine Beschreibung der Unterrichtsziele und des geplanten Stundenablaufs.

[104] Wenn auch unter der kaum vermeidbaren Einschränkung, dass solche Reflexionen auf Grundlage der Veranlassung von außen möglicherweise einen „defensiven Charakter“ (Klewin, Köker & Störtländer, 2020, S. 111) aufweisen; dies gilt insbesondere dann, wenn diese in Zusammenhang mit Prüfungssituationen erstellt werden und weitere Einflüsse, z. B. durch soziale Erwünschtheit, nicht ausgeschlossen werden können.

[105] Beschreibungen des Testinstruments finden sich z. B. in Kempin, Kulgemeyer und Schecker (2018); Vogelsang et al. (2019) oder Kulgemeyer et al. (2021).

Anschließend wird der Unterrichtsgang anhand von 13 geskripteten Videovignetten nachgezeichnet, die einzeln reflektiert werden. Jede Vignette wird dabei (a) durch einen kurzen Vorpromt eingeleitet, in dem Robert einordnet, was in der kommenden Szene im Unterricht geschieht (z. B. „Ich zeige dir gleich, wie ich den Unterricht begonnen habe. Kannst du mir was dazu sagen?“). Anschließend wird (b) die Vignette (Dauer von durchschnittlich ca. 2 Minuten) angesehen, in der jeweils verschiedene problematische und gelungene Aspekte beobachtet werden können. In einem jeweils auf die Vignette folgenden Nachprompt von wenigen Sekunden Dauer wiederholt und konkretisiert Robert seine Bitte um Feedback (z. B. „Kannst du dir einen Reim auf die falschen Äußerungen der Schülerinnen und Schüler machen?“[106] oder „Denkst du, dass ich das Thema Impuls interessant eingeführt habe?“). In Vor- und Nachprompt werden die Lehrkräfte im Sinne der simulierten kollegialen Reflexionssituation außerdem direkt angesprochen. Für die Reflexion der einzelnen Videovignetten (jeweils im Anschluss an den Nachprompt) stehen den Lehrkräften verschiedene Hilfestellungen zur Verfügung, die im Sinne der Simulation Rückfragen an Robert darstellen. Die Hilfestellungen ermöglichen es stets, die Informationen bezüglich der oben beschriebenen Rahmenbedingungen nachzulesen (z. B. „Kannst du mir die Unterrichtsziele noch einmal nennen?“), die Videovignette erneut anzusehen („Kann ich den Unterrichtsausschnitt noch einmal sehen?“) oder in der entsprechenden Vignette genutztes Material näher zu betrachten (z. B. „Kann ich die PowerPoint-Folien noch einmal sehen?“). Abschließend bittet Robert die Lehrkräfte mit vier Fragen um ein globales Feedback zur Unterrichtsstunde:

- *Gesamtitem 1:* „Denkst du, ich konnte meine Ziele erreichen? Vielleicht hast du ja einen Tipp für mich, was ich berücksichtigen könnte, wenn ich die Stunde nochmal durchführe.“
- *Gesamtitem 2:* „Denkst du, dass der Unterrichtsverlauf so stimmig war oder sollte ich etwas anders machen, wenn ich die Doppelstunde nochmal durchführe?“
- *Gesamtitem 3:* „Findest du, dass die Stunde überladen war?“
- *Gesamtitem 4:* „Wenn du dir den Unterricht als Ganzes anschaust, welche Schulnote würdest du mir dafür geben?“

Die Videovignetten enthalten verschiedene problembehaftete Situationen, um Reflexionen zu bestimmten Reflexionspunkten zu evozieren. Insgesamt ergeben sich so über die 13 Videovignetten 77 Reflexionspunkte (diese entsprechen den Items des Testinstruments; sie werden zusammengefasst in insgesamt 48 Kategorien) zu fachlichen, fachdidaktischen und pädagogischen Aspekten. Dazu gehören auf fachlicher Ebene die fachlich falsche Darstellung von Inhalten z. B. in einem Merksatz zum 3. Newtonschen Axiom oder im Rahmen der Diskussion von Mes-

[106] Diese und nachfolgendende mündliche Formulierungen aus dem Testinstrument sind bei Bedarf zur besseren Lesbarkeit sprachlich leicht geglättet.

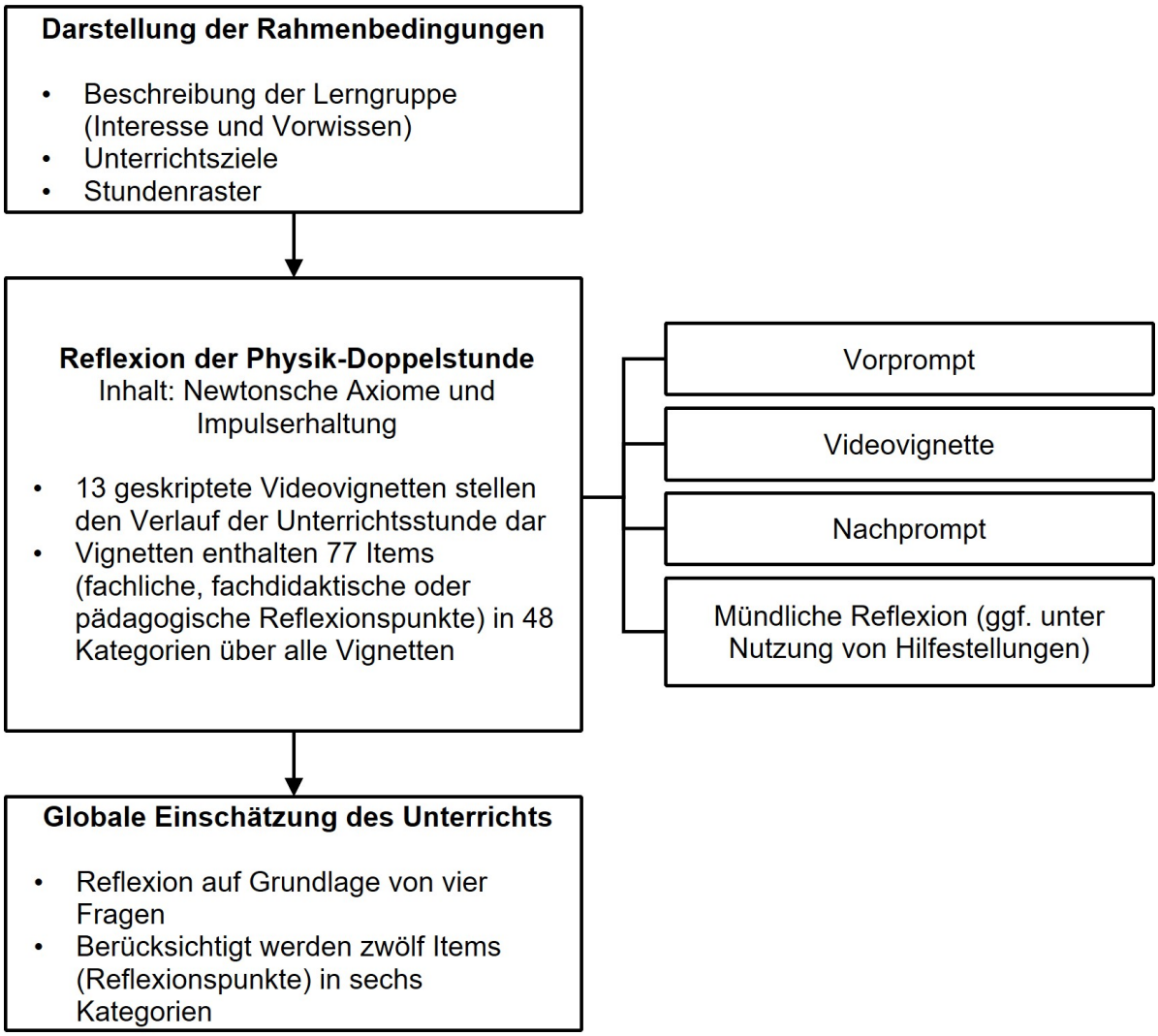


Abbildung 7.2: Struktur des Reflexionsperformanztests (Darstellung in Anlehnung an Kulgemeyer et al., 2021, S. 3044)

sunsicherheiten. Fachdidaktische Reflexionspunkte betreffen z. B. den Umgang mit auftretenden Schülervorstellungen, das geschlechtsstereotypische Verhalten der Lehrkraft oder den Einsatz von Kontexten und Experimenten. Allgemein pädagogische (d. h. nicht für den Physikunterricht spezifische) Reflexionspunkte sind z. B. die Vorwissensaktivierung im Stundeneinstieg oder die (inadäquate) Sicherung von Unterrichtsinhalten (hier: der Impulserhaltung) in einem Merksatz. Im Rahmen der Gesamtitems zur globalen Einschätzung der Unterrichtsstunde werden in sechs Kategorien insgesamt zwölf wiederum fachliche, fachdidaktische und pädagogische Reflexionspunkte berücksichtigt.

Die Reflexion der einzelnen Videovignetten erfolgt im Sinne der simulierten Peer-Reflexion jeweils mündlich im Anschluss an die Nachprompts zu den Vignetten bzw. auf die Gesamtitems. Das Feedback bzw. die Reflexionen der Lehrkräfte

werden für die Auswertung audiografiert. Zur Auswertung wurde ein Kodiermanual entwickelt (Kulgemeyer, 2020, S. 36ff.) Im Rahmen der Auswertung wird für jeden Reflexionspunkt kodiert, ob dieser angesprochen wird[107] und welche Reflexionselemente (ggf. begründet) auftreten[108]. Für jeden Reflexionspunkt (also jedes Item) ergibt sich so ein Score zwischen Null und Sieben (s. Tabelle 7.2). Die erreichten Scores je Reflexionspunkt werden aufsummiert und so in ein quantitatives Maß für die Reflexionsfähigkeit überführt.

Zentrale Validitätsargumente zum Reflexionsperformanztests

In Bezug auf die valide Interpretation der Testwerte des Reflexionsperformanztests als Maß für die Reflexionsfähigkeit lassen sich einige Argumente anführen. In Bezug auf die inhaltliche Validität (in diesem Falle die inhaltlich adäquate Abbildung der (kollegialen) Reflexion von Physikunterricht durch das Testinstrument) wurde mit Hilfe von Curriculums- und Lehrwerksanalysen sichergestellt, dass der fachliche Inhalt der Unterrichtsstunde ein zentrales Thema des schulischen Physikunterrichts darstellt (Vogelsang et al., 2019). Die Darstellung problembehafteter Situationen in den Videovignetten wurde ebenso wie das Kodiermanual zur Auswertung der mündlichen Reflexionen auf Grundlage einer qualitativen Inhaltsanalyse sowie in einer Expert*innenrunde in mehreren Feedbackschleifen kommunikativ validiert (Kulgemeyer et al., 2021; Vogelsang et al., 2019; Kempin, Kulgemeyer und Schecker, 2019). Im Rahmen einer Interviewstudie wurde zudem festgestellt, dass die angehenden Lehrkräfte die simulierte Peer-Feedback-Situation als authentisch einschätzten (Kulgemeyer et al., 2021).

Hinsichtlich der Auswertung der Bearbeitungen ergab sich eine reliable Skala für das Testinstrument ($\alpha_{\text{Cronbach}} = 0,74$ (Kulgemeyer et al., 2021, S. 3047)). Zusätzlich wurden 17 % der Aufnahmen zu den ersten drei Videovignetten von drei unabhängigen Rater*innen kodiert, wobei sich eine sehr gute Interraterreliabilität (von durchschnittlich $Gwet'sAC_1 = 0,904$) ergab (Vogelsang et al., 2019, S. 487).

Außerdem zeigt ein Vergleich der Ergebnisse von angehenden Lehrkräften im Studium ($N = 65$) und im Referendariat ($N = 10$), dass die Referendar*innen etwas bessere Testwerte erreichten ($t(73) = 4,709; p = 0,000; d = 1,6$; 95 %-KI $[0,89; 2,32]$; (Kulgemeyer et al., 2021, S. 2047)), was vor dem Hintergrund der zu unterstellenden umfangreicheren Lerngelegenheiten der Referendar*innen plausibel erscheint.

[107] Die Anzahl der angesprochenen Reflexionspunkte kann als Maß für die Reflexionsbreite interpretiert werden.

[108] Hier werden die beiden Dimensionen *Element der Reflexion* und *Begründung* kombiniert, um die Reflexionstiefe zu bestimmen.

erreichte Stufe	Score	Beispiel
Nicht angesprochen	0	*Reflexionspunkt wird nicht angesprochen*
Beschreibung	1	Es haben sich zwei Schüler*innen gemeldet und du hast sofort eine Schülerin drangenommen.
Bewertung	2	Es ist nicht gut, dass du die Schülerin so schnell drangenommen hast.
Bewertung mit Begründung	3	Es ist nicht gut, dass du die Schülerin so schnell drangenommen hast, weil die anderen Schüler*innen so nicht ausreichend Zeit hatten, über die Aufgabe nachzudenken.
Alternative	4	Du solltest länger warten oder den Schüler*innen eine Murmelphase ermöglichen, bevor du jemanden drannimmst.
Alternative mit Begründung	5	Du solltest länger warten oder den Schüler*innen einen Austausch untereinander ermöglichen, bevor du jemanden drannimmst, weil es Überwindung kostet, über die eigenen Fehlvorstellungen zu sprechen. Es ist wichtig, ihnen zu zeigen, dass auch andere Schüler*innen diese Fehlvorstellungen haben.
Konsequenz	6	Du solltest dir überlegen, zukünftig verstärkt kooperative Methoden einzusetzen, probiere es doch in der nächsten Stunde aus.
Konsequenz mit Begründung	7	Du solltest dir überlegen, zukünftig verstärkt kooperative Methoden einzusetzen, probiere es doch in der nächsten Stunde aus. Kooperative Methoden folgen nämlich üblicherweise einer Art „Think - Pair - Share“-Verfahren, das genau so einen Austausch zwischen den Schüler*innen beinhaltet.

Tabelle 7.2: Bepunktung einer Reflexion im Reflexionsperformanztest anhand der erreichten Stufe (Kempin, Kulgemeyer & Schecker, 2019, S. 358); Beispiele übersetzt aus Kulgemeyer et al. (2021, S. 2046)

Zentrale Ergebnisse aus dem Projekt *ProfiLe-P+* mit Bezug zur Reflexionsfähigkeit

Ziel des Projekts *ProfiLe-P+* war die Untersuchung der Entwicklung professioneller Fähigkeiten - darunter auch die Reflexionsfähigkeit - im Praxissemester im Fach Physik, um Einblicke in den bislang weitgehend ungeklärten Zusammenhang zwischen professionellem Wissen und Handeln zu erlangen (Kempin, Kulgemeyer und Schecker, 2019; Vogelsang et al., 2019). In diesem Rahmen wurde der Reflexionsperformanztest zu zwei Messzeitpunkten (je einmal zu Beginn und zum Ende des Praxissemesters im Abstand von etwa fünf Monaten) eingesetzt. Die Erhebung fand an vier Universitäten in Deutschland statt und stellt für diese Standorte und den Erhebungszeitraum nahezu eine Vollerhebung dar ($N = 95$), sodass von einer gewissen Übertragbarkeit der Ergebnisse auch auf andere Universitäten in Deutschland ausgegangen werden kann (Vogelsang, Kulgemeyer & Riese, 2022). Für die Bearbeitung des Instruments standen jeweils 70 Minuten zur Verfügung. Zusätzlich wurden Tests zu den Bestandteilen des Professionswissens sowie Performanztests zur Unterrichtsplanung und zum Erklären sowie Fragebögen zu demographischen Aspekten und Lerngelegenheiten u. a. zur Reflexion im Rahmen des Praxissemesters eingesetzt (Vogelsang et al., 2019).

Ein Vergleich der Bearbeitungen des Reflexionsperformanztests zu beiden Messzeitpunkten zeigt, dass zwischen beiden Messzeitpunkten keine signifikanten Änderungen in der Reflexionsfähigkeit festgestellt werden konnten (Kulgemeyer et al., 2021; Kempin, Kulgemeyer und Schecker, 2020) Eine detailliertere Betrachtung zeigt, dass sowohl die Reflexionsbreite (Anzahl der angesprochenen Reflexionspunkte) als auch die Reflexionstiefe (erreichte Stufe zu einem angesprochenen Reflexionspunkt) keine signifikanten Veränderungen aufweisen; die Studierenden sprechen durchschnittlich knapp 30 % der Reflexionspunkte an (29 % vor dem Praxissemester und 28 % nach diesem) und erreichen in angesprochenen Reflexionspunkten zu beiden Messzeitpunkten durchschnittliche Scores von $M = 0,31$ ($SD_{\text{Prä}} = 0,44$, $SD_{\text{Post}} = 0,42$), was dem Score einer begründeten Bewertung entspricht (s. Tabelle 7.2) (Kulgemeyer et al., 2021, S. 3049)). Da die meisten Reflexionspunkte gar nicht angesprochen werden, vermuten Kulgemeyer et al. (2021) mangelnde Fähigkeiten zum Noticing auf Seiten der Studierenden.

Kempin, Kulgemeyer und Schecker (2020) betrachten die Reflexionspunkte mit Bezug zu den verschiedenen Wissensbasen getrennt ($N_{\text{Prä}} = 67$ und $N_{\text{Post}} = 57$). Dabei zeigt sich für die Bereiche des fachlichen Wissens (14 Reflexionspunkte) und des fachdidaktischen Wissens (26 Reflexionspunkte) eine nicht signifikante Abnahme der Reflexionsqualität, wohingegen die Qualität der Reflexionen mit Bezug zum pädagogischen Wissen (19 Reflexionspunkte) signifikant zunimmt ($d_{\text{Cohen}} = 0,52$ (Kempin, Kulgemeyer & Schecker, 2020, S. 440)). Analoge Ergebnisse ergeben sich aus der Untersuchung der Phase des Unterrichtseinstiegs an einer Teilgruppe, allerdings mit einer höheren Effektstärke der Zunahme im pädagogischen Wissen

von $d_{\text{Cohen}} = 0,81$ (bei $N = 33$ Prä-Post-Datensätzen für das fachliche Wissen und $N = 30$ Prä-Post-Datensätzen für das fachdidaktische und pädagogische Wissen (Kempin, Kulgemeyer & Schecker, 2019, S. 358)). Die Autor*innen vermuten hier eine Fokusverschiebung auch aufgrund der Reflexionsgespräche mit mentorierenden Lehrkräften an den Praktikumsschulen, da diese überwiegend pädagogische Aspekte thematisierten (Kempin, Kulgemeyer & Schecker, 2019). Futter (2017, S. 237) gibt zu bedenken, dass die mentorierenden Lehrkräfte möglicherweise nicht verbalisierbares (implizites) fachdidaktisches Wissen besitzen und es deswegen in gemeinsamen Reflexionen weniger stark fokussiert wird. Ursächlich für die Akzentuierung des pädagogischen Wissens gegenüber den anderen Wissensbereichen könnte auch sein, dass die Studierenden im Praxissemester pädagogische Themen z. B. aufgrund von überfordernden Situationen in der Durchführung des Unterrichts als besonders relevant wahrnehmen (Kempin, Kulgemeyer & Schecker, 2020) und (übrige) universitäre Inhalte wie das Professionswissen von den Studierenden insgesamt als wenig hilfreich für das Unterrichten eingeschätzt werden (Vogelsang, Kulgemeyer & Riese, 2022). Auch die vergleichsweise starke gemessene Zunahme im pädagogischen Wissen gegenüber dem fachlichen und fachdidaktischen über das Praxissemester hinweg könnte hier Einfluss nehmen oder aber selbst eine Folge der Fokussierung auf pädagogische Aspekte in Reflexionsgesprächen sein (Kempin, Kulgemeyer & Schecker, 2020). Dazu passend kann vor dem Praxissemester ein signifikanter Zusammenhang mittlerer Höhe zwischen der Reflexionsfähigkeit und dem fachdidaktischen Wissen festgestellt werden ($r = 0,48$, $p < 0,01$), der nach dem Praxissemester so nicht mehr beobachtet werden kann; nun bestand ein Zusammenhang zwischen Reflexionsfähigkeit und pädagogischem Wissen ($r = 0,34$, $p < 0,05$ (Kulgemeyer et al., 2021, S. 3048)).

Außerdem werden Zusammenhänge zwischen der Reflexionsfähigkeit zu beiden Messzeitpunkten und dem professionellen Wissen bzw. der Fähigkeit, Unterricht zu planen (verstanden als „antizipatorische Reflexion“; übersetzt aus Vogelsang, Kulgemeyer und Riese (2022, S. 2)), untersucht. Hier zeigt sich in Strukturgleichungsmodellen, dass

- die Reflexionsfähigkeit zu Beginn des Praxissemesters Einfluss auf das Wachstum von fachlichem und pädagogischem Wissen nimmt (Kulgemeyer et al., 2021). Dies passt zum Ansatz, Reflexion würde es ermöglichen, Wissen aus Erfahrung zu generieren (McAlpine et al., 1999) bzw. (passend zum Refined Consensus Model of PCK (Carlson et al., 2019)) personal PCK auf Grundlage von möglicherweise im Reflexionsperformanztest genutztem enacted PCK zu entwickeln (Kulgemeyer et al., 2021).
- sich gleichzeitig aber auch hohes fachliches und fachdidaktisches Wissen „positiv auf den Zuwachs der Reflexionsfähigkeit im [Praxissemester] auswirkt“ (Kempin, Kulgemeyer & Schecker, 2020, S. 441).

- „die Anzahl der hospitierten Stunden zentral für das Wachstum der Reflexionsfähigkeit ist (zur Erinnerung: hier wurde Fremdreflexion gemessen)." (Kulgemeyer et al., 2021, S. 3048).
- Reflexionsgespräche mit Universitätsdozierenden bedeutsam für das Wachstum der Reflexionsfähigkeit sind, wohingegen solche mit mentorierenden Lehrkräften keinen Einfluss zu haben scheinen (Kulgemeyer, Kempin & Weißbach, 2021).
- der Reflexionsfähigkeit als Mediator des fachdidaktischen Wissens (unter Kontrolle des Einflusses vom fachlichen und fachdidaktischen Wissen) kein signifikanter Einfluss auf die Fähigkeit zur Unterrichtsplanung zugesprochen werden kann (Vogelsang, Kulgemeyer & Riese, 2022). Ursächlich hierfür könnte die Operationalisierung der Reflexionsfähigkeit als auf das Handeln von Lehrkräften bezogen sein, während das Verständnis der Unterrichtsplanung als antizipatorische Reflexion eine stärkere Fokussierung auf die eigene Person (eigenes Handeln sowie Professionalität) verlangt und damit möglicherweise andere gedankliche Prozesse zugrunde liegen (Vogelsang, Kulgemeyer & Riese, 2022).
- die Fähigkeiten zur Unterrichtsplanung und -reflexion vor dem Praxissemester je etwa ein Drittel der Varianz der Planungsfähigkeit (31,7 %) bzw. Reflexionsfähigkeit (35,6 %) nach dem Praxissemester erklären (Vogelsang, Kulgemeyer & Riese, 2022, S. 15). Der Beitrag der Planungsfähigkeit vor dem Praxissemester zur Varianzaufklärung der Reflexionsfähigkeit nach dem Praxissemester ist allerdings sehr gering; allein die Reflexionsfähigkeit vor dem Praxissemester erklärt bereits 32,4 % dieser Varianz.

Diese Ergebnisse verdeutlichen, dass das Praxissemester zwar durchaus zur Entwicklung von Professionswissen beiträgt (Kempin, Kulgemeyer & Schecker, 2020), allerdings nicht per se eine Lerngelegenheit für Reflexion von Physikunterricht darstellt. Einerseits wird eine Zunahme der Reflexionsfähigkeit nur für pädagogische Aspekte beobachtet, andererseits scheint die Entwicklung der Reflexionsfähigkeit ein hohes Professionswissen vor dem Praxissemester vorauszusetzen (Kempin, Kulgemeyer & Schecker, 2020) und durch Reflexionsgespräche mit Universitätsdozierenden begünstigt zu werden (Kulgemeyer, Kempin & Weißbach, 2021). Diese Ergebnisse passen zu Befunden von Leonhard und Rihm (2011), die keinen pauschalen Beitrag von universitären Begleitveranstaltungen zum Praxissemester auf die Qualität von Reflexionstexten feststellen konnten.

7.8 Ansätze zur Förderung der Reflexionsfähigkeit

Während diverse Studien, die Reflexionen (angehender) Lehrkräfte erfassen und bewerten, übereinstimmend zu dem Schluss kommen, dass deren Reflexionsfähigkeit insgesamt eher gering ausgeprägt ist (s. Kapitel 7.7.2), liegen vereinzelt Befunde dazu vor, dass erfahrene Lehrkräfte zumindest komplexer oder inhaltlich breiter reflektieren (z. B. Weber, Elstrodt-Wefing und Hoge, 2022[109]; Peterson und Comeaux, 1987) und inhaltliche Zusammenhänge zwischen (erlerntem) Wissen und Reflexionen bestehen (Wulff et al., 2023). Folgerichtig ist die Überzeugung weit verbreitet, dass die Fähigkeit, (Unterricht) zu reflektieren, entwickelbar ist bzw. gefördert werden kann (Damköhler, Elsholz und Trefzger, 2023; Marcos, Sanchez und Tillema, 2011; Herzog, 1995; Day, 1993).[110] Weder empirisch noch theoretisch ist allerdings abschließend geklärt, wie genau Reflexion erlernt wird, ob bzw. welche Entwicklungsstufen in welchem Zeitraum durchlaufen werden und welche Fördermaßnahmen vielversprechend erscheinen (Reintjes und Kunze, 2022; Riel, 2022, S. 158f.; Baer et al. (2007) (letztere zur Entwicklung von Kompetenzen allgemein)).

Inzwischen sind diverse Ansätze zur Anregung höherwertiger Reflexionen benannt, entwickelt und teilweise auch evaluiert, bei denen zwischen einmaligen bzw. unterstützenden und längerfristigen (z. B. in universitäre Veranstaltungen eingegliederte) Maßnahmen unterschieden werden kann.

Eine Möglichkeit, Studierende unmittelbar zu höherwertigen Reflexionen anzuregen, stellen angepasste Prompts bzw. Aufträge zur Reflexion dar, die z. B. explizit Aufschluss über Erwartungen an die Reflexionen oder Strukturierungshilfen bereitstellen (Dyment und O'Connell, 2010; Hall, 2018). Positive Effekte der strukturierten Anleitung von Studierenden auf die Qualität ihrer Reflexionen werden auch in anderen Studien festgestellt: Kost (2019) beobachtet die Adressierung von höheren Reflexionsniveaus, d. h. der vermehrten Ableitung von Ursachen und Konsequenzen bei Reduktion von Beschreibungen und Deutungen aufgrund nur „geringer Einhilfe (Hinweis auf die Prozessschritte)“ (Kost, 2019, S. 111). Ähnlich stellt auch Pultorak (1996) eine Zunahme der Reflexionsqualität durch Interviewfragen fest und Davis (2003) in Bezug auf den Einsatz von (eher generischen) Prompts.

Auch dem Einsatz von Unterrichtsvideos als Reflexionsgrundlage wird pauschal Potenzial für die Unterstützung von Reflexion zugesprochen, wobei hier häufig

[109] In einem Vergleich der schriftlichen Reflexionen von Bachelor-, Masterstudierenden und Lehrkräften für die Primarstufe zu Sprachförderstrategien stellten Weber, Elstrodt-Wefing und Hoge, 2022 zwar ebenfalls fest, dass alle Reflexionstexte vorrangig deskriptive Elemente beinhalten, beobachteten aber zumindest eine Entwicklung hinsichtlich der Komplexität dieser Beschreibungen mit zunehmender Expertise.

[110] Eine Gegenposition nehmen Labott und Reintjes (2022) ein. Sie hinterfragen die Erlernbarkeit von Reflexion und halten insbesondere eine objektive Erfassung für nicht möglich.

das Noticing in die Reflexion inkludiert ist (z. B. Rogge, 2023; Körkkö, 2021; Coffey, 2014). Hieran anschließend erscheint auch in Bezug auf die Auswahl der reflektierten Inhalte die Unterstützung der Studierenden z. B. durch mentorierende Lehrkräfte hilfreich, um eine gewisse Reflexionsbreite zu erreichen (Futter, 2017, S. 236).[111] Bain, Ballantyne, Packer und Mills (1999) stellen hingegen keine Verbesserung der Reflexionsqualität durch gemeinsame Reflexion (in diesem Fall mit Forschenden) fest. Unklar ist, inwieweit die Forschenden für die Begleitung der Reflexion geschult sind, was gemäß den Ergebnissen von Crasborn et al. (2008) Einfluss auf die Reflexion nehmen kann.

Zusätzlich wird untersucht, ob Unterschiede in der Reflexionsqualität ausgemacht werden können, abhängig davon, ob die Anlässe zur Reflexion selbst erlebt (Selbstreflexion) oder beobachtet bzw. im Rahmen einer Vignette bereitgestellt (Fremdreflexion) wurden. Hierbei ist allerdings keine der beiden Vorgehensweisen eindeutig vorzuziehen, sondern es bedarf einer stimmigen Auswahl abhängig vom Einsatzszenario, wie Göbel und Gösch (2019) zusammenfassen:

> „Während Videos von eigenem Unterricht eine stärkere persönliche Betroffenheit und damit einhergehend eine höhere Motivation zur Auseinandersetzung mit dem Material fördern, regen fremde Videos eher zu einer distanzierteren, kritischeren Haltung gegenüber den Handlungen im Video an.“ (Göbel & Gösch, 2019, S. 278)

Für die Messung der Reflexionsfähigkeit kann die Berücksichtigung dieser Faktoren zu Validitätsproblemen führen, wenn z. B. die Anweisungen zur Reflexion bereits (Qualitäts-)Merkmale für bestimmte Stufen der Reflexionsfähigkeit einfordern (Riel, 2022, S. 103):

> „Es ist aber ein unzulässiger Schluss, beispielsweise in Pro- und Kontraabwägungen ein Zeichen für eine höherwertige Reflexion zu sehen, wenn sie explizit mit der Aufgabenstellung eingefordert wurden.“ (Riel, 2022, S. 103)[112]

Zusätzlich zu den Einschränkungen, die bezüglich der Validität festzustellen sind, ist unklar, inwieweit die beschriebenen Möglichkeiten zur Anregung höherwertiger Reflexionen tatsächlich Einfluss auf Reflexionsfähigkeit auch unabhängig vom konkreten Setting der Erhebung bzw. Bearbeitung des Auftrags zur Reflexion nehmen. Zusammenfassend stellen Marcos, Sanchez und Tillema (2011) in einem

[111] Gleichzeitig äußern Dozierende in einer Interviewstudie, dass für Studierende ein „Sinn im Reflektieren“ (Wyss & Mahler, 2021, S. 22) deutlich werden muss, was bspw. durch den Einbezug der Studierenden in die Themenauswahl erreicht werden kann.

[112] Exemplarisch sei hier die Studie von Körkkö, Kyrö-Ämmälä und Turunen (2016) genannt: Sie attestieren eine Entwicklung der Reflexionsfähigkeit im Verlauf eines vierjährigen Lehramtsstudiums, da die an jährliche Praktikumsphasen anschließenden Reflexionen in ihrer Komplexität (insb. in Bezug auf Reflexionsbreite und -tiefe) zunehmen. Sie setzen allerdings auch zunehmend komplexe Reflexionsaufträge ein, deren Einfluss auf die Qualität der Reflexionen nicht kontrolliert wird.

Literaturreview fest, dass zwar die Qualität von Reflexionen häufig diskutiert wird, allerdings oft keine adäquate Anleitung zur Umsetzung von qualitativ hochwertigen Reflexionen erfolgt. Solche Unterstützungsmaßnahmen wären allerdings nötig, wie z. B. Russell (2005) zu bedenken gibt:

> „Fostering reflective practice requires far more than telling people to reflect and then simply hoping for the best." (Russell, 2005, S. 203)

Als in der Umsetzung eher niedrigschwellige Maßnahmen können die von Korthagen (1985) entlang der Schritte seines ALACT-Modells vorgeschlagenen allgemeinen Verhaltensweisen verstanden werden, die Dozierende zur Unterstützung von Studierenden bei der Reflexion übernehmen sollten (z. B. die Unterstützung bei der Auswahl der reflektierten Erfahrungen oder der Findung von Lösungen). Für die Studierenden werden Leitfragen zur Verfügung gestellt, die wiederholt bei der Durchführung einer Reflexion herangezogen werden können (Korthagen & Vasalos, 2005)

Die explizite, längerfristige Förderung der Reflexionsfähigkeit im Rahmen der universitären Lehrkräfteausbildung erfolgt üblicherweise im Rahmen von Seminaren mit Praxisbezug, oft im Rahmen der Begleitung von Praktika. Als gewinnbringend evaluiert wurde dabei u. a. der Einsatz von Portfolios (d. h. wiederholten schriftlichen Reflexionsaufträgen, die Begleitend zum Praktikum bearbeitet werden) (Leonhard & Rihm, 2011)[113] sowie allgemein der Austausch über Reflexionen mit Peers und über Kriterien zur Qualität von Reflexionen sowie das Erhalten von Feedback zu diesen (Sluijsmans, Brand-Gruwel, van Merriënboer und Bastiaens, 2003; Bain, Ballantyne, Packer und Mills, 1999). Analog zu Methoden der Erfassung (studentischer) Reflexionen zeigt sich auch für deren Förderung ein Fokus auf schriftliche Reflexionen (Oner & Adadan, 2011). Zwei aktuelle Arbeiten der Naturwissenschaftsdidaktik, die die Entwicklung der Reflexionsfähigkeit im Verlaufe eines Semesters mit expliziten Lerngelegenheiten untersuchen, stammen von Kobl (2021) und Nowak (2023).

Zur Förderung der Reflexionskompetenz von Studierenden des Lehramts im Fach Chemie entwickelte Kobl (2021) ein Seminar, in welchem die Studierenden nach einem fachlichen Input zur Unterrichtsplanung und Reflexion Unterrichtsstunden planen und im Verlauf des Semesters drei Zyklen aus Umsetzung der Stunde in einem Lehr-Lern-Labor und Reflexion der Stunde durchlaufen. In einem experimentellen Setting vergleicht sie dabei zwei Interventionsgruppen und eine Kontrollgruppe: Eine Gruppe, die ihren eigenen Unterricht nur mit Hilfe der Unterrichtsvideos reflektiert, eine Gruppe, die zusätzlich zu den Videos Feedback

[113] Leonhard und Rihm (2011) untersuchten die Entwicklung der Qualität schriftlicher Reflexionen (in Bezug auf die Reflexionsbreite (verstanden als Einbezug verschiedener Perspektiven in die Reflexion) und Reflexionstiefe angelehnt an Hatton und Smith (1995)) von Studierenden über das Praxissemester abhängig von der Art der Begleitveranstaltung. Während sie insgesamt für beide Kriterien einen Qualitätsverlust feststellten, zeigte sich für ein Begleitformat, das Portfolios einbezog, eine Steigerung der Reflexionstiefe.

zur gehaltenen Unterrichtsstunde erhält und die Kontrollgruppe, die hospitierten Unterricht mit Hilfe der Unterrichtsvideos reflektiert. Zur Evaluation dieser Fördermaßnahme greift Kobl (2021) neben den Ergebnissen aus einem schriftlichen Testinstrument zur Erfassung der Reflexionskompetenz (Kobl, 2021, S. 107ff.) (eingesetzt zu Beginn und Ende des Seminars sowie als Follow-Up Test etwa zwei Monate nach Ende des Seminars (Kobl, 2021, S. 95)) auf die Bewertung schriftlicher Reflexionen mit Hilfe eines Kodiermanuals zu drei Messzeitpunkten (jeweils anschließend an die Durchführung bzw. Hospitation der Unterrichtsstunden) zurück (Kobl, 2021, S. 109ff.). Das Kodiermanual umfasst dabei die Reflexionsbreite (u. a. Vermittlung von chemischem Fachwissen, Schüleraktivierung und Medien (Kobl, 2021, S. 266)) und die Reflexionstiefe (in Bezug auf das Vorliegen von Beschreibungen, Bewertungen, Alternativen, Konsequenzen, Perspektivübernahmen und Literaturbezügen (Kobl, 2021, S. 267)). Dabei stellt sie einen nachhaltigen Wissenszuwachs im deklarativen Wissen über Reflexionskompetenz fest (Kobl, 2021, S. 162) sowie im Vergleich von erster und letzter verfasster Reflexion eine Zunahme der Qualität in beiden Interventionsgruppen, die auf die Steigerung der Reflexionstiefe zurückzuführen ist (in Bezug auf die Reflexionsbreite wurden in keiner Gruppe signifikante Änderungen festgestellt (Kobl, 2021, S. 162ff.)). Die erwarteten positiven Effekte des Feedbacks zur gehaltenen Unterrichtsstunde auf die Reflexionskompetenz in einer Interventionsgruppe können allerdings nur im Reflexionskompetenztest (einem Wissenstest nach Miller, 1990), nicht aber in der Bewertung der schriftlichen Reflexionen bestätigt werden.

Nowak (2023) untersucht die Entwicklung von schriftlichen Selbstreflexionen im Verlauf des physikdidaktisch begleiteten Praxissemesters mit einer qualitativen Inhaltsanalyse von je zwei Reflexionstexten zu drei vorgegebenen Reflexionsanlässen (Unterrichtseinstieg, Einsatz von Experiment und Einsatz von Aufgaben) von $N = 22$ Studierenden (insgesamt 132 Reflexionen) (Nowak, 2023, S. 43ff.). Im Rahmen des Begleitseminars erhalten die Studierenden eine Instruktion über theoretische Grundlagen zum Reflexionsbegriff (u. a. zur Definition), Leitfragen zur Anregung des Reflexionsprozesses sowie jeweils ein nicht-standardisiertes und nicht näher beschriebenes Feedback zu den eingereichten Reflexionen (Nowak, 2023, S. 42f.). Insgesamt stellt sie fest, dass die Reflexionstexte tendenziell fokussierter werden, was sich in einer abnehmenden Textlänge bei gleichbleibender Reflexionstiefe sowie einer abnehmenden Zahl an diskutierten Reflexionsauslösern äußert; die relative Häufigkeit vorliegender Begründungen scheint sich nicht zu verändern (Nowak, 2023, S. 146f.). Darüber hinaus gehen stärker durch Überschriften strukturierte Texte mit höherer Reflexionstiefe einher (Nowak, 2023, S. 148). Auf inhaltlicher Ebene bilden passend zum physikdidaktischen Seminar für den Physikunterricht spezifische Aspekte einen Schwerpunkt (Nowak, 2023, S. 149). Die Untersuchung von Zusammenhängen zwischen der Reflexionstiefe als Maß für die Qualität und Eigenschaften der Reflexionstexte liefert die Erkenntnisse, dass qualitativ hochwertigere Texte eher negativ sind; positive Auslöser werden häufig

nur beschrieben (Nowak, 2023, S. 152f.). Ebenfalls sind die Vollständigkeit der Texte, formulierte Lernziele sowie die Antizipation von Problemen im Vorfeld des Unterrichts Indikatoren für die Reflexionstiefe, während umfangreiche didaktische Vorüberlegungen zu den Rahmenbedingungen ansonsten tendenziell mit geringerer Reflexionstiefe einhergehen (Nowak, 2023, S. 153ff.).

Auch Mientus, Wulff, Nowak und Borowski (2021) streben die Förderung der Reflexionskompetenz im Rahmen der physikdidaktischen Begleitveranstaltung zum Praxissemester an. Nach der Bereitstellung von Informationen zum Reflexionsmodell von Nowak, Kempin, Kulgemeyer und Borowski (2019), das den Studierenden zur Orientierung dienen soll, werden drei schriftliche Reflexionen eingereicht. Den Studierenden wird ein automatisiert erstelltes Feedback zu strukturellen Aspekten sowie ein durch Dozierende verfasstes inhaltliches Feedback bereitgestellt. Während die Qualität der Reflexionen bzw. Entwicklung der Reflexionskompetenz nicht explizit evaluiert wird, stellen Wulff, Mientus, Nowak und Borowski (2021) zu diesem Vorgehen fest, dass die computerbasierten Rückmeldungen zwar geeignet seien, unmittelbares Feedback zur Struktur von Reflexionstexten zu geben, allerdings durch die Studierenden nur teilweise positiv wahrgenommen werden. Sie ziehen daraus den Schluss, dass es weiterhin Bedarf für zusätzliche „menschengemachte Rückmeldung mit Bezug auf die inhaltliche Umsetzung“ (Wulff, Mientus, Nowak & Borowski, 2021, S. 97) gibt, da die inhaltliche Rückmeldung bislang nicht bzw. noch unzureichend (Wulff, Mientus, Nowak & Borowski, 2023a) automatisiert erfolgen kann.

Aus den beschriebenen Maßnahmen und Seminarkonzeptionen lassen sich zwei zentrale, überwiegend geteilte Merkmale für die Förderung der Reflexionsfähigkeit extrahieren, die in einer engeren Begleitung der Studierenden bei der Reflexion münden:

- Im Rahmen des Seminars und bzw. oder in den verwendeten Aufträgen zur Reflexion werden die Studierenden über das zugrundeliegende Reflexionsmodell bzw. Bewertungsschema, also u. a. über erwartete Prozessschritte wie das Entwickeln von Alternativen oder über Qualitätsmerkmale der Reflexion (z. B. der Einbezug verschiedener (theoretischer) Perspektiven), aufgeklärt.
- Die Studierenden erhalten Feedback zu den (üblicherweise schriftlich verfassten) Reflexionen, das Auskunft darüber gibt, inwieweit die jeweils angelegten Kriterien zur Bestimmung der Qualität der Reflexionen erfüllt wurden. Zur theoretischen Fundierung dieses Aspekts wird nachfolgend in Kapitel 8 ein Überblick über die Gestaltung effektiven Feedbacks gegeben.

Vor dem Hintergrund, dass die hier untersuchten Reflexionen häufig auch eine im Seminar zu erbringende Leistung (und damit eine veranlasste Reflexion (Häcker, 2022)) darstellen, ist die transparente Darstellung der Bewertungskriterien (und Bereitstellung von Feedback auf Grundlage der Bewertungskriterien) notwendig. Ohne eine Klärung der Anforderungen ist unklar, inwiefern das Verständnis

eines Reflexionsauftrags durch Studierende den Erwartungen von Dozierenden entspricht (Meier, 2022, S. 55f.; Russell, 2005). Gleichzeitig bleibt offen, inwieweit die Studierenden authentische Reflexionen verfassen oder ihre Reflexionen zur Erfüllung einer (bewerteten) Aufgabenstellung anpassen (vgl. Klewin, Köker und Störtländer, 2020; Häcker, 2019; Gebhardt und Jenert, 2011).

Abschließend sei darauf hingewiesen, dass die Förderung der Reflexionsfähigkeit in der Literatur nicht ohne Einschränkungen gefordert wird. Obwohl sie Reflexion als bedeutsam für Lehrkräfte einstufen, geben von Aufschnaiter, Hofmann, Geisler und Kirschner (2019) zu bedenken, dass zu häufige Reflexion (kurzfristig) destabilisierend wirken kann. „[D]auerreflexive Zustände" (Häcker, 2019, S. 92) seien weder sinnvoll noch erstrebenswert, sondern würden Handeln einschränken. Häcker (2019, 2022) nimmt eine kritischere Haltung gegenüber der Reflexion ein. Er stellt zwar ihre Bedeutung für die Professionalisierung nicht per se in Frage, schränkt aber ein, dass Reflektieren „kein positiver Wert an sich" (Häcker, 2019, S. 92) sei. Spricht man der Reflexion einen Selbstzweck zu, besteht die Gefahr, dass Aufträge zur Reflexion nicht ausreichend begründet werden (Häcker, 2019).[114] Häcker (2019; 2022) betont, dass es nicht darum gehen kann, eine grundsätzliche, nicht domänenspezifische Reflexionsfähigkeit zu erlernen (für die ohnehin keine etablierte Definition und Operationalisierung vorliegt (Häcker, 2019)), sondern vielmehr domänenspezifische „inhaltliche Bezugnahmen, bestimmte konkrete Denkfiguren, -verfahren, -stile, -techniken oder -strategien respektive Fragestellungen" (Häcker, 2022, S. 108) von Bedeutung sind.

7.9 Zwischenfazit: Reflexion von Physikunterricht

Der Reflexionsbegriff wird vielfältig eingesetzt und definiert (s. Kapitel 7.1); es besteht aber gewisse Einigkeit in Bezug darauf, dass die Reflexion ein mehrschrittiger Denkprozess ist (bzw. diesen beinhaltet), der mit einem bestimmten Ziel verbunden ist. Die Relevanz von Reflexion für Lehrkräfte und ihr (unterrichtliches) Handeln ist trotz eines uneinheitlichen Begriffsverständnisses unumstritten (s. Kapitel 7.2). Ein Mehrwert wird sowohl eher abstrakt für die Professionalisierung der Lehrpersonen als auch konkret für die verbesserte Gestaltung von Unterricht unterstellt, indem eine Vermittlung von Theorie und Praxis erfolgt und (konkrete) Handlungsoptionen generiert und berücksichtigt werden (können).

Vor dem Hintergrund des variierenden Begriffsverständnisses überrascht es nicht, dass diverse verschiedene Modellierungen zum Reflexionsprozess (Prozessmodelle) bzw. zur Qualität von Reflexionen (üblicherweise in Stufenmodellen) vorliegen (s. Kapitel 7.5). Das Reflexionsmodell von Nowak, Kempin, Kulgemeyer und Borowski

[114] Weitere kritische Aspekte in Bezug auf die unhinterfragte Betonung der Reflexion fasst Häcker (2022) zusammen (z. B. die Gefahr, durch die Reflexion des eigenen Handelns u. a. „vorgefasste Meinungen zu bestätigen [und] Stereotype zu verfestigen" (Häcker, 2022, S. 100)).

(2019) (s. Kapitel 7.5.4) stellt den Versuch einer Synthese verschiedener Modelle in Bezug auf die berücksichtigten Bestandteile einer Reflexion (die Reflexionselemente sowie das Vorliegen einer Begründung) dar. Es ist grundsätzlich als Stufenmodell einzuordnen, weist aber im Sinne eines integrativen Modells durch die Stufung von Prozessschritten zusätzlich Elemente eines Prozessmodells auf.

Uneinigkeit in der Modellierung des Reflexionsprozesses besteht auch hinsichtlich des Einbezugs des Erlebens bzw. Erkennens einer Situation. In dieser Arbeit wird - passend zum Reflexionsmodell von Nowak, Kempin, Kulgemeyer und Borowski (2019) - unter Reflexion lediglich der Denkprozess verstanden, das vorausgehende Erkennen relevanter Situtationen wird als Noticing konzipiert. Auf Grundlage der in Kapitel 7.6 vorgenommenen Begriffsklärung ist dementsprechend von einem engen sowohl inhaltlichen als auch zeitlichen Zusammenhang zwischen Noticing und Reflexion auszugehen. Da Noticing allerdings überwiegend im Kontext der professionellen Unterrichtswahrnehmung operationalisiert wird, ist dieses Verhältnis bislang empirisch kaum geklärt.

Empirische Ergebnisse zur Erfassung der Reflexionskompetenz legen nahe, dass (angehende) Lehrkräfte Reflexion zwar durchaus einen Mehrwert zusprechen, die Reflexionsfähigkeit insgesamt aber eher gering ausgeprägt ist, sodass ein Bedarf zur Förderung dieser Fähigkeit auch schon bei Studierenden besteht, um bereits in der Ausbildung z. B. im Rahmen von Praxisphasen von Reflexion zu profitieren (s. Kapitel 7.7). Hierzu hat sich insbesondere die Bereitstellung von Wissen über Reflexion sowie die Bereitstellung von Feedback als wirksam erwiesen (s. Kapitel 7.8). Auf Feedback wird daher im nächsten Kapitel näher eingegangen. Zusammenhänge zwischen der Reflexionsfähigkeit und jener zum Noticing werden in Teil IV dieser Arbeit betrachtet.

Die in diesem Kapitel dargestellten theoretischen Grundlagen und empirischen Erkenntnisse sind überwiegend fachunabhängig bzw. -übergreifend, wenn z. B. die Reflexionsfähigkeit (angehender) Lehrkräfte mit unterschiedlichen fachlichen Hintergründen untersucht wurde. In Bezug auf vorliegende empirische Befunde zur Erfassung und Förderung der Reflexionsfähigkeit (s. Kapitel 7.7 und 7.8) wurden Erkenntnisse aus dem Bereich der Physikdidaktik fokussiert. Zusätzlich wurde mit der Arbeit von Kobl (2021) eine Arbeit aus der Chemiedidaktik ausführlicher dargestellt, die fächerübergreifende Bewertungskriterien mit gewisser Überschneidung zu den von Nowak (2023) berücksichtigten Reflexionselementen anlegt, sodass insgesamt von der Übertragbarkeit der Ergebnisse auf Studierende des Physik-Lehramts ausgegangen werden kann.

8 Feedback

Im Kontext der Lehrkräftebildung sowie Lehr-Lern-Kontexten generell beschreibt Feedback allgemein eine Rückmeldung über eine ggf. vorliegende Diskrepanz zwischen Ist- und Sollzustand (z. B. bezüglich bestimmter (professioneller) Fähigkeiten oder ganz konkret bei der Lösung einer konkreten Aufgabe) mit dem Ziel, diese zu überwinden oder zumindest zu reduzieren (z. B. Müller und Ditton, 2014; Hattie, 2009, S. 175; Hattie und Timperley, 2007; Narciss, 2006, S. 18).[115] Feedback kann also erst auf eine Handlung folgen und wird daher auch als „‚consequence' of performance" (Hattie, 2009, S. 174) beschrieben.

Feedback gilt als eine der bedeutsamsten Möglichkeiten, einen Lernprozess zu unterstützen (Casanova, Alsop und Huet, 2021; Hattie, 2009, S. 173). Dabei kann es unterschiedliche Funktionen erfüllen (z. B. Fengler, 2010; Narciss, 2006, S. 76ff.), die sich in Bezug auf den Lernprozess in die drei (nicht trennscharfen) Kategorien Motivation, Verhaltensanpassung und Kompetenzzuwachs zusammenfassen lassen (Behnke, 2016, S. 14f.), indem z. B. bestimmte Antworten oder Verhaltensweisen positiv verstärkt werden oder die Korrektur von Fehlern unterstützt wird (Kobl, 2021, S. 62). Feedback kann somit auch metakognitive Strategien umfassen (Narciss, 2006, S. 79). (Wiederholtes) Feedback kann sogar eine tutorielle Funktion erfüllen, Lernende also bei der Bearbeitung eines Problems (schrittweise und ggf. adaptiv) unterstützen (Narciss, 2006, S. 19ff.; Shute, 2008).

Die Forschung zu Feedback ist in der Lehrkräftebildung - und darüber hinaus - sehr umfangreich. In diesem Kapitel wird daher der Fokus auf solches Feedback gelegt, das auf der Erhebung von Fähigkeiten in (formativen) Assessments (im Sinne von Testinstrumenten, die mit dem Ziel eingesetzt werden, ein Feedback zur Unterstützung des Lernprozesses zu geben; vgl. Hattie und Timperley (2007)) basiert und nicht z. B. auf einer kontinuierlichen Begleitung von Lernenden bei der Bearbeitung einzelner Aufgaben bzw. Probleme. Umgekehrt ist die Bereitstellung von Feedback zentral, wenn Assessments als Lerngelegenheit eingebettet werden sollen (Sippel, 2009).[116]

[115] Einige Autoren, wie z. B. Müller und Ditton (2014) differenzieren Feedback und Rückmeldung darüber hinaus z. B. in Bezug auf die Verhaltensnähe bzw. den Auflösungsgrad der zurückgemeldeten Informationen. Hierfür besteht in der Literatur allerdings kein Konsens. Daher wird auf eine stärkere Ausdifferenzierung der beiden Begriffe in dieser Arbeit verzichtet.

[116] Formatives Assessment beschreibt die in den Lehr-Lern-Prozess integrierte Erfassung bzw. Diagnose von Lern- oder Fähigkeitsständen zur Optimierung dieser Prozesse (Black und

8.1 Feedback-Modell von Hattie und Timperley (2007)

Auf Grundlage des oben benannten Verständnisses, dass Feedback dazu dient, eine Diskrepanz zwischen Ist- und Sollzustand bzw. „erbrachter Leistung und gesetztem Ziel“ (Hosenfeld, 2010, S. 54) zu reduzieren, stellen Hattie und Timperley (2007) ein Modell für effektives Feedback im schulischen Kontext vor (s. Abbildung 8.1; Zielformulierung in Kasten 1). Sie formulieren Maßnahmen (Kasten 2; s. auch z. B. Kluger und DeNisi (1996)), die Lernende und Lehrende ergreifen können, um diese Diskrepanz zu reduzieren, zu denen auch die Unterstützung der Lernenden durch Feedback der Lehrenden gehört. Die Effektivität eines Feedbacks bemisst sich laut Hattie und Timperley (2007, S. 88ff.) daran, dass es die folgenden drei zentralen Fragen (auch *Dimensionen* (Hattie & Timperley, 2007, S. 88)) adressiert (Kasten 3):

- Feed Up: Wo will ich hin? Was sind die Ziele?
  Die Ziele, die den Lernenden transparent gemacht werden, können unterschiedlicher Natur sein und sich sowohl auf das Lösen konkreter Aufgaben als auch auf abstraktere Konzepte (z. B. das Erreichen einer bestimmten Fähigkeit) beziehen und werden als besonders effektiv angesehen, wenn sie herausfordernd sind und sich Lernende stärker engagieren, um das Ziel zu erreichen (Hattie & Timperley, 2007, S. 88f.).
- Feed Back: Bin ich auf dem richtigen Weg zum Erreichen der Ziele? „Wie nahe bin ich dem Ziel schon?“ (Hosenfeld, 2010, S. 54)
  Diese Komponente bezieht sich i. d. R. auf vorausgegangene Handlungen (z. B. Aufgabenbearbeitungen); für eine effektive Rückmeldung werden Informationen über das Erfüllen von Anforderungen oder Erreichen von Standards bzw. Abweichungen zwischen den Handlungen und vorgegebenen Anforderungen oder Standards gegeben (Hattie & Timperley, 2007, S. 89).

Wiliam, 2018; Pellegrino, Chudowsky und Glaser, 2001, S. 37ff.). Es wird unterschieden vom summativen Assessment, das zur abschließenden Einschätzung der Erreichung von Lernzielen eingesetzt wird und in formalisierten Lehr-Lern-Kontexten häufig mit Benotung verbunden ist (Black und Wiliam, 2018; Pellegrino, Chudowsky und Glaser, 2001, S. 37ff.). Zwei zentrale Bestandteile des formativen Assessments, die zumindest in ähnlicher Form auch im Rahmen der in dieser Arbeit entwickelten Lernumgebung zur Unterrichtsreflexion zum Einsatz kommen, sind das Bereitstellen von (formativem) Feedback durch Lehrende zur Ausnutzung des oben beschriebenen Beitrags für erfolgreiche Lernprozesse (Black & Wiliam, 2018) und das Self-Assessment, das Lernende dazu anleiten soll, die eigenen Fähigkeiten (u. a. auf Grundlage transparenter Kriterien) fundiert einzuschätzen (Buholzer et al., 2020; Andrade und Valtcheva, 2009) und sich (auch in Kombination mit Feedback durch Lehrende (Panadero, Jonsson & Botella, 2017)) als positiv für den Lernerfolg erwiesen hat (z. B. Zulliger, Buholzer und Ruelmann, 2022). Formatives Assessment umfasst also grundsätzlich Handeln unterschiedlicher Akteur*innen (Cizek, 2010).

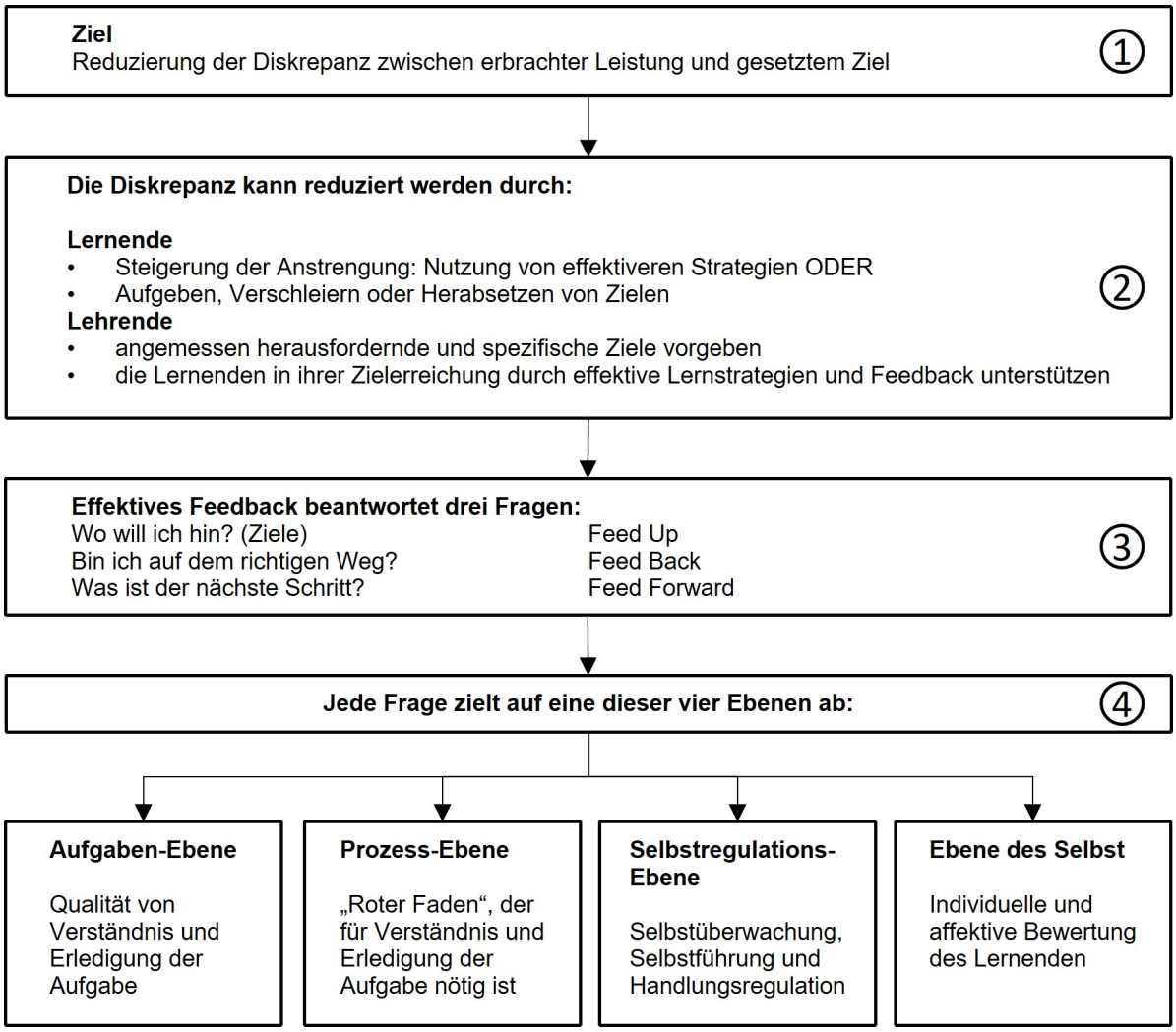


Abbildung 8.1: Feedback-Modell nach Hattie und Timperley (2007, S. 87) (Übersetzung von Hosenfeld, 2010, S. 55)

- Feed Forward: Was sind (sinnvolle) nächste Schritte, um dem Ziel näher zu kommen?
  Die Beantwortung dieser Frage zeigt den Lernenden Möglichkeiten zur Weiterentwicklung auf, die - abhängig von angestrebten Ziel - wieder unterschiedlich konkret bzw. abstrakt ausfallen können (Hattie & Timperley, 2007, S. 90).

Zusätzlich unterscheiden sie vier Ebenen, auf die sich ein Feedback beziehen kann: die Aufgaben-, Prozess- und Selbstregulations-Ebene sowie die Ebene des Selbst. Feedback auf der Ebene des Selbst (z. B. generisches Loben einer Person ohne inhaltlichen Bezug zu einem formulierten Ziel) wird grundsätzlich als ineffektiv angesehen; im Modell enthalten ist sie nur aufgrund ihrer (zu starken) Präsenz im Klassenraum (Hattie & Timperley, 2007). Die Effektivität einer Ansprache der übrigen drei Ebenen hängt davon ab, wozu Feedback gegeben werden soll. Eine Adressierung der Prozess- und Selbstregulationsebene erscheint dann

sinnvoll, wenn Lernende bei der Bewältigung konkreter Probleme unterstützt werden sollen, während Feedback auf Aufgabenebene vorrangig dann eingesetzt werden sollte, wenn daraus Strategien zur Aufgabenbearbeitung oder Prozesse der Selbstregulation unterstützt werden können (Hattie & Timperley, 2007).

8.2 Empfehlungen für effektives Assessment-Feedback

Mit dem Feedback-Modell von Hattie und Timperley (2007) sind bereits einige Empfehlungen für effektives Feedback verbunden, wie die Berücksichtigung von Feed Up, Back und Forward (Kasten 3) und das Adressieren adäquater Ebenen im Feedback - insbesondere das Auslassen von Bewertungen der Person wie Lob oder Bestrafung (Ebene des Selbst, Kasten 4 (Hattie & Timperley, 2007, S. 84); vergleichbar auch Shute (2008)). Darüber hinaus lassen sich diverse weitere Kriterien für die adressatengerechte Gestaltung des Feedbacks auf inhaltlicher, aber auch formaler Ebene finden.[117] Feedback sollte elaboriert sein und nicht nur Informationen über die Korrektheit einer Aufgaben- oder Testbearbeitung bieten, sondern einen möglichst präzisen Abgleich zwischen Ist- und Sollzustand bereitstellen, z. B. indem insbesondere Musterlösungen bereitgestellt[118] oder falsche Lösungen diskutiert werden (Sippel, 2009; Shute, 2008; Hattie und Timperley, 2007) und sich auf ein erreichbares, übergeordnete Ziel beziehen (Shute, 2008; Coe, 1998). Auf interindividuelle Vergleiche sollte zugunsten des Sachbezugs verzichtet werden (Sippel, 2009). Gleichzeitig ist davon auszugehen, dass zu umfangreiches oder besonders komplexes Feedback von Lernenden nicht (umfangreich) berücksichtigt wird (Shute, 2008). In diesem Sinne sollte bspw. auf unübliche statistische Kennwerte verzichtet werden, wenn diese nicht bekannt sind (Hosenfeld, 2010, S. 41). Sowohl positives als auch negatives Feedback kann grundsätzlich zielführend sein, indem entweder motivationale Effekte ausgenutzt werden bzw. der Entwicklungsbedarf aufgezeigt wird (Coe, 1998). Besonders einseitiges Feedback ist allerdings problematisch, da zu positives Feedback wenig Anlass und Unterstützung für Weiterentwicklung bietet, während zu negatives Feedback demotivierend wirken kann (Körkkö, Kyrö-Ämmälä und Turunen, 2016; Sippel, 2009).

Zum zeitlichen Bezug zwischen Assessment und Feedback sind die Ergebnisse differenzierter zu betrachten: Shute (2008) fasst zusammen: „many field studies demonstrate the value of immediate feedback [...], whereas many laboratory studies show positive effects of delayed feedback“ (Shute, 2008, S. 165f.). Insgesamt erscheint unmittelbares Feedback vielversprechend, wenn es sich auf schwierige

[117] Eine umfangreiche Auflistung von Kriterien für die adäquate Gestaltung von computerbasiertem Feedback stellt Shute (2008) zusammen.

[118] Im Kontext handlungsnaher Fähigkeiten sollte entsprechend optimales Verhalten beschrieben werden (Hosenfeld, 2010, S. 42).

Aufgaben oder solche zum prozeduralen und konzeptuellen Wissen bezieht (Shute, 2008; Hattie und Timperley, 2007; Scheeler, Ruhl und McAfee, 2004). Verzögertes Feedback sollte hingegen eingesetzt werden, wenn Feedback zu einfachen Aufgaben gegeben wird oder es sich auf Transferwissen bezieht (Shute, 2008; Hattie und Timperley, 2007). Außerdem scheinen insbesondere wenig fortgeschrittene Lernende von unmittelbarem Feedback zu profitieren, während verzögertes Feedback eher für fortgeschrittene Lernende geeignet scheint (Shute, 2008). Henderson et al. (2019) beschreiben mehrfaches, aufeinander aufbauendes Feedback (jeweils auf Grundlage eines bzw. verschiedener Assessments) als vorteilhaft, da so die Wahrscheinlichkeit zunähme, dass zentrale Aspekte des Feedbacks verstanden und umgesetzt werden.

Die Berücksichtigung der oben aufgeführten Kriterien für die Bereitstellung von gutem bzw. effektivem Feedback begünstigt zwar, dass Lernende dieses Annehmen und Maßnahmen umsetzen, um sich einem gewissen Ziel anzunähern (Kasten 2), bspw. ihre Reflexionsfähigkeit auszubauen, darf allerdings nicht als Garantie dafür verstanden werden (Hattie & Timperley, 2007). Vielmehr ist davon auszugehen, dass das Feedback von Lernenden nicht umfangreich genutzt und teilweise sogar gar nicht gelesen wird (Li & De Luca, 2014). Die Bereitstellung eines Feedbacks sollte demnach als Angebot an Lernende verstanden werden, das abhängig von individuellen Faktoren, wie Motivation, Vorerfahrungen mit Feedback usw., aber auch von externen Faktoren (bspw. zeitlichen Einschränkungen), rezipiert und interpretiert werden und daraufhin zu Aktionen der Lernenden führen kann (z. B. Sippel, 2009; Hattie und Timperley, 2007; Coe, 1998).[119] Es bestehen aber deutliche Überschneidungen zwischen den oben beschriebenen Kriterien und Wünschen zur Gestaltung von Feedback von Lernenden (u. a. Zielorientierung, Transparenz, Nachvollziehbarkeit, Ausgewogenheit zwischen positiven und negativen Aspekten und Feed Forward) (z. B. Li und De Luca, 2014; Ferguson, 2011). Auch der aktive Einbezug der Studierenden, z. B. indem ihre Erwartungen an das Feedback abgefragt werden oder sie Feedback zum Feedback geben können, fördert die Nutzung von Feedback (Henderson et al., 2019; Sippel, 2009).

8.3 Zum Verhältnis von Feedback zum und der Reflexion von Unterricht

Die Bereitstellung von Feedback zu (beobachtetem) Unterricht[120] weist deutliche Überschneidungen zur gemeinsamen Reflexion von Unterricht (z. B. mit Peers oder Mentor*innen) auf (Szogs, Kobl, Volmer & Korneck, 2019). Hilzensauer (2017, S. 63) verwendet die Begriffe Peer-Feedback und Reflexion sogar quasi

[119] Im letzten Jahrzehnt werden die Fähigkeiten, Feedback anzunehmen und umzusetzen zunehmend unter dem Konzept der *Feedback literacy* adressiert (Nieminen und Carless, 2023; Molloy, Boud und Henderson, 2020).

[120] Hier also nicht mehr beschränkt auf das oben fokussierte Assessment-Feedback

synonym. In beiden Fällen wird in einem kommunikativen (oft dialogischen) Setting der Unterricht analysiert und evaluiert. Unterschiede können im Verhältnis der beteiligten Personen festgestellt werden: In einer gemeinsamen Reflexion des Unterrichts (insbesondere, wenn diese mit Peers durchgeführt wird) lassen sich die beteiligten Personen als gleichberechtigt verstehen. Der Unterricht wird gemeinsam analysiert, wobei die verschiedenen Perspektiven der beteiligten Personen berücksichtigt werden (können) und neue Erkenntnisse generiert werden (Szogs, Kobl, Volmer & Korneck, 2019). Dahingegen zeichnet sich Feedback dadurch aus, dass eine Person Rückmeldung zu z. B. ihrem Handeln im Unterricht erhält, während andere Personen Gelungenes, Entwicklungsbedarfe usw. rückmelden. Die Feedback gebenden Personen nehmen in gewisser Weise eine Expertenrolle ein.[121] Eine Feedback-Situation geht also mit einer gewissen Hierarchie zwischen den beteiligten Personen in der spezifischen Situation einher. In einem kollegialen Setting kann in diesem Sinne der Bedarf für Feedback allerdings durchaus Anlass für die (Fremd-)Reflexion des Unterrichts durch die Feedback gebende Person sein.

Der Unterschied zwischen gemeinsamer Reflexion und Feedback erscheint umso geringer, je stärker eine gemeinsame Reflexion von einer Person (z. B. Mentor*innen) gesteuert wird (Szogs, Kobl, Volmer & Korneck, 2019). Gleichzeitig ist die Anleitung zur Reflexion grundsätzlich wünschenswert, da sie als bedeutsam für die Entwicklung der Reflexionsfähigkeit angenommen wird (Day, 1993).

8.4 Zwischenfazit: Feedback

Wie bereits exemplarisch anhand empirischer Befunde zur Förderung der Reflexionsfähigkeit beschrieben (s. Kapitel 7.8), kommt der Bereitstellung von Feedback bei der Unterstützung von Lernprozessen zentrale Bedeutung zu. Allgemeines Ziel von Feedback ist die Reduktion einer Diskrepanz zwischen Ist- und Sollzustand. Für die Förderung der Reflexionsfähigkeit, für die ein Ziel bzw. Sollzustand bislang weder normativ (z. B. aus Standards für die Lehrerbildung (von Aufschnaiter, Hofmann, Geisler & Kirschner, 2019)) noch empirisch abgeleitet werden kann (Häcker, 2022), kann der Sollzustand gegenüber dem Istzustand eher abstrakt als *stärker* ausgeprägte Reflexionsfähigkeit verstanden werden.

Zur Umsetzung von effektivem Feedback empfehlen Hattie und Timperley (2007) die Berücksichtigung von Feed Up (die Darstellung des Ziel- bzs. Sollzustands),

[121] Die „Expertenrolle“ bezieht sich in diesem Fall nicht auf z. B. umfangreicheres Wissen; gerade in Peer-Feedback Situationen (z. B. mit Kommiliton*innen) kann dieses nicht vorausgesetzt werden. Vielmehr geht es darum, dass die Perspektive und Einordnung der Feedback gebenden Person im Fokus steht. Dennoch gibt es deutliche Unterschiede zwischen Feedback von Peers und (Fach-)Expert*innen u. a. in Bezug darauf, ob das Feedback angenommen wird (Hilzensauer, 2017, S. 63).

Feed Back (die Darstellung des Istzustands) und Feed Forward (die Darstellung möglicher nächster Schritte zum Erreichen des Zielzustands) (s. Kapitel 8.1). Ergänzend liegen abhängig vom konkreten Einsatzszenario für Feedback weitere Empfehlungen für dessen effektive Umsetzung vor (s. Kapitel 8.2), wobei zu berücksichtigen ist, dass die Umsetzung des Feedbacks durch Lernende auch unter Berücksichtigung dieser Kriterien nicht als gesichert angenommen werden kann. Die Berücksichtigung der Lernendenperspektive erscheint also zentral für die Bewertung des Feedbacks.

9 Konsequenzen für die Entwicklung der Lernumgebung zur Reflexion von Physikunterricht sowie für weitere Analysen zur Reflexionsfähigkeit

In diesem Teil der Arbeit wurde bisher ein Überblick über die theoretischen Grundlagen gegeben, die die Grundlage für die Messung und Rückmeldung der Reflexionsfähigkeit darstellen. Bevor im Folgenden die entwickelte Lernumgebung vorgestellt wird, wird in diesem Kapitel zunächst das der Lernumgebung zugrundeliegende Begriffsverständnis der Reflexionsfähigkeit konkretisiert (s. Kapitel 9.1). Anschließend werden die Anforderungen beschrieben, denen die entwickelte Lernumgebung gerecht werden soll (s. Kapitel 9.2), sowie das ergänzende Erkenntnisinteresse zur Entwicklung der Reflexionsfähigkeit, das sich auf Grundlage bislang weitgehend offener Forschungsdesiderate ergibt (s. Kapitel 9.3).

9.1 Bestimmung und Modellierung des Begriffs der Reflexionsfähigkeit für diese Arbeit

Wie in Kapitel 7.3 ausgeführt, wird der Reflexionsbegriff insgesamt vielfältig definiert und modelliert. Es bedarf also einer präzisen, idealerweise anschlussfähigen begrifflichen Klärung für diese Arbeit, die eine Einordnung vor dem Hintergrund dieser Breite ermöglicht (Aeppli und Lötscher, 2016; Abels, 2011, S. 49).

9.1.1 Definition von Unterrichtsreflexion

Im Rahmen dieser Arbeit liegt der Fokus auf der Reflexion von (Physik-)Unterricht bzw. unterrichtlichen Situationen, deren Gestaltung Kernaufgabe von (Physik-)Lehrkräften ist (Tenorth, 2006). Es wird also neben fachunabhängigen pädagogischen Aspekten ein fachlicher Bezug zum physikalischen und physik-

didaktischen Wissen hergestellt. Anschlussfähig an den Projektzusammenhang *ProfiLe-P*, wird weiterhin die von Kempin, Kulgemeyer und Schecker (2020) (auch u. a. Kulgemeyer et al. (2021)) formulierte Definition zugrunde gelegt, die die Fokussierung auf Unterricht als übergeordneten Denkgegenstand (ohne weitere Eingrenzungen) bereits beinhaltet:

> Unterrichtsreflexion ist die „theoriegeleitete Analyse von Unterricht mit dem Ziel der Verbesserung der Unterrichtsqualität und [bzw. oder] der Entwicklung der Professionalität von Lehrkräften“ (Kempin, Kulgemeyer & Schecker, 2020, S. 439).

Zur Darstellung der Anschlussfähigkeit an das Forschungsfeld wird diese Definition auch in Bezug auf die übrigen oben diskutierten Aspekte eingeordnet. Das Verständnis des Begriffs Unterrichtsreflexion umfasst mit der Integration der Weiterentwicklung der Unterrichtsqualität sowie der Weiterentwicklung von Professionalität sowohl eine externale als auch eine internale Zielstellung, was anschlussfähig u. a. an die Definition von Aeppli und Lötscher (2016) ist. Während von Aufschnaiter, Fraij und Kost (2019) im Einbezug der externalen Zielstellung die Gefahr sehen, (angehende) Lehrkräfte könnten durch die Fokussierung auf unterrichtliche Kontextfaktoren die vertiefte Beschäftigung mit der eigenen Professionalität vermeiden, wird die Unterrichtsreflexion mit externaler Zielstellung zwar nicht als Reflexion klassifiziert, aber trotzdem als analytische Aktivität anerkannt (Lohse-Bossenz, Schmitt, Lenske und Gold, 2023; von Aufschnaiter, Fraij und Kost, 2019). Eine Eingrenzung zum zeitlichen Zusammenhang zwischen Unterrichtssituation und Reflexion erfolgt in dieser Definition zwar nicht, im Rahmen dieser Arbeit wird aber auf die reflection-on-action fokussiert. Diese ermöglicht eine Reduktion der Komplexität der Reflexionssituation und die Distanzierung von der Situation, die den Einbezug wissenschaftlicher Theorien in die Reflexion begünstigen kann (Leonhard und Abels, 2017; Aeppli und Lötscher, 2016). Auch in Bezug auf die übrigen Unterscheidungsmerkmale ist die hier eingesetzte Definition offen. Unterrichtsreflexion kann (individuell) als Selbst- und Fremdreflexion oder (gemeinsam) in kombinierter Form durchgeführt werden. Auch bezüglich der Offenheit von Unterrichtsreflexion, ihrem Auslöser (z. B. die Initiation durch eine Irritation oder Veranlassung durch Lehrende) sowie dem Einbezug von Emotionen werden über den Theoriebezug hinaus keine weiteren Vorgaben gemacht. Ebenso ist die mediale Umsetzung der Unterrichtsreflexion prinzipiell nicht festgelegt; im Forschungskontext ist eine Darstellung nach außen (in mündlicher oder schriftlicher Form) allerdings notwendig (Leonhard & Rihm, 2011).

9.1.2 Fähigkeit zur Unterrichtsreflexion

Gemäß der in der Definition zum Begriff Unterrichtsreflexion getroffenen Beschränkung der Denkgegenstände auf solche mit unmittelbarem unterrichtlichen Bezug wird im Folgenden auch der Begriff der Reflexionsfähigkeit auf eine unterrichtsbezogene Reflexionsfähigkeit eingegrenzt, die aufbauend auf die begriffliche Ausdifferenzierung in Kapitel 7.4 die kognitiven Fähigkeiten zur Ausführung einer Unterrichtsreflexion umfasst:

> Die Fähigkeit zur Unterrichtsreflexion ist die Gesamtheit der zur Ausführung einer *Unterrichts*reflexion erforderlichen kognitiven personalen Bedingungen.

Im weiteren Verlauf der Arbeit werden die Begriffe Reflexion und Reflexionsfähigkeit im Sinne der Fokussierung auf die Reflexion von Unterricht und synonym zu den Begriffen der Unterrichtsreflexion und der Fähigkeit zur Unterrichtsreflexion verwendet.

9.1.3 Abgrenzung zum Begriff der (Unterrichts-)Analyse

Wie schon die obige Definition des Reflexionsbegriffs nahelegt, stehen die Begriffe Reflexion und Analyse in engem Zusammenhang und werden teilweise sogar synonym verwendet. Beide Begriffe werden grundsätzlich als analytische Tätigkeiten eingeordnet (Lohse-Bossenz et al., 2018; von Aufschnaiter, Fraij und Kost, 2019). Eine Abgrenzung nehmen verschiedene Autor*innen anhand unterschiedlicher Kriterien vor:

- Unter Bezugnahme auf Mühlhausen (2006) unterscheidet Wyss (2013, S. 46) zwischen der geschlossenen Reflexion (Mühlhausens (2006) *Analyse*), die die Untersuchung nach vordefinierten Kriterien beschreibt, und der offenen Reflexion (Mühlhausens (2006) *Reflexion*), der Beurteilung nach selbst gewählten Kriterien.
- Im Verständnis von Lohse-Bossenz, Schmitt, Lenske und Gold (2023) weisen Unterrichtsanalyse und -reflexion weitgehende Übereinstimmung auf. Der zentrale Unterschied liegt darin, dass bei der Unterrichtsreflexion ein (zusätzlicher) Selbstbezug hergestellt werden muss. Die Unterscheidung zwischen Reflexion und Analyse entspricht der in Kapitel 7.3 beschriebenen Unterscheidung zwischen internaler und externaler Zielstellung (von Aufschnaiter, Fraij & Kost, 2019).

Im Rahmen dieser Arbeit wird die Unterrichtsreflexion ebenfalls als analytische Tätigkeit verstanden, die sich durch das in der Definition vorgegebene Ziel auszeichnet: die Verbesserung der Unterrichtsqualität und bzw. oder Entwicklung der

Professionalität von Lehrkräften. Die konkreten Analysetätigkeiten werden - wie im nächsten Abschnitt erläutert - entsprechend dem in Kapitel 7.5.4 vorgestellten Reflexionsmodell gefasst.

9.1.4 Modellierung von Unterrichtsreflexion

Der dieser Arbeit zugrunde liegende Reflexionsbegriff versteht die Unterrichtsreflexion als analytischen Denkprozess. Das Erleben bzw. Beobachten einer Situation und bzw. oder das Umsetzen von Alternativen oder Konsequenzen sind demnach kein Teil der Reflexion.

Zur näheren Konkretisierung des Begriffsverständnisses wird auf das integrative Reflexionsmodell von Nowak, Kempin, Kulgemeyer und Borowski (2019) (s. Kapitel 7.5.4) zurückgegriffen. Die Qualität einer Reflexion wird demnach nicht allein am Auftreten eines Reflexionselements festgemacht, sondern bedarf zusätzlich einer inhaltlichen Bewertung ihrer Angemessenheit. Zur (quantifizierten) Bewertung von Reflexionen unter Einbezug der verschiedenen Dimensionen des Reflexionsmodells liegen allerdings keine expliziten Empfehlungen vor.[122] Während das entwickelte Reflexionsmodell im Sinne eines Stufenmodells die Elemente einer Reflexion (*Beschreibung*, *Bewertung*, *Alternative*, *Konsequenz*) auch hierarchisch ordnet, ist für diese Arbeit vorrangig bedeutsam, dass diese Elemente als inhaltlich aufeinander aufbauende Prozessschritte einer Reflexion zu verstehen sind. Eine Gewichtung der Stufen für die Bewertung der Qualität einer Reflexion wird nicht vorgenommen.

9.2 Anforderungen an die zu entwickelnde Lernumgebung

In Anbetracht der beschriebenen Bedeutung von Reflexion für Lehrkräfte, ihre Unterrichtsentwicklung und Professionalisierung (s. Kapitel 7.2) stellt die Fähigkeit zur Unterrichtsreflexion einen bedeutsamen Bestandteil der professionellen Handlungskompetenz dar. Da die Reflexionsfähigkeit (angehender) Lehrkräfte insgesamt nur gering ausgeprägt erscheint, besteht Bedarf zu ihrer Förderung. Diesem Bedarf könnte bereits im Studium begegnet werden. Ausgangspunkt dafür kann die fundierte Erfassung der Reflexionsfähigkeit bilden. Darauf aufbauendes

[122] Bisherige Forschungsarbeiten, in denen das Modell zur Einschätzung der Qualität von Reflexionen herangezogen wurde, gehen unterschiedlich vor: So beschränkt sich z. B. Nowak (2023) auf eine deskriptive Darstellung u. a. in Bezug auf die Art der Aussage (Reflexionselement, das ggf. mit Begründung vorliegt) sowie die zugrundeliegende Wissensbasis, während Kempin, Kulgemeyer und Schecker (2020) einen Testwert bilden, in den sowohl eine Bewertung der inhaltlichen Sinnhaftigkeit der reflektierten Aspekte als auch die jeweils erreichte Stufe einbezogen werden.

Assessment-Feedback kann dann dem Aufzeigen der Relevanz von Unterrichtsreflexion dienen und die fundierte Selbsteinschätzung[123] durch die Ergänzung eines (objektiven) Maßes für die Reflexionsfähigkeit ermöglichen (vgl. Bohndick, Kohlmeyer und Buhl, 2019).[124]

Das Entwicklungsziel dieser Arbeit umfasst daher die Erstellung einer zweiteiligen Lernumgebung bestehend aus einem Testinstrument zur Fähigkeit zur Unterrichtsreflexion und dazugehörigem Assessment-Feedback, das Teilnehmende bei der fundierten Einschätzung der Fähigkeit unterstützt. Die Erfassung ist dabei mit gewissen Herausforderungen verbunden, die sich insbesondere im bereits diskutierten Spannungsverhältnis zwischen Testökonomie und Authentizität des Testformats (s. Kapitel 6.3) niederschlagen. Daher sind für die Entwicklung der Lernumgebung zwei zentrale Ansprüche zu berücksichtigen:

- *Anspruch in Bezug auf die Testökonomie (unter Berücksichtigung des Assessment-Feedbacks)*: Die Auswertung des Testinstruments soll automatisiert erfolgen, um zeitnah Feedback zur fundierten Selbsteinschätzung bereitstellen zu können.
- *Anspruch in Bezug auf die Authentizität*: Das Testformat soll eine authentische Erfassung der Reflexionsfähigkeit ermöglichen.

Die authentische Erfassung der Reflexionsfähigkeit würde in einem (kaum standardisierten) Handlungstest erfolgen, höher standardisierte und damit weniger aufwendige Testformate wie Wissens- und Kompetenztests gehen allerdings mit geringerer Authentizität einher (vgl. Kapitel 6.2). Es ist also ein Kompromiss

[123] Die (fundierte) Selbsteinschätzung weist Überschneidungen zum (formativen) Self-Assessment auf Grundlage von Feedback auf, das die Erfassung des eigenen Lern- bzw. Fähigkeitsstandes durch Lernende u. a. auf Grundlage transparenter Kriterien beschreibt (Buholzer et al., 2020; Andrade und Valtcheva, 2009). Self-Assessment in diesem Sinne wird allerdings üblicherweise im formalisierten Lehr-Lern-Kontexten zur Optimierung von Lehr-Lern-Prozessen eingesetzt und durch Lehrende begleitet (Buholzer und Brovelli, 2023; Black und Wiliam, 2018; Pellegrino, Chudowsky und Glaser, 2001), während das hier untersuchte Assessment-Feedback auch für einen Einsatz unabhängig von (universitären) Veranstaltungen konzipiert wurde. Gleichzeitig wird die Bezeichnung des (Online-Self-)Assessments an deutschen Universitäten vermehrt zur Bezeichnung von Selbsteinschätzungstests bzw. Testinstrumenten verwendet, die Studieninteressierte bei der Entscheidung für ein Studium bzw. spezifische Studiengänge unterstützen sollen (z. B. Buhl, Sagolla und Bohndick, 2022; Mayr und Nieskens, 2015). Um diese Testinstrumente und den allgemeinen Ansatz zur Unterstützung von Lernprozessen zu differenzieren, wird letzterer fortan als (fundierte) Selbsteinschätzung bezeichnet und für die Evaluation der im Rahmen dieser Arbeit vorgestellten Lernumgebung auf die Einschätzung der eigenen Fähigkeiten auf Grundlage von Assessment-Feedback eingeschränkt.

[124] Bohndick, Kohlmeyer und Buhl (2019) verwenden den Begriff des *weiterentwicklungsorientierten Online-Self-Assessments* für online vorliegende „Selbstauskunftsskalen oder Tests, die automatisch ausgewertet werden, so dass sie sofort ein individuelles Feedback“ (Buhl, Sagolla & Bohndick, 2022, S. 120) ermöglichen und die Professionalisierung der Lehramtsstudierenden anregen sollen, allerdings wiederum im Kontext der Förderung überfachlicher Kompetenzen bzw. der Studienorientierung.

herzustellen, der die standardisierte und automatisiert auswertbare Erfassung der Reflexionsfähigkeit in einem vor diesem Hintergrund möglichst authentischen Testformat - auch für einen individuellen Einsatz des Testinstruments nach Bedarf - ermöglicht und ein für die Teilnehmenden zur fundierten Selbsteinschätzung geeignetes Feedback zulässt.

Nachfolgend wird die im Spannungsfeld dieser Ansprüche entwickelte Lernumgebung zur Messung und Rückmeldung der Reflexionsfähigkeit beschrieben (Teil II) und evaluiert (Teil III). Gemäß dem argumentbasierten Ansatz zur Validierung nach Kane (u. a. 1992; 2013; s. Kapitel 6.1.3) wird für die Evaluation der Lernumgebung ein INA formuliert, das den Bezugsrahmen für die Validierung darstellt und die Grundlage für die Beurteilung der Eignung der entwickelten Lernumgebung für die Messung und Rückmeldung der Reflexionsfähigkeit bildet.

9.3 Schlussfolgerungen für die explorative Untersuchung der Reflexionsfähigkeit und ihrer Entwicklung

Nach erfolgreicher Evaluation steht mit der entwickelten Lernumgebung ein Instrument zur Erfassung und fundierten Selbsteinschätzung der Reflexionsfähigkeit zur Verfügung. So kann ein Ausgangspunkt für die individuelle Weiterentwicklung der Reflexionsfähigkeit bereitgestellt werden. Gleichzeitig besteht Bedarf für die Fortentwicklung der Theorie in diesem Bereich: Auch wenn der Förderbedarf für Reflexionsfähigkeit bereits vielfach attestiert wurde, ist bislang unklar, wie diese sich genau entwickelt und inwiefern z. B. Stufenmodelle inhaltlich mit Entwicklungsstufen der Reflexionsfähigkeit korrespondieren oder welche Voraussetzungen oder Rahmenbedingungen eine solche Entwicklung begünstigen (Reintjes & Kunze, 2022). Ein zweites, damit verbundenes, Forschungsdesiderat liegt in Bezug auf das Verhältnis von Noticing, d. h. dem Erleben einer Situation und der Identifikation zentraler Elemente im Unterricht, und der darauf aufbauenden Unterrichtsreflexion vor (nach Amador und Weston (2024) gilt dies allgemein für den Zusammenhang zwischen Noticing und verschiedenen Aspekten professionellen Handelns). Offen ist, inwieweit die Fähigkeiten zum Noticing und zur Unterrichtsreflexion zusammenhängend sind und ggf. sogar wechselseitig Einfluss auf die Entwicklung der Fähigkeiten nehmen.

Basierend auf den im Rahmen der Validierungsstudien zur entwickelten Lernumgebung gewonnenen Daten werden in Teil IV beide Desiderate adressiert und Analysen zur Generierung von Hypothesen zur zielgerichteten Förderung der Reflexionsfähigkeit durchgeführt. Spezifische Ziele für beide Bereiche sind die folgenden:

- Zur Exploration unterschiedlicher Ausprägungen der Reflexionsfähigkeit wird eine latente Profilanalyse durchgeführt. Identifizierte Profile könnten künftig als Ausgangspunkt für eine stärker individuell zugeschnittene Förderung dienen.
- Die Untersuchung des Zusammenhangs zwischen den Fähigkeiten zum Noticing und zur Unterrichtsreflexion soll Aufschluss darüber geben, inwiefern beide Fähigkeiten unmittelbar und in Bezug auf ihre Entwicklung zusammenhängen und somit auch im Rahmen der Förderung gemeinsam betrachtet werden sollten.

Teil II

Entwicklung der Lernumgebung zur Reflexion von Physikunterricht

10 Reflexionstest

Kempin, Kulgemeyer und Schecker (2018) entwickelten einen Performanztest zur Messung der Reflexionsfähigkeit von Physik-Lehramtsstudierenden (s. Kapitel 7.7.3). Ziel dieser Arbeit ist daran anschließend die Entwicklung eines weniger ressourcenintensiven Tests zur Messung der Reflexionsfähigkeit von Physik-Lehramtsstudierenden, der mit einem Assessment-Feedback für Studierende bzw. Teilnehmende (und Dozierende) verbunden wird, um den Teilnehmenden die fundierte Selbsteinschätzung der eigenen Reflexionsfähigkeit zu ermöglichen. Um ein zeitnahes Feedback zu ermöglichen, wurden geschlossene MC-Aufgaben entwickelt, die automatisiert ausgewertet werden können. Der Test wurde in der Software *LimeSurvey* umgesetzt (LimeSurvey GmbH, n. d.). In diesem Kapitel werden der Reflexionstest vorgestellt (s. Kapitel 10.1) und das Vorgehen bei dessen Entwicklung erläutert (s. Kapitel 10.2) sowie die gewählte Bepunktungsvorschrift dargelegt und diskutiert (s. Kapitel 10.3). Abschließend wird ein Überblick über die Passung des Reflexionstests und des zugrundeliegenden Reflexionsperformanztests zu vorliegenden empirischen Befunden mit Bezug zur Erfassung der Reflexionsfähigkeit gegeben (s. Kapitel 10.4).

10.1 Aufbau des Reflexionstests

Das Setting des Reflexionstests ist weitgehend identisch zu dem ursprünglichen Reflexionsperformanztest: Proband*innen werden in eine kollegiale Beratungssituation versetzt und sollen ihrem fiktiven Mitpraktikanten (der Persona Robert) an der Praktikumsschule Feedback zu seinem Unterricht (einer Doppelstunde zur Wiederholung der Newtonschen Axiome und Einführung von Impuls und Impulserhaltung) geben. Abweichend vom Performanztest wird hier auf ein beliebiges Praktikum (nicht explizit das Praxissemester) verwiesen, um der breiteren Zielgruppe des Instruments gerecht zu werden.

Auch die Grundstruktur des Instruments wurde beibehalten (eine Übersicht über den Ablauf des Instruments ist in Abbildung 10.1 dargestellt): Nach der Bitte, sich in die kollegiale Beratungssituation zu versetzen, erhalten die Proband*innen zunächst Informationen zu den Rahmenbedingungen des Unterrichts. Dazu gehören (a) eine Beschreibung der Lerngruppe bezüglich Jahrgangsstufe, Schulform, Interesse und fachlichem Leistungsstand sowie das Vorwissen aus vorangegangenem

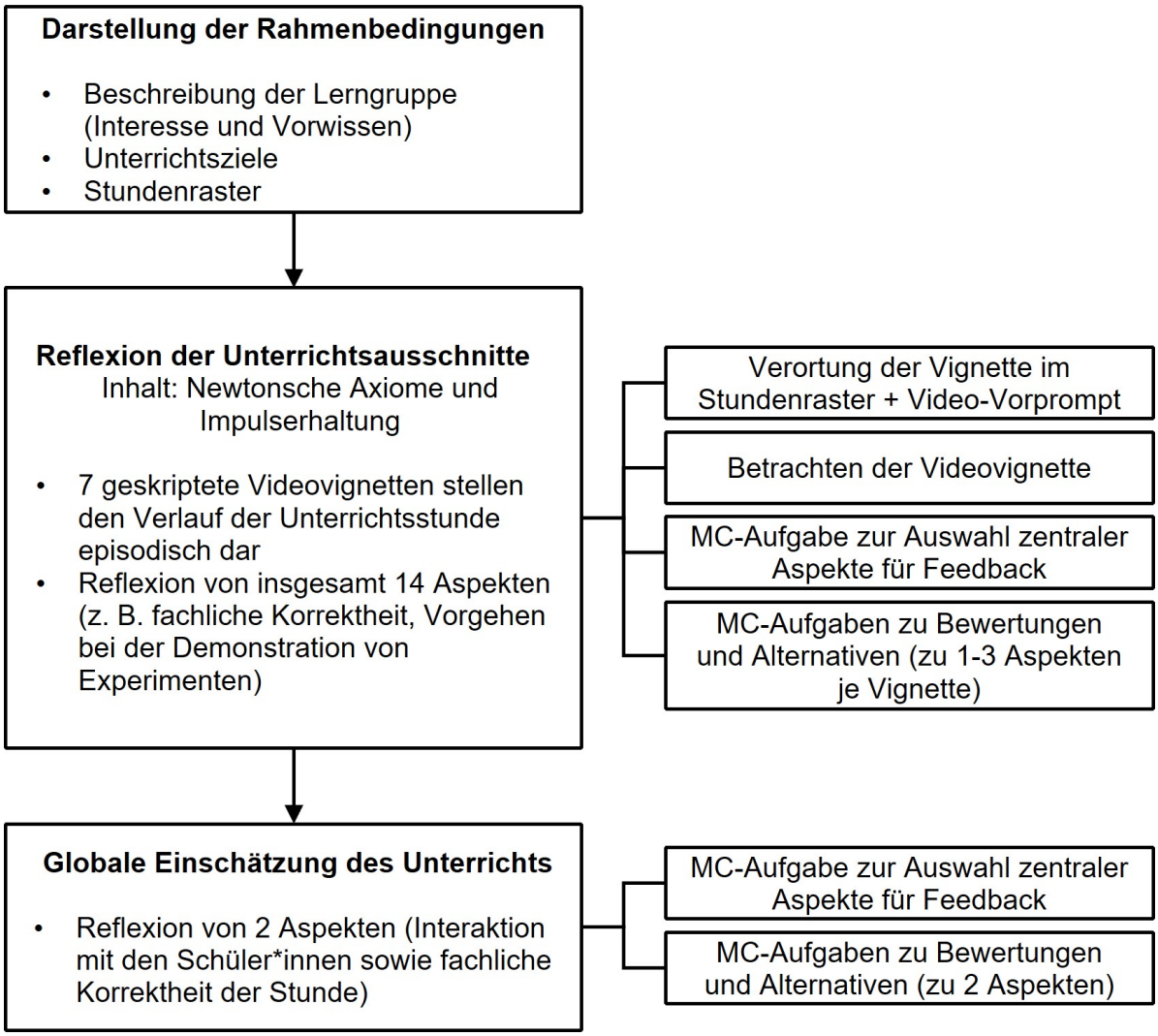


Abbildung 10.1: Struktur des Reflexionstests (Darstellung aufbauend auf Kulgemeyer et al., 2021, S. 3044)

Unterricht in Bezug auf das Stundenthema, (b) die Nennung der Unterrichtsziele für die folgende Doppelstunde sowie (c) der in einem Stundenraster dargestellte geplante Stundenablauf. Anschließend reflektieren die Proband*innen eine Auswahl von sieben Videovignetten, die inhaltlich den Verlauf der Doppelstunde nachzeichnen und zuletzt die Doppelstunde insgesamt.

Die Reflexion der sieben Vignetten erfolgt vorstrukturiert durch Prompts und MC-Aufgaben. Jede Vignette wird durch einen Vorprompt eingeleitet, in dem Robert die folgende Vignette in einem kurzen Video einleitet (z. B. „[...] Ich bin mir nicht ganz sicher, ob ich die Einführung [des Impulsbegriffs] sinnvoll gestaltet habe."). Diese Video-Prompts sind identisch zu denen im Reflexionsperformanztest. Zusätzlich erfolgt eine Einordnung der Vignette in das Stundenraster, indem die jeweilige Phase im Stundenraster farblich hervorgehoben wurde.

Da die Zahl der Vignetten gegenüber dem Reflexionsperformanztest reduziert wurde, aber weiterhin die gesamte Doppelstunde nachvollzogen werden soll, werden diesen Video-Prompts außerdem kurze schriftliche Prompts vorangestellt, sofern im Reflexionsperformanztest vorangegangene Vignetten nicht in den Reflexionstest übernommen wurden (auf Kriterien für die Auswahl der Vignetten wird in Kapitel 10.2.1 eingegangen). In diesen schriftlichen Prompts werden insbesondere die Aspekte aus den gestrichenen Szenen erläutert, die für das Verständnis der jeweiligen Vignetten und das Nachvollziehen der Doppelstunde insgesamt notwendig sind. Im Anschluss an die Vorprompts werden die entsprechenden Vignetten angesehen, die jeweils problematische und bzw. oder gelungene Aspekte beinhalten.

An jede Vignette schließt sich eine erste MC-Aufgabe zum Noticing an: „Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung (Lob und/oder Kritik) erhalten, um etwas für sich bzw. seinen Unterricht zu lernen?“ Die Proband*innen werden also dazu aufgefordert, auszuwählen, welche der nachfolgenden Aspekte der Vignette besonders relevant für die Rückmeldung bzw. Reflexion sind.

Dann erfolgt die Erfassung der Reflexionen der Proband*innen mit Hilfe von weiteren MC-Aufgaben. Je Vignette werden ein bis drei Aspekte reflektiert, indem jeweils (a) eine MC-Aufgabe zur Bewertung bearbeitet wird („Welches Feedback möchten Sie Robert [zum jeweiligen Aspekt] geben?“; hier können aus verschiedenen begründeten Bewertungen des Aspekts diejenigen ausgewählt werden, die der Einschätzung der Proband*innen entsprechen) und (b) eine MC-Aufgabe zu alternativen Handlungsoptionen bezüglich des Aspekts vorgeschlagen werden („Welche der unten stehenden Hinweise [zum jeweiligen Aspekt] möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben?“; auch hier können die Proband*innen aus verschiedenen begründeten Alternativen wählen). Die zuvor im Rahmen der Aufgabe zum Noticing getroffene Auswahl relevanter Aspekte hat dabei keinen Einfluss auf die nachfolgend zu bearbeitenden Aufgaben. Das Testinstrument weist eine lineare Struktur auf, sodass alle Teilnehmenden dieselben Aufgaben bearbeiten. Um für die Proband*innen transparent darzustellen, dass die nachfolgenden MC-Aufgaben nicht unbedingt die von ihnen ausgewählten Aspekte thematisieren, wird angegeben, dass Robert bereits einige Rückmeldungen erhalten hat und sich deswegen Rückmeldungen zu spezifischen Aspekten wünscht (z. B. zur vierten Vignette: „Robert hat schon einige Rückmeldungen erhalten. Besonders wichtig ist ihm jetzt noch eine Rückmeldung zur Wahl des Experiments und zur Ergebnissicherung nach dem Experiment.“). Alle Aufgaben sind in Kapitel A (ab S. 417) angehängt.[125]

Über die sieben Vignetten hinweg werden insgesamt 16 Aspekte z. B. zur fachlichen Korrektheit formulierter Merksätze, der Nutzung von Fachsprache oder dem geschlechtsstereopytischen Umgang mit den Schüler*innen reflektiert. Ta-

[125]Das vollständige Testinstrument kann unter https://umfragen.physikdidaktik.uni-bremen.de/index.php/516586?lang=de eingesehen werden.

belle 10.1 stellt eine Übersicht über die Vignetten und die jeweils reflektierten Aspekte dar. Eine Ausnahme bezüglich der Aufgabenstruktur stellen die Aufgaben zu aufgetretenen Schülervorstellungen zur zweiten Vignette dar. Anstelle einer Bewertung des Handelns der Lehrkraft stehen hier die Schüler*innen und ihre Vorstellungen im Fokus. Daher wird abweichend zunächst erfragt, welche Schülervorstellungen die Proband*innen in der Vignette erkannt haben („Welche der folgenden Schülerfehlvorstellungen sind während des Unterrichtsausschnitts sichtbar geworden?“) und anschließend im Sinne eines alternativen Vorgehens, wie Robert mit ihnen hätte umgehen sollen („Welche der unten stehenden Hinweise zum Umgang mit Schülerfehlvorstellungen möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben?“; in den Vignetten greift Robert die Schülervorstellungen nicht auf).

Nach der Reflexion der sieben Vignetten erfolgt eine globale Betrachtung des Unterrichts mit analogen MC-Aufgaben: Auch hier wird nach einer Noticing-Aufgabe zur Auswahl der relevanten Aspekte die Interaktion mit den Schüler*innen (u. a. stereotypisches Verhalten, fehlende Verständnisversicherung) sowie die fachliche Korrektheit der Stunde (u. a. fachliche Fehler, fehlende Thematisierung der Schülervorstellungen) bewertet und es werden Alternativen vorgeschlagen.

Während der Bearbeitung aller MC-Aufgaben stehen den Proband*innen Hilfestellungen zur Verfügung, die im Sinne des Testsettings als Rückfragen an Robert formuliert sind. So können die Proband*innen die Informationen bezüglich der beschriebenen Rahmenbedingungen nachlesen (z. B. „Kannst du mir noch einmal sagen, welches Vorwissen die Schülerinnen und Schüler haben?“), die Videovignetten erneut ansehen („Kann ich den Unterrichtsausschnitt noch einmal sehen?“) oder in der jeweiligen Vignette eingesetztes Material aufrufen (z. B. „Kannst du mir die PowerPoint-Folien noch einmal zeigen?“).

Vignette	Inhalt der Videovignette	Reflektierte Aspekte
1	Besprechung von Quizfragen zur Wiederholung der Newtonschen Axiome	• fachlich falsche und zu komplexe Formulierung des Merksatzes zum 3. Newtonschen Axiom
2	Diskussion eines Comics zum 3. Newtonschen Axiom	• aufgetretene Schülervorstellungen zum Kraftbegriff
3	Einführung des Impulsbegriffs am Beispiel der Titanic	• Verwendung des alltagssprachlichen Begriffs „Schwung“ in der Einführung des Impulsbegriffs • abrupte Gestaltung des Phasenübergangs zwischen Wiederholung der Newtonschen Axiome und Einführung des Impulsbegriffs
4	Durchführung eines Freihandexperiments (Werfen von zwei Tennisbällen mit unterschiedlicher Masse bzw. unterschiedlicher Geschwindigkeit) zur Veranschaulichung der Zusammenhänge zwischen Masse bzw. Geschwindigkeit und Impuls	• Wahl des Freihandexperiments zur Veranschaulichung der Zusammenhänge zwischen Masse bzw. Geschwindigkeit und Impuls • Sicherung der Einführung des Impulsbegriffs mit einem vorgegebenem, fachlich problematischen Merksatz
5	Durchführung eines Demonstrationsexperiments zur Impulserhaltung (Zusammenstoß zweier Wagen auf der Luftkissenbahn)	• stereotypischer Umgang mit den Schüler*innen • fachlich problematischer Umgang mit Messunsicherheiten im Demonstrationsexperiment • kognitive Aktivierung der Schüler*innen während des Demonstrationsexperiments

6	Experimentelle Überprüfung der Impulserhaltung mit den Messwerten des Demonstrationsexperiments	• stereotypischer Umgang mit den Schüler*innen • problematische Umgang mit Messunsicherheiten in der Berechnung der Impulserhaltung auf Rückfrage einer Schülerin • fehlende Thematisierung der Einheit des Impulses
7	Sicherung der Erkenntnisse aus dem Demonstrationsexperiment zur Impulserhaltung mit einem vorgefertigten Merksatz	• zu komplexe Formulierung des Merksatzes zur Impulserhaltung • Umgang mit den von Schüler*innen formulierten Merksätzen
G	*Globale Reflexion der Doppelstunde; keine Videovignette*	• Interaktion mit den Schüler*innen im Verlauf der Unterrichtsstunde • fachliche Korrektheit in der Unterrichtsstunde

Tabelle 10.1: Übersicht über die Videovignetten und reflektierten Aspekte im Reflexionstest

10.2 Entwicklung und Ausgestaltung des Reflexionstests

Das Ziel, Teilnehmenden ein Assessment-Feedback zu ihrem Abschneiden im Reflexionstest zur Verfügung zu stellen, wird bereits bei der Entwicklung des Reflexionstests berücksichtigt. Um ein zeitnahes Feedback zu ermöglichen, werden geschlossene MC-Aufgaben entwickelt, die automatisiert ausgewertet werden können. Bei der Gestaltung des Tests werden abwägend Überlegungen zum resultierenden Umfang des Testinstruments, zur Fokussierung auf für den Physikunterricht spezifische Situationen und zu authentischen Reflexionen ohne die zu starke Reduktion der inhaltlichen Breite des abgebildeten Unterrichts berücksichtigt.

10.2.1 Auswahl der Videovignetten

Die Bearbeitung der MC-Aufgaben weist gegenüber dem freien Sprechen im Performanztest eine höhere Bearbeitungsdauer auf. Um trotzdem eine angemessene Bearbeitungsdauer des Testinstruments zu erreichen, wurde die Anzahl der Vignetten gegenüber dem Reflexionsperformanztest reduziert. Die Auswahl der aus dem Reflexionsperformanztest zu übernehmenden Vignetten erfolgte entlang folgender Kriterien:

1. Die verbleibenden Vignetten sollen ausreichen, um den Verlauf der Doppelstunde insgesamt nachvollziehbar darzustellen.
2. Gleichzeitig sollen die verbleibenden Vignetten möglichst vielfältige Situationen des Physikunterrichts darstellen und die verschiedenen dargestellten Unterrichtsphasen (Wiederholung der Newtonschen Axiome als kognitiv aktivierender Einstieg in die Unterrichtsstunde, schrittweise Erarbeitung von Impulsbegriff und Impulserhaltung sowie Sicherung der neuen Konzepte) abbilden, um die Generalisierbarkeit der Testwertinterpretation über verschiedene Situationen des Physikunterrichts gegenüber dem Reflexionsperformanztest möglichst geringfügig einzuschränken.
3. Der inhaltliche Fokus liegt auf der Reflexion von *Physik*unterricht, sodass vorrangig die Vignetten beibehalten werden sollen, die zur Reflexion mit einem fachlichen bzw. fachdidaktischen Fokus anregen sollen (d. h. zu denen das Kodiermanual vorrangig fachliche und fachdidaktische Reflexionspunkte enthält).

Insgesamt wurden sieben der 13 Videovignetten aus verschiedenen Phasen des Unterrichts in den Reflexionstest übernommen.

10.2.2 Entwicklung der Aufgaben

Die Entwicklung der Antwortoptionen erfolgte auf Grundlage der Bearbeitungen des Reflexionsperformanztests aus dem Projekt *ProfiLeP+*. Insgesamt liegen 167 Bearbeitungen des Testinstruments vor. Da im Reflexionsperformanztest die mündlichen Reflexionen zu einzelnen Vignetten ausgelassen werden können, sind nicht alle Bearbeitungen vollständig. Es liegen jeweils zwischen 85 und 158 Reflexionen zu den Vignetten vor. Da für die Bearbeitungen außerdem maximal 70 Minuten zur Verfügung standen und die Bearbeitung des Tests nach Ablauf der Erhebungszeit ggf. abgebrochen werden musste, liegen insbesondere zu den späteren Vignetten weniger Reflexionen vor.

Die Konstruktion der Antwortoptionen erfolgte dann in mehreren Schritten. Zunächst wurden alle Reflexionen der Studierenden item- bzw. vignettenweise angehört und alle enthaltenen Aussagen (u. a. die Elemente des Reflexionsmodells, aber auch eigene Erfahrungen, formulierte Fragen, ...) paraphrasiert. Inhaltlich ähnliche (paraphrasierte) Aussagen wurden geclustert und identische Aussagen zusammengefasst. Aus den Aussagenclustern wurden induktiv Oberkategorien formuliert, die möglichst prägnant benennen, auf welche Inhalte der Vignette sich die Aussagen beziehen. Nicht nachvollziehbare Aussagen der Studierenden wurden ausgelassen. Die Oberkategorien dienen als Antwortoptionen in den Noticing-Aufgaben, die jeweils zwischen der Ansicht der Videovignette und der Bearbeitung der Reflexionsaufgaben verortet sind.

Aus den so geclusterten Aussagen wurden entlang des Reflexionsmodells vorläufige Antwortoptionen formuliert und in jedem Cluster in Antwortoptionen zu Bewertungen und Alternativen (inklusive Konsequenzen)[126] gruppiert. Die Antwortoptionen werden jeweils mit Begründung formuliert, um aus testökonomischen Gründen auf mehrstufige Fragen zu verzichten. Nun werden die Antwortoptionen ein erstes Mal mit dem Kodiermanual zum Reflexionsperformanztest abgeglichen, um jeweils festzuhalten, ob sie relevante Rückmeldungen enthalten (also Attraktoren darstellen) oder als Distraktoren einzustufen sind (s. Kapitel 10.2.3). Auf dieser Grundlage wurden die vorläufigen (Gruppen von) Antwortoptionen überarbeitet, wobei sichergestellt wurde, dass

- alle Antwortoptionen Begründungen enthalten (ggf. wurden fehlende Begründungen ergänzt),
- zu jeder Oberkategorie, d. h. in jeder Aufgabe, sowohl positive, als auch negative Antwortoptionen vorliegen; für Aufgaben zu Bewertungen wurde ein möglichst ausgeglichenes Verhältnis positiver und negativer Bewertungen angestrebt, für Aufgaben zu Alternativen wurde sichergestellt, dass jeweils

[126] Da zu vielen Aspekten keine Konsequenzen genannt werden, werden diese in die Aufgaben zu Alternativen inkludiert. Im Folgenden werden sie kurz als Aufgaben zu Alternativen benannt.

eine positive Alternative vorliegt („positive Alternativen“ entsprechen der Empfehlung, ein bestimmtes Verhalten beizubehalten),

- die Formulierungen der Antwortoptionen möglichst präzise und verständlich sind (es sollten also z. B. keine komplexen Satzstrukturen wie doppelte Verneinungen vorliegen oder in einer Antwortoption mehrere Bewertungen, Alternativen oder Begründungen vorkommen; z. B. Busker, 2014; Jonkisz, Moosbrugger und Brandt, 2012). Eine gewisse sprachliche und inhaltliche Komplexität, die durch die Kombination von Bewertung bzw. Alternative und Begründung sowie ggf. Beschreibung zur Verortung des Aspekts in der Unterrichtsvignette verursacht wird, kann allerdings nicht vermieden werden.

Abschließend wurde die Zahl der Antwortoptionen über die Aufgaben hinweg vereinheitlicht. Ziel war die Begrenzung der Zahl der Antwortoptionen auf fünf bis sieben, was in etwa der Anzahl an vorläufigen Antwortoptionen für viele Aufgaben entsprach. Lagen zu viele vorläufige Antwortoptionen vor, wurden ähnliche Optionen zusammengefasst bzw. solche herausgenommen, die sich gemäß dem Kodiermanual zum Reflexionsperformanztest nicht auf für das Feedback relevante Aspekte des Unterrichts bezogen, sodass dieser Rahmen in 30 der 32 Aufgaben unkompliziert erreicht wurde. Jeweils eine Aufgabe umfasst vier (Aufgabe R7Bew1)[127] bzw. acht Antwortoptionen (Aufgabe RGBew1). Hier konnte auf Grundlage der vorliegenden Reflexionen sowie des Kodiermanuals keine zusätzliche Antwortoption gefunden werden bzw. hätte die weitere Reduktion der Antwortoptionen dazu geführt, dass die inhaltliche Breite der Rückmeldung deutlich reduziert worden wäre. Zusätzlich wurde sichergestellt, dass das Verhältnis von korrekten bzw. inkorrekten Antwortoptionen in jeder Aufgabe möglichst ausgeglichen ist (zur Bewertung der Antwortoptionen s. Kapitel 10.2.3).

Im Rahmen der bisher beschriebenen Entwicklung der Aufgaben und Antwortoptionen wurden zunächst deutlich mehr Aufgaben entwickelt, als in das Testinstrument inkludiert werden konnten. Eine Reduktion der Aufgaben, die je Videovignette bearbeitet werden, war - ebenso wie schon die Reduktion der eingebundenen Videovignetten - notwendig, um die Bearbeitungsdauer auf ein handhabbares Maß zu reduzieren. Um eine einheitliche Struktur im Testinstrument herzustellen, werden dabei jeweils die inhaltlich zusammenhängenden Aufgabenpaare aus Bewertungen und Alternativen zu einem Aspekt entweder gemeinsam ins Testinstrument eingebunden oder gemeinsam ausgelassen. Die Auswahl der Aspekte erfolgte auf Grundlage des Kodiermanuals zum Reflexionsperformanztest nach folgenden Kriterien:

[127] Alle Testaufgaben sind vollständig im Anhang in Kapitel A ab Seite 417 abgedruckt, wobei die Kürzel die Zuordnung der Aufgaben zur Videovignette (*R1* bis *R7* bzw. *RG* für die globale Reflexion) bezeichnen, *Bew* bzw. *Alt* angibt, ob es sich um eine Aufgabe zu Bewertungen oder zu Alternativen handelt und die abschließende Ziffer sich auf die je Vignette reflektierten Aspekte bezieht.

- Wie oben in Bezug auf die Auswahl der Vignetten für das Testinstrument beschrieben, weist das Testinstrument eine Fokussierung auf den *Physik*-unterricht auf. Daher wurden vorrangig die Aufgabenpaare inkludiert, die fachliche oder fachdidaktische Aspekte des Kodiermanuals abbilden. Da eine trennscharfe Zuweisung der Aufgaben(-paare) zu den Wissensbasen nicht möglich ist, sind aber durchaus alle drei Wissensbasen (fachliches, fachdidaktisches und pädagogisches Wissen) vertreten.
- Im Kodiermanual zum Performanztest werden außerdem besonders wichtige Reflexionspunkte als sogenannte „Kernindikatoren“ herausgestellt. Die Aufgabenpaare, die Antwortoptionen zu diesen Aspekten beinhalten, wurden vorrangig in das Testinstrument aufgenommen. Dazu gehören auch pädagogische Aspekte, wie z. B. die kognitive Aktivierung der Schüler*innen (während der Durchführung eines Demonstrationsexperiments in der fünften Vignette), die nicht für den Physikunterricht spezifisch sind.
- Pauschal ausgeschlossen wurden (unabhängig von den beiden zuvor genannten Kriterien) diejenigen Aufgabenpaare, die gemäß dem Kodiermanual zum Reflexionsperformanztest zwar als relevant eingestuft werden, in ihm allerdings so neutral bewertet werden, dass alle zugehörigen Antwortoptionen „korrekte“ Bewertungen oder Alternativen darstellen würden. Diese Aufgaben sind inhaltlich kaum dazu geeignet, die Fähigkeit zur Unterrichtsreflexion differenziert zu erfassen.

10.2.3 Bewertung der Antwortoptionen

Die Klassifizierung der Antwortoptionen in „korrekt“ bzw. „inkorrekt“ basiert auf dem Kodiermanual zum Reflexionsperformanztest. Das Kodiermanual beschreibt diejenigen Aspekte des Unterrichts, die idealerweise in einer Reflexion zur Unterrichtsstunde berücksichtigt werden sollten.

Bewertung der Antwortoptionen zur Unterrichtsreflexion

Die formulierten Antwortoptionen wurden vor dem Hintergrund des Kodiermanuals evaluiert: Als korrekt eingestuft wurden diejenigen Antwortoptionen, die als (Teil einer) Reflexion im Performanztest als Bewertung bzw. Alternative kodiert wurden; die übrigen Antwortoptionen werden als inkorrekt eingestuft. Es ist anzumerken, dass nicht jede als inkorrekt bewertete Antwortoption tatsächlich unangebrachtes bzw. (fachlich) falsches Feedback darstellt; die eindeutige Bewertung von Antwortoptionen als richtig oder falsch erscheint gerade für fachdidaktische Fragestellungen schwierig (Tepner & Dollny, 2014).[128] Vielmehr geht es darum,

[128] Sie bevorzugen daher z. B. Aufgabenformate, die ein Ranking verschiedener Antwortoptionen erfordern, gegenüber MC-Formaten (Tepner & Dollny, 2014).

die bedeutsamen Hinweise herauszufiltern. Bspw. werden die Schüler*innen in der Durchführung des Demonstrationsexperiments zur Impulserhaltung (Szene 5) nicht ausreichend kognitiv aktiviert (sie erhalten z. B. keinen Beobachtungsauftrag). Die Antwortoption „Es ist gut, dass du Schüler*innen in das Experiment einbeziehst, um sie im Unterricht zu aktiveren." stellt so zwar kein grundsätzlich falsches Feedback dar; die reine Aktivierung (z. B. durch den Anschrieb der von der Lehrkraft diktierten Werte) ist allerdings nicht ausreichend und wird demnach in dieser Situation als inkorrekt bewertet.

Aufgrund der neutralen Formulierung einzelner Aspekte im Kodiermanual kommt es vereinzelt vor, dass positive und negative Antwortoptionen zu einem Aspekt als korrekt gelten. Ein Beispiel dafür betrifft die Verwendung des alltagssprachlichen Begriffs „Schwung" bei der Einführung des Impulsbegriffs in der vierten Videovignette des entwickelten Reflexionstests dar. Die Formulierung im Kodiermanual lautet: „Es wird erkannt, dass im Dialog Alltagssprache genutzt wird (z. B. ‚Schwung', ‚schwer')"; Kulgemeyer, 2020, S. 56)

Bewertung der Antwortoptionen zum Noticing

Die Bewertung der Antwortoptionen in den Noticing-Aufgaben erfolgte grundsätzlich analog: Eine Antwortoption gilt dann als korrekt, wenn sie einen Aspekt (bzw. mehrere Aspekte) anspricht, deren Reflexion gemäß dem Kodiermanual bedeutsam ist. Passend zur Fokussierung auf Physikunterricht im Testinstrument weisen inkorrekte Antwortoptionen kaum Bezug zum Fach Physik bzw. zur Physikdidaktik auf und betreffen z. B. die Gesprächsführung (Szene 3 und 4) sowie das Auftreten der Lehrkraft (globale Reflexion).

10.3 Bestimmung der Testwerte des Reflexionstests

Das optimale Vorgehen bei der Bepunktung von Reflexionen ist nicht klar aus der Literatur zur Unterrichtsreflexion (insb. zu deren Qualitätskriterien) ableitbar. Aus diesem Grund wurde zur Bestimmung der Testwerte ein Vorgehen gewählt, das möglichst wenig Annahmen zur Struktur der Reflexionsfähigkeit impliziert. So werden bspw. Aufgaben zu Bewertungen und Alternativen eines Aspekts nicht als gekoppelte Aufgabenpaare, sondern unabhängig voneinander verstanden. Innerhalb einer Aufgabe erfolgt das Scoring auf Ebene der angewählten Antwortoptionen und es wird z. B. nicht die Passung verschiedener ausgewählter Antwortoptionen zueinander berücksichtigt.

10.3.1 Bestimmung der Testwerte zur Unterrichtsreflexion und zum Noticing

In die Bestimmung des Testwerts zur Unterrichtsreflexion gehen alle 32 Aufgaben zu Bewertungen und Alternativen ein. Maßgebend für den Aufgabenscore[129] ist jeweils die Anzahl der angewählten korrekten Antwortoptionen; die Anwahl inkorrekter Optionen nimmt keinen Einfluss auf den Aufgabenscore bzw. den Testwert insgesamt. Da die Zahl der Antwortoptionen und der Anteil der korrekten Antwortoptionen über die Aufgaben schwanken, wird die Maximalpunktzahl jeder Aufgabe auf eins normiert. Der Aufgabenscore entspricht somit:

$$\text{Aufgabenscore} = \frac{\text{Anzahl der } \textit{angewählten} \text{ korrekten Antwortoptionen}}{\text{Anzahl der korrekten Antwortoptionen}}$$

Der maximale Testwert liegt damit bei 32 Punkten. In diesem Verfahren wird keine Gewichtung zwischen den Aufgaben zu Bewertungen und Alternativen vorgenommen. Auch wird keine Gewichtung der Relevanz verschiedener reflektierter Aspekte vorgenommen, die im Reflexionsperformanztest durch die Hervorhebung von Kernindikatoren angelegt ist.

Analog zum Testwert zur Unterrichtsreflexion wird auch der zum Noticing bestimmt, indem die auf eins normierten Scores aller acht Aufgaben zum Noticing (jeweils eine zu jeder Videovignette und zur globalen Reflexion der Unterrichtsstunde) aufsummiert werden. Hier können maximal acht Punkte erreicht werden.

10.3.2 Diskussion weiterer Bepunktungsvorschriften

Diskussionswürdig ist die Tatsache, dass die angewählten inkorrekten Antwortoptionen in der Bepunktung nicht berücksichtigt werden. Im Extremfall würde dies dazu führen, dass eine Person, die stets alle Antwortoptionen ankreuzt, den maximal möglichen Testwerte erreicht. Dem könnte mit verschiedenen Maßnahmen entgegengewirkt werden. Beispielsweise könnte die Anwahl von inkorrekten Antwortoptionen mit einem Punktabzug einhergehen (bei minimal 0 Punkten je Aufgabe). Unklar ist dann allerdings, ob zwischen inkorrekten und tatsächlich unangebrachten bzw. falschen Antwortoptionen sowie inkorrekten, da (nur) in der Situation nicht relevanten, Antwortoptionen unterschieden werden müsste. Ein Vergleich dieser beiden Bepunktungsvorschriften (unter Gleichbehandlung aller inkorrekten Antwortoptionen) für die Gesamtstichprobe von $N = 192$ Bearbeitungen des Reflexionstests verdeutlicht, dass die Testwerte deutlich niedriger sind, wenn für falsche Antwortoptionen Punkte abgezogen werden (im Mittel werden dann nur noch $M = 32,4\,\%$ ($SD = 7,2\,\%$) der Maximalpunktzahl anstelle von $M = 66,3\,\%$

[129] Der Begriff des Scores wird für die je Aufgabe erreichten Punktzahlen verwendet und ist damit abzugrenzen von aggregierten Testwerten, die als Maß für Fähigkeiten interpretiert werden.

($SD = 16,1$ %) erreicht). Es zeigt sich außerdem, dass beide Scoring-Systeme insgesamt stark miteinander korrelieren ($r_{\text{Pearson}} = 0,74$; $p < 0,01$). Da die Testwerte im Rahmen des Assessment-Feedbacks zurückgemeldet werden, erscheint das Beibehalten der höheren Testwerte unter motivationalen Gesichtspunkten sinnvoll (vgl. Narciss, 2006, S. 32ff.).

Eine andere Möglichkeit zur Anpassung der Bestimmung der Testwerte wäre die Begrenzung der anwählbaren Antwortoptionen schon in der Bearbeitung des Reflexionstests. So würde die Anwahl inkorrekter Antwortoptionen indirekt Berücksichtigung finden: Sobald eine inkorrekte Antwortoption ausgewählt wird, ist es nicht mehr möglich, die volle Punktzahl in einer Aufgabe zu erreichen. In der Umsetzung birgt dieses Vorgehen die Schwierigkeit, dass die Anzahl der korrekten Antwortoptionen über die Reflexionsaufgaben variiert, sodass die Zahl der anwählbaren Option u. U. nicht mit der Zahl der korrekten Antworten übereinstimmt - zumindest dann nicht, wenn die Vorgabe im ganzen Testinstrument einheitlich sein soll. Unter diesen Umständen könnte z. B. die Begrenzung auf die Auswahl von drei Antwortoptionen sinnvoll sein, da damit in der Hälfte der Aufgaben alle korrekten Antwortoptionen angewählt werden können (gleiches gilt für die Begrenzung auf vier Antwortoptionen, diese stellt allerdings insbesondere in Aufgaben mit nur fünf oder sechs Antwortoptionen nur eine geringe Reduktion des Umfangs an anwählbaren Optionen dar). In jeder Aufgabe könnten so einheitlich 0 bis 3 Punkte (bzw. normiert 0 bis 1 Punkte) erreicht werden abhängig davon, wie viele korrekte Antworten angewählt wurden.

Mit allen hier beschriebenen Vorgehensweisen geht die Schwierigkeit einher, dass ein Umgang mit den unterschiedlichen Anzahlen an korrekten Antwortoptionen und insgesamt für die Aufgabe vorliegenden Antwortoptionen gefunden werden muss. Vor dem Hintergrund des Einsatzes des Reflexionstests als Werkzeug für die fundierte Selbsteinschätzung der Reflexionsfähigkeit erscheint die Begrenzung der Anzahl der anwählbaren Antwortoptionen praktisch nicht nötig. In der Bearbeitung der Reflexionsaufgaben wählen die insgesamt $N = 192$ Teilnehmenden nur in den seltensten Fällen alle Antwortoptionen einer Aufgabe an; insgesamt bei ca. $3,2$ % aller einzelnen Aufgabenbearbeitungen.

10.4 Erläuterungen zur Konzeption des Reflexionstests auf Grundlage des Reflexionsperformanztests von Kempin, Kulgemeyer und Schecker (2018)

Bereits bei der Konstruktion des Reflexionsperformanztests wurden diverse Entscheidungen getroffen, die in Einklang mit der vorliegenden Literatur zum Bereich der Unterrichtsreflexion und zum Einsatz von Unterrichtsvideos stehen. Ein Überblick hierzu, der die Eignung des Reflexionsperformanztests als Grundlage für den Reflexionstest (ergänzend zu den in Kapitel 7.7.3 dargestellten Validitätsargumenten) stützt, wird nachfolgend gegeben. Ergänzend wird dargestellt, inwieweit die vorliegenden Befunde auf den hier vorgestellten Reflexionstest übertragbar sind.

Die Entwicklung beider Testinstrumente ist mit dem Ziel verbunden, die Reflexionsfähigkeit (angehender) Lehrkräfte in einem - wenn auch unterschiedlich stark - standardisierten Setting zu erfassen und dennoch in diesem Rahmen möglichst authentische Bedingungen herzustellen. Zur Herstellung einer standardisierten Reflexionssituation wird eine Fremdreflexion durchgeführt. So reflektieren alle Proband*innen den selben Unterricht. Auf Grundlage bisheriger Befunde ist weder die Selbst- noch die Fremdreflexion eindeutig zu bevorzugen (Riel, 2022, S. 101). Während z. B. Weber, Elstrodt-Wefing und Hoge (2022) beobachten, dass Reflexionen des eigenen Unterrichts (im Vergleich zur Fremdreflexion) eher zur Formulierung von alternativen Handlungsoptionen und Ableiten von Konsequenzen führen (s. auch Hamel und Viau-Guay, 2019), stellen u. a. Seidel et al. (2011) das Gegenteil fest: Studierende, die eigene Unterrichtsvideos reflektierten, waren weniger kritisch und ihre Reflexionen beinhalteten weniger alternative Handlungsoptionen und Konsequenzen.

Der Einsatz von Videovignetten ermöglicht eine vergleichsweise authentische Darstellung des Unterrichts (gegenüber bspw. Textvignetten). Damit einher geht auch, dass Unterricht eher in seiner Komplexität (z. B. aufgrund gleichzeitig ablaufender Situationen) abgebildet werden kann (s. Kapitel 6.4). Die Entscheidung, geskriptete Videovignetten einzusetzen, ermöglicht außerdem die Steuerung der dargestellten Komplexität der Unterrichtssituationen, indem bspw. zwischenzeitlich nur die Lehrperson oder einzelne Schüler*innen gezeigt werden. Zusätzlich kann so ein expliziter inhaltlicher Fokus gelegt werden. Beides ermöglicht die Steuerung der Aufmerksamkeit der Teilnehmenden (Neuweg, 2015b).

Um auch für die Erhebungssituation einen Kompromiss zwischen einer möglichst authentischen Reflexionssituation bei gleichzeitig starker Standardisierung zu finden, wird eine Gesprächssituation im Rahmen einer Peer-Reflexion im Praxissemester simuliert. Diese wird im Reflexionsperformanztest auch mündlich ausgeführt. Damit wird der Tatsache Rechnung getragen, dass Reflexionsgespräche in „all-

täglichen, beruflichen Situationen meist ist einer mündliche[n] Form“ (Vogelsang et al., 2019, S. 481) stattfinden. Auch das Sekretariat der Kultusministerkonferenz (2004) beschreibt die „gegenseitige Hospitation und gemeinsame Reflexion“ (Sekretariat der Kultusministerkonferenz, 2004, S. 6) als gewinnbringend für die Professionalisierung angehender Lehrkräfte. Für den hier entwickelten Reflexionstest muss aufgrund der Verwendung von MC-Aufgaben von einer Einschränkung der Authentizität der Gesprächssituation ausgegangen werden. Weiterhin gilt aber, dass die Simulation der Situation mit einem *fiktiven* Mitpraktikanten dazu beitragen kann, die Äußerung von Kritik niedrigschwellig zu ermöglichen, da weder ein Hierarchieverhältnis besteht noch ein reales z. B. freundschaftliches oder antipathisches Verhältnis, das die Reflexion beeinflussen könnte (Raaflaub, Wyss und Hüsler, 2019; Gebhardt und Jenert, 2011).

Durch den Einsatz der Vor- und Nachprompts im Reflexionsperformanztest, in denen Robert Fragen zu seinem Unterricht stellt, sollen die Proband*innen zur Reflexion animiert werden, ohne diese explizit einzufordern, was die valide Interpretation der Testwerte gefährden könnte (Riel, 2022, S. 103). Darin werden mögliche Ausgangspunkte für die Reflexion der Vignette benannt (z. B. der Unterrichtseinstieg oder die Präsentation eines Demonstrationsexperiments). Da Studierende zumindest in Nachbesprechungen ihres eigenen Unterrichts mit mentorierenden Lehrkräften nur zu einem geringen Teil neue Themen in die Reflexion einbringen (Futter (2017, S. 236) stellt fest, dass lediglich 22 % der Themen durch Studierende eingebracht wurden), könnte das Aufwerfen neuer Aspekte in den Nachprompts die Proband*innen bei einer breiteren Reflexion unterstützen. Verschiedene andere Studien deuten darauf hin, dass der unterstützende Einsatz von Prompts (z. B. Fragen, die die Reflexion anstoßen sollen oder zu vervollständigende Satzanfänge (Blunk, 2020, S. 59)) die Reflexion bereicherten (Blunk, 2020, S. 81ff.; Hamel und Viau-Guay, 2019). Die Prompts ermöglichen außerdem eine gewisse (Vor-)Strukturierung der Unterrichtsreflexion im Testinstrument, die Kost (2019, S. 106f.) gerade bei verbalen Reflexionsanlässen für bedeutsam erachtet. Um außerdem vorrangig die Unterrichtsplanung durch die Lehrkraft und ihr Verhalten im Unterricht zu reflektieren und nicht z. B. die Schüler*innen in ihrem Lernprozess zu diagnostizieren, wird auch in den Vor- und Nachprompts ein Fokus auf die Lehrkraft und ihr Handeln gelegt (z. B. „[...] Was sagst du dazu, wie *ich* versucht habe, die Schüler*innen auf den richtigen gedanklichen Weg zu bringen?“ oder „Wie findest du den theoretischen Einschub, den *ich* da gemacht habe?“). Das Verhalten oder Äußerungen der Schüler*innen können im Rahmen der Reflexion zurate gezogen werden, allerdings eher für eine fundierte Einschätzung des Lehrerhandelns. Das ansonsten offene Antwortformat ermöglicht gleichzeitig die Erfassung der individuellen Überlegungen der Proband*innen ohne Einschränkungen (Kost, 2019, S. 106f.; Leonhard und Rihm, 2011).

Für den geschlossenen Reflexionstest wurden die Vorprompts übernommen. Sie dienen weiterhin der Einleitung der Videovignetten und Angabe möglicher Aus-

gangspunkte für die Beobachtung und anschließende Reflexion der Vignetten. An die Stelle des Nachprompts treten die Aufgabenstellungen zum Noticing, Bewerten und Vorschlagen von Alternativen. Durch sie erfolgt eine stärkere Vorstrukturierung der Reflexion und möglicherweise eine inhaltliche Bereicherung der Reflexionen durch die Teilnehmenden, da weiterhin nicht davon ausgegangen werden kann, dass sie die relevanten Aspekte der Reflexion selbstständig identifizieren.

Die Möglichkeit, während der mündlichen Reflexion im Reflexionsperformanztest bzw. der Bearbeitung der MC-Aufgaben im Reflexionstest eine bestimmte Auswahl an Rückfragen zu stellen (um z. B. die eingesetzten Materialien oder den Verlaufsplan der Stunde erneut einzusehen), bietet ein gewisses Maß an Interaktivität in der Durchführung des Testinstruments, die konstituierend für Performanztests ist (Bartels, 2018, S. 30). Inwiefern diese Form der Interaktion ausreichend ist, um die positiven Effekte bzw. Einstellungen gegenüber echten Interaktionen z. B. im Rahmen gemeinsamer Reflexionen (Hamel & Viau-Guay, 2019) zu erreichen und inwieweit sich dies für beide Testinstrumente unterscheidet, ist unklar. Grundsätzlich wird die (insbesondere gemeinsame) mündliche Unterrichtsreflexion, die im Reflexionsperformanztest zum Einsatz kommt, schriftlichen Reflexionsaufträgen vorgezogen (Hamel & Viau-Guay, 2019). Das Verzichten auf verschriftlichte Reflexionen im offenen Antwortformat des Reflexionsperformanztests hat außerdem den Vorteil, dass Teilnehmende, die mit der Formulierung selbstständiger schriftlicher (Reflexions-)Produkte Schwierigkeiten haben, in der Bewertung nicht benachteiligt werden (vgl. Wyss, 2013, S. 94f.). Auch im Reflexionstest müssen Teilnehmende ihre Reflexionen nicht verschriftlichen, ein Einfluss durch unterschiedlich ausgeprägte Lesefähigkeit kann allerdings nicht ausgeschlossen werden.

11 Assessment-Feedback

Die Auswertung der Bearbeitungen des Reflexionstests sowie die Erstellung des Assessment-Feedback erfolgt halbautomatisiert mit Hilfe eines selbst geschriebenen Programms in *R* (R Core Team, 2022), das auf Grundlage der von *LimeSurvey* zur Verfügung gestellten Ergebnistabelle PDF-Dokumente mit dem individuellen Assessment-Feedback erstellt.[130] So kann das Feedback mit nur wenig Verzögerung zur Verfügung gestellt werden. Um die Anonymität der Teilnehmenden zu gewährleisten, wird das Assessment-Feedback anschließend - sofern die Teilnehmenden dem zugestimmt haben - online auf Servern der Universität Bremen zur Verfügung gestellt und kann von ihnen mit Hilfe eines persönlichen Codes heruntergeladen werden. Vorgestellt wird hier die überarbeitete Version des Assessment-Feedbacks nach Interviews mit Studierenden zu dessen Evaluation (s. Kapitel 18). Nachfolgend wird das Assessment-Feedback für Teilnehmende am Reflexionstest insgesamt vorgestellt (s. Kapitel 11.1), das sich aus einer individuellen Rückmeldung (dem bereitgestellten PDF-Dokument) sowie dem Selbstlernmaterial (s. Kapitel 11.2) zusammensetzt. Abschließend wird ein Einblick in das Assessment-Feedback für Dozierende gegeben, das bereitgestellt wird, wenn die Bearbeitung des Reflexionstests in eine universitäre Veranstaltung integriert wird (s. Kapitel 11.3).

11.1 Aufbau des Assessment-Feedbacks

Das Assessment-Feedback[131] beginnt mit einer kurzen, überblicksartigen Einordnung für die Teilnehmenden. Darin wird auf die Bearbeitung des Reflexionstests verwiesen und der individuelle Testwert zur Reflexionsfähigkeit mit einer Verortung in die Vergleichsgruppe (auf Grundlage der Einordnung in Quartile) verbunden, um einen ersten Bezug zu den individuellen Ergebnissen herzustellen. Außerdem werden einleitend das dem Reflexionstest zugrundeliegende Verständnis von Unterrichtsreflexion beschrieben sowie insbesondere der Fokus auf Bewertungen und Alternativen im Testinstrument transparent gemacht.

[130] Das R-Skript zur Erstellung der Rückmeldung befindet sich im Anhang in Kapitel B ab Seite 439. Auf Grundlage der aus *LimeSurvey* exportierten Ergebnisse werden die Rückmeldungen automatisiert erstellt.

[131] Mit Assessment-Feedback wird die Kombination aus Rückmeldebogen und Selbstlernmaterial beschrieben, bei Bedarf wird im Verlauf der Arbeit zwischen den Bestandteilen Rückmeldung und Selbstlernmaterial differenziert.

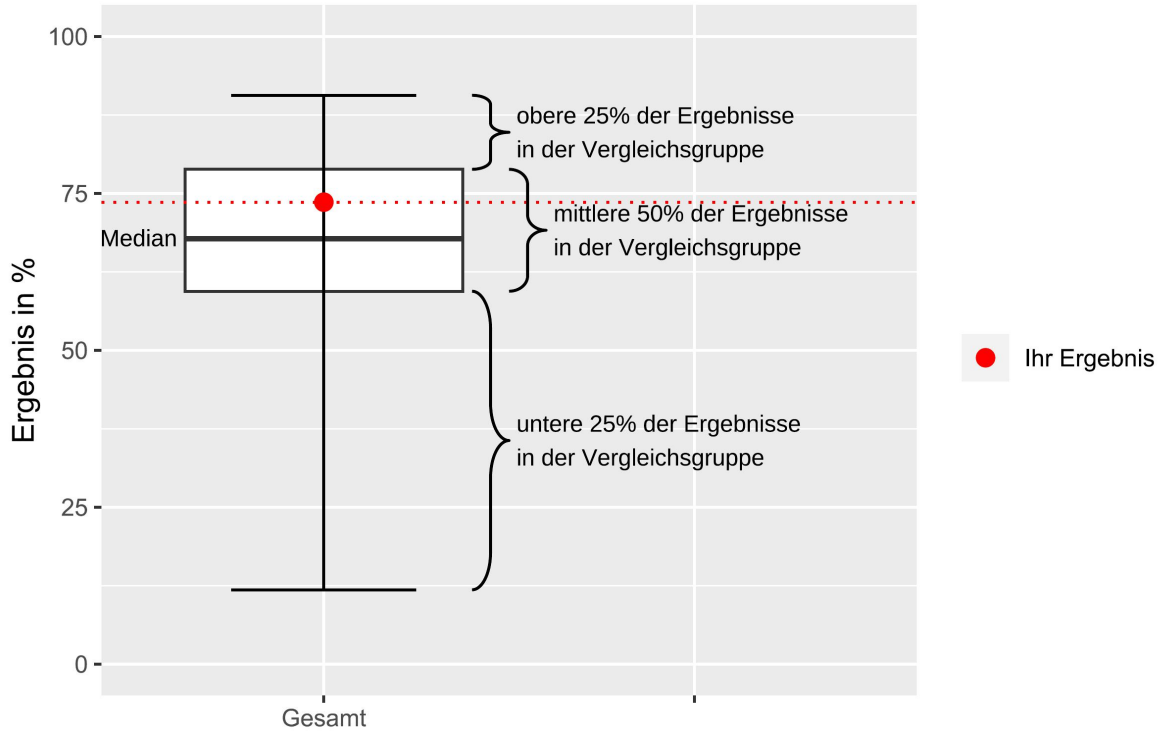


Abbildung 11.1: Exemplarische Darstellung des Boxplots zur Rückmeldung des Gesamtergebnisses

Anschließend erfolgt die ausführliche Rückmeldung der individuellen Ergebnisse: Das Gesamtergebnis sowie die Teilergebnisse zu Bewertungen, Alternativen und Schülervorstellungen werden mit Hilfe von Boxplots dargestellt und in die Vergleichsgruppe eingeordnet. Zur besseren Interpretierbarkeit wird die Zusammensetzung der Vergleichsgruppe überblicksartig beschrieben (mit „X Studierenden im durchschnittlich Y. Fachsemester und Z Referendar*innen“). Um den Teilnehmenden bei der Interpretation der Boxplots Hilfestellung anzubieten, ist der erste Boxplot zum Gesamtergebnis beschriftet (s. Abbildung 11.1). Zusätzlich befinden sich am Ende der Rückmeldung schriftliche Erläuterungen zur Interpretation des Boxplots. Die Rückmeldung der Teilergebnisse wird um kurze Beschreibungen der Teilbereiche (Bewertungen, Alternativen, Schülervorstellungen) ergänzt. Außerdem wird die Zuordnung der Testaufgaben zu den einzelnen Teilbereichen transparent gemacht. Eine Beispielrückmeldung befindet sich in Kapitel J.5.7 ab Seite 511.

Abschließend wird darauf eingegangen, wie die Reflexionsfähigkeit geschult werden kann. Hier wird auf die Bedeutung von (u. a. fachdidaktischem) Wissen als Voraussetzung für eine fundierte Reflexion eingegangen sowie für eine weiterführende, selbstständige Beschäftigung mit der Unterrichtsreflexion auf ein Selbstlernmaterial (s. Kapitel 11.2) verwiesen. Abhängig von den individuellen Ergebnissen

wird außerdem eine Fokussierung auf einen der Teilbereiche aus Bewertungen und Alternativen für die Arbeit mit dem Selbstlernmaterial empfohlen.

11.2 Aufbau des ergänzenden Selbstlernmaterials

Das Selbstlernmaterial dient der weiterführenden Beschäftigung mit Unterrichtsreflexion auf Grundlage des erhaltenen Assessment-Feedbacks. Es ist ebenso wie der Reflexionstest in *LimeSurvey* (LimeSurvey GmbH, n. d.) umgesetzt.[132] Kernstück dieses Materials ist die begleitete Reflexion von drei weiteren Videovignetten aus der Doppelstunde, die die Teilnehmenden bereits aus dem Reflexionstest kennen. Als Rahmung enthält das Selbstlernmaterial analog zum Assessment-Feedback einleitend Hinweise zum Begriff der Unterrichtsreflexion, allerdings mit einem Fokus auf das Üben von Unterrichtsreflexion. Dazu werden allgemeine Leitfragen für die Fremdreflexion bereitgestellt (s. Tabelle 11.1). Auch die Rahmenbedingungen zur Unterrichtsstunde von Robert (Unterrichtsziele etc.) werden erneut zur Verfügung gestellt.

Inhaltlicher Fokus der drei eingebundenen Videovignetten ist die experimentelle Bestätigung und theoretische Herleitung der Impulserhaltung in folgenden Schritten:

1. Vorstellung eines Demonstrationsexperiments zur Impulserhaltung durch die Lehrkraft mit mangelnder kognitiver Aktivierung der Schüler*innen
2. Besprechung des Experiments (d. h. Sammlung von Beobachtungen und Formulierung einer Erklärung)
3. Theoretische Herleitung der Impulserhaltung aus dem 3. Newtonschen Axiom im Rahmen eines Lehrervortrags

Die Vignetten werden zunächst - analog zur Einbettung im Testinstrument - jeweils mithilfe des Stundenrasters in den Verlauf der Doppelstunde eingeordnet. Der anschließende Reflexionsauftrag ist zunächst recht offen und zielt lediglich darauf ab, Noticing (Wozu soll Feedback gegeben werden?) und Unterrichtsreflexion (welches Feedback soll gegeben werden?) anzusprechen: „Schauen Sie das folgende Video und überlegen Sie, wozu Sie Robert welches Feedback geben würden.“ Als Hilfestellung kann jederzeit auf Leitfragen zur Unterrichtsreflexion (s. Tabelle 11.1) zurückgegriffen werden. Außerdem ist eine Beispiellösung eingebunden, deren Struktur (Noticing - Beschreibung - Bewertungen und Alternativen) sichtbar ist und als Orientierung dienen kann (s. Abbildung 11.2). Die Beispiellösung kann nach Bedarf (schrittweise) durch Klicken auf die jeweiligen Blöcke zu Noticing, Beschreibung oder Bewertungen und Alternativen geöffnet und mit den eigenen Überlegungen abgeglichen werden. Abschließend werden zu den fachdidaktischen

[132] Das vollständige Selbstlernmaterial ist online einzusehen (s. Kapitel C).

Reflexionselement	Leitfragen
Beschreibungen	• Was haben die Schüler*innen gemacht? • Was ist passiert? *Wenn bekannt/erkenntlich:* • Warum wurde die Methode/das Experiment/... ausgewählt? Was war das Ziel der Stunde? • Welche besonderen Aspekte wurden bei der Planung bedacht?
Bewertungen	• Was war erwartungsgemäß? • Was war besser als erwartet? • Was war schlechter als erwartet? ... Und warum?
Alternativen	• Was hätte die Lehrkraft anders machen können und warum wäre dieses Handeln erfolgreicher gewesen? • Oder warum würde ich der Lehrkraft empfehlen, es genauso wieder zu machen?
Konsequenzen	• Was ziehe ich aus der heutigen Stunde für die Entwicklung der (beobachteten) Lehrperson und warum? • Was ziehe ich aus der beobachteten Stunde für meine persönliche Entwicklung als Lehrperson und warum?

Tabelle 11.1: Leifragen zur Selbstreflexion von Unterricht nach Nowak, Liepertz und Borowski (2018, S. 223)

Inhalten in den Vignetten Literaturhinweise bereitgestellt, die die Hintergründe der Beispiellösung nachvollziehbar machen und die vertiefte Beschäftigung mit den Inhalten ermöglichen sollen.

11.3 Assessment-Feedback für Dozierende

Die Lernumgebung wird komplettiert durch ein Assessment-Feedback, das Dozierende erhalten, sofern sie den Reflexionstest im Rahmen ihrer Veranstaltung einsetzen. Dieses Assessment-Feedback ist nicht Gegenstand der nachfolgend vorgestellten Validierungsstudien, kann allerdings einen Mehrwert der Lernumgebung bieten, wenn es bspw. als Grundlage der Thematisierung reflexionsbezogener Veranstaltungsinhalte dient und wird daher der Vollständigkeit halber hier vorgestellt.

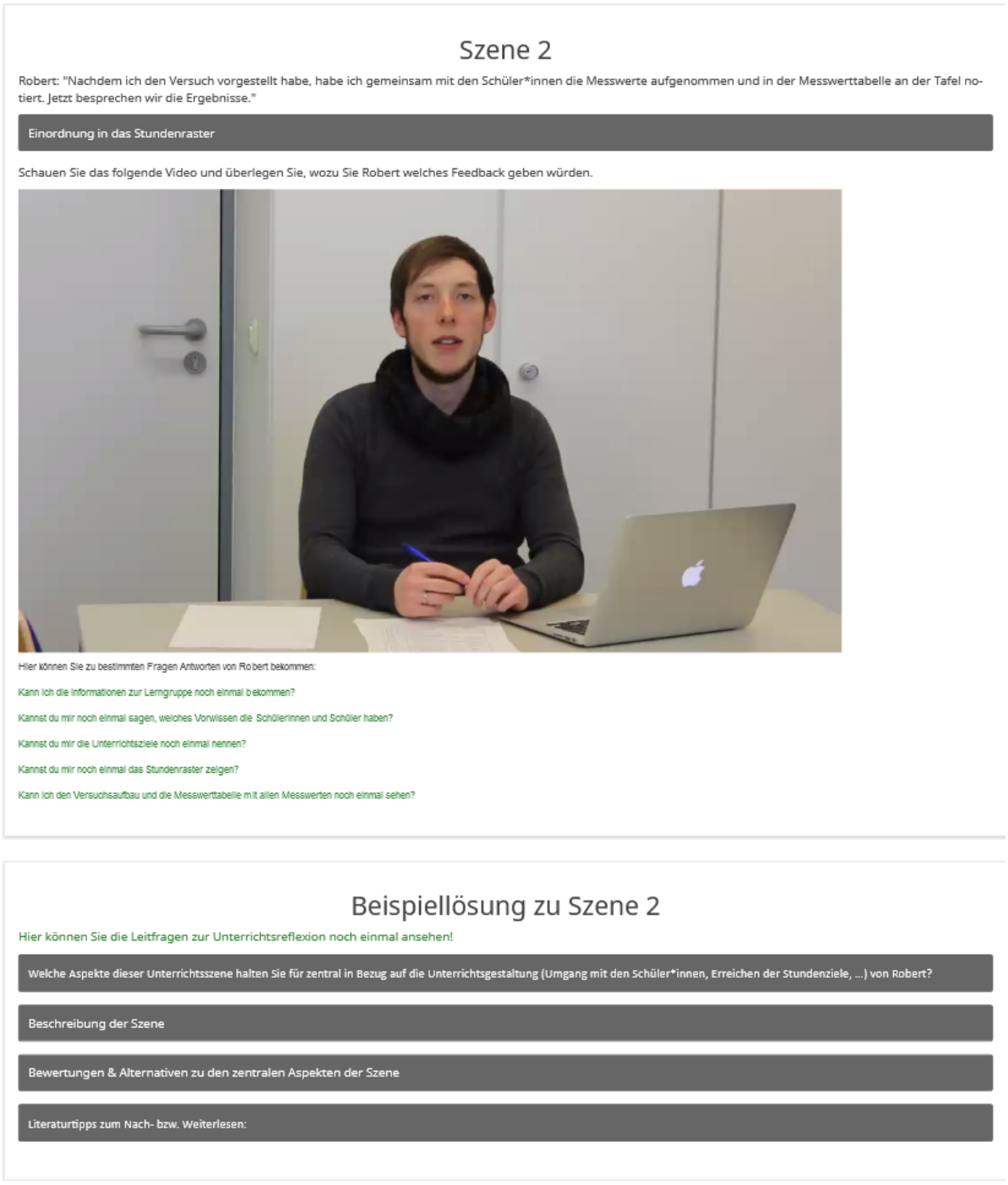


Abbildung 11.2: Screenshot aus dem Selbstlernmaterial: Überblick zur Reflexion der zweiten eingebundenen Videovignette

Zur Zuordnung der Teilnehmenden zu einem Kurs wird ein Kohortencode verwendet, den die Teilnehmenden während der Bearbeitung des Testinstruments angeben. So erhalten die Dozierenden eine Rückmeldung über das Abschneiden ihres Kurses insgesamt, ohne dass die Anonymität der Teilnehmenden gefährdet wird.

Das Assessment-Feedback für Dozierende ist grundsätzlich analog zu dem für Teilnehmende aufgebaut. Zentraler Unterschied ist die Ergebnisdarstellung in Boxplots: Anstelle einer Verortung von individuellen Ergebnissen in den Ergebnissen der Vergleichsgruppe werden die Ergebnisse des Kurses als eigener Boxplot neben dem der Vergleichsgruppe dargestellt. So können die Ergebnisverteilungen beider Gruppen verglichen werden. Zusätzlich kann die Effektstärke (d_{Cohen}) zum Vergleich der Mittelwerte abgelesen werden.[133]

[133]Solch eine Einordnung erscheint für Studierende (insbesondere am Anfang des Studiums) nicht angemessen, da - wie sich in Interviews mit Studierenden zur Evaluation des Assessment-Feedback bestätigt hat (vgl. Kapitel 18.1) - nicht davon ausgegangen werden kann, dass sie mit der Effektstärke vertraut sind.

Teil III

Validierungsstudien zur entwickelten Lernumgebung

12 Das Interpretations-Nutzungs-Argument zum Reflexionstest (Forschungsfrage 1)

Ziel dieser Arbeit ist die Evaluation der Kombination aus dem Testinstrument zur Reflexionsfähigkeit und dem sich anschließenden Assessment-Feedback mit Selbstlernmaterial, wobei die valide Interpretation der Testwerte im Fokus steht. Die Materialien bilden zusammen eine Lernumgebung, in der Teilnehmenden auf Grundlage der Testdurchführung ein Assessment-Feedback bereitstellt wird und so eine fundierte Selbsteinschätzung der Reflexionsfähigkeit ermöglicht werden soll. Somit ist nicht nur sicherzustellen, dass die Testwerte valide als Maß für die Reflexionsfähigkeit interpretiert werden können, sondern auch von den Teilnehmenden als solche interpretiert werden und angemessene Schlussfolgerungen für die eigene Professionalisierung ermöglichen. Als übergeordnete Forschungsfrage für die Validierung der Lernumgebung wird daher nachfolgende Forschungsfrage 1 untersucht:

Forschungsfrage 1: Inwieweit kann die folgende Testwertinterpretation gerechtfertigt werden?

Die Testwerte des Reflexionstests können (a) als Maß für die Reflexionsfähigkeit und (b) als Ausgangspunkt für einen weiteren Professionalisierungsprozess aufgefasst werden.

Als Bezugsrahmen für die Untersuchung der Gültigkeit der Testwertinterpretation wird im Sinne des Argument-Based-Approach (s. Kapitel 6.1.3) ein Interpretations-Nutzungs-Argument (INA) formuliert. Dieses ist analog zu den Arbeiten von Dickmann (2016) und Bartels (2018) an den Übersetzungsschritten orientiert, die ausgehend von der Zieldomäne Reflexionsfähigkeit durchgeführt werden, um die Reflexionsfähigkeit (angehender) Physiklehrkräfte zu erfassen und rückzumelden (s. Abbildung 12.1). An das Produkt jedes Übersetzungsschritts werden dann übergeordnete Qualitätsanforderungen angelegt. Die Bereitstellung des Assessment-Feedbacks ist zentraler Bestandteil des Einsatzszenarios für den entwickelten

Reflexionstest und ist insbesondere hinsichtlich der Bewertung der konsequentiellen Validität relevant. Daher wird zusätzlich zu den in Kapitel 6.1.3 beschriebenen Übersetzungsschritten (vom Konstrukt bis hin zur Zuweisung von Testwerten) die Übersetzung bzw. Einbindung der Testwerte in das Assessment-Feedback berücksichtigt.[134]

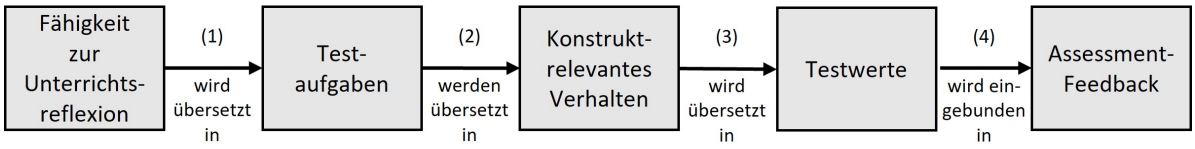


Abbildung 12.1: Darstellung der erforderlichen Übersetzungsschritte, zur Zuweisung und Rückmeldung von Testwerten für die Zieldomäne der Fähigkeit zur Unterrichtsreflexion

Die Anforderungen an das INA lauten demnach (angelehnt an Bartels, 2018, S. 86; Dickmann, 2016, S. 67):

Anforderung an Teil 1 des INA: Die Aufgaben im Reflexionstest bilden relevante und repräsentative Inhalte der Fähigkeit zur Unterrichtsreflexion präzise ab.

Anforderung an Teil 2 des INA: Die Proband*innen zeigen bei der Bearbeitung des Reflexionstests konstruktrelevantes Verhalten.

Anforderung an Teil 3 des INA: Das (Ankreuz-)Verhalten der Proband*innen während der Testbearbeitung wird adäquat in Testwerte übersetzt.

Anforderung an Teil 4 des INA: Das Assessment-Feedback stellt eine geeignete Form der Rückmeldung der Testwerte dar.

Zur Beurteilung des INA werden für jeden Übersetzungsschritt Indikatoren herangezogen, die konkret auf Grundlage theoretischer Überlegungen bzw. empirischer Evidenz geprüft werden. Eine Übersicht über die so zustande kommende Validitätsargumentation liefern die nachfolgenden Abschnitte (Kapitel 12.1 bis 12.4).

[134] Hubley und Zumbo (2011) deuten ähnlich hierzu einen Übersetzungsschritt der Testwerte und ihrer Interpretation in (intendierte) persönliche Konsequenzen bzw. (unintendierte) Nebeneffekte an, diese ergeben sich für den Reflexionstest allerdings auf Grundlage des Assessment-Feedbacks und dessen Interpretation durch die Teilnehmenden.

12.1 Indikatoren zum Teil 1 des INA: Übersetzung Reflexionsfähigkeit in Testinstrument

Die erste Anforderung bezieht sich auf die Übersetzung des Konstrukts der Fähigkeit zur Unterrichtsreflexion in die Gestaltung des Testinstruments:

Die Aufgaben im Reflexionstest bilden relevante und repräsentative Inhalte der Fähigkeit zur Unterrichtsreflexion präzise ab.

Indikatoren für eine positive Evaluation dieser Anforderung beziehen sich auf die inhaltliche Passung der Aufgaben zum Konstrukt der Reflexionsfähigkeit und damit auf die inhaltliche Validität (vgl. Kapitel 6.1.3). Da die Fähigkeit zur Unterrichtsreflexion schon allein auf Grund der Breite möglicher Unterrichtszenarien nicht vollumfänglich in einem Testinstrument abgebildet werden kann, bedarf es einer fundierten Auswahl von für den Physikunterricht inhaltlich relevanten und möglichst repräsentativen Inhalten für das Testinstrument. Dazu wird der Indikator 1.1 formuliert:

Indikator 1.1 Die Auswahl der Szenen und Aufgaben stellt ein umfassendes Abbild der Fähigkeit zur Unterrichtsreflexion in Bezug auf einen inhaltlich bedeutsamen Fachinhalt von Physikunterricht dar.

Die Antwortoptionen in den MC-Aufgaben basieren auf realen Reflexionen von Studierenden zu den Videovignetten (zur Entwicklung der Antwortoptionen s. Kapitel 10.2.2). Die Einschätzung der Qualität einzelner Aussagen bzw. deren Bewertung als (in-)korrekt dient als Grundlage für die Berechnung der Testwerte. Es muss also sichergestellt sein, dass die korrekten Antwortoptionen jeweils die optimalen Rückmeldungen in Bezug auf die beobachtete Videovignette darstellen. Dazu wird Indikator 1.2 formuliert:

Indikator 1.2 Die als korrekt definierten Antworten repräsentieren die zentralen Erkenntnisse der Unterrichtsreflexion zu den dargestellten Szenen.

12.2 Indikatoren zum Teil 2 des INA: Übersetzung Testinstrument in Verhalten

Die Evaluation des zweiten Übersetzungsschritts zielt auf die Passung zwischen den kognitiven Aktivitäten der Teilnehmenden während der Testbearbeitung und dem untersuchten Konstrukt (kognitive Validität) ab. Die an das Verhalten während der Testbearbeitung gestellte Anforderung lautet:

Die Proband*innen zeigen bei der Bearbeitung des Reflexionstests konstruktrelevantes Verhalten.

Um die Anforderung als erfüllt beurteilen zu können, muss sichergestellt sein, dass während der Bearbeitung des Reflexionstests Reflexionsprozesse bei den Teilnehmenden ausgelöst werden und inhaltliche Überlegungen die Bearbeitung der Aufgaben bestimmen, die Testbearbeitung also nicht z. B. durch missverstandene Formulierungen beeinflusst wird. Dazu wird Indikator 2.1 überprüft:

Indikator 2.1 Die Überlegungen der Teilnehmenden (a) beziehen sich auf die Qualität des Unterrichts bzw. die Professionalisierung und (b) passen zur Aufgabenstellung.

Ergänzend ist auch die konzentrierte, ernsthafte Bearbeitung des Testinstruments bedeutsam für die valide Interpretation der Testwerte als Maß für die Reflexionsfähigkeit. Als Indikator dafür kann die Akzeptanz auf Einstellungsebene (s. Kapitel 6.1.5) gegenüber dem Testinstrument durch die Teilnehmenden untersucht werden:

Indikator 2.2 Die Teilnehmenden akzeptieren das Testverfahren.

Eine fehlende Akzeptanz des Testverfahrens und geringe Testmotivation sind als mögliche Gefährdungen der Validität einzuordnen, da sie als Teil der motivational-affektiven Komponente der professionellen Handlungskompetenz die Bereitschaft der Teilnehmenden beeinflussen können, die Reflexionsfähigkeit zum Ausdruck zu bringen.

12.3 Indikatoren zum Teil 3 des INA: Übersetzung Verhalten in Testwerte

Der dritte Übersetzungsschritt fokussiert die angemessene Berechnung der Testwerte:

Das (Ankreuz-)Verhalten der Proband*innen während der Testbearbeitung wird adäquat in Testwerte übersetzt.

Zur Evaluation dieser Anforderung werden verschiedene Validitätsaspekte berücksichtigt, zunächst werden die strukturelle Validität sowie Voraussetzungen für diese untersucht. Als erste Voraussetzung für nähere Analysen wird geprüft, ob der entwickelte Reflexionstest gemäß den Möglichkeiten der zur Bestimmung

der Testwerte verwendeten Skala zwischen unterschiedlichen Ausprägungen der Reflexionsfähigkeit differenziert (Bartels, 2018, S. 88). Dies beschreibt Indikator 3.1:

Indikator 3.1 Der Reflexionstest ist dazu geeignet, verschiedene Ausprägungen der Reflexionsfähigkeit auch innerhalb der verschiedenen Ausbildungsphasen (Bachelor- und Masterstudium) differenziert darzustellen.

Da der Reflexionstest für den Einsatz in verschiedenen Ausbildungsphasen konzipiert ist, ist im Sinne der Generalisierbarkeit sicherzustellen, dass die Bestimmung der Testwerte nicht nur global, sondern auch für die verschiedenen Ausbildungsphasen angemessen ist, sodass zusätzlich zur Untersuchung der Gesamtstichprobe Analysen für die Gruppe der Bachelor- bzw. Masterstudierenden durchgeführt werden.

Zusätzlich ist die möglichst präzise Erfassung der Reflexionsfähigkeit anzustreben. Daher wird mit dem Indikator 3.2 auch die Reliabilität der zur Bestimmung der Testwerte verwendeten Skalen als klassisches Gütekriterium berücksichtigt (vgl. Kapitel 6.1.2). Auch hier wird explizit sichergestellt, dass die Skalen für die verschiedenen Ausbildungsphasen jeweils ausreichende Reliabilitäten aufweisen.

Indikator 3.2 Die zur Bestimmung der Testwerte verwendeten Skalen sind reliabel auch über die verschiedenen Ausbildungsphasen (Bachelor- und Masterstudium) hinweg.

Damit sind bislang allerdings lediglich Vorbetrachtungen der strukturellen Validität abgedeckt. Eine hohe Reliabilität (bzw. interne Konsistenz) wäre bereits ein starker Indikator für die Annahme einer Eindimensionalität des Konstrukts. Um auch die Möglichkeit zu untersuchen, dass den Testwerten unerwartete Strukturen zugrundeliegen, wird ergänzend eine explorative Faktorenanalyse herangezogen:

Indikator 3.3 Die Struktur der Testwerte ist theoretisch erklärbar.

Grundlage für die (theoretische) Interpretation der Faktorenanalyse ist dabei das Verständnis der Fähigkeit zur Unterrichtsreflexion als eine Facette der professionellen Handlungskompetenz, die im Wechselspiel mit verschiedenen anderen Kompetenzen die professionelle Handlungskompetenz einer Lehrkraft bildet. Als solche umfasst sie wiederum „ein Bündel von [...] geistigen Fähigkeiten“ (Frey, 2006, S. 31), wie das Bewerten von vielfältigen Unterrichtssituationen oder das Entwickeln adäquater Handlungsalternativen.

Über diese ersten drei Indikatoren hinweg werden die in die verwendeten Skalen eingehenden Aufgaben im Sinne einer Itemanalyse untersucht. Dazu werden

im Indikator 3.1 die Aufgabenschwierigkeiten (bzw. Lösungshäufigkeiten) näher betrachtet, im Rahmen der Untersuchung der Reliabilität (Indikator 3.2) die Trennschärfen berücksichtigt und abschließend wird eine Dimensionalitätsprüfung (Indikator 3.3) vorgenommen, um zu prüfen, ob die verwendete Skala in verschiedene Faktoren zerfällt (Döring & Bortz, 2016, S. 476).

Weiterhin können externe Faktoren zur Evaluation der Anforderung herangezogen werden (externe Validität). Als Kriterien dienen hier z. B. Testwerte zu Instrumenten, die Konstrukte mit inhaltlichem Bezug zur Reflexionsfähigkeit messen, sowie Lerngelegenheiten für Reflexion. Voraussetzung für den fundierten Einbezug externer Kriterien in die Validitätsargumentation ist das Vorliegen theoretisch oder empirisch begründeter Erwartungen bezüglich der jeweiligen Zusammenhänge (Messick, 1995).

Für die Erfassung des professionellen Wissens ist festgestellt, dass die Testwerte „nicht per se als handlungsvalide angenommen werden können“ (Vogelsang & Reinhold, 2013, S. 123), sich also der theoretisch vermutete Einfluss des Professionswissens nicht unbedingt in der Handlungsqualität niederschlägt. Der entwickelte Reflexionstest ist zwar kein Wissenstest, es ist aber trotzdem davon auszugehen, dass sich die Anforderungen zur realen Handlungssituation unterscheiden. Naheliegend ist daher ein Abgleich der Testwerte des Reflexionstests mit Bewertungen der Fähigkeit zur Unterrichtsreflexion in authentischeren Handlungssituationen. Dazu wird der Reflexionsperformanztest von Kempin, Kulgemeyer und Schecker (2018) herangezogen, der für die Erfassung der Reflexionsfähigkeit von Physik-Lehramtsstudierenden im Kontext des Praxissemesters erprobt ist. Da hierbei das gleiche Konstrukt mit zwei inhaltlich vergleichbaren Testinstrumenten erhoben wird, die sich vorrangig in der Erhebungsmethode (offenes verbales Antwortformat im Reflexionsperformanztest sowie Multiple-Choice-Form im Reflexionstest) unterscheiden, ist grundsätzlich eine hohe Korrelation zwischen den Testwerten beider Instrumente zu erwarten. Hierzu wird Indikator 3.4 überprüft:

Indikator 3.4 Die Testwerte des Reflexionstests weisen Zusammenhänge zur Qualität von Reflexionen in authentischeren Settings (repräsentiert durch die Testwerte des Reflexionsperformanztests) auf.

Verbunden mit Indikator 3.1, der die Notwendigkeit betont, mit Hilfe des Reflexionstests unterschiedliche Ausprägungen der Reflexionsfähigkeit erfassen zu können, ist die Frage nach den Ursachen für die unterschiedlichen Qualitäten. Unterschiedliche Ausprägungen der Reflexionsfähigkeit sollten durch unterschiedliche (konstruktrelevante) Voraussetzungen erklärbar sein. Dazu gehören insbesondere verschiedene Lerngelegenheiten zur Reflexion (z. B. begleitete Praktika), deren Umfang entlang des Ausbildungsfortschritts der Teilnehmenden zunimmt. Gleichzeitig sollten nicht konstruktrelevante Faktoren (z. B. das Geschlecht) keinen Einfluss auf die Ausprägung der Reflexionsfähigkeit nehmen; die Untersuchung

dieser Variablen wird im Rahmen der von Kane (2013, S. 8) empfohlenen gezielten Untersuchung validitätsgefährdender Zusammenhänge in den folgenden Indikator 3.5 inkludiert:

Indikator 3.5 Unterschiedliche Lerngelegenheiten der Teilnehmenden führen zu nachvollziehbaren Unterschieden in den Testwerten.

Da die Fähigkeit zur Unterrichtsreflexion als Facette der professionellen Handlungskompetenz verstanden wird, können auch Zusammenhänge zu anderen Bestandteilen der professionellen Handlungskompetenz im Sinne einer konvergenten Validierung als Hinweise für die valide Interpretation der Testwerte gewertet werden. In dieser Untersuchung wird dafür das fachdidaktische Wissen gemäß Indikator 3.6 betrachtet:

Indikator 3.6 Die Testwerte des Reflexionstests weisen einen Zusammenhang zum fachdidaktischen Wissen auf.

12.4 Indikatoren zum Teil 4 des INA: Übersetzung Testwerte in Assessment-Feedback

An den letzten hier betrachteten Übersetzungsschritt, die Entwicklung des Assessment-Feedbacks (inklusive Selbstlernmaterial) zur Rückmeldung der Testwerte, wird folgende Anforderung gestellt:

Das Assessment-Feedback stellt eine geeignete Form der Rückmeldung der Testwerte dar.

Diese Anforderung bezieht sich vorrangig auf den zweiten Teil der Testwertinterpretation: die Interpretation der Testwerte als Ausgangspunkt für den weiteren Professionalisierungsprozess. Hier ist insbesondere der selten in Validierungsstudien berücksichtigte Aspekt der konsequentiellen Validität angesprochen, da den Teilnehmenden ihre Testwerte im Rahmen des Assessment-Feedbacks transparent gemacht werden.[135] Es ist sicherzustellen, dass das Assessment-Feedback für die Teilnehmenden inhaltlich nachvollziehbar ist und angemessen interpretiert wird. Besonders informativ für die Evaluation des Assessment-Feedbacks erscheint darüber hinaus die tatsächliche Interpretation des Assessment-Feedbacks

[135] Die Vernachlässigung des konsequentiellen Aspekts der Validierung wird z. B. damit begründet, dass die Testwerte nicht explizit zur Bewertung der Teilnehmenden genutzt werden (Bartels, 2018, S. 91). Grundsätzlich besteht dann allerdings weiterhin der Bedarf zur konsequentiellen Validierung für (folgende) Testeinsätze.

durch die Teilnehmenden. Überprüft werden kann dabei u. a. die Verständlichkeit des Assessment-Feedbacks, die Bedeutung, die Teilnehmende dem Assessment-Feedback beimessen sowie die inhaltliche Sinnhaftigkeit der aus dem Feedback gezogenen Schlussfolgerungen. Dies wird in Indikator 4.1 zusammengefasst:

Indikator 4.1 Die Interpretationen der Ergebnisse durch die Teilnehmenden passen inhaltlich zum erhaltenen Feedback.

Zusätzlich zur Diskussion der inhaltlichen Angemessenheit des Assessment-Feedback ist erneut die Akzeptanz der Lernumgebung durch die Teilnehmenden zu berücksichtigen. Es ist davon auszugehen, dass die Bereitschaft, sich mit dem Feedback auseinanderzusetzen und darauf aufbauend Maßnahmen umzusetzen, abhängig davon ist, welchen Mehrwert die Teilnehmenden im Assessment-Feedback bzw. in der entwickelten Lernumgebung insgesamt sehen. Dazu wird Indikator 4.2 formuliert:

Indikator 4.2 Die Teilnehmenden sprechen dem Assessment-Feedback (inklusive Selbstlernmaterial) einen Mehrwert zu.

Weiterhin sollten bereits bei der Konstruktion und Weiterentwicklung des Assessment-Feedbacks theoretische und empirische Erkenntnisse zur Gestaltung von effektivem Feedback berücksichtigt werden. Die entsprechende Argumentation erfolgt zu Indikator 4.3:

Indikator 4.3 Das Assessment-Feedback stellt inhaltlich relevante Testergebnisse unter Berücksichtigung von Qualitätskriterien für effektives Feedback dar.

12.5 Bemerkungen zur Auswahl der Indikatoren für die Validitätsargumentation

Abbildung 12.2 enthält eine vollständige Übersicht über das INA, die daran gestellten Anforderungen sowie die Indikatoren, die zur Einschätzung der Anforderungen geprüft werden.

Die Bewertung der Testwertinterpretation (bzw. Beantwortung der Forschungsfrage) geschieht auf Grundlage der Prüfung aller Indikatoren, wobei die einzelnen Übersetzungsschritte aufeinander aufbauen (Bartels, 2018, S. 90f.): Eine Einschränkung bezüglich der Qualität eines Übersetzungsschritts muss somit auch in der Bewertung der nachfolgenden Schritte berücksichtigt werden.

Mit den oben beschriebenen Indikatoren werden alle sechs von Messick (1995) formulierten Validitätsaspekte in unterschiedlichem Umfang in die Validitätsargumentation einbezogen. Dabei können die Indikatoren nicht immer eindeutig einem

Validitätsaspekt zugeordnet werden. So dienen z. B. die Analysen zur Differenzierungsfähigkeit (Indikator 3.1) und Reliabilität (Indikator 3.2) des Reflexionstests vorrangig der Prüfung von Voraussetzungen für die strukturelle Validität sowie für weiterführende statistische Analysen. Die Authentizität der eingesetzten Videovignetten trägt zur inhaltlichen Validität bei (Indikator 1.1), kann aber auch zur Akzeptanz des Testverfahrens durch die Teilnehmenden beitragen (Indikator 2.2). Zusätzlich generieren die Analysen Evidenz bezüglich der Generalisierbarkeit der folgenden Erkenntnisse, indem jeweils ausdifferenzierte Analysen für die Studienphasen Bachelor- und Masterstudium durchgeführt werden.

Die Untersuchung der konsequentiellen Validität ist beschränkt auf die Interpretation der eigenen Testwerte und Nutzungs*absichten* des Selbstlernmaterials. Basierend auf einem (auf Grundlage des Feedbacks) erkannten Bedarf zur Weiterentwicklung der Reflexionsfähigkeit kann nicht davon ausgegangen werden, dass die Teilnehmenden Maßnahmen zur Weiterentwicklung umsetzen (Bohndick, Kohlmeyer & Buhl, 2019). Die tatsächliche Umsetzung von Konsequenzen durch die Teilnehmenden nach Erhalt des Assessment-Feedbacks ist als zusätzlicher Übersetzungsschritt einzustufen, der in Bezug auf die resultierende Entwicklung der Reflexionsfähigkeit zu evaluieren ist. Eine Weiterentwicklung der Reflexionsfähigkeit auf Grundlage des Assessment-Feedbacks bzw. des Materials insgesamt wäre zwar ein starkes Argument für die Interpretation der Testwerte als Ausgangspunkt für einen weiteren Professionalisierungsprozess. Gleichzeitig kann solch eine Entwicklung aber kaum eindeutig auf das Material zurückgeführt werden, da sie bezüglich diverser weiterer Einflüsse und Lerngelegenheiten konfundiert sein kann. Für die Validitätsargumentation wird daher auf die oben genannten Übersetzungsschritte fokussiert, die vorrangig durch die Entwicklung des Materials beeinflussbar sind.

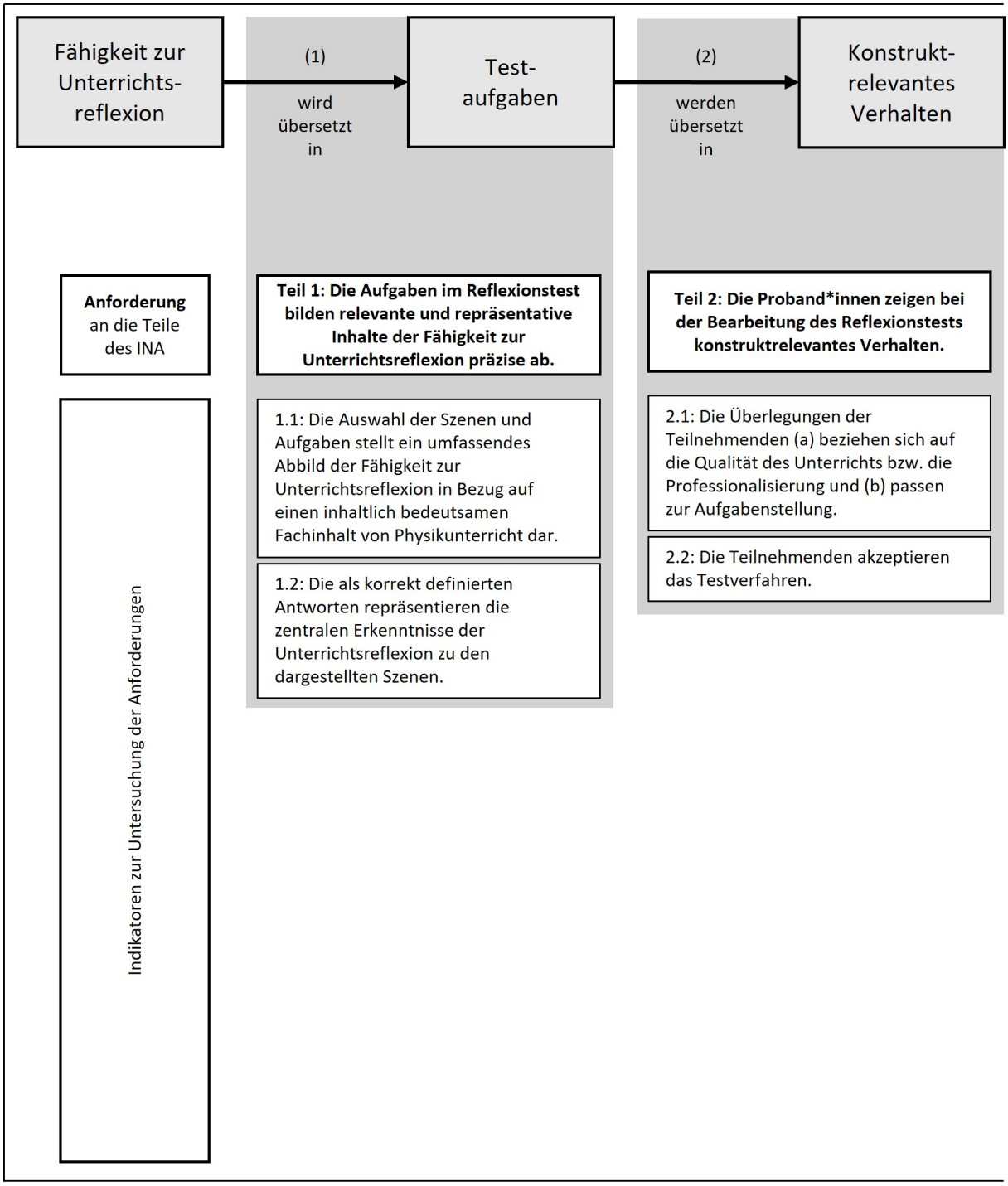
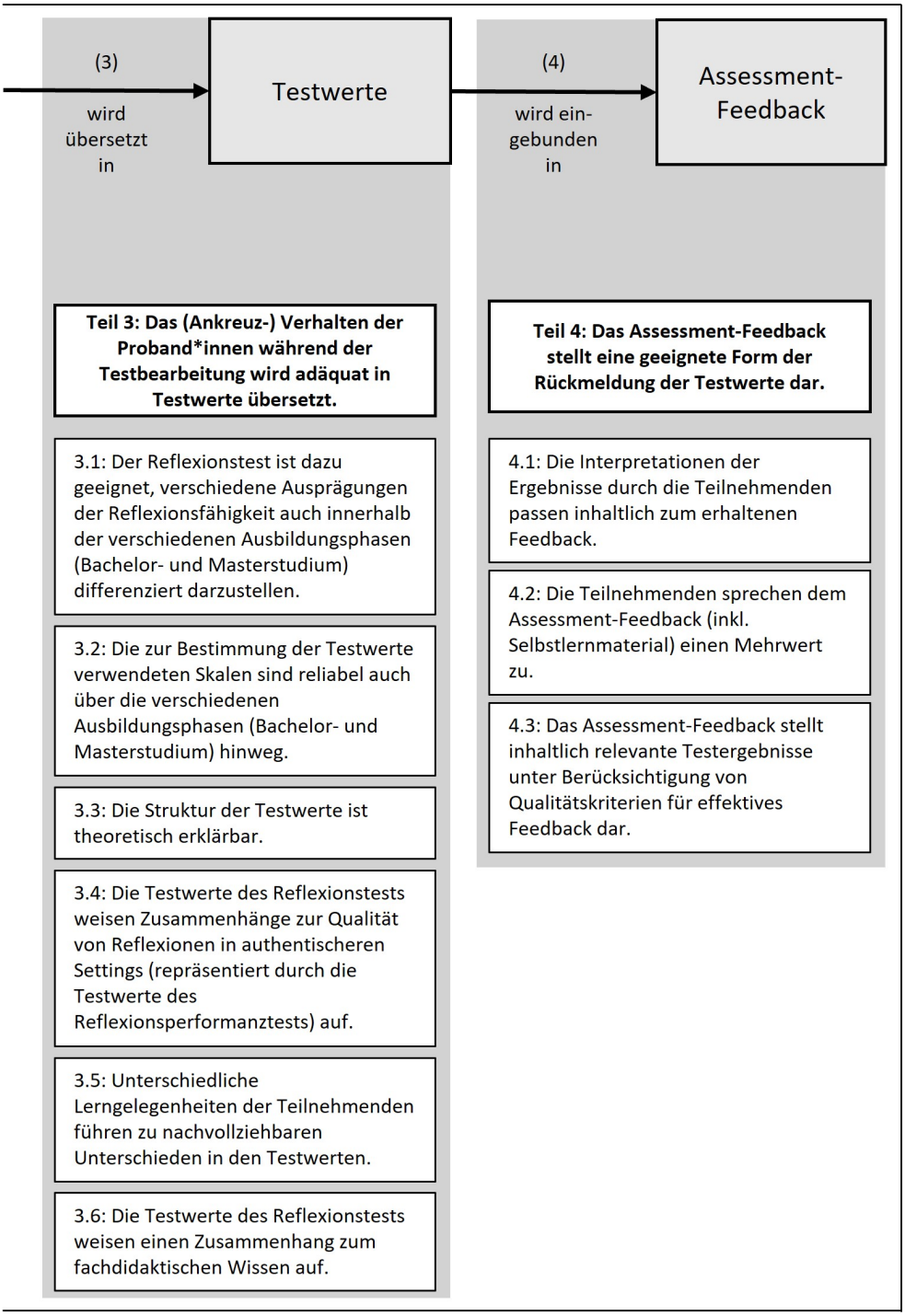


Abbildung 12.2: Darstellung des INA für die Evaluation des Reflexionstests mit Assessment-Feedback ergänzt um die jeweils gestellten Anforderungen und Indikatoren zur Überprüfung dieser


(3)
wird übersetzt in
Testwerte
(4)
wird eingebunden in
Assessment-Feedback
Teil 3: Das (Ankreuz-) Verhalten der Proband*innen während der Testbearbeitung wird adäquat in Testwerte übersetzt.
3.1: Der Reflexionstest ist dazu geeignet, verschiedene Ausprägungen der Reflexionsfähigkeit auch innerhalb der verschiedenen Ausbildungsphasen (Bachelor- und Masterstudium) differenziert darzustellen.
3.2: Die zur Bestimmung der Testwerte verwendeten Skalen sind reliabel auch über die verschiedenen Ausbildungsphasen (Bachelor- und Masterstudium) hinweg.
3.3: Die Struktur der Testwerte ist theoretisch erklärbar.
3.4: Die Testwerte des Reflexionstests weisen Zusammenhänge zur Qualität von Reflexionen in authentischeren Settings (repräsentiert durch die Testwerte des Reflexionsperformanztests) auf.
3.5: Unterschiedliche Lerngelegenheiten der Teilnehmenden führen zu nachvollziehbaren Unterschieden in den Testwerten.
3.6: Die Testwerte des Reflexionstests weisen einen Zusammenhang zum fachdidaktischen Wissen auf.
Teil 4: Das Assessment-Feedback stellt eine geeignete Form der Rückmeldung der Testwerte dar.
4.1: Die Interpretationen der Ergebnisse durch die Teilnehmenden passen inhaltlich zum erhaltenen Feedback.
4.2: Die Teilnehmenden sprechen dem Assessment-Feedback (inkl. Selbstlernmaterial) einen Mehrwert zu.
4.3: Das Assessment-Feedback stellt inhaltlich relevante Testergebnisse unter Berücksichtigung von Qualitätskriterien für effektives Feedback dar.

13 Weitere eingesetzte Testinstrumente

Im Rahmen der Untersuchungen - vor allem zur externen Validität - werden verschiedene weitere Instrumente eingesetzt, die nachfolgend beschrieben werden. Zusätzlich wird der bereits beschriebene Performanztest von Kempin, Kulgemeyer und Schecker (2018) eingesetzt (s. Kapitel 7.7.3).

13.1 Fragebogen zur Demografie

Die Lernumgebung wurde für die breite Zielgruppe (angehender) Physiklehrkräfte (vorrangig für Lehramtsstudierende) entwickelt. Zur Charakterisierung der Stichprobe werden daher demografische Daten der Teilnehmenden erhoben. Dazu gehören für alle Teilnehmenden:

- allgemeine Informationen zu Alter und Geschlecht und
- die Angabe der aktuellen Ausbildungsphase (i. d. R. Bachelor- oder Masterstudium oder Referendariat)

sowie zusätzlich für Studierende

- Angaben zur (schulischen) Ausbildung: abgeschlossene Ausbildungen, Abiturnote, Wahl von Physik bzw. Mathematik als Prüfungsfach im Abitur und die jeweils letzte Physik-, Mathematik- und Deutschnote und
- Angaben zum Studium: angestrebter Abschluss (Lehramts- oder Vollfachstudium), weitere Studienfächer neben Physik, aktuelles Fachsemester sowie Praxiserfahrungen, die während des Studiums gesammelt wurden.

Der Fragebogen zur Demografie wurde zur besseren Vergleichbarkeit der Ergebnisse weitgehend aus den Erhebungen des Projekts *ProfiLeP+* übernommen und aufgrund der erweiterten Zielgruppe an drei Stellen leicht verändert: Da nicht nur Studierende im Praxissemester befragt wurden, wurde zunächst die Abfrage der Ausbildungsphase ergänzt sowie für Studierende die Frage, ob sie im Rahmen ihres Studiums bereits Pflichtpraktika absolviert haben (und ggf. in welchem Umfang). Außerdem wurde die Option *divers* als Geschlechtsangabe ergänzt.[136]

[136]Der vollständige Fragebogen befindet sich im Anhang in Kapitel D ab Seite 455.

13.2 Items zur Motivation während der Testbearbeitung

Es ist davon auszugehen, dass mangelnde Motivation die Bearbeitung des Reflexionstests beeinflussen kann. Daher werden die Teilnehmenden am Ende des Instruments um eine kurze Selbsteinschätzung bezüglich ihrer Motivation während der Testbearbeitung gebeten. Dafür positionieren sie sich auf einer fünfstufigen Likert-Skala von *stimme gar nicht zu* über *teils teils* bis hin zu *stimme völlig zu* (die jeweils dazwischenliegenden Stufen sind nicht beschriftet) zu folgenden drei Aussagen:

- Ich habe die Aufgaben engagiert bearbeitet.
- Ich habe bei der Bearbeitung der Aufgaben mein Bestes gegeben.
- Ich habe mich bei der Bearbeitung der Aufgaben bis zum Schluss konzentriert.

Dieselben Aussagen wurden ebenfalls bereits im Projekt *ProfiLeP+* verwendet, um die Motivation während der Bearbeitung des Performanztests zur Messung der Reflexionsfähigkeit zu erfassen.[137]

Da auch diese Einschätzung anonym durchgeführt wird, kann davon ausgegangen werden, dass Effekte der sozialen Erwünschtheit auf das Antwortverhalten zu diesen Aussagen zumindest reduziert wurden (Döring & Bortz, 2016, S. 439f.). Da es sich um eine nicht weiter untersuchte Selbsteinschätzung handelt, müssen die Ergebnisse dennoch vorsichtig interpretiert werden. Es kann lediglich angestrebt werden, möglicherweise auftretende, sehr starke Einflüsse durch die Testmotivation zu erkennen.

13.3 Fragebogen zur Akzeptanz der Lernumgebung

Die Frage nach der Akzeptanz der Lernumgebung bzw. ihrer Bestandteile ist aus Perspektive der Teilnehmenden sowohl während der Bearbeitung des Instruments bedeutsam, als auch bei der Interpretation des erhaltenen Assessment-Feedbacks und geht über die zuvor beschriebene Erfassung der Motivation hinaus.

Um die Akzeptanz der entwickelten Lernumgebung zu evaluieren, wird aufgrund seiner weiten Verbreitung und vielfach evaluierten Passung für verschiedene technische Systeme das TAM (s. Kapitel 6.1.5) mit ausgewählten Ergänzungen des TAM2 zugrunde gelegt. Die Entscheidung für das TAM und gegen die TPB (und TRA) liegt zunächst in der empirischen Überlegenheit des TAM gegenüber den anderen Theorien begründet (Venkatesh & Davis, 2000). Auch erscheint die Wahl des TAM mit dem Bezug auf technische Systeme und deren Praktikabilität

[137] Der zugehörige Fragebogen befindet sich im Anhang in Kapitel E ab Seite 459.

aufgrund der Umsetzung der Lernumgebung in einem online-basierten Format sinnvoll.

Die Evaluation der Lernumgebung unterscheidet sich zu berichteten Einsätzen des TAM zur Evaluation von technischen Systemen allerdings in zwei zentralen Aspekten:

- Akzeptanzuntersuchungen mit dem TAM beschäftigen sich i. d. R. mit technischen Systemen, die inhaltsunabhängig genutzt werden können (z. B. E-Learning (Olbrecht, 2010)). Die hier betrachtete Lernumgebung aus Testinstrument und Assessment-Feedback ist hingegen ein sehr konkretes, inhaltsgebundenes Material, das für den spezifischen Einsatz der Erfassung, Rückmeldung und Weiterentwicklung von Reflexionsfähigkeit konzipiert ist.[138]
- Damit verbunden ist die Tatsache, dass die untersuchten technischen Systeme i. d. R. für den dauerhaften Einsatz z. B. am Arbeitsort evaluiert werden. Die hier betrachtete Lernumgebung wird punktuell, nicht dauerhaft begleitend, eingesetzt: Auch der einmalige Einsatz der Lernumgebung soll einen Mehrwert liefern, ein mehrfacher Einsatz erscheint zwar möglich, allerdings nur dann sinnvoll, wenn (längerfristige) Entwicklungen begleitet werden sollen. Um Aussagen über die Lernumgebung treffen zu können, müssen die Studierenden sie kennen, sodass die Untersuchung der Akzeptanz nach dem (potenziell abschließenden) Einsatz der Lernumgebung durchgeführt wird.

Diese Unterschiede führen dazu, dass erprobte Modelle und Fragebögen (z. B. von Venkatesh und Davis (2000)) nicht direkt übertragen werden können. Mangels passgenauer alternativer Modelle wird für die Akzeptanzbefragung der Lernumgebung eine Adaption der in Kapitel 6.1.5 vorstellten Modelle vorgenommen: Neben den zentralen Elementen des TAM (1) Nutzungsabsicht/Akzeptanz, (2) wahrgenommene Bedienbarkeit und (3) wahrgenommener Nutzen werden (4) die Relevanz für den Beruf bzw. das Studium und (5) die Qualität der Ergebnisse als kognitive Einflussfaktoren sowie (6) die subjektive Norm als sozialer Einflussfaktor auf den wahrgenommenen Nutzen in Anlehnung an das TAM2 berücksichtigt. Die Aufschlüsselung des wahrgenommenen Nutzens in die verschiedenen Faktoren soll einen detaillierteren Einblick z. B. in Schwierigkeiten bei der Nutzung bzw. in die Bewertung der Lernumgebung ermöglichen, den das TAM so nicht ermöglicht (Mathieson, 1991). Eine Berücksichtigung des tatsächlichen Nutzungsverhaltens - das entspräche hier z. B. des erneuten Einsatzes des Tests - erscheint nicht sinnvoll bzw. realisierbar, da die (erste) Nutzung der Lernumgebung zum Zeitpunkt der Befragung abgeschlossen sein muss und ein kurzfristiger erneuter Einsatz nicht

[138] Im Gegensatz zu Modellen aus der Informationswissenschaft (z. B. DeLone und McLean, 2003) ist die inhaltliche Qualität der entwickelten Materialien kein Bestandteil des TAM und TAM2. Für die Anlage dieser Arbeit stellt dies allerdings keine problematische Einschränkung dar, da die inhaltliche Qualität der Materialien in der Validierung an anderer Stelle berücksichtigt wird.

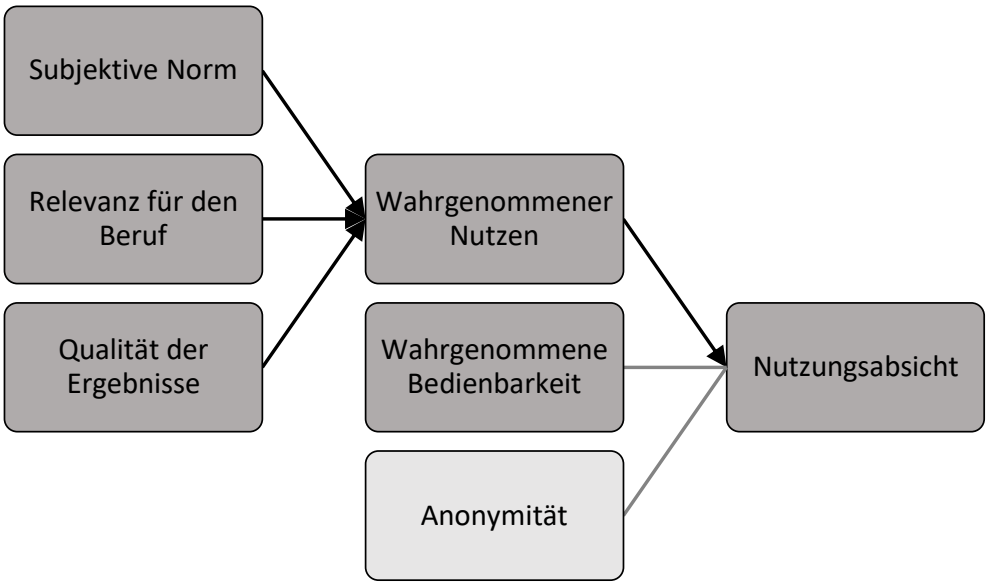


Abbildung 13.1: Adaption des TAM/TAM2 für die Akzeptanzbefragung zur Lernumgebung

erstrebenswert ist. Insofern wird hier auf die positive Haltung bzw. die Nutzungsabsicht als bester Prädiktor für tatsächliches Verhalten (Mathieson, 1991; Turner et al., 2010) fokussiert.

Aufgrund der bereits abgeschlossenen (ersten) Nutzung der Lernumgebung und der potenziell sowohl freiwilligen (aus eigenem Interesse) und vorgegebenen (z. B. im Rahmen einer universitären Veranstaltung) erneuten Nutzung erscheint die Berücksichtigung des Faktors „Freiwilligkeit“ des TAM2 unpassend. Da die Arbeit mit der Lernumgebung individuell ist und insbesondere die Ergebnisse nur individuell zugänglich sind, werden auch die Faktoren „Image“ und „Ergebnisnachweisbarkeit“ nicht berücksichtigt.

Zusätzlich zu den oben aufgeführten Faktoren wird (7) die Wahrung der Anonymität in der Akzeptanzbefragung berücksichtigt. Teilnehmenden wird versprochen, dass ihr individuelles Ergebnis anonym ausgewertet wird. Sollten allerdings z. B. Studierende befürchten, dass Dozierende individuelle Ergebnisse der Studierenden erhalten und bspw. zur Notenvergabe nutzen könnten, könnte dies die Akzeptanz des Testinstruments stark beeinflussen.

Wie bereits erläutert sind erprobte Fragebögen (Operationalisierungen des TAM von Davis (1989) bzw. in der deutschen Übersetzung von Olbrecht (2010, S. 240f.) und des TAM2 von Venkatesh und Davis (2000)) in der vorliegenden Version nicht geeignet, um die Akzeptanz der Lernumgebung zu erheben. In Anlehnung an das Verständnis der einzelnen Einflussfaktoren bzw. Variablen der Modelle werden daher verschiedene Anpassungen vorgenommen:

- Die hypothetische Option „Angenommen, ich habe Zugang zum System, beabsichtige ich, es zu nutzen.“ wird gestrichen, da der Zugang zur Lernumgebung bereits besteht.
- Der Begriff des „Systems“ wird durch „Lernumgebung“ bzw. durch die einzelnen Elemente der Lernumgebung („Testinstrument“, „Rückmeldung“ oder „Selbstlernmaterial“) ersetzt. Es wird also nicht die Lernumgebung pauschal als Einheit evaluiert, da die Akzeptanz jedes der drei Elemente der Lernumgebung Voraussetzung einer umfassenden Akzeptanz der Lernumgebung ist. Hürden für die Akzeptanz können so ggf. genauer identifiziert werden.
- Statt der Passung zu bzw. Relevanz für den „Job“ bzw. „Beruf“ wird die Bedeutung für das „Studium“ als aktuellen Ausbildungsabschnitt der Befragten erfragt.
- Verschiedene Optionen werden für die Lernumgebung konkretisiert, z. B. wird anstelle der Aussage „Der Umgang mit dem System ist für mich klar und verständlich.“ (Olbrecht, 2010, S. 204; ursprünglich Venkatesh und Davis, 2000) eine Unterscheidung zwischen Testdurchführung und Herunterladen der Rückmeldung vorgenommen, da sich die technische Umsetzung beider Vorgänge unterscheidet: (1) „Bei der Testdurchführung traten technische Schwierigkeiten auf.“; (2) „Beim Herunterladen der Rückmeldung traten Schwierigkeiten auf.“
- Zusätzlich werden Aspekte ergänzt, die ein detaillierteres Bild der Einschätzung versprechen (z.B. zum Schwierigkeitsgrad des Testinstruments).

So ergibt sich ein Fragebogen, bei dem die Teilnehmenden sich auf einer vierstufigen Likert-Skala (*stimme gar nicht zu*, *stimme eher nicht zu*, *stimme eher zu*, *stimme voll und ganz zu*) sowie der Option *keine Antwort* zu verschiedenen adaptierten Aussagen positionieren. Der Akzeptanzfragebogen mit der Zuordnung der einzelnen Aussagen zu den Faktoren des adaptierten TAM befindet sich in Kapitel F (ab S. 461).

Aufgrund der Teilung der Lernumgebung in Reflexionstest und Assessment-Feedback mit Rückmeldung und Selbstlernmaterial wird die Einschätzung der Akzeptanz zu den einzelnen Bestandteilen idealerweise an unterschiedlicher Stelle erfragt: (1) Nach Bearbeitung des Reflexionstests, (2) nach Erhalt der Rückmeldung vor der Bearbeitung des Selbstlernmaterials und u. a. für die globale Bewertung (3) nach Bearbeitung des Selbstlernmaterials. Diese Fragmentierung der Akzeptanzbefragung der Teilnehmenden ermöglicht es einerseits, die Akzeptanz der Teilnehmenden unmittelbar nach der Bearbeitung des Reflexionstests bzw. Beschäftigung mit dem Selbstlernmaterial zu erheben. Andererseits ist es im Studiendesign u. a. aufgrund der Anonymität der Erhebungen nach Abschluss des Reflexionstests deutlich erschwert, die Teilnehmenden für eine erneute Befragung zu erreichen, sodass durch das Einbinden eines Teils der Akzeptanzbefragung

in den Reflexionstest möglichst viele Teilnehmende erreicht werden können. Ergänzende Informationen bezüglich der Akzeptanz der Lernumgebung (auch über den Reflexionstest hinaus) liefern die Interviews mit den Studierenden zur konsequentiellen Validität (vor allem zur Interpretation des und zum Umgang mit dem erhaltenen Assessment-Feedback(s)).

13.4 Test zum fachdidaktischen Wissen

Zur Messung des fachdidaktischen Wissens wird auf eine Weiterentwicklung des Testinstruments von Gramzow (2015, S. 102ff.) zurückgegriffen (nachfolgend abgekürzt als FDW-Test). Als theoretische Grundlage dient das Modell des fachdidaktischen Wissens nach Gramzow, Riese und Reinhold (2013) (s. Kapitel 5.2.4). Das Testinstrument erfasst fachdidaktisches Wissen im Inhaltsbereich der Mechanik in Bezug auf die vier Kompetenzfacetten *Instruktionsstrategien*, *Schülervorstellungen*, *fachdidaktische Konzepte* sowie *Experimente und Vermittlung eines angemessenen Wissenschaftsverständnisses (Experimente)* wobei für die unterschiedlichen Facetten jeweils die kognitiven Anforderungen *Reproduzieren*, *Anwenden* und *Analysieren* berücksichtigt werden (Riese, Gramzow & Reinhold, 2017). Umgesetzt wird dies in der ursprünglichen Paper-Pencil-Version des Testinstruments in einer Kombination aus offenen Aufgaben, die mit Hilfe eines Kodiermanuals händisch kodiert werden und MC-Aufgaben (Kulgemeyer et al., 2023). Das Testinstrument wurde in Bezug auf Elemente der inhaltlichen sowie externen und strukturellen Validität (zusammengefasst als Konstruktvalidität; vgl. Gramzow, 2015, S. 171; s. auch Riese, Gramzow und Reinhold, 2017) evaluiert.

Die im Rahmen der hier beschriebenen Untersuchungen eingesetzte Version des Testinstruments wurde für einen niedrigschwelligen Einsatz mit automatisierter Auswertung optimiert. Dazu wurden zunächst die offenen Aufgaben „unter Nutzung und Kategorisierung der in bisherigen Erhebungen erhaltenen frei formulierten Antworten sowie unter Berücksichtigung des Kodiermanuals und aktueller physikdidaktischer Literatur“ (Jordans, Zeller, Große-Heilmann & Riese, 2022, S. 765) in ein MC-Format überführt; die Bewertung dieser Aufgaben wird als Schwellenbepunktung umgesetzt und das Instrument vollständig in *LimeSurvey* digitalisiert (Jordans, Zeller, Große-Heilmann & Riese, 2022). Bei der Evaluation des veränderten Testinstruments zeigte sich, dass beide Versionen das didaktische Wissen nicht identisch abbilden: Für eine Stichprobe von $N = 24$ Physik-Lehramtsstudierenden (bzw. $N = 10$ für die Facette *fachdidaktische Konzepte*), die beide Instrumente bearbeiteten, ergaben sich Zusammenhänge für die verschiedenen Facetten in beiden Instrumenten, die zwischen $r_{\text{Pearson}} = 0,37$ ($p < 0,05$; Facette *Experimente*) und $r_{\text{Pearson}} = 0,81$ ($p < 0,01$; Facette *fachdidaktische Konzepte*) liegen (Kulgemeyer et al., 2023). Ursächlich hierfür könnten umfangreichere Anforderungen sein, die mit der Bearbeitung von offenen Aufgaben im ursprünglichen Format einhergehen

(Kulgemeyer et al., 2023). Gleichzeitig konnten auch für Aufgaben, die in beiden Instrumenten identisch eingebunden sind, nicht immer die zu erwartenden hohen Zusammenhänge beobachtet werden; die Fähigkeiten der Studierenden können hier also nicht als stabil angenommen werden (Jordans, Zeller, Große-Heilmann & Riese, 2022). Der Vorteil der hier eingesetzten geschlossenen Version des Testinstruments liegt in der Möglichkeit der ressourcenschonenden, automatisierten Auswertung, die zusätzlich die zeitnahe Bereitstellung von Assessment-Feedback ermöglicht (Kulgemeyer et al., 2023).

14 Zusammenfassende Übersicht über die durchgeführten Untersuchungen und die Stichprobe

In diesem Kapitel wird zur Orientierung zunächst ein grober deskriptiver Überblick über die vorliegende Stichprobe gegeben (s. Kapitel 14.1), dem ein Überblick über die durchgeführten Untersuchungen zur Evaluation des entwickelten Testinstruments und Assessment-Feedbacks folgt (s. Kapitel 14.2).

14.1 Übersicht über die Stichprobe

Zum Reflexionstest liegen insgesamt $N = 192$ vollständige Bearbeitungen vor, vorrangig von Physik-Lehramtsstudierenden im Bachelor- oder Masterstudium. Im Rahmen der Analysen (insbesondere zu den Indikatoren 3.1 bis 3.6) werden verschiedene Teilstichproben herangezogen, die an den jeweiligen Stellen näher beschrieben werden. Tabelle 14.1 gibt einen ersten Überblick über die Zusammensetzung der Stichprobe hinsichtlich studienbezogener Merkmale. Unvollständige Bearbeitungen wurden für die Auswertung nicht berücksichtigt, da die Gründe, die zum Abbruch der Bearbeitung führten, unklar sind. Für die anschließenden explorativen Untersuchungen zur Reflexionsfähigkeit und ihrer Entwicklung (in Teil IV) wird eine um die drei Personen mit sonstigem Berufsziel (außerhalb des Lehramts) reduzierte Stichprobe von $N = 189$ als Grundgesamtheit herangezogen.

Insgesamt handelt es sich hierbei - ebenso wie bei den Teilstichproben, die zur Untersuchung einiger Indikatoren herangezogen wurden - um eine Gelegenheitsstichprobe, die aufgrund ihrer Zusammensetzung über verschiedene Studienstandorte und bzw. oder -phasen hinweg eine gewisse Breite aufweist, die Anhaltspunkt für die Generalisierbarkeit der Ergebnisse sein kann.

(Teil-)Stichprobe		**N**
Gesamtstichprobe		192
Ausbildungsphase	Bachelor	90
	Master	86
	Referendariat	11
	Sonstiges (inklusive Staatsexamen)	5
Berufsziel	Lehramt Physik (Sek I/II)	172
	Lehramt Sachunterricht mit Wahlfach Physik (Primarstufe)	17
	Sonstiges	3

Tabelle 14.1: Übersicht über die Zusammensetzung der Stichprobe

14.2 Übersicht über die im Rahmen der Validierung durchgeführten Untersuchungen

Wie in Kapitel 12 beschrieben, erfolgt die Validierung der Testwertinterpretation anhand der Prüfung verschiedener Indikatoren zu den Anforderungen an die Übersetzungsschritte entlang des INA. Zur Prüfung der Indikatoren werden theoretische Argumentationen sowie empirische Evidenz herangezogen. Die durchgeführten jeweils Untersuchungen und Ergebnisse werden in den Kapiteln 15 bis 18 berichtet und diskutiert; in Kapitel 19 werden die Ergebnisse zur Validitätsargumentation zusammengeführt. Die Abbildungen 14.1 und 14.2 stellen die vollständige Übersicht über die Indikatoren zur Prüfung der Anforderungen an das INA sowie ergänzend die jeweils herangezogene Bewertungsgrundlage (theoretische Argumentation bzw. durchgeführte Erhebungen und zentrale Analysen) dar. Insbesondere ist jeweils der Einsatz zusätzlicher Testinstrumente angegeben.

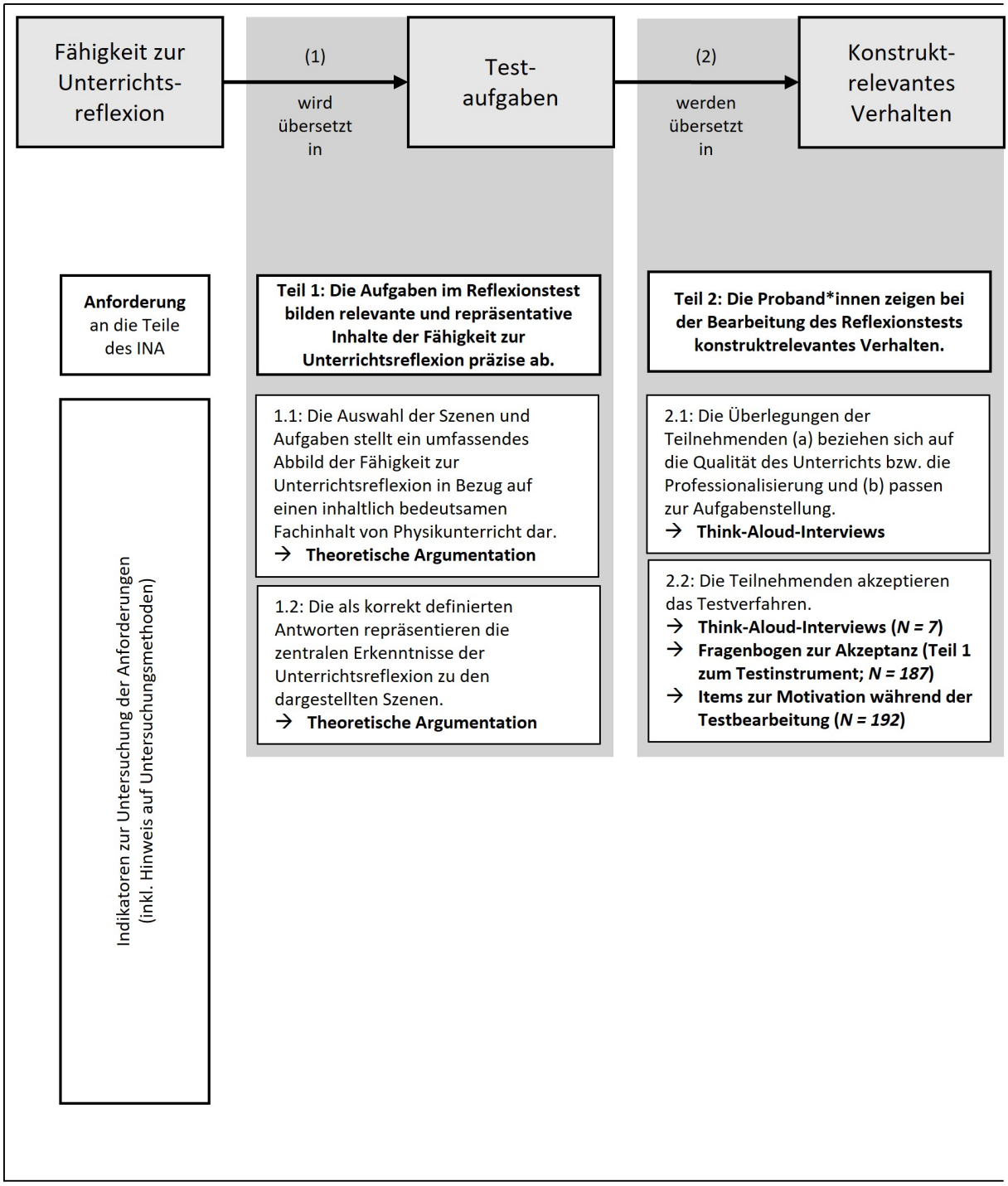
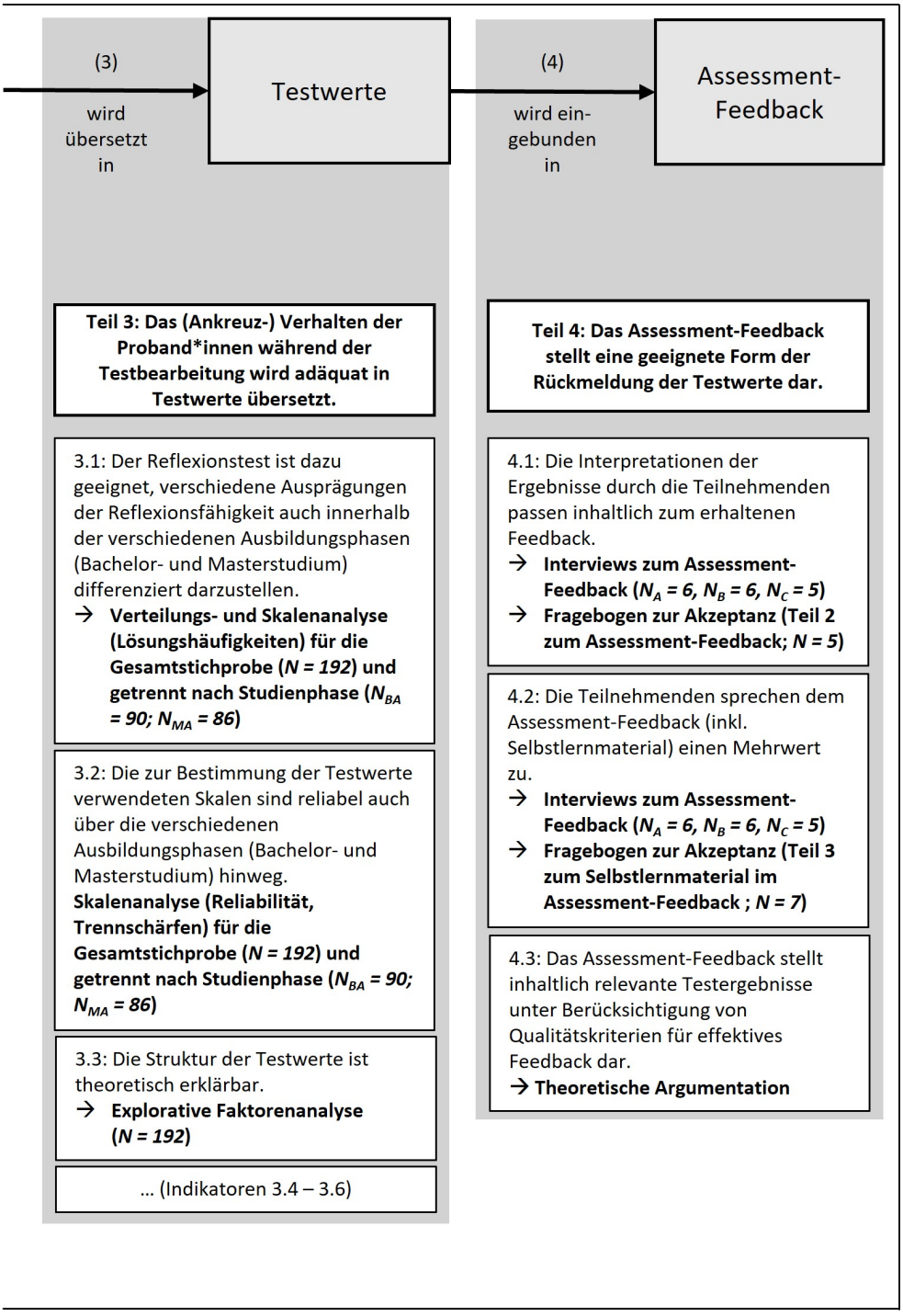


Abbildung 14.1: Darstellung der Untersuchungen entlang des INA mit den jeweiligen Anforderungen und Indikatoren sowie (fett gedruckt) den Bewertungsgrundlagen zu deren Überprüfung (Teil 1/2)


(3)
wird übersetzt in
Testwerte
(4)
wird eingebunden in
Assessment-Feedback
Teil 3: Das (Ankreuz-) Verhalten der Proband*innen während der Testbearbeitung wird adäquat in Testwerte übersetzt.
3.1: Der Reflexionstest ist dazu geeignet, verschiedene Ausprägungen der Reflexionsfähigkeit auch innerhalb der verschiedenen Ausbildungsphasen (Bachelor- und Masterstudium) differenziert darzustellen.
→ Verteilungs- und Skalenanalyse (Lösungshäufigkeiten) für die Gesamtstichprobe (N = 192) und getrennt nach Studienphase (N_{BA} = 90; N_{MA} = 86)
3.2: Die zur Bestimmung der Testwerte verwendeten Skalen sind reliabel auch über die verschiedenen Ausbildungsphasen (Bachelor- und Masterstudium) hinweg.
Skalenanalyse (Reliabilität, Trennschärfen) für die Gesamtstichprobe (N = 192) und getrennt nach Studienphase (N_{BA} = 90; N_{MA} = 86)
3.3: Die Struktur der Testwerte ist theoretisch erklärbar.
→ Explorative Faktorenanalyse (N = 192)
... (Indikatoren 3.4 – 3.6)
Teil 4: Das Assessment-Feedback stellt eine geeignete Form der Rückmeldung der Testwerte dar.
4.1: Die Interpretationen der Ergebnisse durch die Teilnehmenden passen inhaltlich zum erhaltenen Feedback.
→ Interviews zum Assessment-Feedback (N_A = 6, N_B = 6, N_C = 5)
→ Fragebogen zur Akzeptanz (Teil 2 zum Assessment-Feedback; N = 5)
4.2: Die Teilnehmenden sprechen dem Assessment-Feedback (inkl. Selbstlernmaterial) einen Mehrwert zu.
→ Interviews zum Assessment-Feedback (N_A = 6, N_B = 6, N_C = 5)
→ Fragebogen zur Akzeptanz (Teil 3 zum Selbstlernmaterial im Assessment-Feedback ; N = 7)
4.3: Das Assessment-Feedback stellt inhaltlich relevante Testergebnisse unter Berücksichtigung von Qualitätskriterien für effektives Feedback dar.
→ Theoretische Argumentation

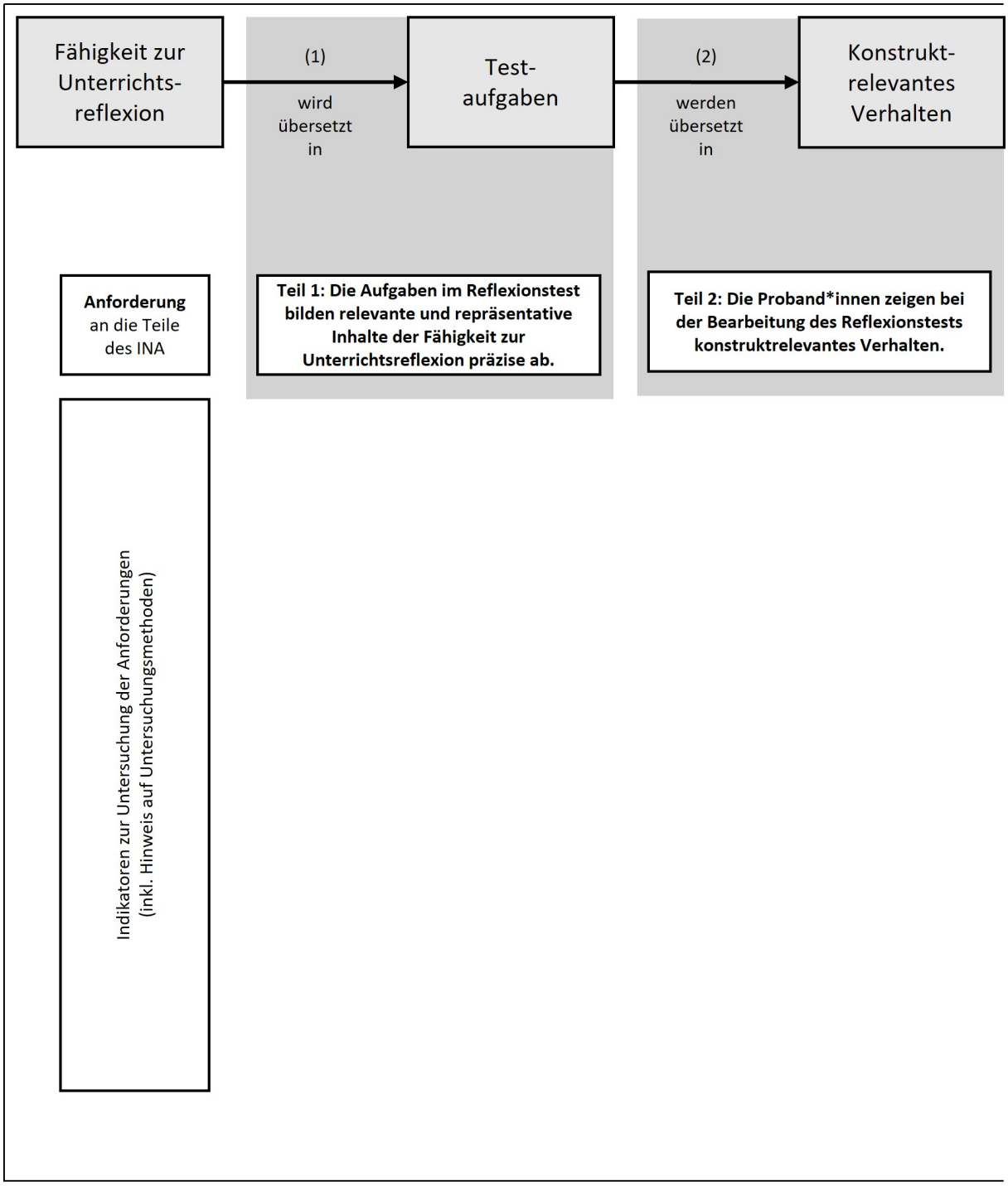
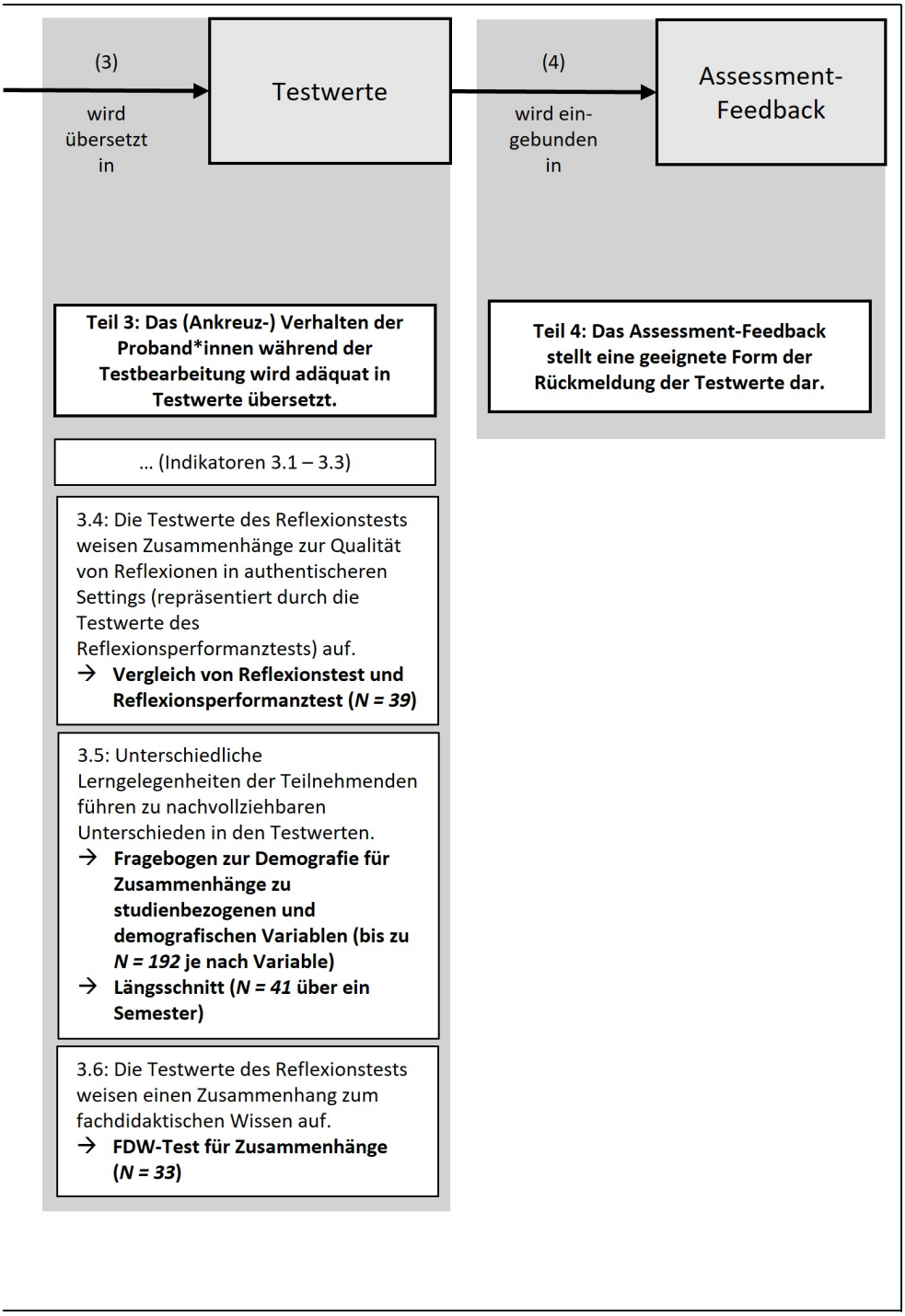


Abbildung 14.2: Darstellung der Untersuchungen entlang des INA mit den jeweiligen Anforderungen und Indikatoren sowie (fett gedruckt) den Bewertungsgrundlagen zu deren Überprüfung (Teil 2/2)


(3)
wird übersetzt in
Testwerte
(4)
wird eingebunden in
Assessment-Feedback
Teil 3: Das (Ankreuz-) Verhalten der Proband*innen während der Testbearbeitung wird adäquat in Testwerte übersetzt.
Teil 4: Das Assessment-Feedback stellt eine geeignete Form der Rückmeldung der Testwerte dar.
... (Indikatoren 3.1 – 3.3)
3.4: Die Testwerte des Reflexionstests weisen Zusammenhänge zur Qualität von Reflexionen in authentischeren Settings (repräsentiert durch die Testwerte des Reflexionsperformanztests) auf.
→ Vergleich von Reflexionstest und Reflexionsperformanztest (N = 39)
3.5: Unterschiedliche Lerngelegenheiten der Teilnehmenden führen zu nachvollziehbaren Unterschieden in den Testwerten.
→ Fragebogen zur Demografie für Zusammenhänge zu studienbezogenen und demografischen Variablen (bis zu N = 192 je nach Variable)
→ Längsschnitt (N = 41 über ein Semester)
3.6: Die Testwerte des Reflexionstests weisen einen Zusammenhang zum fachdidaktischen Wissen auf.
→ FDW-Test für Zusammenhänge (N = 33)

15 Untersuchung der Anforderung an Teil 1 des INA: Die Aufgaben im Reflexionstest bilden relevante und repräsentative Inhalte der Fähigkeit zur Unterrichtsreflexion präzise ab

Die Untersuchung der Anforderung des ersten Teils des INA, der inhaltlichen adäquaten Abbildung der Fähigkeit zur Unterrichtsreflexion, erfolgt in zwei Schritten (eine Übersicht der zugehörigen Indikatoren ist in Abbildung 14.1 auf Seite 186 zum ersten Übersetzungsschritt dargestellt): Zunächst wird in Kapitel 15.1 diskutiert, inwieweit die eingesetzten Videovignetten und Aufgaben im Reflexionstest das Konstrukt der Fähigkeit zur Unterrichtsreflexion angemessen abbilden. Anschließend wird in Kapitel 15.2 die Beurteilung der unterschiedlichen Antwortoptionen und die Angemessenheit des gewählten Vorgehens beim Scoring näher beleuchtet. In Kapitel 15.3 wird abschließend im Sinne der inhaltlichen Validität beurteilt, inwieweit Unterrichtsreflexion im Reflexionstest angemessen abgebildet ist.

15.1 Evaluation der Abbildung der Fähigkeit zur Unterrichtsreflexion im Reflexionstest (Indikator 1.1)

Die Untersuchung des Indikators 1.1 *„Die Auswahl der Szenen und Aufgaben stellt ein umfassendes Abbild der Fähigkeit zur Unterrichtsreflexion in Bezug auf einen inhaltlich bedeutsamen Fachinhalt von Physikunterricht dar.“* erfolgt überwiegend theoretisch bzw. auf Grundlage der Ergebnisse der Entwicklung und Evaluation des Reflexionsperformanztests von Kempin, Kulgemeyer und Schecker (2018). Die inhaltliche Validität ist für die Testwertinterpretation zum Reflexionstest auf

zwei Ebenen zu diskutieren: Einerseits ist sicherzustellen, dass die eingesetzten Videovignetten eine inhaltliche angemessene Grundlage für die Reflexion darstellen. Andererseits ist zu betrachten, ob die Aufgaben die verschiedenen Facetten bzw. von Unterrichtsreflexion adäquat abbilden. Dazu ist in beiden Fällen zu prüfen, dass der Reflexionstest Unterricht und dessen Reflexion in allen relevanten Aspekten abbildet, ohne Irrelevantes darzustellen und diese Aspekte in einem angemessenen Verhältnis zueinander abgebildet werden (Hartig, Frey & Jude, 2012).

Bezüglich der Auswahl des in den Videovignetten dargestellten Unterrichts muss einschränkend festgehalten werden, dass Physikunterricht aufgrund der Vielfalt an möglichen Inhalten, Methoden, beteiligten Personen etc. nicht vollständig abgebildet werden kann, sodass die Auswahl der dargestellten Situationen zu begründen ist. Physikalischer Inhalt der Doppelstunde sind die Newtonschen Axiome und die Impulserhaltung. Da die Mechanik üblicherweise zu Beginn des Studiums unterrichtet wird, ermöglicht diese Themenwahl den Einsatz des Reflexionstests schon in unteren Semestern. Zudem erweist sich das Fachwissen zur klassischen Mechanik als geeigneter Prädiktor für das physikalische Fachwissen generell (Riese und Reinhold, 2008; Friege und Lind, 2004), sodass mit der Fokussierung auf dieses Inhaltsgebiet tendenziell günstige Voraussetzungen für die Generalisierbarkeit der Ergebnisse geschaffen werden. Im Reflexionsperformanztest wird die inhaltliche Breite des dargestellten Unterrichts erreicht, indem die Videovignetten insgesamt eine Physik-Doppelstunde nachbilden. Bei der Auswahl der Vignetten für den Reflexionstest wird versucht, diese Breite überwiegend beizubehalten, indem der Unterrichtsverlauf weiterhin in möglichst vielfältigen Situationen nachgezeichnet wird, der Fokus aber stärker auf für den *Physik*unterricht spezifische Situationen gelegt wird (s. Kapitel 10.2.1).

Theoretische Grundlage für die Aufgabenstruktur im Reflexionstest ist das Reflexionsmodell nach Nowak, Kempin, Kulgemeyer und Borowski (2019) (s. Kapitel 7.5.4). Eine (vollständige) Reflexion umfasst demnach die Elemente Beschreibung einer Situation und bzw. oder von Rahmenbedingungen, Bewertung der Situation, Entwicklung von Alternativen und Ableiten von Konsequenzen; idealerweise in dieser Reihenfolge strukturiert. Der Umgang mit den einzelnen Bestandteilen des Reflexionsmodells wird nachfolgend beschrieben.

Auf die explizite Einbindung von Beschreibungen im Rahmen von Testaufgaben wird verzichtet. Die Umsetzung der Erfassung von Beschreibungen in MC-Aufgaben entspräche insbesondere unter Nutzung von Distraktoren einer Überprüfung der Aufmerksamkeit der Proband*innen und stellt keine selbständige analytische Tätigkeit wie bspw. die selbstständige Verknüpfung einzelner beobachteter Elemente und damit keinen Teil einer Reflexion dar (vgl. Plöger, Scholl und Seifert, 2015). Dennoch kommen Beschreibungen im Testinstrument vor. Sie sind Bestandteil diverser Antwortoptionen in den MC-Aufgaben zu Bewertungen und Alternativen, um explizite, nachvollziehbare Bezüge zu beobachteten Situationen herzustellen. So

enthalten z. B. Bewertungen zur sechsten Vignette, die Roberts Umgang mit Messunsicherheiten kritisieren, die Beschreibung, „dass du [Robert] die Ungleichheit der Messwerte mit ‚Rundungsfehlern' begründest". Auch in den Noticing-Aufgaben werden durch die stichwortartigen Nennungen von beobachtbaren Aspekten der Vignetten - ähnlich der Funktion von Situationsbeschreibungen - Bezüge zu den Videovignetten hergestellt. Die Elemente Bewertungen und Alternativen werden im Testinstrument umfangreich und in ausgewogenem Verhältnis abgebildet, indem zu jedem reflektierten Aspekt je eine Aufgabe zu beiden Elementen in der modellierten Reihenfolge (bzw. Stufung) bearbeitet wird. Konsequenzen als letztes Element des Reflexionsmodells werden nicht in eigenen Aufgaben abgebildet, sondern als Erweiterung von bzw. als längerfristige Handlungsoptionen zur Professionalisierung oder Sicherstellung der Unterrichtsqualität verstanden (z. B. die Empfehlung "Du solltest das Thema Messunsicherheiten für dich selbst noch einmal aufbereiten [...].") und in die Aufgaben zu Alternativen inkludiert. Diese Lösung wurde vorrangig aus dem Grund gewählt, dass auf Grundlage der Bearbeitungen des Reflexionsperformanztests davon auszugehen ist, dass angehende Lehrkräfte insgesamt deutlich weniger Konsequenzen als Bewertungen und Alternativen formulieren (s. Kapitel 10.2.2). Eine Abbildung von Konsequenzen in eigenen MC-Aufgaben entspräche also einer übermäßigen Repräsentation dieses Reflexionselements.

Als zweite Dimension im Reflexionsmodell wird das Vorliegen von Begründungen berücksichtigt. In den Aufgaben zu Bewertungen und Alternativen werden alle Antwortoptionen begründet, wobei sich die Begründungen üblicherweise auf physikalische Aspekte (z. B. „Es ist nicht gut, dass du sagst, du könntest ‚exakt' oder ‚genau' messen, da du so die Schülervorstellung beförderst, man könnte Größen ohne Messunsicherheiten bestimmen." zur fünften Vignette), physikdidaktische Aspekte (z. B. „Es ist nicht gut, dass die Kurzformel ‚Kraft = Gegenkraft' im Merksatz vorkommt, da Schüler*innen diese oft mit dem Kräftegleichgewicht assoziieren." zur ersten Vignette) oder pädagogische Aspekte (z. B. „Es ist gut, dass du die beiden Schüler in das Experiment einbezogen hast. So motivierst und aktivierst du sie." zur vierten Vignette) beziehen und ggf. die Rahmenbedingungen einbeziehen. Einige Begründungen, insbesondere zur Empfehlung, bestimmte Aspekte für eine erneute Durchführung der Unterrichtsstunde beizubehalten, sind eher allgemein gehalten (z. B. „Es ist gut, wie du mit Messunsicherheiten umgegangen bist. Das kannst du erneut so machen." zur fünften Vignette). Die Begründung liegt hier implizit durch die positive Bewertung des Handelns vor, die wiederum auf Grundlage der oben genannten physikalischen, physikdidaktischen oder pädagogischen Überlegungen getroffen worden sein kann. Lediglich bei Aspekten der fachlichen Korrektheit (z. B. „Deine Formulierung des 3. Axioms ist fachlich falsch.") wird keine Begründung für die Bewertung mit aufgeführt, um Proband*innen durch eine erläuternde (fachlich korrekte) Begründung nicht auf Fehler aufmerksam zu machen, da die fachliche Korrektheit eindeutig als richtig oder falsch bewertet wer-

den kann und im Gegensatz zu den übrigen Antwortoptionen kontextunabhängig erfolgen kann (vgl. Kapitel 15.2).[139]

Die Einbettung der Begründungen in die Antwortoptionen zu Bewertungen und Alternativen ermöglicht die Konkretisierung dieser Bewertungen und Alternativen. Für eine Bewertung ohne Begründung, wie z. B. „Der Merksatz ist gut gelungen“ bleibt in der Auswirkung unklar, warum diese Bewertung getroffen wurde, was u. U. die Aussagekraft des Instruments reduziert. Die Dimension der Begründungen wird im Reflexionsmodell als dichotom verstanden, was durch die Vorgabe von Begründungen im Testinstrument nicht abgebildet wird. Um dies zu erreichen, bedürfte es eines zweistufigen Aufgabenformats, in dem zunächst Bewertungen und Alternativen erfasst und jeweils anschließend Begründungen erfragt werden. Das Testinstrument würde so bezüglich der Anzahl der Aufgaben noch umfangreicher und aufgrund der kleinschrittigeren Vorgehensweise repetitiver, sodass einerseits die Bearbeitungsdauer und andererseits die Gefahr eines Motivationsverlusts im Laufe der Testbearbeitung erhöht wird.

Die dritte Dimension des Reflexionsmodells umfasst die Wissensbasen physikalisches, physikdidaktisches und pädagogisches Wissen. Alle 16 Aspekte, die im Rahmen der Bearbeitung des Reflexionstests reflektiert werden, bzw. die zu diesen Aspekten vorliegenden Antwortoptionen, sind dabei stets (mindestens) einer dieser Wissensbasen zuzuordnen (der Umgang mit Schülervorstellungen bspw. dem fachdidaktischen Wissen). Zugrunde liegen dieser Zuordnung i. d. R. die in den Antwortoptionen enthaltenen Begründungen, die zum Großteil auf Wissen aus den unterschiedlichen Wissensbasen Bezug nehmen. Gleichzeitig ist die Zuordnung der Aspekte zu den Wissensbasen oft nicht eindeutig möglich, da in Aufgaben Antwortoptionen mit Bezug zu verschiedenen Wissensbasen zusammengefasst werden können. Insgesamt werden so im Reflexionstest alle drei Wissensbasen berücksichtigt, wobei im Sinne der Reflexion von Physikunterricht auf fachliches und fachdidaktisches Wissen fokussiert wird (diese beiden Wissensbasen werden in neun bzw. zwölf der 16 Aspekte adressiert).

Bislang in dieser Diskussion nicht berücksichtigt sind die Aufgaben zum Noticing nach jeder Videovignette. Diese werden in der Bestimmung der Testwerte zur Reflexionsfähigkeit nicht berücksichtigt. Somit ist einzuschätzen, ob sie den Reflexionstest um zusätzliche, irrelevante Aspekte ergänzen und die inhaltliche Validität des Testinstruments gefährden könnten. Vor dem Hintergrund, dass Noticing als direkte Voraussetzung für bedeutsame Reflexion verstanden werden kann (s. Kapitel 7.6) und Teil realer (kollegialer) Reflexionssituationen ist, ist davon auszugehen, dass diese Aufgaben zur Authentizität der Reflexionssituation insgesamt beitragen.

[139] Prinzipiell wäre es mit einer Formulierung wie „Deine Formulierung des 3. Axioms ist nicht gut gelungen, da sie fachlich falsch ist.“ möglich, auch für diese Antwortoptionen die Struktur aus Bewertung und Begründung zu verwenden.

Auch wenn die bisherigen Ausführungen Argumente dafür enthalten, dass die eingesetzten Videovignetten eine inhaltlich angemessene Grundlage für die Reflexion darstellen und die Bestandteile einer Reflexion gemäß dem Reflexionsmodell von Nowak, Kempin, Kulgemeyer und Borowski (2019) adäquat abgebildet sind, ist kritisch zu hinterfragen, inwieweit die Einschränkung der Handlungsfreiheit, die mit der Verwendung von MC-Aufgaben einhergeht, Einfluss auf die valide Interpretation der Testwerte nimmt. So deutet ein Vergleich der kognitiven Prozesse bei der Bearbeitung von offenen bzw. geschlossenen Aufgaben bei der Erfassung der professionellen Unterrichtswahrnehmung darauf hin, dass je nach Aufgabenformat unterschiedliche kognitive Prozesse ablaufen und insbesondere der Anteil der interpretativen Aussagen bei der Bearbeitung offener Aufgaben größer ist (Eßling, Sunder, Steffensky & Meschede, 2024). Im Reflexionstest geht die Einschränkung der Handlungsfreiheit mit einer Einschränkung der Authentizität der simulierten kollegialen Reflexionssituation einher: Während die Teilnehmenden im Reflexionsperformanztest von Kempin, Kulgemeyer und Schecker (2018) ihre Reflexionen mündlich und in einem offenen Antwortformat äußern, das eine Gesprächssituation mit dem fiktiven Mitpraktikanten simulieren soll, liegt mit der Entwicklung des hier evaluierten Reflexionstests ein Wechsel hin zu einem geschlossenen und schriftlichen Antwortformat vor; die Reflexion wird nicht unmittelbar erfasst, stattdessen müssen die Studierenden ihre Reflexion mit den vorliegenden Antwortoptionen abgleichen, um die passenden auszuwählen. Gleichzeitig reduziert die Vorgabe der zu reflektierenden Aspekte die Komplexität der Reflexionssituation für die Teilnehmenden, da die Reflexion unabhängig vom Noticing erfolgen kann.[140] Damit werden also sowohl die Authentizität der Handlungssituation als auch die Komplexität der Reflexionssituation reduziert, was als Verschiebung des Testtyps von einem Performanz- zu einem Kompetenztest nach Miller (1990) bewertet werden kann. Vergleichbare Verschiebungen stellen auch Kulgemeyer et al. (2023) für Testinstrumente fest, die analog - z. B. um zeitnahes Feedback zu ermöglichen - stärker standardisiert wurden.

Zusammenfassend kann zum Indikator 1.1, der auf die adäquate Auswahl von Szenen und Aufgaben zur Abbildung von Unterrichtsreflexion abzielt, festgehalten werden, dass Physikunterricht im Rahmen der Videovignetten zwar beschränkt auf einen Fachinhalt, aber dennoch in gewisser Breite bezüglich der Unterrichtsphasen, eingesetzter Methoden, Beispiele etc. abgebildet wird. Die inhaltliche Auswahl der zu reflektierenden Aspekte spricht insgesamt alle drei Wissensbasen (fachliches, fachdidaktisches und pädagogisches Wissen) des Reflexionsmodells an, allerdings mit einem Fokus auf für den Physikunterricht spezifische Inhalte, weshalb Antwortoptionen zum pädagogischen Wissen nur in geringerem Anteil im Instrument vorkommen (erste Dimension des Reflexionsmodells). Die zweite Dimension des

[140] Da das Noticing als der Reflexion vorausgehend verstanden wird (s. Kapitel 7.6; empirische Hinweise darauf werden in Teil IV diskutiert), ist dies nicht als unmittelbare Einschränkung für die Erfassung der Reflexionsfähigkeit zu verstehen.

Reflexionsmodells, die Elemente der Reflexion, werden in komprimierter Form abgebildet. Bearbeitet werden jeweils aufeinander folgende MC-Aufgaben zu Bewertungen und Alternativen, wobei letztere teilweise Konsequenzen inkludieren, die als längerfristige Handlungsoptionen verstanden werden, die einen Beitrag zur Professionalisierung leisten bzw. die Unterrichtsqualität sicherstellen. Dass Konsequenzen nicht systematisch in alle Aufgaben zu Alternativen eingebunden sind, entspricht inhaltlich den realen Reflexionen Studierender, die die Grundlage für die Entwicklung der Antwortoptionen darstellen. Beschreibungen des Unterrichts werden nicht explizit eingefordert, beschreibende Elemente sind allerdings in den Antwortoptionen der Aufgaben zu Bewertungen enthalten. Insgesamt kann so also davon ausgegangen werden, dass die Reflexionselemente in einem angemessenen Verhältnis zueinander im Reflexionstest abgebildet sind. Die dritte Dimension des Reflexionsmodells, das Vorliegen von Begründungen, wird im Reflexionstest nicht differenziert berücksichtigt. Da aber alle Antwortoptionen begründet werden, wird diese Dimension im Sinne einer vollständigen Reflexion abgebildet. Die eingebundenen Noticing-Aufgaben stellen zwar nicht unmittelbar Teile von Unterrichtsreflexion dar, tragen aber zur Authentizität der Reflexionssituation und damit inhaltlichen Validität bei.

15.2 Evaluation der Bewertung der Antwortoptionen (Indikator 1.2)

Um den Bearbeitungen der Teilnehmenden fundiert Testwerte zuweisen zu können, müssen die Antwortoptionen im Testinstrument als korrekt oder inkorrekt definiert werden. Zur Beurteilung der Angemessenheit dieser Zuweisung wird Indikator 1.2 *„Die als korrekt definierten Antwortoptionen repräsentieren die zentralen Erkenntnisse der Unterrichtsreflexion zu den dargestellten Szenen.“* diskutiert.

Für handlungsnahe Fähigkeiten, wie die hier untersuchte Fähigkeit zur Unterrichtsreflexion, ist die Klassifizierung der Antwortoptionen in richtige und falsche aufgrund der vielfältigen Möglichkeiten der Unterrichtsgestaltung und Perspektiven zur Bewertung nicht trivial (vgl. Kaiser et al., 2015) und erfolgt unter Einbezug des Kontexts (Dollny, 2011, S. 44f.). Grundlage für die im Reflexionstest vorgenommene Klassifikation der Antwortoptionen in korrekte oder inkorrekte ist das Kodiermanual zum Reflexionsperformanztest von Kempin, Kulgemeyer und Schecker (2018). Eine Antwortoption im Reflexionstest wird als korrekt eingestuft, sofern sie im Reflexionsperformanztest kodiert wird, d. h. eine relevante Reflexion dargestellt. Die im Kodiermanual berücksichtigten Aspekte wurden von einer Expert*innenrunde aus neun Wissenschaftler*innen mit einem Forschungsschwerpunkt in der physikdidaktischen Professionsforschung im Rahmen des Projekts *ProfiLeP+* in mehreren Schleifen kommunikativ validiert (Kempin, Kulgemeyer

& Schecker, 2019). Es ist daher anzunehmen, dass sie die inhaltlich relevanten Rückmeldungen zum in den Videovignetten dargestellten Unterricht abbilden. Als Limitation dieser Vorgehensweise ist stets die begrenzte Anzahl und konkrete Auswahl an Expert*innen zu verstehen, da unklar ist, ob beispielsweise erfahrene Lehrkräfte die gleichen Einschätzungen zu den dargestellten Situationen treffen würden.

15.3 Beurteilung des Teils 1 des INA: Übersetzung Unterrichtsreflexion in Testinstrument

Die Prüfung der Indikatoren zur Anforderung an den ersten Teil des INA stellt sicher, dass die Fähigkeit zur Unterrichtsreflexion im Testinstrument angemessen abgebildet ist. Bei der Abbildung des Konstrukts der Unterrichtsreflexion können Einschränkungen der Repräsentativität der gewählten Unterrichtsausschnitte hinsichtlich der konkret dargestellten Situationen (eingebundene Methoden, fachliche Ungenauigkeiten etc.) sowie der dargestellten Rahmenbedingungen (Jahrgangsstufe, Klassengröße etc.) nicht vermieden werden. Bei der Auswahl der Szenen im Testinstrument wurde angestrebt, trotz der Einschränkung auf einen Fachinhalt möglichst vielfältige (problembehaftete) Situationen des Physikunterrichts abzubilden (der Fokus liegt auf fachlichen und fachdidaktischen Aspekten) und Rahmenbedingungen zu wählen, die Relevanz für eine breite Zielgruppe aufweisen. Die Auswahl der Aufgaben und Antwortoptionen im Testinstrument basiert auf realen Reflexionen von Studierenden und ist theoretisch am Reflexionsmodell von Nowak, Kempin, Kulgemeyer und Borowski (2019) orientiert, wodurch Unterrichtsreflexion sowohl umfassend und präzise als auch authentisch abgebildet wird. Die Handlungssituation selbst weist aufgrund der erhöhten Standardisierung durch MC-Aufgaben gegenüber einem offenen verbalen Antwortformat und aufgrund der Vorgabe der Aspekte zur Reflexion eine eher geringe Authentizität und Komplexität auf, sodass der Reflexionstest als Kompetenztest im Sinne Millers (1990) einzustufen ist. Limitierend ist außerdem festzuhalten, dass - entsprechend den realen Reflexionen der Studierenden - im Testinstrument Bewertungen und Alternativen fokussiert werden. Die Teilfähigkeit, Konsequenzen aus dem beobachteten Unterricht zu ziehen, die im Reflexionsmodell als selbstständiges Element einer Reflexion aufgeführt ist, ist nur eingeschränkt und als Teil der Aufgaben zu Alternativen abgebildet. Insgesamt kann also davon ausgegangen werden, dass im Reflexionstest Unterrichtsreflexion für einen bedeutsamen Fachinhalt der Mechanik, insbesondere vor dem Hintergrund realer studentischer Reflexionen, gelungen abgebildet ist (Indikator 1.1).

Die Einordnung bzw. Bewertung der Antwortoptionen in korrekte und inkorrekte als Grundlage für die Berechnung der Testwerte basiert auf dem Kodiermanual

zum zugrundeliegenden Reflexionsperformanztest, welches wiederum in einer Expert*innenrunde kommunikativ validiert wurde. Insofern kann davon ausgegangen werden, dass die korrekten Antwortoptionen die inhaltlich relevanten Erkenntnisse der Unterrichtsreflexion zu den dargestellten Szenen reflektieren (Indikator 1.2).

Die grundsätzliche Bestätigung der beiden Indikatoren stützt die inhaltlich valide Interpretation der Testwerte: Der Reflexionstest bildet die Fähigkeit zur Unterrichtsreflexion an repräsentativen Inhalten des Physikunterrichts ab. Einschränkend ist neben der reduzierten Handlungsfreiheit festzuhalten, dass eine inhaltliche Fokussierung notwendig ist, die in diesem Fall zu einer Beschränkung auf einen Fachinhalt des Themengebiets der Mechanik führte.

16 Untersuchung der Anforderung an Teil 2 des INA: Die Proband*innen zeigen bei der Bearbeitung des Reflexionstests konstruktrelevantes Verhalten

Die Untersuchung dieses Teils des INA fokussiert das Verhalten der Teilnehmenden während der Bearbeitung des Reflexionstests (eine Übersicht über die Anforderung an diesen Teil des INA und die zugehörigen Indikatoren und durchgeführten Untersuchungen ist in Abbildung 14.1 auf Seite 186 zum zweiten Übersetzungsschritt dargestellt). Dazu werden in Kapitel 16.1 zunächst die kognitiven Prozesse der Teilnehmenden mit einer Think-Aloud-Interviewstudie erhoben und bezüglich ihrer Passung zum Konstrukt der Unterrichtsreflexion ausgewertet. Anschließend wird die Akzeptanz des Reflexionstests durch die Teilnehmenden betrachtet (Kapitel 16.2), da davon auszugehen ist, dass diese eine Voraussetzung für die ernsthafte Bearbeitung des Testinstruments ist. Auf dieser Grundlage wird abschließend in Kapitel 16.3 die dem zweiten Teil des INA inhärente Anforderung eingeschätzt.

16.1 Evaluation der kognitiven Validität des Reflexionstests (Indikator 2.1)

Ziel der Untersuchung der kognitiven Prozesse während der Testbearbeitung ist eine Evaluation des Indikators 2.1 *„Die Überlegungen der Teilnehmenden (a) beziehen sich auf die Qualität des Unterrichts bzw. die Professionalisierung und (b) passen zur Aufgabenstellung.“* Um einen Einblick in die Überlegungen der Teilnehmenden während der Testbearbeitung zu erhalten, wurden $N = 7$ Physik-Lehramtsstudierende in einer Interviewstudie zum lauten Denken während der Testbearbeitung aufgefordert (*Think-Aloud-Interviews*). Die Interviews wurden transkribiert und inhaltsanalytisch ausgewertet.

16.1.1 Die Methode der Think-Aloud-Interviews

Think-Aloud (auch *lautes Denken*) beschreibt im Allgemeinen das Verbalisieren von Gedanken während der Bearbeitung einer (beliebigen) Aufgabe mit dem Ziel, einen Einblick in „Denkvorgänge und Prozesse bei der Bearbeitung“ (Konrad, 2010, S. 373) dieser Aufgabe zu erhalten. Grundlage der Methode ist das Drei-Speicher-Modell menschlicher Informationsverarbeitung (Konrad, 2010): Von den Sinnesorganen aufgenommene Informationen werden zunächst im sensorischen Register „modalitätsspezifisch gespeichert“ (Konrad, 2010, S. 377); ein geringer Anteil der Informationen gelangt in das Kurzzeitgedächtnis und von da aus potenziell (z. B. durch ausreichende Wiederholung) in das Langzeitgedächtnis (Konrad, 2010). Informationen im Kurzzeitgedächtnis sind direkt zugänglich und verbalisierbar, während Informationen im Langzeitgedächtnis dazu zunächst ins Kurzzeitgedächtnis transferiert werden müssen (Ericsson & Simon, 1993, S. 11).

Ericsson und Simon (1993, S. 16ff.) unterscheiden zwischen drei verschiedenen Leveln der Verbalisierung:

1. Level 1 („Talk aloud“ (Ericsson & Simon, 1993, S. 17), auch „Introspektion“ (Konrad, 2010, S. 376)): Die verbalisierten Informationen liegen bereits verbal encodiert vor und können direkt (ohne zusätzlichen Übersetzungsschritt) laut ausgesprochen werden.
   Beispielaufforderung: „Sprechen Sie alles laut aus, was Sie während der Bearbeitung des Testinstruments leise zu sich selbst sagen.“ (angelehnt an Dickmann (2016, S. 85))
2. Level 2 („Think aloud“ (Ericsson & Simon, 1993, S. 17), auch „unmittelbare Retrospektion“ (Konrad, 2010, S. 376)): Die zu verbalisierenden Informationen liegen nicht in verbal encodierter Form vor, sondern müssen vor der Verbalisierung in diese übersetzt werden.
   Beispielaufforderung: „Sprechen Sie alles laut aus, was Ihnen während der Bearbeitung des Testinstruments durch den Kopf geht, ohne die Überlegungen zu erklären oder zu strukturieren.“ (angelehnt an Dickmann (2016, S. 85))
3. Level 3 („reflection prompts“ oder „verzögerte Retrospektion“ (Konrad, 2010, S. 376ff.)): Es sollen lediglich bestimmte Informationen verbalisiert werden (z. B. Erklärungen für ein bestimmtes Handeln). Die vorliegenden Informationen werden also durch die Teilnehmenden gefiltert bzw. sie müssen über diese reflektieren.
   Beispielaufforderung: „Sprechen Sie nur das laut aus, was Ihnen durch den Kopf geht, wenn Sie ihre angewählten Antwortoptionen abändern.“ (angelehnt an Dickmann (2016, S. 85))

Ergänzend dazu unterscheiden Ericsson und Simon (1993) zwischen dem gleichzeitigen Think-Aloud („concurrent verbal reports“ (Ericsson & Simon, 1993, S. 16) entsprechend Level 1 und 2) und dem nachträglichen Think-Aloud („retrospec-

tive verbal reports“ (Ericsson & Simon, 1993, S. 16), auch „stimulated recall“ (Sandmann, 2014, S. 181), entsprechend Level 3). Weiterhin können Einsätze von Think-Aloud-Interviews in Bezug auf ihre Strukturiertheit, den Ort der Erhebung (Einsatz im *Labor* oder *Feld*) oder einen Medieneinsatz unterschieden werden (Buber, 2009; Konrad, 2010).

Neben dem Einsatz z. B. in der Problemlöseforschung oder in Usability-Studien (Konrad, 2010) kann die Methode der Think-Aloud-Interviews im Rahmen der Entwicklung von Testinstrumenten „zur Prüfung der Aufgabengüte und Item-Validierung“ (Sandmann, 2014, S. 182) zum Einsatz kommen: Proband*innen führen das Testinstrument durch und werden dabei (bzw. im Anschluss) darum gebeten, ihre Gedankengänge zu verbalisieren. Eine Auswertung der Think-Aloud-Protokolle kann dann Aufschluss über „missverständliche Formulierungen und unerwünschte Verstehensprobleme [sowie] Unklarheiten in der Instruktion“ (Sandmann, 2014, S. 182) liefern.

Zur Auswertung wird das aufgenommene Think-Aloud-Interview üblicherweise transkribiert (*Think-Aloud-Protokolle*) und anschließend kodiert, wobei jeweils passend zum Forschungsinteresse Transkriptionsregeln bestimmt und ein Auswertungsmanual bzw. Analysekategorien entwickelt werden (Noushad, van Gerven und de Bruin, 2024; Sandmann, 2014).

Für die Durchführung von Think-Aloud-Interviews werden folgende zentrale Hinweise gegeben:

- Zu Beginn der Erhebung sollte eine Trainings- bzw. Beispielaufgabe gestellt werden, damit die Teilnehmenden das Think-Aloud vorab erlernen bzw. sich an die Situation gewöhnen können (Bilandzic, 2017; Sandmann, 2014). Mögliche Trainingsaufgaben stellen z. B. die Multiplikation von (mehrstelligen) Zahlen oder das Finden von Reimwörtern dar (Ericsson & Simon, 1993, S. 377).
- Als Instruktionen zum Think-Aloud sollte eine unkomplizierte, direkte Aufforderung wie „Bitte denken Sie laut“ (Bilandzic, 2017, S. 407) eingesetzt werden, die um eine Erläuterung ergänzt werden kann, wie: „Stelle dir vor, du bist allein im Raum und führst ein Selbstgespräch. Es geht nicht darum, deine Gedanken zu erklären.“ (angelehnt an Bilandzic (2017, S. 409)) oder „Wichtig ist, dass du möglichst immer redest.“ (Sandmann, 2014, S. 184). Eine inhaltliche Beeinflussung der Teilnehmenden durch die Formulierung der Aufforderung ist zu vermeiden (Bilandzic, 2017).
- Falls Teilnehmende aufhören, ihre Gedanken zu verbalisieren, sollten sie nach einer gewissen Zeit an das laute Denken erinnert werden (Ericsson & Simon, 1993, S. 376).

Während Ericsson und Simon (1993, S. 16) annehmen, dass - im Gegensatz zum nachträglichen Think-Aloud - zumindest das gleichzeitige Think-Aloud keinen Ein-

fluss auf die individuellen kognitiven Prozesse während der Aufgabenbearbeitung hat und die Verbalisierungen demnach Aufschluss über die realen Überlegungen liefern, sieht Bilandzic (2017) in der Übersetzung der Gedanken in gesprochene Sprache einen potenziellen Einfluss auf die kognitiven Prozesse während der Bearbeitung und somit eine mögliche Gefährdung der Validität von Schlussfolgerungen, die auf Grundlage von Think-Aloud-Protokollen gezogen werden. Weitere Gefährdungen stellen die etwaige Selektion der laut ausgesprochenen Gedanken durch die Teilnehmenden und daraus folgenden möglicherweise unvollständigen Think-Aloud-Protokolle oder Effekte, wie soziale Erwünschtheit, die bei der Durchführung von Interviews stets zu berücksichtigen sind, dar (Bilandzic, 2017; Sandmann, 2014). Darüber hinaus ist bei retrospektivem Think-Aloud zu beachten, dass Gedanken z. B. vergessen oder verwechselt werden können oder interviewte Personen ihre Überlegungen im Nachhinein rationalisieren oder rechtfertigen (Bilandzic, 2017).

16.1.2 Umsetzung der Think-Aloud-Interviews

Für die Untersuchung der kognitiven Prozesse von Teilnehmenden während der Bearbeitung des entwickelten Reflexionstests wurden im Wintersemester 2021/22 $N = 7$ Think-Aloud-Interviews mit Physik-Lehramtsstudierenden im Bachelor- ($N = 4$) und Masterstudium ($N = 3$) der Universität Paderborn via Videotelefonie durchgeführt, in denen sie dazu aufgefordert wurden, ihre Gedanken während der Bearbeitung des Testinstruments laut auszusprechen (Level 2 nach Ericsson und Simon (1993, S. 18)). Da an dieser Stelle neben etwaigen Verständnisschwierigkeiten in Formulierungen des Testinstruments insbesondere die Gedanken während der Testbearbeitung von Interesse sind, wurde eine Form des gleichzeitigen Think-Aloud gewählt (vgl. Dickmann, 2016, S. 86).

Die Verbalisierungen wurden audiografiert. Um die permanente Verbalisierung der Gedanken sicherzustellen, war die Interviewleitung durchgehend anwesend. Parallel dazu wurden die Bildschirminteraktionen der Teilnehmenden aufgezeichnet, um die Zuordenbarkeit von Aussagen zu den Elementen des Tests sicherstellen zu können. Im Anschluss an die Testbearbeitung wurde ein kurzes Stimulated-Recall-Interview eingeplant, welches Nachfragen z. B. zu unklaren Aussagen der Studierenden durch die Interviewleitung ermöglichte.

Da die Interviews im Rahmen von Videotelefonaten durchgeführt wurden, wurden die Teilnehmenden vorab gebeten, ein möglichst ruhiges und ungestörtes Setting herzustellen. Zusätzlich wurde ihnen die Möglichkeit gegeben, die Kamera auszuschalten, was zur Wahrung der Anonymität beitragen und die Teilnehmenden beim lauten Denken unterstützen kann (Noushad, van Gerven & de Bruin, 2024).

16.1.3 Aufbereitung der Think-Aloud-Interviews

Die insgesamt $N = 7$ Think-Aloud-Interviews weisen eine Länge von 76 bis 145 Minuten (durchschnittlich 114 Minuten) auf. Die Aufnahmen wurden für die Auswertung vollständig wörtlich transkribiert, die an den Interviews teilnehmenden Studierenden mit LD1 bis LD7 benannt. Der Leitfaden zur Transkription befindet sich im Anhang in Kapitel G.1 (ab Seite 463), die vollständigen, segmentierten und kodierten Transkripte befinden sich im Anhang in Kapitel O (ab Seite 531).

Die Transkripte (Think-Aloud-Protokolle) wurden zunächst eventbasiert segmentiert. Der Vorteil der eventbasierten Segmentierung gegenüber der Orientierung an Zeitintervallen liegt in der Übereinstimmung von Segmenten und Sinneinheiten (Brückmann & Duit, 2014). In der Segmentierung der vorliegenden Interviews waren folgende Events ausschlaggebend für den Abschluss bzw. Beginn eines neuen Segments:

- Sprecher*innenwechsel (ausgenommen sind lediglich kurze Einwürfe, wie „Mhm“, die den Redefluss der anderen Person nicht unterbrechen und somit keinen inhaltlichen Einfluss nehmen; eine Teilung der Aussage in zwei Segmente würde deren Nachvollziehbarkeit mindern und wäre insofern nicht zielführend; vgl. Kuckartz (2018, S. 43)),
- das Vorlesen eines neuen Elements aus dem Testinstrument (z. B. eine Antwortoption),
- die Bezugnahme auf eine neue Antwortoption im Testinstrument, auch wenn hier ggf. inhaltliche Verknüpfungen bestehen (z. B. bezieht sich die Aussage „Die Auswertung des Experiments, dazu kann man jetzt irgendwie nichts sagen“ (LD7, Z. 135) auf die Antwortoption „Auswertung des Experiments“, die direkt anschließende Aussage „und Messunsicherheiten hat er zum Beispiel auch nicht thematisiert“ (LD7, Z. 136) auf die Antwortoption „Messunsicherheiten“) sowie
- die Äußerung einer inhaltlich neuen und eigenständigen Überlegung (z. B. wenn eine eigene Erfahrung beschrieben wird, nicht aber eine Begründung für die Auswahl einer Antwortoption)

Falls inhaltliche Bezüge zu Antwortoptionen im Testinstrument unklar waren, wurde zusätzlich die Bildschirmaufnahme zu Rate gezogen. Die Bildschirmaufnahme bildet so zusammen mit dem Interviewtranskript die Kontexteinheit, die zum Verständnis der Segmente herangezogen werden kann (vgl. Kuckartz, 2018, S. 44). Das vollständige Manual zur Auswertung (zur Segmentierung sowie der nachfolgend beschriebenen Kodierung) der Think-Aloud-Protokolle befindet sich im Anhang in Kapitel G.

16.1.4 Kodierung der Think-Aloud-Interviews

In einem ersten Durchgang wurde das Material vollständig gelesen und es wurden alle Segmente markiert, die Hinweise auf Hürden im Verständnis des Testinstruments beinhalten. Dazu gehören neben expliziten Äußerungen von Unverständnis (z. B. „Das Satz, der ist schwierig, vielleicht verstehe ich ihn auch gerade einfach nicht."; LD2, Z. 72) auch implizite Hinweise wie das wiederholte (Vor-)Lesen von Antwortoptionen. Dabei wurden auch solche Stellen kodiert, die zwar auf ein korrektes Verständnis der Teststelle verweisen, bei denen aber z. B. durch formulierte Fragen eine gewisse Unsicherheit deutlich wird (z. B. „Du solltest hinterfragen, woher die geschlechterspezifischen Zuschreibungen stammen, die du vertrittst, um sie zu überwinden. (..) Ah da geht es um Mädels und sowas, ne?"; LD2, Z. 260). Auch wenn solche Fragen nicht automatisch ein anderes als das intendierte Verständnis der Teststelle implizieren, würde das wiederholte Kodieren derselben Teststelle in den verschiedenen Interviews auf eine ungünstige Formulierung oder fehlende Informationen hindeuten und somit einen Bedarf zur Überarbeitung des Testinstruments kennzeichnen.

In einem zweiten Durchgang wurden die Aussagen in den Segmenten vollständig kodiert. Dazu wurde im Sinne der qualitativen Inhaltsanalyse nach Kuckartz (2018, S. 44ff., 88ff.) induktiv am Material ein Kodiermanual entwickelt. Insgesamt wurden die acht in Tabelle 16.1 aufgeführten Kategorien unterschieden. Verbunden mit dem Ziel, die kognitiven Prozesse während der Bearbeitung des Reflexionstests zu analysieren, wurden dabei vorwiegend Kategorien gebildet, die die Art der Überlegung abbilden (z. B. Bewertungen von Teilen des Unterrichts oder Schilderungen eigener Erfahrungen) und nicht die angesprochenen Inhalte clustern. Die Sinnhaftigkeit der Aussagen (z. B. inwiefern ein vorgeschlagenes alternatives Vorgehen tatsächlich eine Verbesserung des Unterrichts bewirkt) wird somit in der Auswertung nicht berücksichtigt.

Kategorie	Erläuterung	Beispiel
Vorlesen	Text wird ggf. inkl. Füllwörtern wie „äh" oder Verhaspeln vorgelesen.	*Vorlesen mit Verhaspeln:* „Messunsicherheiten, Umgang mit der- mit dem Einwand von Marie. [...]" (LD6, Z. 271)
Beschreibung	Es wird Unterricht (Lehrer- oder Schüler*innen handeln) oder der Unterrichtsablauf beschrieben.	*Beschreibung des Unterrichtsablaufs:* „Okay. (..) Grober Verlauf also: Quiz, dieser kleine Comic, (...) und zwei kleine- zwei Experimente? Ein Experiment. Ja." (LD2, Z. 8)

Bewertung von Unterricht (mit-/ohne Begründung)	Es wird Lehrerhandeln bzw. die Unterrichtsgestaltung (z. B. die Fülle an Unterrichtsinhalten) als gut/nicht gut/... bewertet.	*Bewertung:* „[...] ich finde, Beispiele sind da ganz gut.“ (LD5, Z. 266) *Bewertung mit Begründung:* „Das Experiment fand ich irgendwie auch nicht so (.) super gewählt, weil es dann irgendwie was ganz anderes war, als mit der Titanic, das hat dann auch sehr voneinander abgewichen.“ (LD7, Z. 97)
Alternativen zum Vorgehen (mit/ohne Begründung)	Es werden alternative Vorgehensweisen zum Lehrerhandeln bzw. zu Unterrichtsinhalten oder dem Unterrichtsablauf formuliert.	*Alternative mit 2 Begründungen:* „[...] jedoch hätte ich es, glaube ich, besser gefunden, wenn er von Anfang auch sagt, gut, je schneller der Ball ist, das bedeutet, er hat eine höhere Geschwindigkeit, schwerer bedeutet, dass die Masse größer ist, und nicht so sehr noch in der Alltagssprache geblieben wäre. (.)“ (LD3, Z. 101) *Alternative (in Kombination mit der Überlegung, wieso die Lehrkraft wie beobachtet vorgeht)*: „Wieso stellt er das jetzt alleine vor und nicht mit dem Schüler?“ (LD5, Z. 10)

Zustimmung bzw. Ablehnung einer Antwortoption (mit/ohne Begründung)	Vorgegebenen Antwortoptionen wird zugestimmt bzw. diese werden abgelehnt. Das kann sowohl durch kurze Bemerkungen wie „Ja“ oder „stimmt nicht“ als auch durch elaborierte Aussagen geschehen.	*Zustimmung mit Begründung:* „[...] es könnten Rundungsfehler sein bei den Schülern, es könnten aber auch eben Mess- Messunsicherheiten sein, die sollte er erwähnen, deswegen würde ich das ankreuzen.“ (LD5, Z. 226) *Ablehnung einer Antwortoption (trotz Signalwort „Ja“ für Zustimmung):* „[...] Ja, das kann nicht sein, ansonsten wären sie sich ja alle einig. (..)“ (LD1, Z. 81)
Eigene Erfahrung	Es werden eigene Erfahrungen aus der Perspektive als Lernende*r oder Lehrende*r geschildert bzw. es wird auf andere Erfahrungen Bezug genommen (z. B. geführte Gespräche).	„[...] Wenn ich mir vorstelle, ich hätte jetzt so ein Experiment in der Uni gemacht, dann hätte ich ja auch beide Wagen messen müssen.“ (LD7, Z. 166)
Weitere Überlegungen zum Unterricht bzw. den Inhalten	Es werden sonstige Überlegungen (ggf. offene Fragen) zum beobachteten Unterricht geäußert, die sich z. B. auf das Vorwissen der Schüler*innen beziehen oder Vermutungen der handelnden Personen beinhalten können.	*Schülergemäßheit:* „[...] Ähm die Formulierung fand ich an einigen Stellen, [...] aber weiß ich nicht, wie sehr Schülerinnen und Schüler das verstehen würden.“ (LD2, Z. 29) *Vermutung:* „[...] ein Beispiel kommt wahrscheinlich auch noch später.“ (LD5, Z. 133)
Eigene Professionalisierung	Es werden Einschätzungen zur eigenen Professionalisierung vorgenommen (Einordnung des eigenen Wissens, Verhaltens o. ä. in Bezug auf den Lehrberuf).	„[...] da müsste ich nochmal nachgucken, werde ich aber höchstwahrscheinlich nicht machen, weil ich heute noch genügend zu tun habe.“ (LD1, Z. 86)

Sonstiges	Aussagen, die keiner der bisherigen Kategorien zugeordnet werden können (u. a. Aussagen zur Methode des Think-Aloud oder Aussagen zur Gestaltung und zu (potenziellen) Überarbeitungsbedarfen des Testinstruments die in der ersten Durchsicht der segmentierten Transkripte identifiziert werden)	*Methode:* „Die Aufgaben muss ich jetzt aber nicht laut vorlesen, die also das zählt dann nicht zum lauten Denken, ne?“ (LD7, Z. 4) *Testinstrument:* „[...] Also dieser Test Tex- Limesurvey oder wie auch immer das nennt ist definitiv zu lang für meinen Geschmack.“(LD1, Z. 310)

Tabelle 16.1: Übersicht über die Kategorien zur Kodierung der Think-Aloud-Interviews

Kodiert wurden alle in einem Segment auftretenden Aussagen. Üblicherweise wird einem Segment genau eine Kodierung zugewiesen. Mehrfachkodierungen sind dann möglich, wenn verschiedene Aussagetypen innerhalb eines Segments vorkommen, das Segment aber nicht inhaltlich nachvollziehbar in mehrere Segmente aufgeteilt werden kann, da die Aussagen stark verwoben sind (z. B. beinhaltet die Aussage „auf jeden Fall zum Umgang mit den Schüler- innen, weil er die irgendwie so komisch getrennt, die Mädchen haben eine schöne Handschrift, dann dürfen die das aufschreiben und Jungs können besser ablesen, dann dürfen die das machen, (.) das fand ich nicht gut.“ (LD7, Z. 133) neben der Zustimmung zur Antwortoption mit Begründung auch die Bewertung „das fand ich nicht gut“).

16.1.5 Ergebnisse

Nachfolgend werden die Ergebnisse der Auswertung der Think-Aloud-Interviews dargestellt und diskutiert. Dabei werden zunächst die Segmente betrachtet, die in der ersten Durchsicht der segmentierten Transkripte potenzielle Hürden im Verständnis des Testinstruments bzw. einzelner Antwortoptionen erkennen lassen und potenzielle Änderungsbedarfe kennzeichnen. Anschließend werden die vollständig kodierten Transkripte bezüglich der Art der auftretenden Überlegungen charakterisiert und interpretiert, um abschließend die Bewertung der kognitiven Validität des Testinstruments zu ermöglichen.

Identifikation von Hürden bei der Testbearbeitung

In der ersten Durchsicht der segmentierten Think-Aloud-Protokolle wurden solche Segmente identifiziert (und als *Sonstige** kodiert), die auf mögliche Hürden beim Verständnis des Testinstruments oder aufgetretene Missverständnisse hindeuten. Die über alle sieben Interviews aufgetretenen insgesamt 48 Kodierungen verweisen auf 44 unterschiedliche Stellen (z. B. Antwortoptionen) im Testinstrument. Eine nähere Analyse der Protokolle zeigt auf, dass die als mögliche Hürden markierten Stellen überwiegend individuell sind. Ein Fehler in der Orthographie, auf den zwei Studierende verwiesen (LD1, Z. 177; LD4, Z. 77), konnte unmittelbar korrigiert werden.

Im Reflexionstest ist vorgegeben, welche Aspekte reflektiert werden (d. h. zu welchen Aspekten Aufgaben zu Bewertungen und Alternativen bearbeitet werden). Um dies zu verdeutlichen, wird im Testinstrument die Reflexion jeweils mit einem Zwischenprompt eingeleitet: „Robert hat schon einige Rückmeldungen erhalten. Besonders wichtig ist ihm jetzt noch eine Rückmeldung zu ...“. Im Think-Aloud-Protokoll von LD1 wird an sechs Stellen deutlich, dass nicht verstanden wurde, dass die zu reflektierenden Aspekte vorgegeben sind und nicht der individuellen Auswahl der bedeutsamen Aspekte (im Rahmen der vorangestellten Aufgabe zum Noticing) entsprechen (z. B. „das überspringe ich jetzt einfach, das habe ich nicht ausgewählt, ich habe keine Ahnung, warum ich hier bin“; LD1, Z. 451). Erst zum Ende der Bearbeitung des Testinstruments stellt LD1 sich die Frage, „[...] oder sollte das so?“ (LD1, Z. 453). Da LD1 offenbar während der Bearbeitung der ersten beiden Videovignetten auch überlesen hat, dass im Reflexionstest positives und negatives Feedback gegeben werden kann (dies stellt LD1 während der Bearbeitung der dritten Vignette fest: „Oh kann hier auch Lob äh geben, das hab ich gar nicht gewusst“ (LD1, Z. 118); in der vorigen Vignette wurde die Fragestellung nicht vollständig gelesen: „Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung abababababa.“ (LD1, Z. 74)), könnte vermutet werden, dass auch die Zwischenprompts zur Auswahl der reflektierten Aspekte nicht sorgfältig gelesen wurden. Da keine andere Person hier Schwierigkeiten oder Unklarheiten äußert, ist nicht von einem Änderungsbedarf auszugehen, zumal unklar ist, inwiefern nachträglich vorgenommene Änderungen (z. B. Hervorhebungen der Möglichkeit, positives Feedback zu geben) eine Beeinflussung des Fokus während der Testbearbeitung zur Folge hätten.

Zusätzlich zeigt sich in den Think-Aloud-Transkripten von LD1 und LD2, dass acht Antwortoptionen beim ersten Lesen von jeweils einer Person nicht oder nicht wie intendiert verstanden wurden. Darunter gibt es einige Fälle, in denen die Unklarheit anschließend durch ein erneutes Lesen der Option aufgelöst wurde (z. B. „Ich glaube, hier ist eine Aussage in der Kategorie verwechselt. [...] Ach, es geht um die Verwendung der Einheit, ah jetzt verstehe ich, was der Satz soll“ (LD1, Z. 342f.) oder „Ach Rundungsfehler der Kinder [...]. Ich habe die ganze Zeit

Rundungsfehler und Messunsicherheiten durcheinandergebracht.“ (LD2, Z. 267)). Für drei Optionen wurde zwar direkt die korrekte Interpretation der Antwortoption geäußert, allerdings eine gewisse Unsicherheit deutlich (z. B. „ich interpretiere das jetzt mal so, der Satz will mir sagen, dass es doof ist, dass er erst (.) etwas lobt, dass für das Geschlecht bezogen ist und dann für etwas kritisiert, was nicht auf das Geschlecht bezogen ist und somit die Schülerin verwirrt, ob jetzt ihre Handlungen auf ihr Geschlecht bezogen sind oder nicht. Angenommen, es funktioniert so, (.) sage ich nein.“ (LD21, Z. 277)). Weiterhin können verschiedene potenzielle Hürden ausgemacht werden, die offenbar die Testschwierigkeit für die Studierenden beeinflussen. Dazu gehören elf (wiederum verschiedene) Stellen in denen den Studierenden die Unterschiede zwischen verschiedenen Antwortoptionen entweder nicht erkennen oder als sehr gering einschätzen (z. B. „Ist das nicht das selbe, wie das davor?“ (LD1, Z. 85) oder „es geht aber beides eigentlich in ein über (unv.), da würde glaube ich ein Ankreuzaspekt reichen“ (LD5, Z. 177)) sowie zwölf Aussagen bzw. Fragen, die die inhaltliche Auseinandersetzung mit den Antwortoptionen aufzeigen und offenbar nicht selbst gelöst oder beantworten werden können (z. B. zur Antwortoption „Es ist nicht gut, dass du die Schüler*innen in der Benutzung des Begriffs ‚Schwung‘ bestärkst. So förderst du ihre Fehlvorstellungen (z. B. dass beide Begriffe gleichgesetzt werden können und dass Körper mehr Kraft haben, wenn sie mehr Schwung besitzen).“: „wo kommt denn hier jetzt noch die Kraft her? Hätte ich jetzt da gar nicht reingedeutet“ (LD2, Z. 9) oder zur Antwortoption „Es ist nicht gut, dass die Schüler*innen dir ohne Arbeitsauftrag zusehen, da die Schüler*innen so keinen Anreiz haben, aufmerksam zu sein.“: „Warte, da wurde kein Arbeitsauftrag gegeben? (.) Ich dachte, der wurde einfach nur aus Zeiteffizienz weggeschnitten.“ (LD1, Z. 315)). Da hier jeweils unterschiedliche Antwortoptionen angesprochen werden und die Bearbeitung des Testinstruments nicht durch falsche Interpretationen beeinflusst wird, besteht für diese Optionen kein Änderungsbedarf.

Problematisch ist eine für vier Studierende (LD1, LD2, LD5 und LD6) vorliegende Unklarheit im Testinstrument. Die unklare Antwortoption ist ein Distraktor zur Erkennung aufgetretener Fehlvorstellungen zu Vignette 2. Die ursprüngliche Formulierung „Die Schüler*innen sehen nicht, dass der Apfel über den Baum mit der Erde verbunden ist und der Baum durch das ‚Festhalten‘ an der Erde eine Gegenkraft zum am Baum ‚ziehenden‘ Apfel ausübt.“ wird umformuliert zu „Die Schüler*innen erkennen nicht, dass der Apfel mit seinem Gewicht am Baum zieht und der Baum sich an der Erde festhalten muss, um die Gegenkraft auszuüben.“ Die überarbeitete Formulierung wurde im Nachgang informell vier weiteren Studierenden vorgelegt und diese wurden um die Beschreibung ihres Verständnisses der Antwortoption bzw. die Benennung von Unklarheiten gebeten. Dabei wurden keine Unklarheiten deutlich, sodass diese Unzulänglichkeit als gelöst angesehen werden kann.

Näher zu betrachten sind außerdem die beiden nachfolgenden Unklarheiten. Sie scheinen die Testbearbeitung zwar nicht grundsätzlich zu stören, da sie jeweils nur einer Person aufgefallen sind, weisen allerdings auf eine mangelnde Passung zwischen einzelnen Antwortoptionen und dem Aufbau des Testinstruments bzw. der Darstellung des Unterrichts in den Videovignetten hin:

- Die Antwortoption „Es wäre besser, vor dem Einstieg in das neue Thema auch die Tafel zu wischen [...]“ (zu Videovignette 3) löst bei einer Person Irritation aus, da sie die beschriebene Tafel zuvor übersah (LD6, Z. 149). Dies ist hier zu berücksichtigen, da die beschriebene Tafel im Rahmen der Vignette lediglich zu Beginn der Vignette für etwa 55 Sekunden (und damit über die Hälfte der Vignette) zu sehen ist. In den letzten etwa 45 Sekunden der Vignette ist aufgrund eines Aufnahme- bzw. Schnittfehlers allerdings eine gewischte Tafel sichtbar. Die Antwortoption stellt kein zentrales Feedback zur Vignette und somit einen Distraktor für die Bearbeitung der Aufgabe dar, sodass Studierenden, die die Skizze an der Tafel übersehen bzw. von einer gewischten Tafel ausgehen und die Antwortoption daher nicht anwählen, kein Nachteil in der Bestimmung der Testwerte entsteht. Auf die Entfernung der Antwortoption wird zugunsten des ausgeglichenen Verhältnisses zwischen als korrekt oder inkorrekt bewerteten Antwortoptionen verzichtet.
- Zur Antwortoption „Es ist nicht gut, dass du die Masse des Wagens und insbesondere die Geschwindigkeit selbst bestimmst, da du den Schüler*innen vorher schon erklärt hast, wie sie das selbst machen könnten.“ (zu Videovignette 5; im Kontext der Auswertung eines Demonstrationsexperiments) merkt LD1 (Z. 316) richtigerweise an, dass die eingebundene Vignette diese Erklärung nicht enthält. Die Erklärung zur Bestimmung der Geschwindigkeiten ist im Performanztest in einem früheren Video dargestellt, dass nicht in den Reflexionstest übernommen wurde. Diese Information wird daher explizit in der Einleitung der Vignette im Testinstrument ergänzt.

Untersuchung der kognitiven Prozesse während der Testbearbeitung

Anschließend erfolgte die vollständige Kodierung der Think-Aloud-Protokolle für die Testbearbeitung (zuvor aufgenommene Inhalte wie z. B. das Einholen des Einverständnisses für die Aufnahme oder Erläuterungen zum Ablauf des Interviews vor Beginn der Testbearbeitung werden hier nicht berücksichtigt).[141] Das Vorlesen von Bestandteilen des Testinstruments ist als Artefakt der Interviewmethode einzuordnen und trägt zur Verständlichkeit der Think-Aloud-Protokolle bei, lässt aber keine unmittelbaren Rückschlüsse auf die kognitiven Prozesse

[141] Eine vorläufige und weniger detaillierte Auswertung zur Untersuchung der kognitiven Prozesse während der Testbearbeitung, die zu vergleichbaren Schlüssen führt, befindet sich in Weißbach und Kulgemeyer (2022a).

während der Testbearbeitung zu. Für die folgende Beurteilung des Indikators 2.1 werden entsprechende Kodierungen daher nicht berücksichtigt. Unter Auslassung dieser Kodierungen wurden in den Think-Aloud-Protokollen zwischen 276 und 531 (durchschnittlich $M = 381$; $SD = 102$) Kodierungen vorgenommen.[142] Durchschnittlich entfallen etwa die Hälfte ($M = 51$ % ; $SD = 12$ %) der Kodierungen auf konstruktrelevante Überlegungen, d. h. Reflexionen, die inhaltliche Auseinandersetzen mit Antwortoptionen im Testinstrument oder anderweitig auf den Unterricht bezogene Aussagen. Die übrigen Kodierungen sind der Kategorie *Sonstiges* zuzuordnen. Dies betrifft in sechs der sieben Interviews unter zwanzig Prozent der Kodierungen. Im siebten Interview (LD1) ist es mit 43 % und ein auffallend großer Anteil. Eine Übersicht über die Anteile der einzelnen Kategorien an den Kodierungen befindet sich in Tabelle 16.2.

Zur Prüfung der Beurteilerübereinstimmung wurden 347 Segmente (14, 5 % der insgesamt vorliegenden Segmente) durch eine zweite Person kodiert und als Maß der Übereinstimmung für die verschiedenen Kategorien κ_{Cohen} bestimmt.[143] Insgesamt ergaben sich Werte zwischen $\kappa_{\text{Cohen}} = 0,54$ für die Kategorie *Bewertung mit Begründung* und $\kappa_{\text{Cohen}} = 0,91$ für die Kategorie *Vorlesen* (für die übrigen Werte s. Tabelle 16.2). Die Übereinstimmung in den Kategorien *Beschreibung*, *Bewertung* und *Bewertung mit Begründung* ist mit $\kappa_{\text{Cohen}} < 0,60$ knapp als „mittelmäßig“ und noch nicht ausreichend zu bewerten (Döring & Bortz, 2016, S. 346). Für die übrigen Kategorien werden mindestens ausreichende ($\kappa_{\text{Cohen}} > 0,60$), teilweise sogar sehr gute Werte (von $\kappa_{\text{Cohen}} > 0,75$) für die Übereinstimmung erzielt (Döring & Bortz, 2016, S. 346).[144] Die beiden Kategorien *Erfahrung* und *Professionalisierung* kommen in den Interviews kaum, in den doppelkodierten Segmenten gar nicht vor. Um dennoch eine Qualitätsprüfung für die Kodierung dieser Kategorien durchzuführen, wurden die (Erst-)Kodierungen der zweiten kodierenden Person vorgelegt und die Angemessenheit der Kodierungen durch sie überprüft.

Zur Beurteilung des Indikators 2.1 bedarf es zunächst einer inhaltlichen Bewertung der unterschiedlichen Kategorien bezüglich ihrer Passung zur Unterrichtsreflexion bzw. zum Reflexionstest. Dazu sind die Kategorien zu identifizieren, die inhaltlich relevante Überlegungen zum Unterricht sowie zur eigenen Professionalisierung bzw. der Professionalisierung der beobachteten Lehrperson beinhalten (Indikator 2.1a).

[142] Im Rahmen der vollständigen Kodierung der Think-Aloud-Protokolle, die der Arbeit angehängt ist (s. Kapitel O), wurde auch das Vorlesen kodiert. In den verschiedenen Interviews entfallen insgesamt zwischen 23 % und 48 % ($M = 39$ % ; $SD = 7$ %) der Kodierungen auf das Vorlesen von Inhalten des Testinstruments (unter Berücksichtigung der Kodierungen zum Vorlesen wurden in den Think-Aloud-Protokollen zwischen 366 und 865 (durchschnittlich $M = 649$; $SD = 175$) Kodierungen vorgenommen).

[143] Das zugehörige *R*-Skript befindet sich im Anhang in Kapitel Q ab Seite 877. Die zweiten Kodierungen entfallen auf Segmente der Interviews mit LD4 und LD5 und sind ebenfalls gemeinsam mit den Transkripten angehängt (s. Kapitel O).

[144] Die Bestimmung der Beurteilerübereinstimmung für die Kategorie *konstruktrelevante Überlegungen insgesamt* basiert auf der Zusammenfassung der aufgeführten Kategorien ohne *Sonstiges*.

Mit Ausnahme der Kategorie *Sonstiges* weisen alle Kategorien inhaltlich Bezüge zum durchgeführten Unterricht, zu dessen Inhalten oder zur Professionalität der eigenen Person oder der Persona Robert, dessen Unterricht reflektiert wird, auf. Der Anteil der Kodierungen in der Kategorie *Sonstiges* ist insgesamt gering und liegt in sechs der sieben Interviews unter 20 % (im siebten (LD1) bei 43 %). Es kann also konstatiert werden, dass sich der Großteil der Äußerungen auf den Unterricht beziehungsweise die Professionalisierung als Lehrperson bezieht.

Entsprechend Indikator 2.1b *„Die Überlegungen der Teilnehmenden passen zur Aufgabenstellung.“* werden außerdem die vorliegenden konstruktrelevanten Kodierungen auf Ebene der Aufgaben zu Bewertungen und Alternativen beleuchtet, um zu untersuchen, inwiefern sich die Überlegungen abhängig von den konkreten Aufgabenstellungen unterscheiden. Insgesamt verweisen durchschnittlich $M = 50$ % ($SD = 11$ %) der Kodierungen unmittelbar auf geäußerte Zustimmung oder Ablehnung einer Antwortoption (teilweise mit Begründungen), sodass jeweils ein direkter Bezug zur Aufgabenstellung vorliegt. Für die übrigen konstruktrelevanten Kodierungen (Bewertungen des Unterrichts, Vorschläge für alternative Handlungsoptionen oder fachliche Überlegungen) ist davon auszugehen, dass sie durch die betrachtete Videovignette oder jeweils vorliegende Aufgabenstellung bzw. Antwortoption ausgelöst wurden und demnach einen inhaltlichen Bezug zur Aufgabe aufweisen.

Mit der Zweiteilung der Aufgaben in solche zur Bewertung und zum Vorschlagen alternativer Handlungsoptionen werden zwei im Sinne des Reflexionsmodells von Nowak, Kempin, Kulgemeyer und Borowski (2019) unterschiedliche Elemente der Reflexion unterschieden. Es könnte erwartet werden, dass auch die Teilnehmenden entsprechend dieser Reflexionselemente Überlegungen anstellen und passend dazu während der Bearbeitung von Aufgaben zu Bewertungen (vorwiegend) Bewertungen äußern und analog (überwiegend) Alternativen bei der Bearbeitung von Aufgaben zu Alternativen formulieren. Dies lässt sich in den Interviews allerdings nicht beobachten. Bei der Bearbeitung aller Aufgaben (zu Bewertungen und Alternativen) wurden Alternativen häufiger genannt als Bewertungen. Das Verhältnis von geäußerten Alternativen und Bewertungen variiert hier allerdings stark zwischen den einzelnen Studierenden, die Anzahl an Alternativen beträgt das $1,5$-fache bis hin zum 15-fachen (beides LD3) der Anzahl an Bewertungen (eine Person (LD4) äußert bei der Bearbeitung von Aufgaben zu Alternativen ausschließlich Alternativen und keine Bewertungen).[145] Es ist also davon auszugehen, dass die (konstruktrelevanten) Überlegungen der Teilnehmenden zwar inhaltliche Zusammenhänge zur jeweiligen Aufgaben aufweisen, die entlang der Reflexionselemente Bewertungen und Alternativen strukturierten Aufgaben aber nicht zu unterschiedlichen kognitiven Prozessen während der Testbearbeitung führen. Diese Beobachtung muss allerdings nicht als Gefährdung für die valide Testwertin-

[145] Eine Übersicht über die entsprechenden Anteile der Kategorien an den Kodierungen für beide Aufgabengruppen befindet sich im Anhang in Tabelle H.1 auf Seite 478.

terpretation bewertet werden. Bewertungen und Alternativen können inhaltlich zusammenhängen, insbesondere sind Alternativen im Sinne des Reflexionsmodells nach Nowak, Kempin, Kulgemeyer und Borowski (2019) als höherwertiges Reflexionselement einzustufen, die auf eine vorausgegangene (ggf. nicht geäußerte) Bewertung aufbauen. Die inhaltliche Passung der Überlegungen zum reflektierten Aspekt wird hierdurch also nicht berührt.

Im Rahmen der Kodierung wurde überdies unterschieden, ob zu den Bewertungen, den Alternativen, aber auch zur Zustimmung oder Ablehnung von Antwortoptionen explizit Begründungen ausgeführt wurden. Insgesamt werden, wie in Tabelle 16.2 ersichtlich, in jedem Think-Aloud-Protokoll insgesamt mehr unbegründete als begründete Aussagen (Bewertungen, Alternativen sowie Zustimmungen und Ablehnungen) kodiert. Das gilt auch ausdifferenziert für die meisten Paare an Kategorien (mit und ohne Begründung) in unterschiedlicher Ausprägung über die verschiedenen Kategorien und Think-Aloud-Protokolle. Ausnahmen bilden lediglich die Think-Aloud-Protokolle der Interviews mit LD3 und LD7 (im Verhältnis von begründeten und unbegründeten Zustimmungen zu bzw. Ablehnungen von Antwortoptionen).

Auffallend wenige Äußerungen fallen in die Kategorien *eigene Erfahrung* und *Professionalisierung*. Beide Kategorien werden nur selten (*eigene Erfahrung* maximal sechs Mal und *Professionalisierung* maximal vier Mal jeweils von LD1) und nicht von allen Studierenden (*eigene Erfahrung* von vier und *Professionalisierung* von drei Studierenden) angesprochen. Dies könnte als Hinweis darauf gewertet werden, dass kaum Bezüge zwischen dem beobachteten Unterricht und eigenen Erfahrungen hergestellt werden, bzw. das Gesehene kaum abstrahiert wird, um daraus Konsequenzen für die eigene (oder Roberts) Professionalisierung zu ziehen.

Bewertung des Indikators 2.1

Auf Grundlage der Think-Aloud-Protokolle wird deutlich, dass der Großteil der Kodierungen einen Bezug zum in den Unterrichtsausschnitten dargestellten Unterricht oder zur Professionalisierung der Studierenden bzw. des beobachteten Mitpraktikanten aufweisen. In Bezug auf die geäußerten Reflexionselemente (im Sinne des Stufenmodells von Nowak, Kempin, Kulgemeyer und Borowski (2019)) können keine Unterschiede abhängig von der Bearbeitung von Aufgaben zu Bewertungen und Alternativen ausgemacht werden, die Studierenden äußern stets mehr Alternativen. Für die Bewertung der inhaltlichen Sinnhaftigkeit der Überlegungen ist dies allerdings unproblematisch, da sowohl Bewertungen als auch Alternativen zum Unterricht konstruktrelevante Überlegungen darstellen. Der überwiegende Anteil konstruktrelevanter Überlegungen bei der Testbearbeitung kann also als Hinweis für die Validität der Testwertinterpretation gewertet werden, der weiter dadurch gestützt wird, dass die identifizierten Hürden während der Be-

arbeitung des Testinstruments mit wenigen Ausnahmen individuell sind und keine Änderungsbedarfe kennzeichnen. Lediglich eine für mehrere Studierende unklar formulierte Antwortoption wurde angepasst sowie eine aufgrund der Kürzungen gegenüber dem Performanztest im Rahmen der Videovignetten fehlende Information schriftlich ergänzt. Der in der dritten Vignette vorliegende Fehler, dass die zuvor beschriebene Tafel in der zweiten Hälfte der Vignette nicht beschrieben ist, ist nicht grundsätzlich gefährdend für die Validität der Testwertinterpretation, da sich hieraus keine Konsequenzen für die Bewertung der Testbearbeitung ergeben, da die zugehörige Antwortoption kein zentrales Feedback darstellt und insofern nicht als korrekt gewertet wird. Auch der nur geringe Anteil an hergestelltem Bezug zu eigenen Erfahrungen bzw. Schlussfolgerungen in den Think-Aloud-Protokollen, die für die Professionalisierung der beteiligten Lehrkräfte gezogen werden, ist nicht als gefährdend für die valide Testwertinterpretation zu bewerten, da dies keine Anforderung für die Entwicklung des Testinstruments als Abbild der Fähigkeit zur Unterrichtsreflexion darstellte.

16.2 Evaluation der Akzeptanz des Testverfahrens (Indikator 2.2)

Mit der Untersuchung von Indikator 2.2 *„Die Teilnehmenden akzeptieren das Testverfahren.“* soll sichergestellt werden, dass die vorliegenden Bearbeitungen als ernsthafte Auseinandersetzungen mit den Videovignetten im Reflexionstest interpretiert werden können. Dazu wird der in den Reflexionstest integrierte Fragebogen zur Akzeptanz ausgewertet und durch Erkenntnisse aus Kommentaren der Teilnehmenden und der Selbsteinschätzung der Motivation während der Testbearbeitung ergänzt.

16.2.1 Evaluation des Akzeptanzfragebogens im Reflexionstest

Um Informationen über die Perspektive der Studierenden auf das Testinstrument zu erhalten, enthält der Reflexionstest elf Items zur Akzeptanz (s. Kapitel 13.3), die auf einer vierstufigen Likert-Skala beantwortet werden. Zudem kann die Option *keine Antwort* ausgewählt werden, wenn Teilnehmende das entsprechende Item nicht einschätzen können oder wollen. Insgesamt liegen zu $N = 187$ Bearbeitungen Einschätzungen zumindest zu einem Teil des Items vor, während fünf Bearbeitungen des Reflexionstests abgeschlossen wurden, ohne dass Angaben zur Akzeptanz gemacht wurden. Aufgrund der Breite der berücksichtigen Faktoren (u. a. Anonymität und Relevanz für den Beruf) lassen sich die Items nicht zu

einer Skala zusammenfassen[146], sondern werden einzeln betrachtet. Daher werden nachfolgend jeweils prozentuale Zustimmungswerte dargestellt.

Eine Voraussetzung dafür, den Reflexionstest überhaupt als relevant anzusehen, ist, der Unterrichtsreflexion per se eine Bedeutung zuzusprechen. Dem Item *„Ich finde es wichtig, Unterricht zu reflektieren.“* stimmen die Teilnehmenden insgesamt deutlich zu (97 % der Teilnehmenden stimmen eher oder voll und ganz zu; $N = 186$).

Auch die Videovignetten werden von den Studierenden bezüglich ihrer (a) Authentizität (*Die „Unterrichtsausschnitte zeigen authentische Situationen.“*) und (b) Eignung zur Reflexion von Unterricht (*„Die Unterrichtsausschnitte sind gut geeignet, um über Unterricht zu reflektieren.“*) insgesamt positiv eingeschätzt (es stimmen jeweils eher oder voll und ganz zu: (a) 75 %; $N = 186$; (b) 94 %; $N = 186$).

In Bezug auf die zur Auswahl stehenden Antwortoptionen geben die Teilnehmenden überwiegend an, *„eine Antwortoption gefunden [zu haben], die zu [den eigenen] Überlegungen gepasst hat“* (89 % stimmen eher oder voll und ganz zu; $N = 187$). Der Schwierigkeitsgrad der Fragen zu den Vignetten wird tendenziell als angemessen eingeschätzt (der Aussage *„Die Fragen zu den Unterrichtsausschnitten haben einen angemessenen Schwierigkeitsgrad.“* stimmen 94 % eher oder voll und ganz zu; $N = 183$) und die Fragen werden eher nicht als zu einfach bewertet (nur 22 % stimmen der Aussage *„Die Fragen zu den Unterrichtsausschnitten waren zu einfach.“* eher oder voll und ganz zu; $N = 181$). Auch der Umfang der Fragen wird tendenziell nicht als zu hoch bewertet (der Aussage *„Die Fragen zu den Unterrichtsausschnitten waren zu umfangreich.“*; stimmen 32 % eher oder voll und ganz zu; $N = 187$). Gleichzeitig besteht aus Perspektive der Teilnehmenden ein Bedarf, die Antwortoptionen um zusätzliche Überlegungen zu ergänzen (57 % stimmen der Aussage *„Bei der Bearbeitung der Fragen fehlte mir die Möglichkeit, selbst weitere Antwortoptionen zu ergänzen.“* eher oder voll und ganz zu; $N = 185$).

Den Teilnehmenden soll trotz anonymer Erhebung der Reflexionsfähigkeit ein Assessment-Feedback mit Hilfe eines persönlichen Codes zur Verfügung gestellt werden. Wenn die Teilnehmenden nicht davon ausgehen, dass ihre Anonymität gewährleistet ist, muss befürchtet werden, dass dies die Testbearbeitung beeinflussen kann (z. B. indem zusätzliche Hilfsmittel genutzt werden, um ein besseres Ergebnis zu erzielen). Der Aussage *„Ich befürchte, dass meine Testergebnisse mir im Nachhinein zugeordnet werden können.“* stimmt knapp ein Drittel (29 %; $N = 171$) der Teilnehmenden eher oder voll und ganz zu, allerdings fällt auf, dass die Einschätzung dieses Items am häufigsten ausgelassen wurde. Die Ursache hierfür ist unklar, möglicherweise fällt den Teilnehmenden die Einschätzung

[146] Dafür spricht auch die niedrige Reliabilität von $\alpha_{\text{Cronbach}} = 0,55$ unter Berücksichtigung aller Items.

schwer, da sie ihre Rückmeldung noch nicht erhalten haben und somit das damit verbundene Verfahren (noch) nicht einschätzen können.

Dass der Großteil der Teilnehmenden (82 %; $N = 185$) sich für die eigene Rückmeldung interessiert und in lediglich zwölf Fällen dem Verfahren zur Er- und Bereitstellung der Rückmeldung widersprochen wurde, spricht zusätzlich für die Akzeptanz des Testverfahrens.

Das Auftreten technischer Schwierigkeiten wurde zwar in der *LimeSurvey*-Umgebung auch mit derselben vierstufigen Likert-Skala erhoben, die inhaltliche Interpretation der Abstufung ist allerdings ohne Erläuterung der Teilnehmenden unklar. Außerdem wurden technische Schwierigkeiten teilweise nur im Kommentarfeld angegeben bzw. nicht erläutert, sodass die Einschätzung technischer Schwierigkeiten auf Grundlage der Kommentare der Teilnehmenden erfolgt, auf die nachfolgend eingegangen wird.

16.2.2 Auswertung der Kommentare zum Testinstrument

Am Ende des Testinstruments besteht für Teilnehmende die Möglichkeit, die Testbearbeitung zu kommentieren. Dafür steht ein Freitextfeld nach folgender Aufforderung zur Verfügung: „Hier finden Sie Platz für Anmerkungen, Kommentare und Fragen. Sollten es technische Schwierigkeiten in der Testdurchführung gegeben haben, spezifizieren Sie diese bitte hier.“ Das Kommentarfeld wurde in insgesamt 41 Fällen genutzt, wobei einige Teilnehmende mehrere, inhaltliche verschiedene Aspekte anmerkten, die dann einzeln berücksichtigt wurden. Die Kommentare der Teilnehmenden können als ergänzende Hinweise bezüglich der Akzeptanz des Testinstruments herangezogen werden. Dazu wurden die Kommentare inhaltlich in sieben Cluster gruppiert, auf die nachfolgend eingegangen wird.

In elf Fällen wird Bedarf für die Ergänzung weiterer eigener Überlegungen oder Anmerkungen deutlich. Dies kann sich entweder konkret auf einzelne Situationen und die (ggf. als fehlend empfundene) Passung der vorliegenden Antwortoptionen beziehen (z. B. „fachliche Korrektheit: Merksatz zur Impulserhaltung sollte vektoriell formuliert werden oder es sollte eine Einschränkung des Merksatzes auf getestete Situation (Bewegung auf einer Linie) genannt werden“ (REE26[147])) oder allgemeiner im Wunsch nach mehr Freitextfeldern für Anmerkungen niederschlagen (z. B. „Evtl. wären solche Kommentarfunktionen auch schon vorher schön gewesen“ (EAI29)). Insbesondere gibt eine Person an, zu einer Frage keine passende Antwortoption gefunden zu haben („Hier hätte ich nichts angekreutzt [*sic*], leider war dies nicht möglich. [...] “ (VAR12)).

[147]Die den Zitaten in diesem Kapitel zugeordneten Kürzel entsprechen dem im Reflexionstest angegebenen persönlichen Code der Teilnehmenden.

Gleichzeitig kommentieren die Teilnehmenden in acht Fällen, dass sich die Fragen oder Antwortoptionen im Testinstrument zu häufig wiederholten, was die Motivation während der Testbearbeitung beeinflussen würde (z. B. „Manche Antwortmöglichkeiten oder Fragestellungen traten so häufig auf, dass ich teilweise die Konzentration und Motivation verloren habe und in der Beantwortung inkonsistent wurde“; HAL18). Zusätzlich geben drei Teilnehmende an, das Testinstrument als lang zu empfinden, bzw. länger, als die veranschlagte Zeit von einer Stunde benötigt zu haben (z. B. „Während ich den Test gemacht habe, habe ich nebenbei andere Au[f]gaben noch gemacht, die ich erledigen musste (Grund: Test hat sich in die Länge gezogen).“ (OAE09)).

In sechs Fällen geben die Teilnehmenden explizit an, einen Mehrwert im Reflexionstest zu sehen (z. B. „Es war ein gelungenes Beispiel, um mein Interesse zu wecken. Ich musste mich in gewissen Situationen zurückhalten, da es schon kontroverse Äußerungen waren. Um die Unterrichtseinheit zu reflektieren, war es meiner Meinung nach super.“ (OOR03)).

Ebenfalls in sechs Kommentaren wird die Authentizität der Videovignetten angesprochen. Die Teilnehmenden unterscheiden sich hier allerdings in ihrer Einschätzung. Während die Authentizität in drei Fällen als unzureichend kritisiert wird (z. B. „die Geschlechterrollen waren zu plump dargestellt“ (HAL18)), bewertet eine andere Person sie als „[i]nteressante und realistische Unterrichtsinszenierungen“ (WNR30) und in zwei Kommentaren wird hinterfragt, ob die Vignetten überspitzt waren, was implizit die Möglichkeit der Realitätstreue zulässt.

In insgesamt zehn Kommentaren werden verschiedene technische Schwierigkeiten beschrieben, die den Teilnehmenden während der Testdurchführung begegnet sind. Zwei Teilnehmende geben an, das Testinstrument von vorn beginnen zu müssen, wenn die Bearbeitung z. B. aufgrund mangelnder zeitlicher Ressourcen nicht am Stück durchgeführt werden konnte. Dies könnte dazu führen, dass das Testinstrument oder zumindest dessen Anfang in der zweiten Bearbeitung nicht mit der gleichen Ernsthaftigkeit bearbeitet wird. Diese Hürde für die konzentrierte und ernsthafte Bearbeitung konnte ausgeräumt werden, indem in der *LimeSurvey*-Umfrage die Option aktiviert wird, die Bearbeitung des Testinstruments zu unterbrechen und zu einem späteren Zeitpunkt fortzuführen. Da die einzelnen Vignetten jeweils in das Stundenraster eingeordnet werden und im Rahmen der Hilfestellungen zu den Aufgaben jeweils die Grundlegenden Informationen zur Unterrichtssituation nachgelesen werden können, kann davon ausgegangen werden, dass den Teilnehmenden alle notwendigen Informationen zur Verfügung stehen, um die ausstehenden Vignetten auch nach einer Unterbrechung des Instruments fundiert bearbeiten zu können. In vier Kommentaren geben Teilnehmende an, dass die Rückkehr zum Reflexionstest nach dem Öffnen einer Hilfestellung „nur über Umwege“ (OII03) möglich war. Hier bestand Bedarf zur Optimierung der *LimeSurvey*-Einstellungen, sodass die Hilfestellungen in einem neuen Tab

oder Fenster geöffnet werden und die Umfrage nicht verlassen wird. Die übrigen Kommentare mit Bezug zu technischen Schwierigkeiten sind jeweils nur einfach genannt und scheinen somit keine geteilten Hürden darzustellen, die Änderungen im Testinstrument bedürfen bzw. sind teilweise zu allgemein beschrieben, als dass behebbare Ursachen ausgemacht werden könnten (z. B. „[...] manchmal hängt es sich auf, wenn man weiter gehen möchte“ (HOC17)). Gleiches gilt für die 15 übrigen Kommentare, die als *Sonstiges* zusammengefasst wurden, da sie jeweils individuelle Überlegungen darstellen.

16.2.3 Evaluation des Fragebogens zur Motivation während der Bearbeitung des Reflexionstests

Es ist zu vermuten, dass sich fehlende Akzeptanz des Reflexionstests in niedriger Motivation während der Testbearbeitung niederschlägt. Darauf deuten auch einzelne Kommentare der Teilnehmenden hin (z. B. „Die Fragen doppeln sich häufig, dass mindert die Motivation während der Durchführung; man überfliegt dann nur noch Fragen und Antworten.“ (NOE26)). Daher werden an dieser Stelle zusätzlich die Selbstauskünfte der Teilnehmenden zur Motivation während der Testbearbeitung berücksichtigt.

Die Studierenden geben insgesamt eine hohe Testmotivation an: Auf einer fünfstufigen Likert-Skala (s. Kapitel 13.2) geben sie an, die Aufgaben eher engagiert bearbeitet zu haben (23 % stimmen teilweise, 63 % (völlig) zu) und bei der Bearbeitung der Aufgaben ihr Bestes gegeben zu haben (15 % stimmen teilweise, 70 % (völlig) zu). Die Zustimmung zur Aussage *„Ich habe mich bei der Bearbeitung der Aufgaben bis zum Schluss konzentriert.“* fällt etwas geringer aus (30 % stimmen teilweise, 54 % (völlig) zu).

Bewertung des Indikators 2.2

Zur Einschätzung der Akzeptanz der Teilnehmenden gegenüber dem Testverfahren wird vorrangig auf den Akzeptanzfragebogen zurückgegriffen, da dieser von fast allen Teilnehmenden ausgefüllt wurde, sowie auf Angaben zur Testmotivation. Zusätzlich werden die verfassten Kommentare einbezogen, die daraus resultierenden Einschätzungen sind allerdings lediglich als exemplarisch zu verstehen, da der Anteil der Teilnehmenden, die einen Kommentar verfasst haben, sehr gering ist.

Die Angaben im Akzeptanzfragebogen machen deutlich, dass die Teilnehmenden der Unterrichtsreflexion eine Relevanz zusprechen und die eingesetzten Videovignetten als authentisch und geeignete Grundlage für Unterrichtsreflexion bewerten. In drei der sechs Kommentare der Teilnehmenden, die sich auf die Authentizität des Testinstruments beziehen, werden die Szenen allerdings als überspitzt bewertet. Die vorliegenden Antwortoptionen wurden ebenfalls insgesamt positiv bewertet,

auch, wenn offenbar einige Teilnehmende das Bedürfnis haben, weitere eigene Überlegungen einzubringen, wie im Akzeptanzfragebogen sowie im Rahmen der Kommentare deutlich wird.

Elf Teilnehmende bewerten das Testinstrument in den Kommentaren als zu umfangreich und geben an, dass in ihm zu viele Wiederholungen enthalten seien. Dieser Eindruck wird durch die Angaben im Akzeptanzfragebogen eher nicht bestätigt, sodass hier nicht pauschal davon ausgegangen werden muss, dass die Testbearbeitung durch den Umfang des Instruments beeinträchtigt wird und somit keine grundlegende Gefährdung für die valide Testwertinterpretation vorliegt.

Bedeutsam für die Akzeptanz des Reflexionstests durch die Teilnehmenden sowie dessen reibungslosen Einsatz ist die Behebung der in den Kommentaren beschriebenen technischen Schwierigkeiten. Beide mehrfach aufgetretenen Probleme können durch Änderungen der *LimeSurvey*-Einstellungen behoben werden: Einerseits wird die Möglichkeit gegeben, das Testinstrument bei Bedarf zu pausieren und zu einem späteren Zeitpunkt fortzuführen. Andererseits werden die Hilfestellungen in einem neuen Tab oder Fenster geöffnet, sodass die Umfrage nicht verlassen wird, wenn diese in Anspruch genommen werden.

Die Einschätzung der meisten Teilnehmenden in den Items zur Testmotivation, den Reflexionstest überwiegend engagiert und konzentriert bearbeitet zu haben, stützt zusätzlich die Einschätzung, dass die Akzeptanz des Testverfahrens insgesamt hoch ist und Bereitschaft zur Anwendung der Reflexionsfähigkeit durch die Teilnehmenden vorliegt.

16.3 Beurteilung des Teils 2 des INA: Übersetzung Testinstrument in Verhalten

Der zweite Teil des INA fokussiert das Verhalten der Teilnehmenden während der Testbearbeitung. Anzustreben ist, dass die Teilnehmenden zum Instrument passende Überlegungen anstellen, d. h. sich zur Beantwortung der Aufgaben mit dem dargestellten Unterricht auseinandersetzen und diesen reflektieren.

Die inhaltliche Auswertung der Think-Aloud-Protokolle, die Aufschluss über die Überlegungen der Teilnehmenden während der Testbearbeitung liefern, verdeutlicht, dass überwiegend inhaltlich passende Überlegungen angestellt wurden (Indikator 2.1). Die Studierenden verbalisieren vorrangig Teile von Unterrichtsreflexionen oder sonstige Überlegungen mit direktem Bezug zum Unterricht. Ein Zusammenhang zwischen dem Verhältnis von verbalisierten Bewertungen und Alternativen zur jeweils bearbeiteten Aufgabe konnte nicht festgestellt werden, der Anteil der Alternativen überwiegt durchgehend deutlich. Aufgrund des inhaltlichen Bezugs der Überlegungen zum Unterricht und des möglichen engen

Zusammenhangs zwischen Bewertungen und dazu passenden Alternativen ist dies nicht als Widerspruch zum Indikator 2.1 einzustufen. Als Limitation dieser Studie bleibt die Auswahl der Stichprobe als Gelegenheitsstichprobe festzuhalten, die sich zwar aus Bachelor- und Masterstudierenden des Lehramts im Fach Physik zusammensetzt, allerdings eine relativ geringe Stichprobengröße umfasst und nur Physik-Lehramtsstudierende beinhaltet, was Fragen bezüglich der Generalisierbarkeit der Ergebnisse insbesondere auch auf Studierende des Grundschullehramtes oder Referendar*innen offenlässt. Auch der Detailgrad der Auswertung ist aufgrund der Beschränkung auf Verbalisierungen begrenzt. Auf zusätzliche Informationen über die kognitiven Prozesse könnte möglicherweise auf Grundlage von längeren Sprechpausen (u. a. stellvertretend für Zögern) oder der zeitlichen Dauer der Auseinandersetzung mit den einzelnen Bestandteilen des Testinstruments geschlossen werden.

Die Akzeptanz des Reflexionstests kann auf Grundlage der Rückmeldungen der Teilnehmenden außerdem als grundsätzlich hoch eingeschätzt werden (Indikator 2.2): Positiven Aussagen (z. B. bezüglich der Authentizität und Eignung der Videovignetten zur Reflexion) wird (eher) zugestimmt, kritischen Aussagen, z. B. zum unangemessenen Schwierigkeitsgrad des Testinstruments, (eher) nicht. Einige Teilnehmende nutzten das Kommentarfeld, um dem Testinstrument einen Mehrwert zu attestieren, was diese Einschätzung stützt. Verschiedene aufgetretene technische Schwierigkeiten, die die Akzeptanz des Verfahrens möglicherweise beeinträchtigen, konnten behoben werden (u. a. die Möglichkeit, die Testbearbeitung zu unterbrechen). Dem Bedürfnis der Teilnehmenden, die Möglichkeit für individuelle Rückmeldungen zu den Videovignetten zu schaffen, kann nicht nachgekommen werden, ohne den Vorteil der automatisierten Auswertung für das Assessment-Feedback aufzugeben. Dies ist als Limitation des Antwortformats einzuordnen.[148] Und obwohl einige Studierende im Kommentarfeld angeben, das Testinstrument zumindest teilweise als repetitiv und zu umfangreich wahrzunehmen, was die Motivation zur intensiven Auseinandersetzung mit dem Reflexionstest schmälert, geben die Teilnehmenden eine insgesamt hohe Testmotivation an. Für die Erfassung von Motivation und insbesondere Akzeptanz des Reflexionstests ist einschränkend festzuhalten, dass keine etablierten bzw. erprobten Instrumente eingesetzt wurden, sondern spezifisch auf den Reflexionstest zugeschnittene Items eingesetzt wurden, sodass die Auswertung an dieser Stelle überwiegend auf der Ebene der Einzelitems verbleibt und keine aggregierte Einschätzung im Sinne eines quantifizierbaren Maßes für die Akzeptanz angegeben und für folgende Untersuchungen verwendet werden kann. Auch scheinen nicht alle Teilnehmenden, die im Rahmen der Likert-Skala angeben, technischen Schwierigkeiten begegnet zu sein,

[148] Um diesem Bedürfnis nachzukommen und frei formulierte Antworten der Teilnehmenden einzubeziehen, könnten möglicherweise zukünftig Methoden des maschinellen Lernens zum Einsatz kommen; Untersuchungen zur automatisierten Auswertung schriftlicher Reflexionen mit Sprachmodellen stellen Mientus, Wulff, Nowak und Borowski (2023) und Wulff, Mientus, Nowak und Borowski (2023b) vor.

von der Möglichkeit Gebrauch zu machen, diese anschließend im Kommentarfeld näher zu beschreiben. Hier ist also kritisch zu hinterfragen, inwiefern die fehlenden Informationen über Schwierigkeiten bei der Durchführung des Testinstruments zu einer Verzerrung der Bewertung des Reflexionstests führen könnten und in welchem Umfang Teilnehmende sich die Zeit nehmen, auch sonstige negative Eindrücke zum Testinstrument, die ggf. die Akzeptanz oder Motivation schmälern können, im Kommentarfeld zu berichten.

Insgesamt können die vorliegenden Befunde vorbehaltlich der Einschränkungen bezüglich der Stichprobengröße und ggf. ausbleibender Beschreibungen negativer Erfahrungen bei der Bearbeitung des Testinstruments durch die Teilnehmenden als Argumente für die kognitiv valide Interpretation der Testwerte verstanden werden. Die Anforderung, dass die Teilnehmenden vorwiegend konstruktrelevantes Verhalten bei der Bearbeitung des Reflexionstests zeigen, kann daher auf dieser Basis als erfüllt bewertet werden.

Kategorien der Kodierung	Interviews							κ_{Cohen}
	LD1	LD2	LD3	LD4	LD5	LD6	LD7	
Beschreibung	4 %	9 %	7 %	2 %	9 %	9 %	9 %	0,58
Bewertung	3 %	7 %	1 %	4 %	8 %	8 %	7 %	0,62
Bewertung mit Begründung	1 %	3 %	1 %	2 %	2 %	3 %	1 %	0,54
Alternative	4 %	11 %	14 %	8 %	17 %	11 %	8 %	0,82
Alternative mit Begründung	2 %	3 %	5 %	≈ 0 %	4 %	3 %	1 %	0,57
Zustimmung/ Ablehnung von Antwortoptionen	18 %	25 %	29 %	38 %	24 %	36 %	24 %	0,73
Zustimmung/ Ablehnung von Antwortoptionen mit Begründung	15 %	18 %	33 %	23 %	22 %	11 %	37 %	0,77
Erfahrung	2 %	≈ 0 %	0 %	0 %	0 %	1 %	≈ 0 %	-
Professionalisierung	≈ 0 %	≈ 0 %	0 %	0 %	0 %	≈ 0 %	0 %	-
Andere Überlegungen zum Unterricht	7 %	10 %	2 %	6 %	6 %	8 %	8 %	0,72
Konstruktrelevante Überlegungen insgesamt	57 %	86 %	92 %	83 %	92 %	88 %	96 %	0,80
Sonstiges	43 %	14 %	8 %	17 %	8 %	12 %	4 %	0,75

Tabelle 16.2: Übersicht über Anteile der verschiedenen Kategorien an allen Kodierungen (unter Auslassung der Kategorie *Vorlesen*) je Interview (als „≈ 0“ sind diejenigen Anteile der Kategorien *eigene Erfahrung* und *Professionalisierung* eingetragen, die nur zu sehr geringen Anteilen in den jeweiligen Think-Aloud-Protokollen vorkommen; „0“ beschreibt Kategorien die im jeweiligen Interview nicht angesprochen werden) sowie die Beurteilerübereinstimmung für die Kategorien (κ_{Cohen}); die Kategorie *Konstruktrelevante Überlegungen* fasst alle Kategorien bis auf *Sonstiges* zusammen

17 Untersuchung der Anforderung an Teil 3 des INA: Das (Ankreuz-)Verhalten der Proband*innen während der Testbearbeitung wird adäquat in Testwerte übersetzt

Zur Untersuchung der Übersetzung des Ankreuzverhaltens der Teilnehmenden in Testwerte als dritten Teil des INA werden verschiedene Indikatoren mit unterschiedlichen statistischen Methoden geprüft (eine Übersicht über die Anforderung an diesen Teil des INA und die dazu betrachteten Indikatoren und durchgeführten Untersuchungen ist in den Abbildungen 14.1 und 14.2 auf den Seiten 186 und 188 zum dritten Übersetzungsschritt dargestellt). Dazu wird in Kapitel 17.1 die Eignung des Testinstruments zur Differenzierung zwischen verschiedenen Ausprägungen der Reflexionsfähigkeit sowie anschließend die Reliabilität der Skalen zu den bestimmten Testwerte (Kapitel 17.2) geprüft. In Kapitel 17.3 wird im Sinne der strukturellen Validierung die Struktur der Testwerte in einer explorativen Faktorenanalyse untersucht. Zur externen Validierung werden Zusammenhänge zwischen den Testwerten im Reflexionstest und Testwerten im Reflexionsperformanztest stellvertretend für ein authentischeres Reflexionssetting betrachtet (Kapitel 17.4) und Zusammenhänge zu unterschiedlichen Lerngelegenheiten (z. B. der Studienphase; Kapitel 17.5) sowie zum fachdidaktischen Wissen untersucht (Kapitel 17.6).[149] Abschließend wird in Kapitel 17.7 auf dieser Grundlage beurteilt, inwieweit die Anforderung an den dritten Teil des INA erfüllt ist.

[149] Die statistischen Analysen werden mit *R* (R Core Team, 2022) durchgeführt, das zugehörige Skript befindet sich vollständig im Anhang in Kapitel R ab Seite 881.

17.1 Evaluation der Differenzierungsfähigkeit des Reflexionstests (Indikator 3.1)

Zur Untersuchung des Indikators 3.1 *„Der Reflexionstest ist dazu geeignet, verschiedene Ausprägungen der Reflexionsfähigkeit auch innerhalb der verschiedenen Ausbildungsphasen (Bachelor- und Masterstudium) differenziert darzustellen.“* wird zunächst die Verteilung der Testwerte insgesamt sowie für die Teilfähigkeiten Bewerten und Vorschlagen von Alternativen betrachtet. Zusätzlich werden die Lösungshäufigkeiten der Aufgaben bestimmt. So kann geprüft werden, inwiefern die Breite der möglichen Testwerte ausgeschöpft wird und ob möglicherweise Boden- bzw. Deckeneffekte auftreten.

17.1.1 Analyse der vorliegenden Verteilung

Die $N = 192$ Teilnehmen erreichen Testwerte zwischen rund 12 % und 91 %, wobei mit $N = 165$ Personen der Großteil der Teilnehmenden Testwerte von über 50% erreicht. Die Verteilung der Testwerte ist in Abbildung 17.1 dargestellt. Die Untersuchung der vorliegenden Daten auf Normalverteilung mit dem Kolmogorow-Smirnow-Test ergibt ein knapp nicht signifikantes Ergebnis ($D(192) = 0,10$; $p = 0,053$[150]), sodass von einer Normalverteilung der Testwerte mit Mittelwert von etwa 66 % und Standardabweichung von 16 % auszugehen ist.

Für die Teilstichproben der Bachelor- und Masterstudierenden kann unmittelbar auf Grundlage des Kolmogorow-Smirnow-Tests von normalverteilten Daten ausgegangen werden:[151]

- Bachelorstudierende: $D(90) = 0,13$; $p = 0,097$ mit Mittelwert von etwa 61 % (Standardabweichung etwa 19 %)
- Masterstudierende: $D(86) = 0,09$; $p = 0,555$ mit Mittelwert von etwa 72 % (Standardabweichung etwa 12 %)

Ergänzend werden die Teilfähigkeiten Bewerten und Vorschlagen von Alternativen betrachtet.[152] Die Testwerte für die Teilfähigkeit Bewerten können ebenfalls

[150] Die Ergebnisse der Kolmogorow-Smirnow-Tests werden hier und im Folgenden im Format $D([\text{Stichprobengröße}]) = [\text{Teststatistik}]$; $p = [\text{Signifikanzwert}]$ berichtet.

[151] Auf analoge Untersuchungen für die Stichprobe der Referendar*innen mit dem Ziel der Generalisierbarkeit der Ergebnisse wird hier aufgrund der geringen Stichprobengröße und unklaren Zusammensetzung in Bezug auf die Zugangswege zum Referendariat verzichtet.

[152] Diese werden zwar im Rahmen der nachfolgend berichteten Auswertungen überwiegend nicht explizit aufgeschlüsselt, allerdings im Rahmen des Assessment-Feedbacks einzeln dargestellt, sodass zentrale Bestandteile der Skalenanalyse durchgeführt werden (Untersuchung der Verteilung und internen Konsistenz auf Skalenebene sowie die Lösungshäufigkeiten und Trennschärfen auf Aufgabenebene). Im Rahmen des Assessment-Feedbacks wird zusätzlich ein Testwert für das Erkennen von und den Umgang mit Schülervorstellungen angegeben.

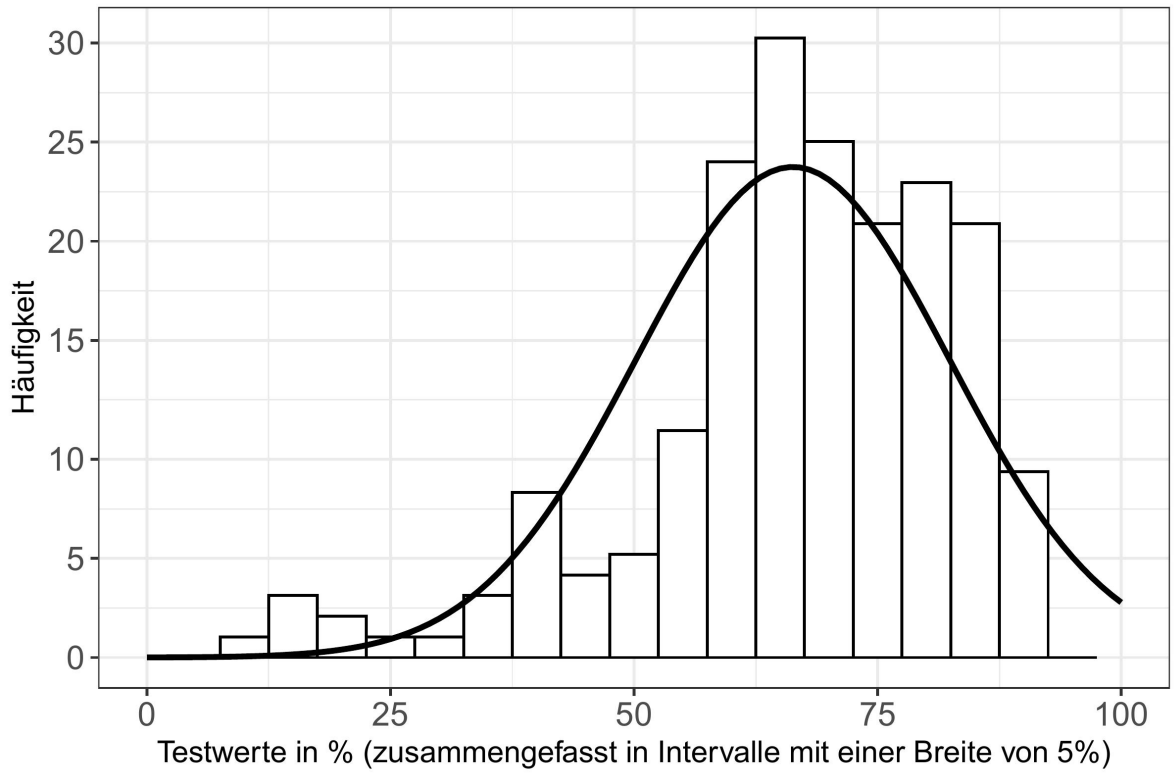


Abbildung 17.1: Histogramm der erreichten Testwerte (in %) mit Dichtefunktion der Normalverteilung ($N = 192$)

insgesamt als normalverteilt mit einem Mittelwert von 64% und einer Standardabweichung von 16 % angenommen werden. Der Kolmogorow-Smirnow-Test liefert ein knapp nicht signifikantes Ergebnis ($D(192) = 0,096$; $p = 0,058$). Für die Teilfähigkeit Vorschlagen von Alternativen kann auf Grundlage des Kolmogorow-Smirnow-Tests keine Normalverteilung unterstellt werden, da dieser signifikant wird ($D(192) = 0,12$; $p < 0,01$). Gleichzeitig erreichen die Teilnehmenden Testwerte zwischen etwa 9 % und 93 % und die Aufgaben zu Alternativen weisen - wie nachfolgend dargestellt wird - überwiegend annehmbare Lösungshäufigkeiten und Trennschärfen auf, sodass trotzdem insgesamt davon ausgegangen werden kann, dass sich auch die Testwerte zu den Teilfähigkeiten Bewerten und Vorschlagen von Alternativen zur Differenzierung eignen.

17.1.2 Analyse der Lösungshäufigkeiten

Die in die verwendeten Skalen eingehenden Aufgaben werden im Rahmen der Prüfung dieses Indikators und der nächsten beiden Indikatoren im Sinne einer Itemanalyse untersucht. Dazu werden hier zunächst die Item- bzw. Aufgabenschwierigkeiten näher betrachtet (Döring & Bortz, 2016, S. 476). Als Auskunft über Aufgabenschwierigkeit werden die Lösungshäufigkeiten betrachtet. Anzustreben sind Lösungshäufigkeiten zwischen $0,20$ und $0,80$, um möglichst keine zu schweren bzw. zu leichten Aufgaben in das Testinstrument zu inkludieren, die nicht geeignet sind, zwischen den Teilnehmenden zu differenzieren (Döring & Bortz, 2016, S. 477).

Für die vorliegende Stichprobe von $N = 192$ ergibt sich für den Reflexionstest eine annehmbare durchschnittliche Lösungshäufigkeit über alle Aufgaben zur Reflexion von $M = 0,66$ ($SD = 0,16$) mit einzelnen Lösungshäufigkeiten zwischen $M = 0,36$ (R4Alt1) und $M = 0,91$ (R6Bew1). Da der Reflexionstest über das Bachelor- und Masterstudium hinweg eingesetzt werden soll, werden die Kennwerte außerdem für diese beiden Teilstichproben ($N_{\text{BA}} = 90$ und $N_{\text{MA}} = 86$) getrennt betrachtet. Dabei ergeben sich für beide Gruppen annehmbare durchschnittliche Aufgabenschwierigkeiten von $M = 0,61$ ($SD = 0,14$ mit einer Spannweite von $0,33$ (R1Bew1) bis $0,86$ (R6Bew1)) für die Bachelorstudierenden und $M = 0,72$ ($SD = 0,17$ mit einer Spannweite von $0,35$ (R3Alt1) bis $0,96$ (R6Bew1)) für die Masterstudierenden, wobei - wie bei der Gesamtstichprobe - einzelne Aufgaben wiederum Lösungshäufigkeiten von über $0,80$ aufweisen.[153] Zusätzlich zeigt sich, dass die Lösungshäufigkeiten für die Gruppe der Masterstudierenden

Da hier allerdings nur zwei Aufgaben eingehen und die Struktur der Aufgaben von der restlichen Erfassung der Reflexionsfähigkeit über Aufgaben zu Bewertungen und Alternativen abweicht, werden die zugehörigen Testwerte bzw. Aufgabenscores an dieser Stelle nicht weiter analysiert.

[153] Konkret betrifft dies in der Gesamtstichprobe acht Aufgaben, in den Gruppe der Bachelor- bzw. Masterstudierenden drei bzw. zwölf Aufgaben.

zwar insgesamt höher liegen als für die Gruppe der Bachelorstudierenden, die für die beiden Gruppen einzeln bestimmten Lösungshäufigkeiten aber mit $r_{\text{Pearson}} = 0,96$ ($p < 0,01$) einen sehr hohen Zusammenhang aufweisen (Cohen, 1988, S. 79ff.).

Insgesamt wird deutlich, dass die Lösungshäufigkeiten zwar eine gewisse Streuung aufweisen, aber - wie bereits die höheren Testwerte insbesondere für die Gruppe der Masterstudierenden nahelegen - eher hoch sind. Entgegen der Empfehlung von Döring und Bortz (2016, S. 277) werden diese Aufgaben aus Gründen der inhaltlichen Validität nicht entfernt: Da zu jedem der reflektierten Aspekte jeweils eine Aufgabe zu Bewertungen und zu Alternativen bearbeitet wird, erscheint die Entfernung von Aufgaben nur in Paaren sinnvoll. Da in keinem Fall beide Aufgaben zu einem Aspekt eine besonders hohe Lösungshäufigkeit aufweisen, verbleiben auch die Aufgaben mit hoher Lösungshäufigkeit im Reflexionstest.

17.1.3 Bewertung des Indikators 3.1

Für die Testwerte zur Reflexionsfähigkeit insgesamt (und für die Teilfähigkeit zum Bewerten) kann von einer Normalverteilung der Daten ausgegangen werden. Auffällig ist, dass die Verteilungen aller drei Skalen linksschief sind, die Teilnehmenden erreichen also tendenziell hohe Testwerte. Dementsprechend liegen tendenziell hohe Lösungshäufigkeiten zu den Testaufgaben vor. Von einem Deckeneffekt muss allerdings noch nicht ausgegangen werden, da die Häufigkeitsverteilung im Bereich der hohen Testwerte abflacht und die Maximalpunktzahl im Test in keinem Fall erreicht wird. Auch werden von einigen Teilnehmenden niedrige Testwerte im Reflexionstest insgesamt und in den Teilfähigkeiten erreicht. Zusätzlich liegen - wie nachfolgend in Kapitel 17.2 berichtet wird - die Trennschärfen für die Testaufgaben durchschnittlich im mittleren bis hohen Bereich, was ebenfalls für die Differenzierungsfähigkeit des Reflexionstests spricht. Insgesamt scheint der Reflexionstest also durchaus in der Lage zu sein, Ausprägungen der Reflexionsfähigkeit und der betrachteten Teilfähigkeiten zu differenzieren. Der Indikator 3.1 kann damit als bestätigt bewertet werden.

17.2 Evaluation der Reliabilität der verwendeten Skalen (Indikator 3.2)

Die Evaluation der Reliabilität der verwendeten Skalen adressiert Indikator 3.2 *„Die zur Bestimmung der Testwerte verwendeten Skalen sind reliabel auch über die verschiedenen Ausbildungsphasen (Bachelor- und Masterstudium) hinweg.“* Dazu wird als Maß für die interne Konsistenz jeweils Cronbachs α bestimmt und um die Betrachtung der Trennschärfe der eingehenden Aufgaben ergänzt. Als Maß

für die Trennschärfe wird jeweils die korrigierte Item-Skala-Korrelation bestimmt (Döring & Bortz, 2016, S. 478).

Für den Reflexionstest liegen insgesamt $N = 192$ Bearbeitungen vor, die zur Bestimmung von α_{Cronbach} zugrunde gelegt werden können. Um die Angemessenheit der Skalen über die Studienphasen hinweg zu beurteilen, werden reduzierte Datensätze von $N_{\text{BA}} = 90$ und $N_{\text{MA}} = 86$ herangezogen. Bearbeitungen von Referendar*innen und Studierenden mit Studienziel erstes Staatsexamen sowie Personen, die als Ausbildungsphase „Sonstiges“ angaben, sind hier ausgeschlossen.

17.2.1 Betrachtung des Testwerts zur Reflexionsfähigkeit

Zur Berechnung des Testwerts zur Reflexionsfähigkeit werden alle 32 Aufgaben (je 15 Aufgaben zu Bewertungen und Alternativen sowie zwei Aufgaben zu Schülervorstellungen) berücksichtigt. Für die Gesamtstichprobe ergibt sich ein sehr hoher Reliabilitätskoeffizient von $\alpha_{\text{Cronbach}_{\text{Ges}}} = 0,95$ (Döring & Bortz, 2016, S. 443). Die Aufgaben weisen Trennschärfen zwischen $0,27$ (R3Bew1) und $0,75$ (RGAlt1) (durchschnittlich $0,58$) auf. Gemäß dem von Döring und Bortz (2016, S. 478) beschriebenen Kriterien sind Trennschärfekoeffizienten ab $0,30$ als mittelmäßig und ab $0,50$ als hoch zu bewerten, wobei insbesondere bei der Erhebung komplexer Konstrukte niedrigere Trennschärfen „zugunsten einer größeren inhaltlichen Breite der Items“ (Döring & Bortz, 2016, S. 478) zugelassen werden können. Für die Gesamtstichprobe weisen fast alle Aufgaben ausreichende Trennschärfen auf, eine Aufgabe (R3Bew1) verfehlt die Empfehlungen für anzustrebende Trennschärfen nur knapp.

Eine Aufteilung der Stichprobe in Bachelor- und Masterstudierende führt zu ähnlichen Ergebnissen.[154] Während die Reliabilitätskoeffizienten für die beiden Teilstichproben weiterhin hoch sind ($\alpha_{\text{Cronbach}} > 0,90$) und alle Trennschärfen für die Bachelorstudierenden im annehmbaren Bereich liegen, weisen für die Masterstudierenden drei Aufgaben niedrige Trennschärfen von unter $0,3$ auf (R3Bew1, R3Alt1 und R4Bew1).

17.2.2 Betrachtung der Testwerte für die Teilfähigkeiten zum Bewerten und Vorschlagen von Alternativen

Im Rahmen des Assessment-Feedbacks werden getrennte Testwerte für die Aufgaben zu Bewertungen und Alternativen sowie Schülervorstellungen rückgemeldet. Es ist auch hier zu prüfen, ob die zugehörigen Skalen eine ausreichende Reliabilität aufweisen. Die Skalen für die Testwerte zu Bewertungen und Alternativen bestehen

[154] Eine Übersicht über die Kennwerte für die betrachteten Stichproben befindet sich in Tabelle I.1 in Anhang auf Seite 479.

jeweils aus 15 Aufgaben (s. Kapitel 10.1) und weisen sowohl für die Gesamtstichprobe, als auch für die Teilstichproben der Bachelor- und Masterstudierenden hohe bis sehr hohe Reliabilitäten von $\alpha_{\text{Cronbach}} > 0,8$ auf (s. Tabelle I.1 im Anhang auf Seite 479). Die übrigen beiden Aufgaben beziehen sich auf das Erkennen von und den Umgang mit Schülervorstellungen (R2Bew1 und R2Alt1). Die Ergebnisse zu diesen Aufgaben werden zwar im Rahmen des Assessment-Feedback rückgemeldet, aufgrund der geringen Aufgabenzahl wird aber keine Berechnung eines Reliabilitätskoeffizienten vorgenommen. Für beide Aufgaben liegen annehmbare Trennschärfen und Lösungshäufigkeiten aus der Betrachtung des Testwerts zur Reflexionsfähigkeit insgesamt vor (vgl. Kapitel 17.1.2).

Genauso wie für den Testwert zur Reflexionsfähigkeit insgesamt werden auch hier für alle Stichproben mindestens ausreichende und meist auch hohe Reliabilitätskoeffizienten von über 0,9 erreicht (Döring & Bortz, 2016, S. 443). Auch die durchschnittlichen Trennschärfen liegen jeweils in einem mittelmäßigen bis guten Bereich, während einzelne Aufgaben wieder eher niedrige Werte von unter 0,3 aufweisen.

17.2.3 Bewertung des Indikators 3.2

Die verwendeten Skalen zur Bestimmung von Testwerten zur Reflexionsfähigkeit insgesamt sowie für die Teilfähigkeiten Bewerten und Vorschlagen von Alternativen weisen sowohl für die Gesamtstichprobe der Teilnehmenden als auch getrennt für Teilnehmende der beiden Ausbildungsphasen Bachelor- und Masterstudium mindestens ausreichende, zumeist hohe Reliabilitäten auf.

Die Trennschärfe der einzelnen Aufgaben liegt durchschnittlich jeweils im mittleren bis hohen Bereich, variiert aber deutlich zwischen den verschiedenen Aufgaben. Eine mögliche Ursache hierfür könnten Unterschiede in Bezug auf den Umfang der in den Aufgaben adressierten Aspekte bzw. Wissensbasen sein. Wenige Aufgaben erreichen in der Gesamtstichprobe bzw. der Gruppe der Masterstudierenden niedrige Trennschärfen von unter 0,3. Dies betrifft allerdings Aufgaben mit eher niedrigen Lösungshäufigkeiten, sodass die niedrigen Trennschärfen hier eher schwierigere Aufgaben zu kennzeichnen scheinen, die vor dem Hintergrund der Komplexität des erhobenen Konstrukts und der insgesamt eher hohen erreichten Testwerte akzeptiert werden. Der Indikator 3.2 wird damit insgesamt als bestätigt bewertet.

17.3 Evaluation der Struktur der Testwerte (Indikator 3.3)

Grundsätzlich wird die Fähigkeit, Unterricht zu reflektieren als eine (eindimensionale) Facette der professionellen Handlungskompetenz aufgefasst, die im Wechselspiel mit verschiedenen anderen Kompetenzen die professionelle Handlungskompetenz einer Lehrkraft bildet. Als solche umfasst sie wiederum „ein Bündel von [...] geistigen Fähigkeiten“ (Frey, 2006, S. 31), wie das Bewerten von vielfältigen Unterrichtssituationen oder das Finden adäquater Handlungsalternativen. Diese Annahme wird bereits durch die hohe Reliabilität der Skala für den Testwert zur Reflexionsfähigkeit von $\alpha_{\text{Cronbach}_{\text{Ges}}} = 0,95$ gestützt. Gleichzeitig kann von einer hohen Reliabilität nicht automatisch auf eine homogene Skala geschlossen werden (Schecker, 2014). Daher wird zur strukturellen Validierung ergänzend eine explorative Faktorenanalyse durchgeführt, die etwaige unerwartete Strukturen aufdecken kann, die den Testwerten zugrundeliegen und einen Beitrag zur Beurteilung des Indikators 3.3 *„Die Struktur der Testwerte ist theoretisch erklärbar.“* liefert.

Ziel der explorativen Faktorenanalyse ist also die Gewinnung von Erkenntnissen über die Struktur bzw. Dimensionalität der Testwerte des Reflexionstests. Durchgeführt wurde eine Maximum-Likelihood-Faktorenanalyse mit dem *R*-Paket *nFactors* (Gilles Raiche & David Magis, 2022). Als Voraussetzung für die explorative Faktorenanalyse ist eine ausreichend hohe (Partial-)Korrelation zwischen den Items zu prüfen. Dazu wurde der Bartlett-Test berechnet, der mit $\chi^2(496) = 3133,765$; $p < 0,001$ ein hochsignifikantes Ergebnis liefert und es wurden die Measure of Sampling Adequacy (MSA-) Koeffizienten bestimmt, die für alle Aufgaben mit über $0,80$ mindestens gut und durchschnittlich mit $0,93$ sehr gut sind (vgl. Bühner, 2011, S. 347f.).

Zur Bestimmung der Anzahl an zu extrahierenden Faktoren wird der Scree-Plot herangezogen. Auf Grundlage des deutlichen Knicks bei zwei Faktoren im Scree-Plot (s. Abbildung 17.2) ist davon auszugehen, dass die eingesetzten Aufgaben einen Faktor darstellen (vgl. Bühner, 2011, S. 322f.).[155]

Dieses Ergebnis ist gerade vor dem Hintergrund der hohen Reliabilität der verwendeten Skala im Reflexionstest plausibel und lässt die Interpretation zu, dass die Reflexionsfähigkeit als eindimensionales Konstrukt erfasst wird. Der Indikator 3.3 kann damit als bestätigt angenommen werden.

[155] Ein Vergleich verschiedener Verfahren, die zur Bestimmung der Anzahl an zu extrahierenden Faktoren angewendet werden können, liefert ein etwas weniger eindeutiges Bild: Während z. B. der Minimum Average-Partial-Test (Bühner, 2011, S. 325f.) ebenfalls die Extraktion einer Komponente nahelegt, könnten auf Grundlage der Parallelanalyse nach Horn (Bühner, 2011, S. 323ff.) zwei Faktoren extrahiert werden. Die einzelnen Aufgaben können auf Grundlage der für diese Lösung bestimmten Faktorladungen allerdings häufig nicht eindeutig einem Faktor zugeordnet werden, sondern weisen ähnliche Höhen auf, was ein Problem für die inhaltliche Interpretation der Faktoren darstellt.

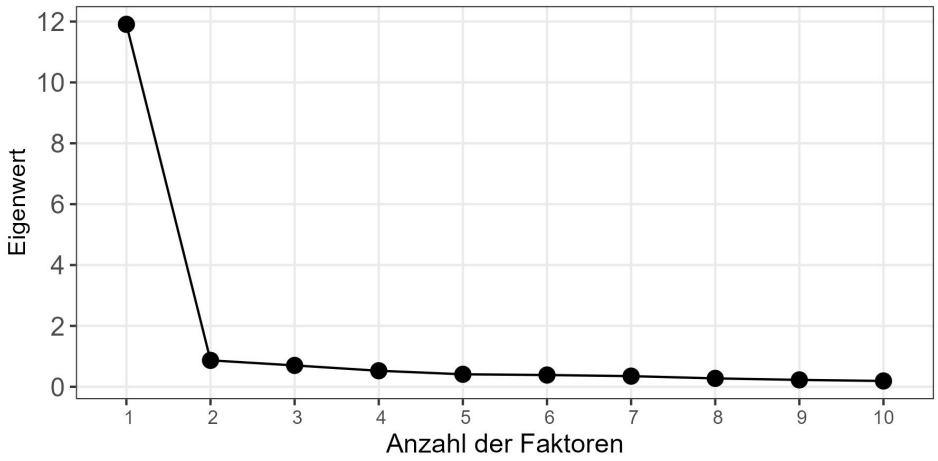


Abbildung 17.2: Scree-Plot zur explorativen Faktorenanalyse von den Aufgaben im Reflexionstest

Während sich die bisherigen Analysen auf die Struktur der Testwerte sowie Kennzahlen der eingehenden Aufgaben beziehen, werden nachfolgend Untersuchungen zu Zusammenhängen der Reflexionsfähigkeit mit externen Variablen angestellt. Dabei wird gestützt durch das Ergebnis dieser Faktorenanalyse nur der Testwert zur Reflexionsfähigkeit insgesamt (und keine Testwerte zu Teilfähigkeiten) berücksichtigt.

17.4 Evaluation des Zusammenhangs zur Reflexion in authentischeren Situationen (Indikator 3.4)

Bei der Erfassung der Fähigkeit zur Unterrichtsreflexion mit dem Reflexionstest muss davon ausgegangen werden, dass die Anforderungen von denen in der realen Handlungssituation (also der Reflexion im kollegialen Gespräch) abweichen, was prinzipiell eine Gefahr für die valide Testwertinterpretation darstellen könnte (wie Untersuchung der inhaltlichen Validität zum ersten Teil des INA in Kapitel 15.1 beschrieben). Insbesondere ist von einem Einfluss des MC-Formats auszugehen, da eigene Überlegungen mit den vorgegebenen Antwortoptionen abgeglichen werden müssen und die Antwortoptionen Überlegungen der Teilnehmenden möglicherweise ergänzen. Um die Testwerte dennoch valide als Ausdruck der Reflexionsfähigkeit interpretieren zu können, wird ihr Zusammenhang zur Reflexion in einem authentischeren Setting untersucht. Dazu wird der Reflexionsperformanztest von Kempin, Kulgemeyer und Schecker (2018) eingesetzt, der die Fähigkeit zur Unterrichtsreflexion zwar weiterhin in einem standardisierten Verfahren erfasst, aber die reale Reflexionssituation authentischer simuliert. Auf dieser Grundlage erfolgt die Beurteilung von Indikator 3.4 *„Die Testwerte des Reflexionstests weisen Zusam-*

menhänge zur Qualität von Reflexionen in authentischeren Settings (repräsentiert durch die Testwerte des Reflexionsperformanztests) auf.“

17.4.1 Studiendesign

Zum Vergleich des (MC-)Reflexionstests und des Reflexionsperformanztests führte eine Stichprobe von $N = 39$ Studierenden beide Testinstrumente durch. Die Stichprobe setzte sich zusammen aus Physik-Lehramtsstudierenden im Bachelor- ($N = 10$) und Masterstudium ($N = 12$) sowie $N = 17$ Studierenden des Sachunterrichts (Grundschullehramt) mit Wahlfach Physik. Die Studie wurde im Wintersemester 2022/23 an der Universität Bremen durchgeführt und die Bearbeitungen der Testinstrumente in Lehrveranstaltungen integriert.[156] Dabei wurde der Reflexionsperformanztest stets zuerst bearbeitet, um zu vermeiden, dass die Auswahl der reflektierten Aspekte sowie die konkreten Antwortoptionen in den MC-Aufgaben des Reflexionstests die Reflexionen der Studierenden im Reflexionsperformanztest beeinflussen. Zwischen den Durchführungen beider Instrumente lagen in den Veranstaltungen drei bis acht Wochen.

Die Testwerte zu den Bearbeitungen des Reflexionstests wurden weiterhin automatisiert mit *R* bestimmt. Der Reflexionsperformanztest wurde mit Hilfe des zugehörigen Kodiermanuals von der Autorin (einer im Projekt *ProfiLeP+* trainierten Kodiererin) ausgewertet.

17.4.2 Ergebnisse

Für den Reflexionstest ergeben sich durchschnittliche Testwerte von rund $M = 67\ \%$ ($SD = 21\ \%$). Für die vorliegende Stichprobe (sowie die Teilstichproben der Bachelor- und Masterstudierenden des Physik-Lehramts und der Studierenden des Grundschullehramtes) kann nach dem Kolmogorow-Smirnow-Test davon ausgegangen werden, dass die Daten normalverteilt sind ($p > 0,05$). Auch für diese Stichprobe liegt eine hohe Reliabilität von $\alpha_{\text{Cronbach}} = 0,93$ vor.[157]

Im Reflexionsperformanztest erreichen die Studierenden deutlich niedrigere Testwerte von durchschnittlich $M = 9\ \%$ ($SD = 4\ \%$). Auch hier kann für die vorliegende Stichprobe nach dem Kolmogorow-Smirnow-Test von einer Normalverteilung der Daten ausgegangen werden ($p > 0,05$) und die Skala zur Bestimmung des Testwerte weist eine ausreichende Reliabilität von $\alpha_{\text{Cronbach}} = 0,82$ auf. Zusätzlich können verschiedene Subskalen für den Performanztest zu Reflexionen mit Bezug

[156] Für die Studierenden bestand dabei stets die Möglichkeit, das Testinstrument zwar zu bearbeiten, aber der Verwendung ihrer Daten zu Forschungszwecken zu widersprechen.

[157] Die Ergebnisse dieser und nachfolgend berichteter Kolmogorow-Smirnow-Tests sind für beide Testinstrumente und alle betrachteten Stichproben in Tabelle I.2 im Anhang auf Seite 480 aufgeführt.

zum fachlichen, fachdidaktischen und pädagogischen Wissen betrachtet werden, indem jeweils die zugehörigen Reflexionspunkte aufsummiert werden. Auch für diese kann auf Grundlage des Kolmogorow-Smirnow-Test von Normalverteilungen ausgegangen werden (es gilt jeweils $p > 0,05$).

Zusammenhänge zwischen den Testwerten beider Instrumente

Zur Untersuchung des Zusammenhangs zwischen beiden Testverfahren zur Messung der Reflexionsfähigkeit wird eine Korrelation zwischen den Testwerten beider Instrumente bestimmt. Aus den vorliegenden Bearbeitungen ergibt sich eine hochsignifikante Korrelation von mittlerer an der Schwelle zu hoher Größe $r_{\text{Pearson}} = 0,49$ ($p < 0,01$) (Cohen, 1988, S. 79ff.).

Im Reflexionstest ist die Darstellung des Unterrichts mit Videovignetten gegenüber dem Performanztest gekürzt. Zusätzlich kann ein Testwert zum Reflexionsperformanztest bestimmt werden, in den lediglich die Teilscores zu den Videovignetten eingehen, die auch im Reflexionstest eingebunden sind. Hier erreichen die Studierenden durchschnittliche Testwerte in vergleichbarer Höhe wie unter Berücksichtigung aller Videovignetten von ebenfalls $M = 9$ % ($SD = 4$ %); die Daten liegen gemäß dem Kolmogorow-Smirnow-Test normalverteilt vor ($D(39) = 0,13; p = 0,57$) und die Skala weist eine Reliabilität von $\alpha_{\text{Cronbach}} = 0,76$ auf, liegt also leicht unter dem üblicherweise als ausreichend bewertetem Wert von $\alpha_{\text{Cronbach}} = 0,80$, was aber vor dem Hintergrund der Komplexität des zu erfassenden Konstrukts der Reflexionsfähigkeit dennoch ausreichend erscheint (Schecker, 2014). Eine Untersuchung der Korrelation zwischen dem Reflexionstest und dem angepassten Testwert im Reflexionsperformanztest liefert eine Korrelation gleicher Größe wie zum ursprünglich bestimmten Testwert von $r_{\text{Pearson}} = 0,49$ ($p < 0,01$).

Diskussion der Gruppenzugehörigkeit

Da die Stichprobe aus Bachelor- und Masterstudierenden der Physik sowie Studierenden des Sachunterrichts zusammengesetzt ist, wird zusätzlich geprüft, inwiefern die Ergebnisse über diese drei unterschiedlichen Gruppen generalisierbar sind. Zunächst fällt auf, dass für keine der drei Gruppen allein eine signifikante Korrelation zwischen den Testwerten beider Instrumente beobachtet werden kann, wobei für die kleinen hier vorliegenden Stichproben nur große Effekte detektiert werden können.[158]

Zusätzlich werden einfaktorielle Varianzanalysen (ANOVAs) für die drei Gruppen berechnet. Als weitere Voraussetzung der ANOVA (neben der bereits geprüften

[158] Die Berechnung des Korrelationskoeffizienten nach Pearson ergibt keine signifikanten Zusammenhänge: $p = 0,38$ für die Bachelor-Physik-Lehramtsstudierenden, $p = 0,10$ für die Master-Physik-Lehramtsstudierenden und $p = 0,05$ für die Sachunterrichtsstudierenden.

Normalverteilung beider Testwerte) wird die Varianzhomogenität der Testwerte in den drei betrachteten Gruppen Bachelor- und Master-Physik-Lehramtsstudierende sowie Studierende des Sachunterrichts mit Wahlfach Physik geprüft (vgl. Rasch, Friese, Hofmann und Naumann, 2014b, S. 30f.), wobei für die nachfolgend betrachteten Testwerte beider Testinstrumente auf Grundlage von durchgeführten Levene-Tests von Varianzhomogenität ausgegangen werden kann.[159]

Die ANOVA liefert für beide Testwerte zur Reflexionsfähigkeit keine signifikanten Unterschiede zwischen den Gruppen ($F(2;36) = 2,38; p = 0,11$ für den Reflexionstest und $F(2;36) = 2,78; p = 0,08$ für den Performanztest). Analoge Analysen zu den Subskalen zeigen, dass ausschließlich für den Testwert zum fachlichen Wissen im Performanztest ein signifikanter Unterschied festgestellt werden kann ($F(2;36) = 12,63; p < 0,01$).[160] Eine Post-hoc-Analyse zum paarweisen Vergleich der Gruppen mit Bonferroni-korrigierten t-Tests zeigt, dass signifikante Unterschiede zwischen der Gruppe der Sachunterrichtsstudierenden und den Bachelor-Physik-Lehramtsstudierenden ($p < 0,01$) sowie den Master-Physik-Lehramtsstudierenden ($p < 0,01$) vorliegen, wobei sowohl die Physik-Lehramtsstudierenden im Bachelor ($M = 6$ %; $SD = 4$ %) als auch im Master ($M = 7$ %; $SD = 2$ %) signifikant höhere Testwerte zur Reflexionsfähigkeit mit fachlichem Bezug erzielen als die Sachunterrichtsstudierenden ($M = 2$ %; $SD = 2$ %). Die Gruppen der Bachelor- und Master-Physik-Lehramtsstudierenden unterscheiden sich auch hier nicht signifikant.

17.4.3 Bewertung des Indikators 3.4

Der Vergleich von Reflexions- und Reflexionsperformanztest zeigt auf, dass die Studierenden im Reflexionstest deutlich höhere Punktzahlen erreichen, das Testinstrument also leichter zu sein scheint. Da beide Instrumente auf die Reflexion der gleichen Unterrichtsstunde abzielen, könnten hier Vorteile durch Erinnerungseffekte vermutet werden. Da die Studierenden im Reflexionsperformanztest frei reflektieren, können sich diese - abgesehen von eigenen Überlegungen - nicht auf konkrete Reflexionen bzw. Rückmeldungen an den fiktiven Mitpraktikanten beziehen, sondern sind beschränkt auf die Inhalte der (ggf. in der Bearbeitung noch folgenden) Szenen. Betrachtet man außerdem die Gesamtstichproben, für die Bearbeitungen von jeweils einem der Testinstrumente vorliegen, wird deutlich, dass dieser Unterschied unabhängig von der konkreten Stichprobe auftritt: Während in den $N = 192$ Bearbeitungen des Reflexionstests durchschnittlich $M = 66$ % ($SD = 16$ %) der Punkte erreicht wurden, wurden im Performanztest (im Projekt *ProfiLeP+*) durchschnittlich $M = 14$ % ($SD = 5$ %; $N = 76$) der Punkte vor bzw. $M = 15$ % ($SD = 6$ %; $N = 57$) der Punkte nach dem Praxissemester erreicht. Außerdem kann argumentiert werden, dass der Reflexionsperformanztest durch

[159] Die Ergebnisse der Levene-Tests befinden sich in Tabelle I.3 im Anhang auf Seite 481).
[160] Die Ergebnisse aller ANOVAs befinden sich in Tabelle I.3 im Anhang auf Seite 481).

das offene Aufgabenformat, in dem im Sinne des Noticing bereits die Themen der Reflexion durch die Studierenden ausgewählt werden müssen, eine komplexere Anforderung enthält. Das unterschiedlich gute Abschneiden der Studierenden in beiden Testinstrumenten könnte also als Artefakt der Erhebungsmethode, darunter u. a. die Einbindung von Noticing und die Möglichkeit im Performanztest, Reflexionen zu einzelnen Videovignetten auszulassen, interpretiert werden, stellt aber kein grundsätzliches Problem für die Untersuchung des Zusammenhangs zwischen beiden Testwerten dar.[161]

Die Untersuchung des Zusammenhangs zwischen beiden Testwerten ergab eine mittlere bis annähernd hohe Korrelation von $r_{\text{Pearson}} = 0,49$ ($p < 0,01$). Da beide Instrumente zur Erfassung der Reflexionsfähigkeit dienen, könnte eine hohe Korrelation zwischen den jeweiligen Testwerten erwartet werden. Gleichzeitig können Einflüsse durch die unterschiedlichen Testformate und die diskutierten, damit verbundenen veränderten Anforderungen nicht ausgeschlossen werden, die im Rahmen der Erfassung der professionellen Unterrichtswahrnehmung zu einer lediglich geringen Korrelation von $r_{\text{Pearson}} = 0,20$ ($p < 0,05$) zwischen offenem und geschlossenem Antwortformat beizutragen scheinen (Eßling, Sunder, Steffensky & Meschede, 2024). Bartels (2018, S. 139ff.) verglich zwei Verfahren zur Erfassung der Erklärfähigkeit von Studierenden: einen Performanztest, in dem die Teilnehmenden trainierten „Schüler*innen“ physikalische Sachverhalte erklärten, und einen auf MC-Aufgaben basierenden „performanznahen“ Test. Er fand für eine Stichprobe von $N = 12$ Teilnehmenden eine Korrelation von $r_{\text{Pearson}} = 0,54$ ($p < 0,05$) (Bartels, 2018, S. 140), die in etwa in der gleichen Größenordnung zur Korrelation zwischen den hier untersuchten Testinstrumenten zur Reflexionsfähigkeit liegt. Die Höhe der vorliegenden Korrelation erscheint vor dem Hintergrund dieser Ergebnisse nachvollziehbar. Inwiefern sich die kognitiven Prozesse bei der Bearbeitung dieser beiden Testformate bzw. -typen unterscheiden oder ggf. durch motivationale Faktoren (z. B. die Gewissenhaftigkeit bei der Bearbeitung (Eßling, Sunder, Steffensky & Meschede, 2024)) konfundiert sind, wäre z. B. im Rahmen einer vergleichenden Think-Aloud-Studie zu untersuchen.

Dass zwischen den verschiedenen Gruppen der Physik-Lehramtsstudierenden sowie Sachunterrichtsstudierenden mit Wahlfach Physik keine signifikanten Unterschiede in Bezug auf die Testwerte und Zusammenhänge zwischen den Testwerten beider Instrumente festgestellt wurden, ist zudem ein Hinweis darauf, dass die Ergebnisse über das Physik-Lehramtsstudium hinaus generalisierbar für Studierende sind, die zumindest Physik als (Wahl-)Fach in ihrem Studium belegen. Der einzige Unterschied zwischen den Gruppen betrifft den Testwert für Reflexionen mit Bezug zum fachlichen Wissen im Performanztest. Hier erreichen die Sachunterrichtsstudierenden signifikant weniger Punkte als die Physik-Lehramtsstudierenden. Vor dem Hintergrund der Studienstruktur beider Studiengänge ist dies inhaltlich nach-

[161] Der Zusammenhang zwischen den Fähigkeiten zum Noticing und zur Unterrichtsreflexion wird in Kapitel 22 näher untersucht.

vollziehbar: Während davon auszugehen ist, dass die Physik-Lehramtsstudierenden im Bachelor und Master bereits fachwissenschaftliche Veranstaltungen zur Physik (insbesondere zur Mechanik) besucht haben, belegten die Sachunterrichtsstudierenden zum Zeitpunkt der Studie ihre erste und inhaltlich weniger umfangreiche Physik-Veranstaltung (u. a. zur Mechanik), sodass vermutet werden kann, dass sie dementsprechend weniger (fachliches) Vorwissen aufweisen. Für die valide Interpretation der Ergebnisse des Reflexionstests ist dies aber nicht gefährdend.

Insgesamt kann also aufgrund des gefundenen Zusammenhangs zwischen den Testwerten beider Testinstrumente, der zumindest an der Grenze zur erwarteten hohen Korrelation liegt, Indikator 3.4 als erfüllt angesehen werden.

17.5 Evaluation des Zusammenhangs zu Lerngelegenheiten (Indikator 3.5)

Gemeinsam mit dem Reflexionstest wurde ein Fragebogen zur Demografie und zu Vorerfahrungen der Teilnehmenden (insb. der Studierenden) eingesetzt (s. Kapitel 13.1). In diesem Kapitel werden Unterschiede in den Testwerten in Abhängigkeit dieser Aspekte untersucht. So sollen Erkenntnisse zur Bewertung von Indikator 3.5 *„Unterschiedliche Lerngelegenheiten der Teilnehmenden führen zu nachvollziehbaren Unterschieden in den Testwerten.“* gewonnen werden. Da der Fragebogen zur Demografie aus dem vorangegangenen Projekt *ProfiLeP+* übernommen wurde, liegen für die einzelnen erhobenen Variablen keine (empirisch) begründeten Vermutungen über erwartete Zusammenhänge vor. Grundsätzlich ist davon auszugehen, dass unterschiedliche (bzw. mehr) Lerngelegenheiten zu unterschiedlichen (bzw. höheren) Testwerten führen. Gleichzeitig sollten Unterschiede in den Testwerten für die valide Interpretation nicht abhängig von für das Konstrukt irrelevanten Faktoren wie z. B. dem Geschlecht sein.

Zusätzlich kann für eine Teilstichprobe, die den Reflexionstest zweifach im Abstand von etwa einem Semester bearbeitet hat, die längsschnittliche Entwicklung der Reflexionsfähigkeit untersucht werden, wobei aufgrund der Teilnahme an physikdidaktischen Grundlagenvorlesungen bzw. dem Absolvieren des Praxissemesters davon ausgegangen werden kann, dass zwischen beiden Messzeitpunkten (MZPs) Lerngelegenheiten für die Studierenden bestanden.

17.5.1 Vorgehen bei den Analysen

Da der Fragebogen zur Demografie für die Teilnehmenden in unterschiedlichen Ausbildungsphasen (Studium, Referendariat bzw. Sonstiges) unterschiedlich umfangreich ist, liegen für die nachfolgenden Untersuchungen von Zusammenhängen

zwischen erreichten Testwerten und demografischen Variablen unterschiedlich große Datensätze vor; es werden allerdings jeweils alle vorliegenden Fälle berücksichtigt und die Stichprobengröße wird jeweils dazu benannt.

Für die Untersuchungen wird jeweils nur der Testwert des Reflexionstests berücksichtigt und keine Aufschlüsselung auf Bewertungen und Alternativen vorgenommen, da nicht davon auszugehen ist, dass einzelne Variablen unterschiedlich Einfluss auf die Teilfähigkeiten nehmen. Abhängig vom Skalenniveau der jeweils betrachteten Variablen werden Pearson-Korrelationen, (Welch-)t-Tests oder (Welch-)ANOVAs (bzw. Spearman-Korrelationen und Wilcoxon-Tests als nicht parametrische Alternativen) sowie bei signifikanten Unterschieden Cohens d als Maß der Effektstärke berechnet.[162] Die Bewertung der Effektstärken erfolgt entsprechend der Richtwerte von Cohen (1988; 1992) bzw. darauf aufbauend Döring und Bortz (2016, S. 820).

Ergänzend zu den quasi-längsschnittlichen Untersuchungen auf Grundlage der erhobenen demografischen Variablen liegen für eine Teilstichprobe von $N = 41$ Studierenden Testbearbeitungen zu zwei MZPs im Abstand von etwa einem Semester vor. Diese Stichprobe setzt sich zusammen aus $N = 25$ Studierenden, die zwischen den Bearbeitungen das Praxissemester absolvieren und $N = 16$ Bachelorstudierenden, die zwischen beiden MZPs physikdidaktische Grundlagenveranstaltungen besuchten. Für beide Gruppen kann also davon ausgegangen werden, dass Lerngelegenheiten für Unterrichtsreflexion bestehen, sodass eine Verbesserung der Reflexionsfähigkeit zu erwarten wäre. Zur Bestimmung von Unterschieden in den durchschnittlichen Testwerten zu beiden MZPs wird ein gepaarter t-Test berechnet und ergänzend als Maß für die Effektstärke ein für gepaarte Stichproben angepasstes $d_{\text{Cohen},z}$ nach Lakens (2013) berechnet.[163]

17.5.2 Zusammenhänge zu demografischen Variablen (Quasi-Längsschnitt)

Den nachfolgenden Analysen liegen abhängig von den betrachteten Variablen jeweils leicht unterschiedliche Stichproben zugrunde. Aufgrund des Umfangs der notwendigen Prüfung von Voraussetzungen für die statistischen Tests werden diese im Anhang vollständig berichtet (Kapitel I.3.1 ab Seite 482). Die Verteilung der

[162] Zur Auswahl der entsprechenden Verfahren wurde auf Übersichtswerke zu quantitativen Methoden zurückgegriffen, die jeweils einen Überblick über Voraussetzungen für die jeweiligen Testverfahren bereitstellen (u. a. Rasch, Friese, Hofmann und Naumann (2014a), Rasch, Friese, Hofmann und Naumann (2014b) und Field (2009)).

[163] Zur Orientierung wird auch die Größe dieser angepassten Effektstärken weiterhin nach Cohen (1988) bewertet (Lakens, 2013). Da die so bestimmte Effektstärke für gepaarte t-Tests durch die Stärke der Abhängigkeit der Ergebnisse beider MZPs beeinflusst wird, ist der direkte Größenvergleich mit den zu anderen Tests bestimmten Effektstärken allerdings nicht zulässig (Rasch, Friese, Hofmann & Naumann, 2014b, S. 63f.).

Testwerte des Reflexionstests kann auf Grundlage des Kolmogorow-Smirnow-Tests - wie schon für die Gesamtstrichprobe zu Indikator 3.2 in Kapitel 17.1 berichtet - i. d. R. nur knapp als normalverteilt angenommen werden. Um Zusammenhänge zwischen den Testwerten und hier betrachteten Variablen mit der nötigen Vorsicht zu betrachten, werden ggf. zusätzlich nicht parametrische Tests berechnet; diese führen zu prinzipiell gleichen Ergebnissen und werden hier nicht zusätzlich berichtet (sie sind ebenfalls im Anhang in Kapitel I.3.1 ab Seite 482 aufgeführt).

Zunächst werden Unterschiede in den Testwerten in Abhängigkeit von der aktuellen Studien- bzw. Ausbildungsphase untersucht. Dazu wird eine einfaktorielle Varianzanalyse für die drei Phasen Bachelorstudium, Masterstudium und Referendariat berechnet ($N = 187$; ausgeschlossen werden die Bearbeitungen von Personen mit dem Studienziel 1. Staatsexamen sowie mit der Angabe „Sonstige“). Voraussetzung für eine ANOVA ist neben der gegebenen Normalverteilung der Testwerte Varianzhomogenität in den Gruppen. Da diese für die drei betrachteten Gruppen nicht vorliegt, wird eine Welch-ANOVA berechnet. Diese liefert mit $F(2; 35,48) = 10,52$; $p < 0,01$ ein signifikantes Ergebnis. In der post-hoc-Analyse mit dem Games-Howell-Test und Bonferroni-Korrektur zeigt sich, dass die Masterstudierenden hochsignifikant besser abschneiden als die Bachelorstudierenden ($\Delta M = 11$ %; 95 %-KI : [5 %; 165 %]); $p < 0,001$).

Für die Teilstichprobe der Studierenden (mit aktuellem Studienziel Bachelor, Master oder Staatsexamen; $N = 178$) kann als ein weiteres Maß für den Studienfortschritt der Zusammenhang zum Fachsemester untersucht werden. Da die Verteilung der Fachsemester nicht als normalverteilt angenommen werden kann, wird ρ_{Spearman} bestimmt. Es ergibt sich eine niedrige Korrelation von $\rho_{\text{Spearman}} = 0,18$ ($p < 0,05$).[164] Unklar ist allerdings, welche Veranstaltungen die Studierenden abgeschlossen haben, sodass der Studienfortschritt nur näherungsweise als Indikator für Lerngelegenheiten gewertet werden kann.

Zusätzlich wird das Vorhandensein von praktischen Erfahrungen im Unterrichten z. B. im Rahmen von Praktika erhoben. Die Teilstichprobe der Studierenden ($N = 178$) setzt sich zusammen aus $N = 47$ Studierenden, die bislang kein und $N = 131$ Studierenden, die (mindestens) ein Pflichtpraktikum im Studium absolviert haben. Zwischen den Testwerten beider Gruppen liegt ein signifikanter Unterschied ($t(176) = -2,70; p < 0,01$) vor, wobei die Gruppe mit Praktikumserfahrung ($M = 68$ %; $SD = 16$ %) gegenüber der Gruppe ohne Erfahrung ($M = 61$ %; $SD = 16$ %) mit kleinem Effekt besser abschneidet ($d_{\text{Cohen}} = 0,46$). Für den Umfang an absolvierten Pflichtpraktika (erfasst anhand der Wochen im Praktikum) und Testwerten ist allerdings kein korrelativer Zusammenhang zu beobachten ($p > 0,05$). Für (zusätzlich) freiwillig absolvierte Schulpraktika kann

[164] Unter Ausschluss der Studierenden des Grundschullehramts, die diese Frage möglicherweise nicht einheitlich beantworten, da sie das Wahlfach Physik zwar im ersten Semester, das übergeordnete Fach Sachunterricht allerdings in fortgeschritteneren Semestern studieren, ergibt sich ein weiterhin niedriger Zusammenhang von $\rho_{\text{Spearman}} = 0,21$ ($p < 0,01$)

kein signifikanter Unterschied festgestellt werden ($t(176) = 1,14; p = 0,26$). Mit $N = 20$ Studierenden hat allerdings auch nur ein geringer Anteil der Stichprobe überhaupt freiwillige Praktika absolviert.

Ein Teil der Studierenden ($N = 51$) gab außerdem an, als Vertretungslehrkraft zu arbeiten bzw. gearbeitet zu haben. Ein Unterschied in den Testwerten zu den Studierenden, die keine Erfahrung als Vertretungslehrkräfte haben, lässt sich nicht feststellen ($t(75,49) = -0,21; p \approx 0,84$; da für die beiden Gruppen keine Varianzhomogenität vorliegt, wird ein Welch-t-Test berechnet).

Zuletzt wird erfasst, ob die Studierenden außerschulische bzw. universitäre Lehrerfahrung haben.[165] In beiden Fällen lassen sich signifikante Unterschiede zwischen den Gruppen mit und ohne Erfahrung ausmachen: Die $N = 153$ Studierenden, die angeben, außerschulische Lehrerfahrung gesammelt zu haben, schneiden mit moderater Effektstärke $d_{\text{Cohen}} = 0,65$ durchschnittlich besser ab ($M = 67$ %; $SD = 15$ %) als die $N = 25$ Studierenden ohne solche Erfahrung ($M = 57$ %; $SD = 19$ %) ($t(29,13) = -2,78; p < 0,01$; aufgrund mangelnder Varianzhomogenität wird ein Welch-t-Test berechnet). Analog gilt für die $N = 43$ Studierenden, die angeben, universitäre Lehrerfahrung gesammelt zu haben ($M = 73$ %; $SD = 12$ %), dass sie gegenüber den $N = 135$ übrigen Studierenden ($M = 64$ %; $SD = 17$ %) besser abschneiden ($t(178) = -2,98; p < 0,01$); ebenfalls mit moderater Effektstärke von $d_{\text{Cohen}} = 0,57$.

Da alle Studierenden Lehramtsstudierende sind, wird keine Analyse zur Abhängigkeit der Testwerte davon durchgeführt, ob der Studiengang ein Lehramts- oder reiner Physikstudiengang ist. Unterschieden werden kann allerdings zwischen Physik-Lehramtsstudierenden (für die weiterführende Schule; $N = 159$) und Grundschullehramtsstudierenden mit Physik als Wahlfach ($N = 19$). Ein Test auf Unterschiede in den Mittelwerten ergibt keine signifikanten Unterschiede zwischen beiden Gruppen ($t(176) = 0,04; p = 0,97$). Zur Untersuchung möglicher Zusammenhänge zum Zweitfach können die Bearbeitungen von $N = 176$ Studierenden herangezogen werden.[166] Für die Analyse werden drei Gruppen unterschieden: $N = 25$ Studierende mit (mindestens) einem naturwissenschaftlichen Zweitfach, $N = 111$ Studierende mit Mathematik als Zweitfach (die größte Gruppe) und $N = 40$ Studierende mit sonstigen Zweitfächern. Eine ANOVA zeigt keine signifikanten Unterschiede zwischen den Testwerten der Gruppen ($F(2,173) = 1,10; p = 0,33$).

Darüber hinaus werden weitere, nicht direkt mit dem Studium oder Ausbildungsziel (Lehrkraft) zusammenhängende Variablen untersucht. Dazu gehören für die Gesamtstichprobe ($N = 192$) das Alter sowie Geschlecht. Während zwischen Alter

[165] Hier ist davon auszugehen, dass universitäre Lehrerfahrung i. d. R. als Teil außerschulischer Lehrerfahrung verstanden wird, da keine explizite Abgrenzung erfolgt.

[166] Vier Studierende gaben kein Zweitfach an. Für zwei der vier Studierenden konnte dieses aus der zweiten Testbearbeitung der jeweiligen Person ergänzt werden; zwei Testbearbeitungen wurden ausgeschlossen.

und Testwerten keine signifikante Korrelation vorliegt ($p > 0,05$), lässt sich ein kleiner signifikanter Unterschied in den Testwerten nach Geschlechtsidentität[167] feststellen ($t(191) = 2,02$; $p < 0,05$; $d_{\text{Cohen}} = 0,30$): Weibliche Teilnehmende ($M = 68$ %; $SD = 14$ %) schneiden hier besser ab als männliche Teilnehmende ($M = 64$ %; $SD = 17$ %).[168]

Für die Studierenden ($N = 178$) können außerdem Variablen der Schullaufbahn untersucht werden. Eine Übersicht über die durchgeführten Untersuchungen liefert Tabelle 17.1.[169] Insgesamt lassen sich Zusammenhänge zwischen dem Abschneiden im Testinstrument und der Abitur- sowie letzten Deutschnote finden[170], wobei der negative Zusammenhang zur Abiturnote dadurch zu erklären ist, dass - den Testwerten entgegengesetzt - eine bessere Note durch niedrigere Werte repräsentiert ist.

[167] Gefragt wurde nach dem „Geschlecht“; neben weiblich und männlich konnte allerdings die Antwortoption „divers“ gewählt werden. Von dieser Option hat keine Person Gebrauch gemacht.

[168] Wie zu Beginn des Kapitels beschrieben, werden - da die Normalverteilung der Daten auf Grundlage des Kolmogorow-Smirnow-Tests nur knapp unterstellt werden kann - zusätzlich nicht-parametrische Tests gerechnet. Die Prüfung von Unterschieden in den Testwerten abhängig vom Geschlecht mit dem Wilcoxon-Test verfehlt die Grenze zur Signifikanz nur knapp ($p = 0,051$), sodass weiterhin tendenziell von einem Unterschied in den Testwerten abhängig von der Geschlechtsidentität auszugehen ist.

[169] Erhoben wurde zusätzlich, welche Ausbildung(en) die Studierenden vor Aufnahme des Studiums abgeschlossen haben. Diese Variablen werden allerdings nicht in die Analyse einbezogen, da diverse Studierende offenbar keine vollständigen Angaben machten (z. B. häufiger angaben, ohne (Fach-)Abitur oder abgeschlossene Ausbildung ein Studium abgeschlossen zu haben), sodass der Informationsgehalt der Abfrage unklar ist. Auch auf die Untersuchung von Mittelwertsunterschieden abhängig davon, ob das Abitur in Deutschland erreicht wurde ($N = 175$ gegenüber $N = 3$ ausländischen Abschlüssen) und Mathematik ein Prüfungsfach im Abitur war ($N = 173$ gegenüber $N = 5$ Personen ohne Mathematik als Abiturprüfungsfach), wird aufgrund der stark unterschiedlichen Gruppengrößen verzichtet.

[170] Da für die Testwerte nur knapp von einer Normalverteilung ausgegangen werden kann, wurde für den Zusammenhang zwischen Testwerten und letzter Deutschnote zusätzlich ρ_{Spearman} bestimmt, wobei die Grenze zur Signifikanz für diesen Zusammenhang mit $p = 0,23$ verfehlt wird; der Zusammenhang zwischen beiden Testwerten und letzter Deutschnote kann also nicht gesichert festgestellt werden.

Variable	Art des Tests	Stichprobe	Ergebnis
Abiturnote	Korrelation (r_{Pearson})	$N = 175$ Studierende mit in Deutschland erlangter Hochschulreife	$r_{\text{Pearson}} = -0,22$ ($p < 0,01$)
Abitur nach Klasse 12 (G8) oder 13 (G9)	Wilcoxon-Test	$N = 165$ Studierende mit in Deutschland erlangter Hochschulreife unter Ausschluss der Option „Sonstiges“	n. s. ($t(163) = 1,36$; $p = 0,17$)
Letzte Physiknote (0-15 Punkte)	Korrelation (ρ_{Spearman})	$N = 178$ Studierende	n. s. ($p = 0,76$)
Letzte Mathematiknote (0-15 Punkte)	Korrelation (ρ_{Spearman})	$N = 178$ Studierende	n. s. ($p = 0,23$)
Letzte Deutschnote (0-15 Punkte)	Korrelation (r_{Pearson})	$N = 178$ Studierende	$r_{\text{Pearson}} = 0,15$ ($p < 0,05$)
Physik als Prüfungsfach	Welch-t-Test	$N = 178$ Studierende ($N = 96$ mit und $N = 82$ ohne Physik als Prüfungsfach)	n. s. ($t(149) = -1,51$; $p = 0,13$)

Tabelle 17.1: Untersuchung der Zusammenhänge zwischen den Testwerten und Variablen der Demografie mit Bezug zur schulischen Ausbildung (*n. s. - nicht signifikant*)

17.5.3 Untersuchung der Testwerte im Längsschnitt über ein Semester

Für insgesamt $N = 41$ Studierende liegen Testwerte zu zwei MZPs im Abstand von etwa einem Semester vor, in dem die Bachelorstudierenden ($N = 15$) eine physikdidaktische Grundlagenvorlesung besuchten und die Masterstudierenden $N = 26$) das Praxissemester (inklusive Begleitveranstaltung) absolvierten. Zur Untersuchung von Unterschieden zwischen beiden MZP wird ein gepaarter t-Test berechnet.[171] Es zeigt sich, dass die Studierenden zum zweiten MZP ($M = 71$ %; $SD = 16$ %) zwar signifikant besser abschneiden als zum ersten ($M = 66$ %; $SD = 16$ %) ($t(40) = -3,70; p < 0,01$), wobei mit Effekt mittlerer Höhe festgestellt werden kann ($d_{\text{Cohen},z} = 0,58$).

Vergleichbare Ergebnisse ergeben sich auch für die Teilstichproben der Bachelor- und Masterstudierenden einzeln: In beiden Gruppen sind die Testwerte zum zweiten MZP höher[172], wobei für die Gruppe der Bachelorstudierenden ein mittlerer Effekt vorliegt (Bachelorstudierende: $t(15) = -2,88; p < 0,05$; $d_{\text{Cohen},z} = 0,72$ und für die Gruppe der Masterstudierenden ein noch kleiner, an der Schwelle zum mittleren, Effekt: $t(24) = -2,38; p < 0,05$; $d_{\text{Cohen},z} = 0,48$).

17.5.4 Bewertung des Indikators 3.5

Ein Zusammenhang zwischen unterschiedlichen Lerngelegenheiten für Unterrichtsreflexion und dem Abschneiden im Reflexionstest stellt ein Argument für die Validität der Testwertinterpretation als Maß für Reflexionsfähigkeit dar und sollte insbesondere auch für den mehrfachen Einsatz des Reflexionstests zur Feststellung individueller Entwicklungen gegeben sein, um mögliche Fortschritte erkenntlich zu machen. Im Rahmen dieser Validierungsstudien wurden demografische Variablen als indirekter Zugang zu möglichen Lerngelegenheiten erhoben und untersucht, welche Zusammenhänge zu den Testwerten bestehen. Hierbei zeigte sich ein Zusammenhang zwischen den Testwerten und verschiedenen Variablen mit Bezug zum Physik-Lehramtsstudium: zum allgemeinen Studienfortschritt erfasst durch die aktuelle Studienphase und das Fachsemester im Physikstudium sowie dem Umfang bisher absolvierter Pflichtpraktika[173]. Dass kein Unterschied zwischen den Test-

[171] Die Prüfung der Normalverteilung der Differenzen der Testwerte zu beiden MZP ist für die Stichproben in Kapitel I.3.2 im Anhang auf Seite 489 dargestellt.

[172] Die Bachelorstudierenden erreichen zum zweiten MZP durchschnittlich Testwerte von $M = 65$ % ($SD = 19$ %) gegenüber durchschnittlich $M = 58$ % ($SD = 20$ %) zum ersten MZP, die Masterstudierenden durchschnittlich $M = 75$ % ($SD = 11$ %) gegenüber $M = 71$ %; $SD = 11$ %).

[173] Hierbei ist von einem Zusammenhang zwischen absolvierten Pflichtpraktika und ihrem Umfang sowie dem Studienfortschritt auszugehen. Es liegen allerdings keine detaillierteren Informationen über die von den Studierenden absolvierten Veranstaltungen vor, sodass mögliche Zusammenhänge hier nicht weiter aufgeschlüsselt werden können.

werten der Masterstudierenden und Referendar*innen festgestellt wurde, könnte in der vergleichsweise kleinen Stichprobengröße der Gruppe der Referendar*innen sowie in ihrer Zusammensetzung begründet liegen, da bekanntermaßen unterschiedliche Zugangswege zum Referendariat vertreten sind (ob hier ein Einfluss besteht, kann allerdings nicht überprüft werden, da der individuelle Zugang zum Referendariat nicht im Reflexionstest erfasst wurde). Auch schnitten Studierende, die außerschulische Lehrerfahrung (z. B. im Rahmen von Nachhilfeunterricht oder universitären Tutorien) haben, besser ab als Studierende ohne diese Erfahrung. Diese beobachteten Zusammenhänge stützen den Indikator 3.5 *„Unterschiedliche Lerngelegenheiten führen zu nachvollziehbaren Unterschieden in den Testwerten.“*, da die Studierenden jeweils besser abschneiden, wenn davon auszugehen ist, dass ihnen mehr Lerngelegenheiten zur Verfügung standen (bei der Erfassung des Studienfortschritts durch den Besuch weiterer universitärer Veranstaltungen sowie im Rahmen praktischer Erfahrungen durch die Möglichkeit, erlebte Situationen im Pflichtpraktikum ggf. begleitet zu reflektieren).

Keine signifikanten Unterschiede ließen sich aufgrund freiwillig absolvierter Schulpraktika und der Arbeit als Vertretungslehrkraft feststellen. Das Vorliegen solcher Erfahrungen gab mit $N = 20$ (freiwillige Schulpraktika) bzw. $N = 51$ (Arbeit als Vertretungslehrkraft) allerdings jeweils nur eine Minderheit der Teilnehmenden an. Beide Erfahrungen könnten durch die Beobachtung und Durchführung von Unterricht Lerngelegenheiten für Unterrichtsreflexionen bieten. Im Gegensatz zu den bisher diskutierten Lerngelegenheiten ist allerdings davon auszugehen, dass sie nicht (universitär) begleitet sind, was vor dem Hintergrund der Relevanz von Reflexionsgesprächen mit universitären Mentorierenden für die Entwicklung der Reflexionsfähigkeit im Praxissemester (Kulgemeyer, Kempin & Weißbach, 2021) auch für andere Praxiserfahrungen als förderlich anzunehmen ist. Auch bleibt u. a. unklar, in welchem Rahmen (z. B. vor oder im Studium) die freiwilligen Praktika absolviert wurden und in welchem Fach die Studierenden als Vertretungslehrkräfte arbeiteten; ein Bezug zu Physikunterricht muss nicht gegeben sein. Es könnte vermutet werden, dass der vorliegende negative Zusammenhang zwischen Testwerten und Arbeit als Vertretungslehrkraft damit zusammenhängt, dass sich der Fokus der Studierenden durch das selbstständige Unterrichten von den zentralen fachlichen und fachdidaktischen Aspekten entfernt, die im Reflexionstest vorrangig reflektiert werden; eine vergleichbare Vermutung für die Entwicklung der Reflexionsfähigkeit hin zu einem pädagogischen Fokus im Praxissemester stellen Kempin, Kulgemeyer und Schecker (2020; 2019) auf.

Dass keine signifikanten Unterschiede zwischen den Physik-Lehramtsstudierenden und den Grundschullehramtsstudierenden mit Wahlfach Physik festgestellt wurden, stützt die Herangehensweise, beide Gruppen in den Analysen nicht zu differenzieren. Zwar liegt eine abweichende Studienstruktur vor; die Sachunterrichtsstudierenden haben zum Zeitpunkt der Erhebung insbesondere die fachlichen und fachdidaktischen Veranstaltungen zur Mechanik nicht abgeschlossen, haben aber im Rahmen

anderer Veranstaltungen (z. B. im Rahmen erziehungswissenschaftlicher Veranstaltungen) möglicherweise weitere und im Studienverlauf zunehmende Lerngelegenheiten für Unterrichtsreflexion durchlaufen, die für die Physik-Lehramtsstudierenden nicht zur Verfügung stehen.

Die übrigen erfassten demografischen Variablen (u. a. Alter, Geschlecht und Abiturnote) erscheinen unabhängig vom Studium und sollten demnach keine Zusammenhänge zum Abschneiden im Testinstrument aufweisen. Entgegen dieser Einschätzung ließen sich allerdings tendenziell Zusammenhänge zur Geschlechtsidentität, der Abiturnote sowie der letzten Deutschnote der Schulzeit feststellen, die als mögliche Hürde für die valide Interpretation der Testwerte festzuhalten sind und deren Einfluss ggf. weiter untersucht werden sollte. Grundsätzlich sind die Zusammenhänge zwischen Abitur- bzw. letzter Deutschnote und den erreichten Testwerten nicht als besondere Bedrohung der Validität herauszuheben; sie weisen niedrige Höhen auf und passen damit grundsätzlich zu anderen Befunden, in denen Schulnoten sich sogar als prädiktiv für den Studienerfolg erwiesen (z. B. Schild, 2021, S. 166f.; Sorge, Petersen und Neumann, 2016 und zur fächerübergreifenden Meta-Analyse von Trapmann, Hell, Weigand und Schuler, 2007). Der vorliegende Zusammenhang zur letzten Deutschnote könnte außerdem vor dem Hintergrund des hohen Textumfangs in den MC-Aufgaben durch Lesefähigkeit geprägt sein.

Die beobachtete Entwicklung der Testwerte im Laufe eines Semesters, das bekanntermaßen (unterschiedliche) Lerngelegenheiten für Bachelor- und Masterstudierende enthält, nämlich zumindest den Besuch einer physikdidaktischen Grundlagenveranstaltung bzw. das Praxissemester inklusive physikdidaktischer Begleitung, deutet außerdem auf eine Generalisierbarkeit der Sensitivität des Testinstruments über theoretische und praktische Lerngelegenheiten hin.

Die durchgeführten Analysen zum Indikator 3.5 *„Unterschiedliche Lerngelegenheiten der Teilnehmenden führen zu nachvollziehbaren Unterschieden in den Testwerten.“* enthalten insgesamt diverse Befunde, die für dessen Gültigkeit sprechen. Dazu gehören insbesondere die Zusammenhänge der Testwerte zu verschiedenen Lerngelegenheiten (u. a. dem Studienfortschritt oder absolvierten Pflichtpraktika). Aufgrund der zusätzlich festgestellten Zusammenhänge zu einigen nicht studienbezogenen Variablen (u. a. Geschlecht und Deutschnote) und insbesondere der nicht weiter aufzulösenden Ursachen für diese Zusammenhänge ist der Indikator 3.5 nicht umfassend als bestätigt anzusehen; es bedarf weiterer Untersuchungen zur Klärung der Ursachen für diese Zusammenhänge und Vermeidung konstruktirrelevanter Varianz.

17.6 Evaluation des Zusammenhangs zum fachdidaktischen Wissen (Indikator 3.6)

Auf theoretischer Ebene ist ein Zusammenhang zwischen der Fähigkeit, Unterricht zu reflektieren und anderen Bestandteilen der professionellen Handlungskompetenz (z. B. dem fachlichen und fachdidaktischen Wissen oder Einstellungen gegenüber der Reflexion) anzunehmen (vgl. z. B. Riese und Reinhold, 2010). Gemäß dem Fokus auf die Reflexion von Physikunterricht wird hier das fachdidaktische Wissen als Facette des Professionswissens herangezogen, auf der auch der Schwerpunkt für die Reflexionen im Reflexionstest liegt. Dazu wird Indikator 3.6 *„Die Testwerte des Reflexionstests weisen einen Zusammenhang zum fachdidaktischen Wissen auf.“* untersucht.

17.6.1 Studiendesign

Zur Untersuchung des Zusammenhangs zwischen Testwerten im Reflexionstest und dem fachdidaktischen Wissen führten Studierende vor bzw. zu Beginn des Praxissemesters und am Ende bzw. nach Ende des Praxissemesters zusätzlich zum Reflexionstest den Test zum fachdidaktischen Wissen durch. Dazu wurde der digitalisierte und in ein MC-Format überführte Test zum fachdidaktischen Wissen (FDW-Test) verwendet (Jordans, Zeller, Große-Heilmann & Riese, 2022), der auf dem Testinstrument von Gramzow (2015, S. 102ff.) basiert (s. Kapitel 13.4). Insgesamt liegen von $N = 33$ Studierenden Daten für den FDW-Test zu mindestens einem der Zeitpunkte vor, wobei beide Tests zu Beginn des Praxissemester von $N_{\text{Prä}} = 18$ Studierenden und zum Ende des Praxissemesters von $N_{\text{Post}} = 23$ Studierenden bearbeitet wurden (insgesamt $N = 41$ Datensätze).[174] Die Daten wurden im Sommersemester 2022 (an den Universitäten Aachen, Bremen und Paderborn) und im Sommersemester 2023 (an den Universitäten Aachen und Bremen) erhoben. Zur Untersuchung der Zusammenhänge werden Pearson-Korrelationen bestimmt, wobei sowohl die die Korrelation für alle vorliegenden Fälle sowie die Korrelationen zu beiden MZPs einzeln bestimmt werden, falls hier veränderte Zusammenhänge auftreten (vgl. Kempin, Kulgemeyer und Schecker, 2019).

17.6.2 Ergebnisse

Die vorliegenden Daten können laut Kolmogorow-Smirnow-Test für beide Testinstrumente sowohl für die MZPs einzeln als auch die vorliegende Fälle als

[174] Darunter sind $N = 14$ vollständige Prä-Post-Datensätze von Studierenden, die beide Instrumente zu beiden Zeitpunkten bearbeiteten. Für die Analysen werden allerdings die vollständigen Prä- ($N = 23$ bzw. Post-Datensätze ($N = 18$) verwendet.

normalverteilt angenommen werden.[175] Die Testwerte im Reflexionstest liegen dabei mit $M = 70$ % zum Prä-MZP und $M = 75$ % zum Post-MZP über dem durchschnittlichen Testwert für die Gesamtstichprobe im Reflexionstest und weisen eine geringere Standardabweichung ($SD = 11$ % für beide MZPs) auf (vgl. Kapitel 17.1), was für die hier untersuchte, im Studium fortgeschrittene und eher homogene Kohorte der Praxissemesterstudierenden angemessen erscheint. Die jeweils als Maß der internen Konsistenz bestimmten Werte für Cronbachs α weisen ähnliche Höhen wie für die Gesamtstichprobe im Reflexionstest ($\alpha_{\text{Cronbach}} = 0,95$; s. Kapitel 17.2) bzw. für eine Referenzstichprobe im Rahmen der Evaluation des Tests zum fachdidaktischen Wissen ($\alpha_{\text{Cronbach}} = 0,67$; Jordans, Zeller, Große-Heilmann und Riese, 2022, S. 765) auf. Die Reliabilität des Tests zum fachdidaktischen Wissen ist damit als niedrig zu bewerten (Döring & Bortz, 2016, S. 443), wird aber im Rahmen der Testentwicklung als zufriedenstellend bewertet (Jordans, Zeller, Große-Heilmann & Riese, 2022).

Zur Untersuchung des Zusammenhangs zwischen den Testwerten im Reflexions- und FDW-Test wurden Korrelationen (r_{Pearson}) berechnet. Es ergeben sich zwar in der Tendenz negative, allerdings weder für die beiden MZPs einzeln noch für alle vorliegenden Datensätze aus Reflexions- und FDW-Testwerten zusammen signifikante Zusammenhänge (für alle vorliegenden Datensätze: $r_{\text{Pearson}} = -0,16$; $p = 0,32$[176]). Auch eine feinere Aufschlüsselung auf die jeweils bestimmbaren Teilskalen (*Instruktionsstrategien*, *Schülervorstellungen*, *Experimente und Wissenschaftsverständnis* sowie *fachdidaktische Konzepte* im Test zum fachdidaktischen Wissen) liefert für alle vorliegenden Datensätze zu beiden MZPs ebenfalls keine signifikanten Zusammenhänge zu den Testwerten des Reflexionstests ($p > 0,05$).[177]

17.6.3 Bewertung des Indikators 3.6

Auf Grundlage der berichteten Ergebnisse kann der Indikator 3.6 nicht bestätigt werden, im Rahmen der hier durchgeführten Untersuchung konnte kein Zusammenhang zwischen der Reflexionsfähigkeit und dem fachdidaktischen Wissen gefunden werden. Damit können bisherige Befunde zum Zusammenhang von fachdidaktischem Wissen und Reflexionsfähigkeit nicht reproduziert werden: Im Rahmen des Projekts *ProfiLeP+* wurden vor und nach dem Praxissemester Zusammenhänge

[175] Eine Übersicht über die berechneten Tests sowie nachfolgend berichtete Kennwerte zum Reflexions- und FDW-Test für beide MZPs sowie die Gesamtstichprobe befindet sich in Tabelle I.5 im Anhang auf Seite 490).

[176] Für die MZPs einzeln ergibt sich für den ersten MZP $r_{\text{Pearson}} = -0,31$ ($p = 0,21$) und für den zweiten MZP $r_{\text{Pearson}} = -0,11$ ($p = 0,63$).

[177] Die vorliegenden Testwerte sind für drei der vier Facetten gemäß dem Kolmogorow-Smirnow-Test als normalverteilt anzunehmen (s. Tabelle I.6 im Anhang auf Seite 490), sodass der Korrelationskoeffizient nach Pearson (r_{Pearson}) berechnet wird. Eine Ausnahme stellt die Facette *Experimente und Wissenschaftsverständnis* dar; hier wird der Koeffizient nach Spearman (ρ_{Spearman}) berechnet.

zwischen der Reflexionsfähigkeit (erhoben mit dem Reflexionsperformanztest) und dem fachdidaktischen Wissen (erhoben mit dem Instrument von Gramzow, 2015, S. 120ff.) untersucht. Dabei wurde vor dem Praxissemester ein signifikanter Zusammenhang ($r_{\text{Pearson}} = 0,48$, $p < 0,01$) gefunden (der nach dem Praxissemester nicht mehr beobachtet wurde: $r_{\text{Pearson}} = 0,12$, n. s. (Kulgemeyer et al., 2021, S. 3048)).

Während eine Reproduktion des Zusammenhangs zwischen Reflexionsfähigkeit und fachdidaktischem Wissen als Argument für die valide Interpretation der Testwerte als Maß für die Reflexionsfähigkeit einzustufen wäre, muss der nicht vorliegende Zusammenhang nicht zwangsläufig zur Interpretation führen, dass der entwickelte Reflexionstest eine andere als die Reflexionsfähigkeit misst. Möglich ist, dass die vergleichsweise geringe Stichprobengröße nicht ausreicht, vorliegende Effekte aufzudecken. Auch ein Einfluss der Verwendung der digitalisierten Version des Tests zum fachdidaktischen Wissen gegenüber der in *ProfiLeP+* genutzten Version kann nicht ausgeschlossen werden. Zur Klärung eines etwaigen Zusammenhangs zwischen fachdidaktischem Wissen und Reflexionsfähigkeit bedarf es also weiterer Untersuchungen, insbesondere, da die vorliegende Studie gegensätzlich zu bisherigen Ergebnissen tendenziell auf negative Zusammenhänge hindeutet.

17.7 Beurteilung des Teils 3 des INA: Übersetzung Verhalten in Testwerte

Die Untersuchung des dritten Teils des INA soll die angemessene Übersetzung des (Ankreuz-)Verhaltens der Teilnehmenden in Testwerte sicherstellen. Dazu wurden verschiedene statistische Analysen der eingesetzten Skalen sowie zu Zusammenhängen zu anderen Testinstrumenten und Lerngelegenheiten durchgeführt.

Auf Grundlage der durchgeführten Untersuchungen kann festgehalten werden, dass die Testwerte zum Reflexionstest normalverteilt vorliegen und somit geeignet sind, verschiedene Ausprägungen der Reflexionsfähigkeit darzustellen (Indikator 3.1). Ergänzend konnten angemessene Lösungshäufigkeiten und Trennschärfen auf Aufgabenebene festgestellt werden (ebenfalls Indikator 3.1) sowie hohe Reliabilitäten für die Skala zur Bestimmung des Testwerts zur Reflexionsfähigkeit sowie für die Teilfähigkeiten des Bewertens und Vorschlagens von Alternativen (Indikator 3.2). Auch die im Assessment-Feedback aufgeschlüsselten Teilfähigkeiten Bewerten und Vorschlagen von Alternativen erscheinen geeignet zur Differenzierung (wenn auch die Testwerte zum Vorschlagen von Alternativen nicht als normalverteilt angenommen werden können) und liegen als reliable Skalen vor. Die durchgeführte explorative Faktorenanalyse stützt die Interpretation der Testwerte als Maß für Reflexionsfähigkeit insofern, als dass auf Grundlage des Scree-Plots die Erfassung eines eindimensionalen Konstrukts angenommen werden kann (Indikator 3.3).

Damit liegen insbesondere Argumente für die strukturelle Validität der Testwertinterpretation vor. Die Stichprobe, die diesen Untersuchungen zugrunde liegt, muss zwar als nicht-repräsentative Gelegenheitsstichprobe eingestuft werden, da sie aber (vorrangig) aus Studierenden im Bachelor- und Masterstudium verschiedener Standorte sowie einigen Referendar*innen zusammengesetzt und somit eher heterogen ist, können diese Ergebnisse zumindest als Hinweise auf die Generalisierbarkeit der Ergebnisse der Validierungsstudien interpretiert werden.

Weiterhin wurden Untersuchungen zur externen Validität durchgeführt. Der vergleichende Einsatz vom entwickelten Reflexionstest und dem zugrundeliegenden Reflexionsperformanztest zeigte einen Zusammenhang zwischen den Testwerten beider Instrumente von mittlerer, an der Schwelle zur hohen, Effektstärke ($r_{\text{Pearson}} = 0,49$; $p < 0,01$), der exemplarisch für einen Zusammenhang der Testwerte des Reflexionstests zu Reflexionen in authentischeren Handlungssituationen verstanden werden kann (Indikator 3.4). Die Untersuchung von Unterschieden in den Testwerten abhängig von Lerngelegenheiten sowie weiteren demografischen Variablen (Indikator 3.5) lieferte differenziertere Evidenz: Während sich Zusammenhänge zwischen Testwerten und theoretischen sowie praktischen Lerngelegenheiten im Quasi-Längsschnitt (u. a. repräsentiert durch den Studienfortschritt oder den Umfang absolvierter Pflichtpraktika) und mit sehr geringer Effektstärke von unter $d_{\text{Cohen},z} = 0,10$ auch im Längsschnitt (bezogen auf physikdidaktische Grundlagenveranstaltungen im Bachelor- und das Praxissemester im Masterstudium) feststellen ließen, blieben Zusammenhänge zu nicht (universitär) begleiteten praktischen schulischen Erfahrungen z. B. im Rahmen von freiwillig absolvierten Praktika, aus. Gleichzeitig ließen sich in Bezug auf wenige erhobene demografische Variablen (z. B. das Geschlecht) vermutlich konstruktirrelevante Varianzen in den Testwerten feststellen, die näherer Untersuchung bedürfen. Ein Einfluss durch die nur quasi-längsschnittlichen Analysen der Zusammenhänge mit den demografischen Variablen ist an dieser Stelle nicht auszuschließen. Dass hier und im Vergleich vom Reflexionstest und dem zugrundeliegenden Reflexionsperformanztest (Indikator 3.4) keine Abhängigkeit der Testwerte und Zusammenhänge (mit Ausnahme eines nachvollziehbaren Unterschieds zum fachlichen Wissen beim Vergleich von Reflexions- und Reflexionsperformanztest) zur angestrebten Schulform (Grundschule bzw. weiterführende Schule) festgestellt wurde, spricht für die Generalisierbarkeit der Ergebnisse auf eine in dieser Hinsicht heterogene Zielgruppe. Zuletzt wurde der Zusammenhang zum fachdidaktischen Wissen untersucht (Indikator 3.6). Zwischen den Testwerten der eingesetzten Instrumente (Reflexionstest und Test zum fachdidaktischen Wissen in der Version von Jordans, Zeller, Große-Heilmann und Riese (2022)) konnte kein Zusammenhang festgestellt werden. Vor dem Hintergrund des beobachteten Zusammenhangs zwischen Reflexionsfähigkeit (erfasst mit dem Reflexionsperformanztest) und fachdidaktischem Wissen (erfasst mit dem Test von Gramzow (2015, S. 102ff.)), wäre ein solcher auch für die hier durchgeführte Untersuchung zu erwarten. Das Ausbleiben dieser

Korrelation stellt potenziell eine Gefährdung der Validität dar; auch hier sollten weitere Untersuchungen durchgeführt werden, um u. a. den Einfluss der geringen Stichprobengröße auf diese Ergebnisse zu evaluieren.

Insgesamt liegen damit tendenziell positive Befunde zur strukturellen Validität vor: Die Interpretation der Testwerte erfolgt für die Gesamtstichprobe sowie Bachelor- und Masterstudierende auf Grundlage einer reliablen (eindimensionalen) und durchaus differenzierenden Skala, wenngleich die Mehrzahl der Teilnehmenden eher hohe Testwerte von über 50 % erreicht. Bezüglich der externen zur Validierung herangezogenen Kriterien sind die beinahe hohe Korrelation zum Reflexionsperformanztest als authentischere Methode zur Erfassung der Reflexionsfähigkeit sowie Zusammenhänge zu Maßen für Lerngelegenheiten (z. B. dem Studienfortschritt) als stützend für die valide Testwertinterpretation anzusehen. Als limitierend muss insbesondere die ausbleibende Korrelation zum fachdidaktischen Wissen eingeordnet werden.

18 Untersuchung der Anforderung an Teil 4 des INA: Das Assessment-Feedback stellt eine geeignete Form der Rückmeldung der Testwerte dar

Der vierte Teil des INA fordert die angemessene Rückmeldung der Testwerte im Rahmen des Assessment-Feedbacks (eine Übersicht über die Anforderung an diesen Teil des INA und die dazu berücksichtigten Indikatoren und durchgeführten Untersuchungen ist ist in Abbildung 14.1 auf Seite 186 zum vierten Übersetzungsschritt dargestellt). Dazu werden in Kapitel 18.1 Ergebnisse aus Interviewstudien zur Evaluation des Assessment-Feedbacks diskutiert. In Kapitel 18.2 wird die Gestaltung des Assessment-Feedbacks mit bekannten Kriterien für effektives Feedback abgeglichen, bevor in Kapitel 18.3 eine Einschätzung der konsequentiellen Validität der Testwertinterpretation vorgenommen wird.

18.1 Evaluation des Assessment-Feedbacks (Indikatoren 4.1 und 4.2)

Ziel der Evaluation des Assessment-Feedback ist die Evaluation der Indikatoren 4.1 *„Die Interpretationen der eigenen Ergebnisse durch die Teilnehmenden passen inhaltlich zum erhaltenen Feedback.“* und 4.2 *„Die Teilnehmenden sprechen dem Assessment-Feedback (inklusive Selbstlernmaterial) einen Mehrwert zu.“*

Um einen Einblick in die Wahrnehmung und inhaltliche Interpretationen des Assessment-Feedbacks und des Selbstlernmaterials durch die Studierenden zu erhalten, wurden mehrfach strukturierte leitfadengestützte Interviewstudien durchgeführt. Zunächst wurde in $N = 6$ Interviews eine vorläufige Version des Assessment-Feedback pilotiert (s. Kapitel 18.1.1). Anschließend wurde das überarbeitete Assessment-Feedback (nun ergänzt um das Selbstlernmaterial) in insgesamt $N = 11$ Interviews evaluiert (s. Kapitel 18.1.2).

Die Interviews sollen Aufschluss darüber geben, inwiefern die jeweilige Version des Assessment-Feedbacks für die Studierenden angemessen und verständlich ist, worauf die Studierenden bei deren Interpretation achten und welche Schlüsse aus dieser gezogen werden. Analog wird in Bezug auf das Selbstlernmaterial untersucht, wie die Studierenden mit diesem arbeiten und inwiefern ihm ein Mehrwert zugesprochen wird. Für beide Bestandteile der Lernumgebung sollen außerdem ggf. bestehende Änderungsbedarfe festgestellt werden.

Die Interviews wurden jeweils transkribiert, inhaltsanalytisch ausgewertet und die Ergebnisse zur Überarbeitung des Assessment-Feedbacks genutzt. Nachfolgend werden diese aufeinander aufbauenden Interviewstudien sowie deren Ergebnisse und daraus gezogenen Schlussfolgerungen chronologisch vorgestellt.

18.1.1 Pilotierung des Assessment-Feedbacks

Auf Grundlage der in Kapitel 8 dargestellten theoretischen Grundlagen zur Gestaltung von Feedback wurde eine erste Version des Assessment-Feedbacks entwickelt. Sie besteht aus einer kurzen Einleitung, in welcher an die Teilnahme am Reflexionstest erinnert und sich für diese bedankt wird. Außerdem wird entlang der vier Reflexionselemente des Modells von Nowak, Kempin, Kulgemeyer und Borowski (2019) ein Einblick in das dem Reflexionstest zugrundeliegende Verständnis der Reflexionsfähigkeit[178] gegeben. Hieran schließt sich das Kernstück an, die Rückmeldung der individuellen Testwerte zur Reflexionsfähigkeit sowie zu den Aufgaben zu Bewertungen, Alternativen und Schülervorstellungen. Zur Einordnung der Ergebnisse werden diese in Form eines Diagramms rückgemeldet, in dem das eigene Ergebnis als roter Punkt vor einem Boxplot, der die Ergebnisse der Vergleichsgruppe (zu diesem Zeitpunkt mit 21 Studierenden noch recht gering) darstellt ist. Abschließend sind Hinweise zur Interpretation der Boxplots inkludiert (z. B. „Im Rahmen der Box liegen die mittleren 50 % der bisherigen Testergebnisse“).[179] Exemplarisch ist diese Version des Assessment-Feedbacks in Kapitel J.1.5 (ab S. 495) angehängt.

Durchführung der Interviews

Im Sommersemester 2021 wurden $N = 6$ Interviews mit Physik-Lehramtsstudierenden im Bachelorstudium der Universität Paderborn durchgeführt. Inhaltlich lassen die Interviews in drei Phasen gliedern: Zunächst wurde erfragt, welche Schlüsse aus dem erhaltenen Assessment-Feedback gezogen wurden. Anschließend

[178] Im Rahmen der zur Pilotierung eingesetzten Version des Assessment-Feedback noch bezeichnet als *Reflexionskompetenz*.

[179] Diese erste Version des Assessment-Feedback besteht lediglich aus der Rückmeldung und beinhaltet noch kein Selbstlernmaterial.

wurde die Interpretation des Assessment-Feedbacks erfragt und dessen Gestaltung in den Blick genommen. Als Grundlage diente eine anonymisierte „Beispielrückmeldung“, sodass die Studierenden ihre eigenen Ergebnisse nicht offenlegen mussten. Die verwendete Beispielrückmeldung zeigte ein Gesamtergebnis etwas über dem Median (im zweitbesten Quartil) und durchmischte Teilergebnisse über, unter bzw. etwa entsprechend dem Median, allerdings immer im Rahmen der mittleren beiden Quartile der Ergebnisse der Vergleichsgruppe im jeweiligen Bereich. Zusätzlich zur Interpretation des Assessment-Feedbacks (u. a. „Welche Schlüsse ziehen Sie aus dieser Rückmeldung zum Reflexionstest?“) wurden mögliche Änderungsbedarfe explizit diskutiert. Dazu wurde z. B. erfragt, inwiefern die Studierenden mit Boxplots vertraut sind, welche Fragen für sie offen blieben und ob bestimmte zusätzliche Informationen (z. B. zum Projekthintergrund oder eine Musterlösung) gewünscht waren. Hierbei wurden insbesondere auch Ergänzungen erfragt, die für effektives Feedback geeignet erschienen (vgl. Kapitel 8.2), den Umfang des Assessment-Feedback allerdings ggf. zulasten der Annahme des Feedbacks weiter erhöhen würden. Abschließend wurden die Passung des Verfahrens zum Herunterladen der Rückmeldung für die Studierenden sowie möglicherweise aufgetretene (technische) Schwierigkeiten in den Blick genommen und für die Studierenden bestand die Möglichkeit, sonstige Kommentare oder Fragen anzusprechen.

Der Interviewleitfaden befindet sich im Anhang in Kapitel J.1 (ab S. 491). Die Interviews wurden via Videotelefonie geführt und für die anschließende Auswertung die Interviews audiografiert.

Auswertung der Interviews

Für die Auswertung wurden die Interviews transkribiert, wobei analog zur Transkription der Think-Aloud-Interviews vorgegangen wurde (s. Kapitel G.1 ab Seite 463).[180] Anschließend wurden die Transkripte kategoriengeleitet analysiert. Dazu wurde dem von Krüger und Riemeier (2014) beschriebenen Vorgehen gefolgt: Nach der Transkription wird eine redigierte Version des Transkripts erstellt, diese wird in Bezug auf die (Haupt-)Auswertungskategorien geordnet - wobei die darin enthaltenen inhaltlich zusammenhängenden Aussagen (häufig jeweils Antworten auf eine Interviewfrage) den nachfolgend beschriebenen Kategorien (ggf. auch mehreren) zugeordnet werden - und anschließend expliziert (Krüger & Riemeier, 2014). Abschließend werden die Interviews inhaltlich auf Grundlage der Kategorien strukturiert (Krüger & Riemeier, 2014), um sie vergleichend betrachten zu können.

Ergänzend wurden auf Grundlage des Interviewleitfadens deduktiv Subkategorien gebildet (Kuckartz, 2018, S. 72), die der Beantwortung der Fragestellung dienen sollen, wie das Assessment-Feedback von den Studierenden wahrgenommen und

[180] Die vollständigen, entlang der Sprechbeiträge segmentierten Transkripte sind der Arbeit angehängt (s. Kapitel P).

interpretiert wird. Bei der Auswertung der Interviews wurden die Kategorien jeweils induktiv um Subkategorien erweitert. Einen Überblick über die Haupt- und Subkategorien liefert Tabelle 18.1.

Rückmeldung			
Kat. 1 Boxplots		**Kat. 4 Gewichtung der Bestandteile**	
AF1a	Nicht bekannt	AF4a	(Eigene) Ergebnisse
AF1b	Bekannt aus der Ausbildung	AF4b	Sonstige Bestandteile
AF1c	Bekannt aus der Schulzeit	**Kat. 5 Änderungsbedarfe Rückmeldung**	
AF1d	Bekannt aus dem Studium	AF5a	Keine
AF1e	Hilfestellungen zum „Lesen“ der Boxplots	AF5b	Umfang der Rückmeldung/-Überflüssige Inhalte
AF1f	Unklarheiten	AF5c	Formulierungen/inhaltliche Struktur
Kat. 2 Interpretation		AF5d	Ergänzen von Informationen zu Zielen des Projekts
AF2a	Orientierung an der Vergleichsgruppe	AF5e	Ergänzen von Informationen zum Zielzustand der Reflexionsfähigkeit
AF2b	Orientierung an Prozentwerten bzw. am Maximalwert von 100%	AF5f	Aufschlüsselung der Ergebnisse auf Aufgabenebene
AF2c	Vergleich der eigenen (Teil-)-Ergebnisse	AF5g	Anpassung der Vergleichsgruppe
AF2d	Fehler in der Interpretation	AF5h	Darstellung in Boxplots
Kat. 3 Schlussfolgerungen		AF5i	Ergänzen einer Musterlösung und bzw. oder von Beispielaufgaben
AF3a	Bedeutung/Mehrwert des Ergebnisses	AF5j	Ergänzen von Literaturhinweisen
AF3b	Bewertung des Ergebnisses (z. B. als positiv)	AF5k	Informationen zur Förderung der Reflexionsfähigkeit
AF3c	Identifikation von Verbesserungspotenzial		
AF3d	Beschreibung möglicher Ursachen für die Ergebnisse		
AF3e	Unklarheiten bzgl. der (Interpretation der) Ergebnisse		
Sonstiges			
Kat. 1 zum Testinstrument		**Kat. 2 Erhalt der Rückmeldung**	

S1a	Gestaltung (Aufbau und Umfang) des Testinstruments	S2a	Problemlos
		S2b	Instantan gewünscht
		S2c	Zusenden (z. B. per Mail) gewünscht

Tabelle 18.1: Übersicht über die Kategorien zur Kodierung der Interviews zur Pilotierung des Assessment-Feedbacks.

Bei der Kodierung der Transkripte wurde nicht dazwischen unterschieden, ob die Aussagen der Studierenden sich aus der Erinnerung auf die eigene Rückmeldung oder auf die im Interview verwendete Beispielrückmeldung beziehen, da einerseits die eigene Rückmeldung nur kurz und primär zum Einstieg explizit thematisiert wurde und andererseits nach erster Durchsicht der Transkripte nicht davon auszugehen war, dass die Studierenden grundlegend verschiedene Herangehensweisen bei den Interpretationen der eigenen und der Beispielrückmeldung zeigen.

Ergebnisse und Schlussfolgerungen

Die Länge der Interviews spannt zwischen 19 und 34 Minuten (durchschnittlich etwa 27 Minuten). Nachfolgend werden die Ergebnisse der Interviews beschrieben und jeweils innerhalb der Abschnitte diskutiert. Die interviewten Studierenden werden dabei mit A1 bis A6 benannt. Voraussetzung für die angemessene Interpretation des Assessment-Feedbacks ist, dass die Studierenden mit der Ergebnisdarstellung in den Boxplots umgehen können. Daher wird zunächst auf die Angemessenheit der Darstellungsform für die Zielgruppe eingegangen, bevor die daraus gezogenen Schlussfolgerungen betrachtet werden. Anschließend werden die Wahrnehmung der einzelnen Bestandteile des Assessment-Feedbacks, explizit geäußerte sowie implizit deutlich gewordene Änderungsbedarfe und der Zugang zum Assessment-Feedback diskutiert. Abschließend wird auf sonstige Aussagen der Studierenden zum Testinstrument eingegangen.

Ist die gewählte Darstellungsform der Ergebnisse in Boxplots für die Studierenden angemessen? Boxplots sind für drei der sechs Studierenden vorher unbekannt (A3, Z. 14; A5, Z. 16, 18; A6, Z. 3; Kat. AF1a). A2 und A4 kennen Boxplots aus der Schule (A2, Z. 16; A4, Z. 17), wobei A4 aber explizit angibt, nicht mit Boxplots gearbeitet zu haben (A4, Z. 17). A1 hat im Rahmen der Ausbildung mit Boxplots gearbeitet (A1, Z. 37, 39; Kat. AF1b). A5 gibt an, dass die Darstellung von Diagrammen in der Rückmeldung untypisch sei, da es „sonst [...] irgendwie immer mehr Rückmeldung in schriftlicher Form (.) und nicht so mit Diagrammen [gab]“ (A5, Z. 20). Zur Interpretation der unbekannten Boxplots

werden unterschiedliche Strategien beschrieben (Kat. AF1e): Während A5 angibt, die Boxplots unter Zuhilfenahme der gegebenen Hinweise zu verstehen (A5, Z. 22), nutzt A3 offenbar externe Quellen, um die Boxplots zu verstehen (A3, Z. 14). A6 sieht eine Parallele zwischen den Antennen der Boxplots und der Darstellung der Unsicherheit von physikalischen Größen mit Fehlerbalken (A6, Z. 3). A1 äußert Unverständnis bezüglich der Beschreibung von als Punkten dargestellten Ausreißern in den Boxplots, da diese (in der vorliegenden Rückmeldung) nicht vorhanden und somit nicht sichtbar sind (A1, Z. 15-27; Kat. AF1f).

Alle sechs Studierenden orientieren sich bei der Einordnung der Ergebnisse an denen der Vergleichsgruppe (dabei insb. dem Median; A1, Z. 41; A3, Z. 5, 7; A4, Z. 3; A5, Z. 3, 11; A6, Z. 7) bzw. stellen einen expliziten Bezug zur Vergleichsgruppe her (A2, Z. 26; beides Kat. AF2a). Zusätzlich berücksichtigen A1, A2, A5 und A6 explizit die gegebenen Prozentwerte (A1, Z. 41; A2, Z. 24; A5, Z. 11; A6, Z. 7; Kat. AF2b); A4 und A5 beschreiben Ergebnisse als „relativ mittig“ (A4, Z. 21) bzw. „relativ weit unten“ (A5, Z. 26), wobei nicht eindeutig festzustellen ist, ob sich diese Einordnung auf die erreichten Prozentwerte oder die Einordnung in Bezug auf die Vergleichsgruppe bezieht. A2 merkt hier außerdem an, dass nicht klar ist, worauf sich die maximal erreichbaren 100 % beziehen, vermutet aber korrekterweise, dass es die Angabe der korrekten Antworten im Reflexionstest ist (A2, Z. 10, 12, 14). Auch A3 äußert Unsicherheit in Bezug darauf, wie sich 0 % oder 100 % in einer Testbearbeitung widerspiegeln (A3, Z. 36). Außerdem stellen zwei der Studierenden Bezüge zwischen den einzelnen Teilergebnissen der Beispielrückmeldung her bzw. vergleichen diese (A1, Z. 41; A3, Z. 7; Kat. AF2c).

Bei der Beschreibung der Boxplots werden teilweise fachlich falsche Aussagen getätigt (Kat. AF2d). Neben dem falschen Verständnis des Medians als Mittelwert bzw. „Durchschnitt“ (A1, Z. 41; A5, Z. 11), äußert A2, „dass man ja in den allermeisten Fällen mit seinem roten Punkt innerhalb der Box liegen wird“ (A2, Z. 56), was allerdings nur für genau die mittleren 50 % der Ergebnisse gilt.

Insgesamt wird deutlich, dass Boxplots für den Großteil der befragten Studierenden zuvor unbekannt waren (Kat. AF1a), sie aber trotzdem die eigenen bzw. die Beispielergebnisse überwiegend sinnvoll in Bezug auf die Vergleichsgruppe (Kat. AF2a) und bzw. oder die Prozentskala (Kat AF2b) einordnen. Gleichzeitig zeigt sich bei einigen Studierenden ein in bestimmten Aspekten falsches Verständnis der Boxplots (Kat. AF2d) z. B. darin, dass sie den Median fälschlicherweise als „Durchschnitt“ (A1, Z. 41; A5, Z. 11) bezeichnen, was die eigene Einschätzung zwar im Detail verändern kann. Die formulierten Tendenzen (z. B. die Einordnung des eigenen Ergebnisses im oberen oder unteren Bereich der Vergleichsgruppe) und damit vermutlich auch die Bewertungen der eigenen Ergebnisse bzw. die daraus gezogenen Schlussfolgerungen bleiben aber die gleichen.

Der von A1 geäußerten Unklarheit bezüglich des Hinweises zur Interpretation der nicht vorhandenen Ausreißer (Kat. AF1f) im Boxplot wird zukünftig entgegenge-

wirkt, indem die Antennen des Boxplots nicht mehr die Länge des eineinhalbfachen Interquartilsabstand haben, sondern alle potenziellen Ausreißer inkludieren. So soll außerdem vermieden (oder zumindest reduziert) werden, dass Ergebnisse, die als Ausreißer unter der Box liegen, als demotivierend wahrgenommen werden.

Welche Schlussfolgerungen ziehen die Studierenden aus der (erhaltenen bzw. Beispiel-) Rückmeldung? Die Studierenden beschreiben, welchen Mehrwert ihnen die eigene Rückmeldung bietet bzw. welches Gewicht sie der Rückmeldung beimessen (Kat AF3a). A1 zieht die positive Schlussfolgerung, „gar nicht so falsch in dem Job“ (A1, Z. 5) zu sein. A2 gibt an, die Rückmeldung zeige auf, welche Teilbereiche der Reflexion es gäbe bzw. welche berücksichtigt wurden und sei aber „wirklich hilfreich und interessant [...] als Reflexionsanlass“ (A2, Z. 24). Passend dazu gleicht A2 die erhaltene Rückmeldung mit der eigenen Einschätzung der Fähigkeiten ab (A2, Z. 5, 18), misst dem eigentlichen Ergebnis aber keine allzu große Bedeutung bei (A2, Z. 24). Ähnlich bewertet A6 die Rückmeldung als „Momentaufnahme“ (A6, Z. 9) in Bezug auf die Einordnung des Ergebnisses in die Vergleichsgruppe. In Bezug auf die Interpretation der Boxplots verweist A2 darauf, dass diese keine absolute Rückmeldung darüber geben, „ob man jetzt wirklich gut oder schlecht ist, sondern eigentlich immer nur, ob man schlechter als andere ist“ (A2, Z. 24). Dies verleite dazu, das eigene Ergebnis nur in Relation zur Vergleichsgruppe zu betrachten, obwohl „das Ziel sein [sollte], da überall deutlich über dem, was auch die Boxplots da jetzt zeigen, zu sein“ (A2, Z. 24). A2 schränkt aber ein, dass die Ergebnisse im Bachelorstudium „noch gar nicht perfekt sein“ können (A2, Z. 24), was auch A1 bei der Interpretation der Beispielrückmeldung berücksichtigt: Beim Bewerten und Vorschlagen der Alternativen „war ich ungefähr so gut wie meine Mitstudenten, was ja auch Sinn ergeben würde, weil wir ja in ähnlichen Semestern sind, haben wir wahrscheinlich schon ähnliches gelernt oder nicht gelernt, also da würde ich mir wahrscheinlich weniger Sorgen drum machen“ (A1, Z. 41). A3 erkennt außerdem, dass der Teilbereich Schülervorstellungen geringer gewichtet wird, als die Bereiche Bewerten und Alternativen aufzeigen („so wie das hier dargestellt ist, macht das für mich den Eindruck als wären die Schülervorstellungen etwas weniger stark gewichtet“, A3, Z. 22), was die Interpretation der Ergebnisse allerdings nicht zu beeinflussen scheint.

Vier der sechs Studierenden ziehen positive oder zumindest eher positive Schlussfolgerungen aus der eigenen bzw. der Beispielrückmeldung (A1, Z. 41; A2, Z. 24; A3, Z. 3, 18; A4, Z. 21; Kat AF3b), z. B. dass die Person, die die Beispielrückmeldung erhalten hat, „nicht bei den schlechtesten aus [der] Vergleichsgruppe“ (A1, Z. 41) abschneidet, „auf einem guten Weg“ (A4, Z. 21) ist bzw. ein „recht gutes [didaktisches] Fachwissen hat“ (A3, Z. 18) und daher in den Bereichen Alternativen und Bewertung gut abschneidet. (Eher) positive Interpretationen der Ergebnisse erfolgen in der Regel dann, wenn die Ergebnisse über dem Median liegen, allerdings ordnet bspw. A4 das eigene, „im Mittel“ (A4, Z. 3) liegende

Ergebnis - unter Berücksichtigung des geringen Studienfortschritts - als „ganz gut“ (ebd.) ein.

Zudem identifizieren fünf Studierende - allerdings unterschiedlich konkretes - Verbesserungspotenzial auf Grundlage der Rückmeldung (A1, Z. 41; A2, Z. 24; A3, Z. 3, 18; A4, Z. 21, 23; A5, Z. 26; Kat. AF3c). A1, A3, A4 und A5 benennen hier u. a. den Bereich der Schülervorstellungen als zu verbessernden Bereich (A1, Z. 41; A3, Z. 18; A4, Z. 21; A5, Z. 26), in den „man sich dann wahrscheinlich [...] reinarbeiten“ (A5, Z. 26) sollte. A4 formuliert diesen Verbesserungsbedarf, obwohl A4 Schülervorstellungen noch nicht bekannt sind, da A4 „dazu jetzt noch nichts gehört“ (A4, Z. 15) hat. Gleichzeitig spricht A2 an, dass das Ableiten von Handlungsschritten aus den Ergebnissen der Rückmeldung „sehr sehr schwierig [ist], gerade, wenn man [...] [vom] Unterrichten selber völlig weit entfernt ist.“ (A2, Z. 18), denn die Idee von A2, sich „auf diesen Bereich speziell mehr [zu] fixieren, [...] finde ich schwierig, weil ja doch in Veranstaltungen [...] sehr sehr eng vorgegeben ist, was gemacht wird“ (ebd.). Auch A5 fragt sich, „was ich mir dann fürs spätere Lehrerleben daraus ziehen kann“ (A5, Z. 38), A6 überlegt, „ob wir in der Uni da [Anm.: über identifizierte Verbesserungspotenziale] nochmal drüber sprechen“ (A6, Z. 23) und A4 sieht „noch nicht, was ich verbessern kann“ (A4, Z. 49).

Mögliche Ursachen, die das Abschneiden im Instrument erklären könnten (Kat. AF3d), liefert nur A3, einerseits in Bezug auf das identifizierte Verbesserungspotenzial im Bereich der Schülervorstellungen mit der Vermutung, „dass diese Person [Anm: die Person, der die im Interview genutzte Beispielrückmeldung zuzuordnen ist] vielleicht schon länger aus der Schule raus ist [...] oder halt sich noch nicht so sehr mit dem Thema Schülervorstellungen überhaupt auseinandergesetzt hat“ (A3, Z. 18). Außerdem vermutet B3 einen Einfluss des didaktischen Fachwissens auf das Abschneiden in den Ergebnissen zum Bewerten und Vorschlagen von Alternativen (A3, Z. 18).

Bei der Interpretation der Rückmeldungen wird eine Unklarheit auf Seiten der Studierenden deutlich (Kat AF3e). So fragt B5 im Interview, „was genau sagt mir jetzt diese Prozentangabe?“ (A5, Z. 5), wobei zu berücksichtigen ist, dass die Rückmeldung, die Studierenden eine Version des Assessment-Feedback erhalten haben, in welcher die Prozentskala fälschlicherweise von 0 % bis 1 % und nicht von 0 % bis 100 % dargestellt ist. Diese falsche Angabe war – wie auch A3 bemerkt (A3, Z. 8) – bereits zum Zeitpunkt der Interviews in der Beispielrückmeldung korrigiert.

Auf einer globalen Ebene formulieren die Studierenden stark variierende Bedeutungen, die sie der Rückmeldung zusprechen (Kat AF3a). Diese reichen von der Bewertung der Rückmeldung als Bestätigung für die angemessene Berufswahl (A1, Z. 24) bis hin zur Bewertung der Rückmeldung als reine „Momentaufnahme“ (A6, Z. 9). Da der Reflexionstest nicht dazu entwickelt wurde, die Berufseignung

zu diagnostizieren und insbesondere in Bezug darauf nicht validiert ist, ist die Einschätzung von A1, „gar nicht so falsch in dem Job“ (A1, Z. 5) zu sein, keine Aussage, die gesichert auf Grundlage der Rückmeldung zu treffen ist. Die übrigen, eher vorsichtigen Gewichtungen der Rückmeldung als „Momentaufnahme“ (A6, Z. 9) oder „Reflexionsanlass“ (A2, Z. 24) sind hingegen auf Basis der Rückmeldung angemessen, da sie zu den Zielen der Messung der Reflexionsfähigkeit und der intendierten Testwertinterpretation passen. Ebenso zeigen die Überlegungen zur Bedeutung der Vergleichsgruppe (A1, Z. 41), des Studienfortschritts (A2, Z. 24) und der Gewichtung der einzelnen Teilbereiche im Gesamtergebnis (A3, Z. 22) bei der Interpretation der Boxplots als Maßstab für ein als angemessen bewertetes Abschneiden eine sinnvolle Einordnung der Ergebnisse unter Berücksichtigung des individuellen Kontextes.

Auf Ebene der einzelnen (Gesamt- und Teil-) Ergebnisse sind sowohl die von vier Studierenden formulierten positiven Einordnungen der Ergebnisse (in der Regel) über dem Median (Kat AF3b) sowie die von fünf Studierenden identifizierten Verbesserungspotenziale (Kat. AF3c) auf Grundlage der Rückmeldung als plausibel zu bewerten und erscheinen angemessen vor dem Hintergrund der intendierten Testwertinterpretation. Auch die von den Studierenden beschriebenen möglichen Ursachen für die in der Beispielrückmeldung dargestellten Ergebnisse (die mangelnde Beschäftigung mit Schülervorstellungen; A3, Z. 18; und das hohe fachdidaktische Wissen; A5, Z. 5; Kat. AF3d) stellen nachvollziehbare und sinnvolle Überlegungen dar. Gleichzeitig äußern die Studierenden Unsicherheiten und Schwierigkeiten, auf Grundlage ihrer Ergebnisse Handlungsoptionen zur Umsetzung des Verbesserungspotenzials zu erkennen (Kat. AF3c). Konkrete, umsetzbare Handlungsoptionen zur Entwicklung der Reflexionsfähigkeit werden von keiner Person benannt. Diese wären aber notwendig, um die Professionalisierung der Studierenden in Bezug auf die Reflexionsfähigkeit voranzutreiben. Die Schwierigkeit, konkrete Handlungen aus der Rückmeldung abzuleiten, wird auch an anderer Stelle von den Studierenden angesprochen und weiter unten im Kontext der Änderungsbedarfe der Rückmeldung diskutiert.

Wie gewichten die Studierenden die Bestandteile der Rückmeldung? Als besonders hilfreiche oder relevante Inhalte der Rückmeldung stellen jeweils vier Studierenden einerseits die dargestellten Ergebnisse heraus (Kat. AF4a), u. a. „die Dinge, die in den Diagrammen visualisiert sind“ (A1, Z. 43), wie die Differenz zwischen eigenem Ergebnis und dem Median (A1, Z. 43; A3, Z. 24), oder „wo man im Vergleich zum Maximum so liegt“ (A3, Z. 24) sowie identifizierte Verbesserungsbedarfe (A2, Z. 24; A4, Z. 23). Zusätzlich wird als „interessant“ (A3, Z. 24) bewertet, dass keine Person 100 % der Punkte erreicht hat (A3, Z. 24) bzw. als „spannend“ (A4, Z. 25), dass die Ergebnisse der Vergleichsgruppe „ja eigentlich im Mittel noch weiter nach oben gehen“ (A4, Z. 25) könnten. Andererseits werden Erklärungen der Teilfähigkeiten (z. B. A5, Z. 30; A6, Z. 11) bzw. die

Hinweise zur Interpretation der Boxplots (als „Hilfestellungen“, A1, Z. 45; A6, Z. 5) als besonders bedeutsam aufgeführt. Darüber hinaus werden von einzelnen Studierenden verschiedene (Text-)Teile der Rückmeldung hervorgehoben (Kat. AF4b): A4 beschreibt, durch die Erläuterungen der Teilfähigkeiten „überhaupt erstmal verstanden“ (A4, Z. 67) zu haben, „was damit gemeint ist“ (A4, Z. 67), und A1 findet die Benennung der Teilfähigkeiten gut, um einen „Einblick darein [zu bekommen], wie denn diese empirische Forschung da funktioniert“ (A1, Z. 49). A1 gibt außerdem an, die Erläuterungen zum Begriff des Reflektierens als „ziemlich wichtig“ (A1, Z. 45) für das Verständnis der Rückmeldung zu bewerten.

Die Aussagen über die Bedeutung der einzelnen Bestandteile der Rückmeldung für die Studierenden zeigen, dass von den meisten Studierenden (vier der sechs Befragten) grundsätzlich die (eigenen) Ergebnisse (Kat. AF4a) sowie von drei bzw. zwei Studierenden die Erläuterungen der Teilfähigkeiten und zu den Boxplots (Kat. AF4b), die für das Verständnis der Rückmeldung notwendig sind, als besonders bedeutsam erachtet werden. Damit wird der Großteil der Bestandteile des Assessment-Feedbacks an dieser Stelle explizit hervorgehoben. Die Einleitung wird nicht als bedeutsam herausgestellt, bezieht sich allerdings auch nicht explizit auf die eigenen Ergebnisse, und die Erläuterungen zum Begriff der Reflexionsfähigkeit werden nur von A1 als „ziemlich wichtig“ (A1, Z. 45) bewertet, „damit man halt weiß, was überhaupt betrachtet wurde“ (A1, Z. 45). Unklar ist, ob und ggf. weshalb die übrigen Studierenden diese begriffliche Klärung als weniger relevant empfinden. Möglicherweise führen die Studierenden die Interpretation der Ergebnisse durch, ohne einen Bezug zum übergeordneten Konstrukt der Reflexionsfähigkeit zu erkennen oder zu benennen.

Wo bestehen aus der Perspektive der Studierenden Änderungsbedarfe in Bezug auf die Rückmeldung? Welche weiteren Missverständnisse auf Seiten der Studierenden werden deutlich, die ggf. einen Änderungsbedarf indizieren?
Im Anschluss an die geführten Interviews wird die Rückmeldung überarbeitet, dabei werden insbesondere die Perspektiven der Studierenden sowie die von Ihnen vorgeschlagenen Änderungen berücksichtigt.

Vier der sechs Studierenden geben an, die Rückmeldung enthalte keine überflüssigen Informationen (A2, Z. 28; A3, Z. 26; A4, Z. 27; A6, Z. 21; Kat. AF5a), lediglich einzelne Hinweise zum Verständnis bzw. zur Interpretation der Boxplots werden als unnötig erachtet (Kat AF5b). Dazu gehört der Hinweis zur Darstellung von – zu diesem Zeitpunkt im Rahmen der Rückmeldung nicht vorhandenen – Ausreißern (A1, Z. 47; ursprüngliche Formulierung des Hinweises: „Etwaige schwarze Punkte zeigen ‚Ausreißer‘, also Ergebnisse, die im Vergleich zu den mittleren 50 % der bisherigen Testergebnisse besonders hoch oder niedrig sind.“) sowie die Beschriftung des Medians (A5, Z. 34). A2 ergänzt, „wenn man weiß, wie ein Boxplots funktioniert, kann man halt den kurzen Absatz [Anm.: Absatz mit

Hinweisen zur Interpretation der Boxplots] überspringen“ (A5, Z. 27). Zusätzlich schlagen einzelne Studierende Änderungen in Bezug auf den Aufbau und konkrete Formulierungen der Rückmeldungen vor (Kat. AF5c). A1 merkt in Bezug auf die vorliegenden Formulierungen an, dass die wiederholte Nutzung des Wortes „können“ irritierend sein könnte (A1, Z. 7) und dass die Beschreibung der Teilfähigkeit der Schülervorstellungen diese fälschlicherweise betont: „vielleicht ist es deshalb nicht so super ausgerechnet einen da besonders hervorzuheben mit ‚die Diagnose ist ein zentraler Aspekt der Physikdidaktik‘“ (A1, Z. 12). Außerdem schlägt A3 zur Struktur des Assessment-Feedbacks vor, das erzielte Ergebnis bereits zu Beginn der Rückmeldung anzugeben (Kat AF5b): „[D]as Ganze noch verschriftlicht einfach was ja was dieses Diagramm aussagt und am besten das auch gleich in dem ersten Absatz, damit man sein Ergebnis irgendwie einordnen kann (.) bevor man überhaupt diesen Plot gesehen hatte“ (A3, Z. 26), denn die Rückmeldung war „erstmal [ein] bisschen überwältigend“ (A3, Z. 83) und „das macht es dann einfacher, sich vielleicht damit zu befassen“ (A3, Z. 83).

Mit Ausnahme des für A1 (Z. 47) überflüssigen Hinweises zu Ausreißern im Boxplot gibt es aus der Perspektive der Studierenden also keinen tatsächlichen Änderungsbedarf in Bezug auf potenziell überflüssige Inhalte der Rückmeldung. Der Hinweis zur Darstellung der Ausreißer als schwarze Punkte in den Boxplots wird entfernt, da – wie im Folgenden zur Darstellung der Diagramme ausführlicher beschrieben – die Antennen bis zu den höchsten bzw. niedrigsten Ergebnissen verlängert werden und somit keine Ausreißer in den Boxplots auftreten können.

Die vorgeschlagenen Änderungen zu konkreten Formulierungen (Kat. AF5c) werden überwiegend übernommen. Insbesondere werden zu Beginn der Rückmeldung (vor theoretischen Erläuterungen zum Reflektieren und der ausführlichen Ergebnisdarstellung) das individuelle Gesamtergebnis (in Prozent) sowie die Einordnung in die Vergleichsgruppe (in die oberen bzw. unteren 25 % bzw. $50 - 25$ %) schriftlich dargestellt. Ein Beispielergebnis aus dem Sommersemester 2022 wird wie folgt dargestellt:

„Insgesamt haben Sie $68,8$ % der maximal möglichen Punktzahl erreicht. Damit liegt ihr Ergebnis im Bereich der oberen $50 - 25$ % der Testergebnisse der Teilnehmenden aus der Vergleichsgruppe (bestehend aus 80 Studierenden im durchschnittlich $4,9$. Fachsemester).“

Die Erläuterung zum Teilbereich der Schülervorstellungen wird auch unter Berücksichtigung des geringeren Gewichts dieses Bereichs im Vergleich zu den Ergebnissen zum Bewerten und Vorschlagen von Alternativen im Gesamtergebnis neu formuliert, das A3 im Rahmen der Interpretation der Rückmeldung anspricht (A3, Z. 22). Dazu wird ein Hinweis ergänzt:

„Unter den Aspekt ‚Schülervorstellungen‘ fallen die beiden Items, in denen Sie beobachtete Schülervorstellungen benennen und Strategien zum Umgang mit diesen vorschlagen sollten. Hinweis: Die Anzahl der Items zu Schülervorstellungen

ist also deutlich geringer, als die zu Bewertungen und Alternativen – diesem Aspekt kommt also deutlich weniger Gewicht in der Gesamtwertung zu."

Auf Nachfrage geben einzelne Studierende den Wunsch an, mehr Informationen zu den folgenden Punkten erhalten zu wollen: Zum Ziel des Reflexionstests (A5, Z. 52, 54) und zum Projekt, in dem das Instrument entstanden ist (z. B., dass das Projekt noch nicht abgeschlossen ist und am eigenen Universitätsstandort durchgeführt wird (A4, Z. 41, 43) oder welche Ziele mit dem Projekt verfolgt werden (A6, Z. 39; A3, Z. 52; Kat. AF5d). A1 und A5 schätzen dies nur für besonders interessierte Personen als interessant ein (A1, Z. 63; A5, Z. 56). Allerdings befürchtet A1, „dass es eventuell zu weit ausufern könnte oder zu kurz kommen könnte, weil wenn man es nur einmal kurz mit rein schreibt [...], hat [man] trotzdem noch keinen Eindruck vom ganzen Projekt und wenn man es im Ganzen beschreibt, hat man noch eine Zusatzseite" (A1, Z. 63), sodass es „vielleicht besser [ist], das deshalb rauszulassen" (A1, Z. 63). A2 und A4 halten eine Ergänzung des Ziels des Instruments nicht für notwendig, da „das Ziel doch relativ (..) klar formuliert [ist]" (A4, Z. 39; ähnlich auch A2, Z. 46) und auch die Ergänzung von Informationen zum Projekt wird von A5 als „nicht unbedingt hilfreich" (A5, Z. 56) erachtet und A4 beschreibt bezüglich der Rahmenbedingungen des Projekts, „das ist alles klar geworden als wir [...] die Auswertung gekriegt haben, aber am Anfang war das halt noch nicht so ganz klar" (A4, Z. 43).

Die Erläuterung des anzustrebenden Zielzustands für die Reflexionsfähigkeit wird von vier Studierenden überwiegend als sinnvoll erachtet (A2, Z. 48; A3, Z. 54; A4, Z. 45; A5, Z. 60; Kat. AF5e). A2 interessiert: „[W]as habt ihr jetzt als Ausgangszustand genommen, also hat Herr N [Anm: Professor für Physikdidaktik und Dozent der Veranstaltung in dessen Rahmen der Reflexionstest bearbeitet wurde] einmal den Test bearbeitet und das sind jetzt hundert Prozent, das wär jetzt zum Beispiel interessant zu wissen, oder ist das jetzt jemand, der grad seinen Bachelor hat oder jemand der grad seinen Master hat oder jemand der jetzt schon etwas länger als Lehrer schon arbeitet, das wäre tatsächlich interessant, weil das ja wahrscheinlich auch einen ziemlichen Unterschied macht, äh aus welchem Blickwinkel man da drauf guckt" (A2, Z. 48). A2 und A3 erläutern außerdem, dass sie davon ausgegangen sind, dass das Erreichen eines Zielzustands mit einem hohen Testwert einhergeht („tatsächlich hab ich als Zielzustand die Hundert Prozent interpretiert ähm hab aber auch gleichzeitig gesehen dass es noch nicht in der Testgruppe erreicht wurde", A3, Z. 56) und dies darauf abgestimmt ist, „was ihr [Anm.: die Studierenden] gelernt haben solltet" (A2, Z. 48). A5 gibt hingegen an, von keinem Ziel ausgegangen zu sein, weil man „sich halt immernoch (.) immer verbessern kann und irgendwas dran ändern kann an seiner Reflexionskompetenz" (A5, Z. 64). Die übrigen beiden Studierenden befürchten, dass diese Information bei einem schlechten Ergebnis (z. B. „im unterdurchschnittlichen Bereich" (A6, Z. 45)) entmutigend" (A1, Z. 67) wirken bzw. „Verunsicherung bringen" (A6, Z. 45) kann. A6 ist daher nicht sicher, ob ein Zielzustand angegeben werden sollte (A6,

Z. 45), A1 kommt zu dem Schluss, dass dies im Erwachsenenalter möglich ist, da „wir schon besser gelernt [haben], diese Selbsteinschätzung zu machen, als jetzt zum Beispiel Schulkinder“ (A1, Z. 67).

Auf eine ausführliche Ergänzung von Informationen zum Ziel des Einsatzes des Reflexionstests, dem Projekt, in welchem das Instrument entwickelt worden ist (Kat. AF5d) sowie zu einem anzustrebenden Zielzustand der Reflexionsfähigkeit (Kat. AF5e) wird verzichtet. Die Studierenden beschreiben insgesamt keine Verständnisschwierigkeiten, die mit den fehlenden Informationen verbunden sind, sondern äußern bzw. vermuten lediglich Interesse, ohne dass hier konkrete oder geteilte Aspekte aufgeführt werden, die ergänzt werden sollten. Da außerdem zwei der Studierenden das Ziel des Reflexionstests als klar formuliert wahrnehmen und wiedergeben (A2, Z. 46; A4, Z. 39), scheint diese Information - zumindest im Rahmen der Einbettung des Instruments in eine Veranstaltung - ausreichend transparent gewesen zu sein. Um die Transparenz des Ziels auch dann zu gewährleisten, wenn das Instrument nicht in gleichem Umfang in eine Veranstaltung eingebettet ist, wird die Einleitung der Rückmeldung um folgenden Satz ergänzt:

„Diese Rückmeldung soll Ihnen dabei helfen, sich in Bezug auf diese Fähigkeit (Anm.: Reflexionsfähigkeit wird im vorangegangenen Satz benannt) realistisch einschätzen zu können.“

Informationen zum zugrundeliegenden Projekt werden nicht ergänzt, um den Umfang der Rückmeldung nicht weiter zu erhöhen, was auch A1 bei einer Ergänzung befürchtet (A1, Z. 63; Kat. AF5b). Gleichzeitig wird aufgrund des neu verwendeten Layouts der Rückmeldung die Zugehörigkeit zum Projekt transparent gemacht, indem Projektname bzw. Projektlogo ebenso wie die am Projekt beteiligten Universitäten und Verantwortlichen auf dem Deckblatt dargestellt sind. Bei Bedarf können also die Ansprechpartner bzw. Verantwortlichen zum Projekt kontaktiert werden.

Die Formulierung eines theoretisch oder empirisch fundierten Zielzustands wird zwar von vier Studierenden klar befürwortet (Kat. AF5e), ist aber im Rahmen der Rückmeldung nicht möglich. Hier fehlen normative Erwartungen bzw. empirische Befunde dazu, welche Ausprägung der Reflexionsfähigkeit als ausreichend oder angemessen anzusehen ist, um bspw. guten Unterricht zu gestalten (z. B. Häcker, 2022). Für die Bearbeitung des Reflexionstests kann festgehalten werden, dass ein Ergebnis umso besser ist, je höher der erreichte Prozentwert ist – wobei auch das Erreichen von 100 % aufgrund der fehlenden Befunde nicht als Bestätigung einer absolut ausreichenden Reflexionsfähigkeit gewertet werden kann. Diese Einschätzung entspricht offenbar auch dem intuitiven Verständnis zweier Studierender (A2, Z. 48; A3, Z. 56). Perspektivisch könnten bspw. die Ergebnisse von Referendar*innen, die den Test bearbeitet haben, als zusätzlicher Vergleichswert dafür dienen, wie Personen mit abgeschlossenem Lehramtsstudium im Reflexionstest abschneiden. Auch ein solcher Vergleich ermöglicht zwar keine

abgesicherte Aussage darüber, ob die Reflexionsfähigkeit der Studierenden für das Berufsleben angemessen ist, kann aber zumindest als Orientierungshilfe dienen, um einzuschätzen, mit welchen Fähigkeiten andere Studierende das Studium beenden bzw. welche Fähigkeiten sie innerhalb der nächsten Ausbildungsphase erreicht haben.

Vier der Studierenden äußern verschiedene Bedarfe nach weiteren Informationen zu den Ergebnissen bzw. einer näheren Aufschlüsselung dieser, die sich in 3 Aspekte zusammenfassen lassen. A6 und A1 geben an, eine Aufschlüsselung der Ergebnisse als nicht sinnvoll zu erachten, da damit die Länge der Rückmeldung zunimmt (Kat. AF5b) „und dann die Konzentration beziehungsweise das Interesse dann [...] abnimmt“ (A6, Z. 49) und da „ein bisschen Zeit zwischen [Bearbeitung des Instruments und Erhalt der Rückmeldung liegt, so] dass man nicht die Erinnerung an die einzelnen Fragen und so weiter hat und es eigentlich besser ist, eine ähm zusammengefasste Rückmeldung zu bekommen, weil die einem in dem Moment mehr bringt und man vielleicht auch nicht die Muße hat, sich da jede einzelne Aufgabe nochmal anzugucken oder sich gar nicht richtig erinnert“ (A1, Z. 71). Vier der Studierenden wünschen sich eine Aufgliederung der individuellen Ergebnisrückmeldung auf Aufgabenebene (A2, Z. 50; A3, Z. 34, 36, 38, 58, 60; A4, Z. 31, 33, 35; A5, Z. 72; Kat. AF5f) (z. B. „fände ich persönlich auf jeden Fall interessant, das alles so einzeln betrachten zu können [...], dass man da nochmal jede einzelne Frage hätte“ (A5, Z. 72)), die mit einer Aufschlüsselung der Zuordnung der einzelnen Aufgaben im Reflexionstest zu den drei Teilergebnissen einhergeht (A2, Z. 50; A3, Z. 34; A4, Z. 31ff.) und so Aufschluss über Ursachen für das Abschneiden gibt („ich würd ganz gerne wissen an welchen Stellen wurde das so bewertet also in diesem (.) Fragebogen quasi“ (A4, Z. 29)). So fehlte A4 „halt so der direkte Bezug zu den Antworten, die ich dann in diesem Test gegeben habe“ (A4, Z. 34). A2 gibt an, „mit diesen Begriffen Alternative, Bewertung, Schülervorstellung, auch mit den Erläuterungen da drunter zwar als Begriff was anfangen [...], aber die nicht wirklich im Test verorten [zu können]“ (A2, Z. 50). Eine solche Aufschlüsselung würde es laut A4 ermöglichen, zu erkennen, „in dem Moment habe ich dann vielleicht falsch gehandelt oder nicht so gut gehandelt und das wäre dann eine Möglichkeit, [...] besser zu handeln“ (A4, Z. 33). A2 und A4 schätzen eine Anpassung der Vergleichsgruppe an den eigenen Studienfortschritt (z. B. nach Fachsemester, A4, Z. 37, 55, oder „vielleicht dann auch gar nicht nach Fachsemester, sondern vielleicht eher nach ordnen, was für Veranstaltungen hat man schon gehört, weil das ist, glaub ich, der interessante Punkt, gerade wenn Leute Veranstaltungen durcheinander hören“ (A2, Z. 30)) als interessant ein (Kat. AF5g). Den Vergleich mit der vorliegenden Vergleichsgruppe findet A2 „schwammig“ (A2, Z. 26). A2 sieht in der Anpassung der Vergleichsgruppe den Vorteil, feststellen zu können, „vielleicht auch so jetzt als Lerneffekt, wie viel nehme ich eigentlich aus meinem Studium mit“ (A2, Z. 30). Außerdem wünscht sich A4 eine ausführlichere Beschreibung der Teilbereiche bspw. mit Beispielen

aus dem Unterricht (A4, Z. 67-71) und A5 die Rückmeldung in Form eines Texts anstelle der vorliegenden Boxplots, da dieser „hilfreicher“ (A5, Z. 70) gewesen wäre.

Insgesamt wird deutlich, dass die Studierenden mehrere, uneinheitliche Vorstellungen dazu äußern, wie die Rückmeldung weiter ausdifferenziert werden könnte. Die von vier Studierenden gewünschte Ausdifferenzierung der Ergebnisse auf Aufgabenebene (Kat. AF5f) wird nicht umgesetzt. Eine solche Ausdifferenzierung würde bei den 32 vorliegenden Aufgaben zu Bewertungen und Alternativen mit jeweils vier bis acht (i. d. R. fünf bis sieben) Antwortoptionen den Umfang der Rückmeldung beträchtlich erhöhen, sodass die zentralen, umfassenden Ergebnisse möglicherweise an Bedeutung verlieren und die Rückmeldungen ggf. weniger aufmerksam gelesen werden (A1, Z. 71; A6, Z. 49). Einer detaillierten Ausdifferenzierung der Ergebnisse auf Ebene der Antwortoptionen würde außerdem einer Musterlösung gleichkommen, die an dieser Stelle nicht ausgegeben werden soll, um weitere Einsätze des Reflexionstests auch bei gleichen Studierenden (z. B. vor und nach dem Praxissemester) zu ermöglichen.

Die von den Studierenden gewünschte Zuordnung der Testelemente zu den Teilbereichen der Rückmeldung (Bewertungen, Alternativen und Schülervorstellungen), um diese im Test verorten zu können (A2, Z. 50; A3, Z. 34; A4, Z. 31, 33, 35), wird ausführlicher als bisher umgesetzt. Dazu wird zunächst im Abschnitt „Reflektieren – Was ist das genau & wie kommt das im Test vor?“ genauer beschrieben, welche Teilbereiche der Reflexion im Testinstrument fokussiert werden (z. B. zu Bewertungen und Alternativen: „Im Test wurden Ihnen zu jedem Unterrichtsausschnitt Aspekte vorgegeben, zu denen Sie in je einer Aufgabe (1) das Vorgehen der Lehrkraft bewerten und (2) Alternativen zum Vorgehen vorschlagen sollten.“). Außerdem werden die Erläuterungen zu den Teilfähigkeiten nach der Darstellung der Ergebnisse in Boxplots jeweils um die Verortung der zur Teilfähigkeit gehörenden Aufgaben im Reflexionstest ergänzt. Für Bewertungen bzw. Alternativen wird jeweils der Hinweis ergänzt, dass diese in der ersten bzw. zweiten Frage zu einem Aspekt verortet sind; für die zwei Aufgaben zum Teilbereich der Schülervorstellungen wird beschrieben, dass hierzu „die beiden Items [gehören], in denen Sie beobachtete Schülervorstellungen benennen und Strategien zum Umgang mit diesen vorschlagen sollten“. Diese klarere Zuordnung könnte es außerdem ermöglichen, die Teilbereiche mit Beispielen zu verknüpfen, wie von A4 gewünscht (A4, Z. 67ff.). Zusätzlich werden die Teilbereiche Bewertungen und Alternativen anhand der Beispiellösungen im ergänzten Selbstlernmaterial mit inhaltlichen Beispielen versehen.

Eine Anpassung der Vergleichsgruppe nach Studienfortschritt (Kat. AF5g) ist aufgrund der bislang geringen Größe des Datensatzes nicht möglich, könnte perspektivisch aber realisiert werden. Die gezielte Wahl der Vergleichsgruppe ermöglicht dann eine präzisere Einordnung der Ergebnisse, wobei aufgrund der

Zusammensetzung des Datensatzes aus verschiedenen Universitäten die von A2 gewünschte Differenzierung nach absolvierten universitären Veranstaltungen (A2, Z. 30) nicht realisierbar ist, sondern auf gröbere Kriterien, wie bspw. den Studiengang (Bachelor oder Master) oder das aktuelle Fachsemester zurückgegriffen werden muss.

Der von A5 geäußerte Wunsch, die Rückmeldung in Textform anzubieten (A5, Z. 70), wird nicht umfassend umgesetzt, da die Boxplots – obwohl die Studierenden die Arbeit mit ihnen nicht gewohnt sind – insgesamt geeignet erscheinen, die individuellen Rückmeldungen sowie die Verortung in der Vergleichsgruppe nachvollziehbar darzustellen. Durch die oben beschriebene Ergänzung des schriftlich formulierten Gesamtergebnisses zu Beginn der Rückmeldung wird dieser Wunsch aber zumindest teilweise umgesetzt, sodass lediglich die Teilergebnisse nur in Form von Boxplots dargestellt sind.

Auch die Eignung der verwendeten Boxplots wird im Rahmen der Interviews explizit thematisiert (Kat. AF5h). Fünf der Studierenden sind der Meinung, dass die Wahl der Boxplots zur Darstellung der Ergebnisse in der Rückmeldung (ggf. unter Zuhilfenahme der Erläuterungen (A4, Z. 57)) sinnvoll ist (A1, Z. 73; A2, Z. A3, Z. 62; A4, Z. 55; A6, Z. 51, 53) und begründen dies damit, dass die Darstellung A1 „in ähnlichen Kontexten irgendwie begegnet ist“ (A1, Z. 73) bzw. - fälschlicherweise damit - „dass man ja in den allermeisten Fällen mit seinem roten Punkt innerhalb der Box liegen wird und [das heißt,] das, was ich mache, ist halbwegs okay [...]“ (A2, Z. 56, 58). A6 findet, die Verwendung anderer Diagrammtypen macht „ja gar keinen Sinn“ (A6, Z. 53). Nur A5 gibt an, dass „andere Diagramme auf jeden Fall gut helfen“ (A5, Z. 76) können, allerdings ohne zu konkretisieren, was für ein Diagramm verwendet werden sollte. A1 bewertet zusätzliche (z. B. Säulen-) Diagramme als „nice to have“ (A1, Z. 77), aber nicht notwendig, wohingegen A2 herausstellt, dass eine Darstellung in Säulendiagrammen „eher das Potenzial hätte, eben, wenn man denn schlechter ist, so ein bisschen zu frustrieren“ (A2, Z. 56), da ein Unterschied zwischen dem eigenen Ergebnis und bspw. einem Vergleichswert „ein bisschen stärker noch in den Vordergrund [tritt], als er das jetzt [im Boxplot] tut“ (A2, Z. 56). Abgesehen von der Wahl des Diagrammtyps kommentieren die Studierenden die konkret dargestellten Boxplots: A1 merkt an, dass Ergebnisse, die auf den Grenzen der Antennen des Boxplots liegen, nicht ausreichend auffallen und stärker hervorgehoben werden sollten (A1, Z. 33). A3 spricht darüber hinaus an, dass laut Achsenbeschriftung an den Boxplots fälschlicherweise 0 % bis 1 % und nicht 0 % bis 100 % dargestellt werden (A3, Z. 8), was allerdings zum Zeitpunkt der Interviews bereits korrigiert war. A2 beschreibt, die Darstellung der Antennen im Boxplot als nicht nötig, allerdings auch nicht störend einzuschätzen („die [Information] bräuchte ich zum Beispiel gar nicht, aber die gehört halt dazu“ (A2, Z. 58)). A3 schlägt vor, die Hinweise zur Interpretation durch die Beschriftung eines Boxplots zu ersetzen: „[M]an könnte ja einfach mal diesen ähm Bereich, also vom oberen Teil der Antenne

bis zum unteren Teil der Antenne, [...] markieren, entweder mit einer geschweiften Klammer, oder halt [...] dranschreiben, dass das eben äh ja quasi alle Ergebnisse sind und dann eben das, was quasi auch unten im Text nochmal dran steht halt einfach nur an die Diagramme beschriften, dass man [...] an den Median dran schreibt, dass es der Median ist und äh an [...] diesen mittleren Teil eben, dass es die mittleren fünfzig Prozent sind und [...] dann [...] könnte [man], ohne, dass man weiß, was ein Boxplot ist, eigentlich schon verstehen was einem dieses Diagramm aussagt“ (A3, Z. 62).

Insgesamt wird also deutlich, dass die Studierenden die Darstellung der Ergebnisse in Boxplots als geeignet für die erhaltene Rückmeldung bewerten, wobei einzuschränken ist, dass A2 eine inhaltlich falsche Begründung für diese Einschätzung abgibt: Man würde „in den allermeisten Fällen mit seinem roten Punkt innerhalb der Box liegen“ (A2, Z. 56). Aufgrund dieser explizit geäußerten Zustimmung zur gewählten Darstellungsform sowie der Ablehnung bzw. zumindest skeptischer Betrachtung anderer Diagrammtypen und der zusätzlich oben beschriebenen angemessenen Einordnung und Interpretation der erhaltenen Ergebnisse durch die Studierenden, wird die Darstellung der Ergebnisse in Boxplots beibehalten. Gleichzeitig werden die von den Studierenden vorgeschlagenen Änderungen umgesetzt, die das Verständnis der Boxplots weiter unterstützen sollen: Die individuellen Ergebnisse werden optisch stärker hervorgehoben, indem der rote Punkt vergrößert wird. Andererseits wird, dem Vorschlag von A3 (Z. 62) folgend, der erste Boxplot mit der Darstellung des Gesamtergebnisses um Hinweise ergänzt, sodass neben der Beschriftung des Medians geschweifte Klammern die Bereiche anzeigen, in denen jeweils die unteren bzw. oberen 25 % (untere bzw. obere Antenne des Boxplots) sowie die mittleren 50 % der Ergebnisse in der Vergleichsgruppe (Box des Boxplots) liegen. Entgegen dem Vorschlag von A3 (Z. 62) werden die schriftlich formulierten Hinweise am Ende der Rückmeldung nicht entfernt, sondern bleiben als zusätzliche, etwas ausführliche Hilfestellung bestehen, die abhängig vom individuellen Bedarf genutzt bzw. übersprungen werden können.

Die Studierenden befürworten insgesamt die Ergänzung der Rückmeldung um Beispielaufgaben sowie einer Musterlösung (A1, Z. 81; A2, Z. 52; A3, Z. 64, 66; A4, Z. 51, 53; A5, Z. 82, 84; A6, Z. 59; Kat. AF5i), A2 bevorzugt eine Variante, in der alle Aufgaben des Instruments nachvollzogen werden können (A2, Z. 52) und Begründungen für die Lösung sowie didaktische Hinweise gegeben werden (A2, Z. 18). Nur A6 findet, eine Beispielaufgabe „müsste meiner Meinung nach nicht unbedingt rein“ (A6, Z. 57). A1 sieht in der Ergänzung einer Aufgabe mit begründeter Musterlösung den Vorteil, dass „man dann eben sehr gut nachvollziehen kann, wie das Ganze funktioniert und wie es zu dem Wert geführt hat, den man erreicht hat“ (A1, Z. 81), also die Bewertung der Testbearbeitung für die Studierenden nachvollziehbar wird. A5 sieht in der Einbindung einer Beispielaufgabe oder mehrerer Beispielaufgaben den Vorteil, dass diese eine „Hilfe zur Erinnerung“ (A5, Z. 82) an den Reflexionstest darstellt. Zwei Studierende formulieren Einschränkungen:

A6 befürchtet, wenn die dargestellte Beispielaufgabe aus dem Testinstrument entnommen ist, könnte es „für Verwirrung dann wieder sorgen, wenn man [...] gerade die Aufgabe [...] komplett falsch angekreuzt hat“ (A6, Z. 59) und daraus schließt, die Aufgaben im Reflexionstest insgesamt falsch bearbeitet zu haben (ebd.) und A5 schlägt vor, die dargestellte Lösung „nicht unbedingt Musterlösung [zu] nennen, weil ja vielleicht [...] nicht nur ein Weg der richtige ist“ (A5, Z. 84).

Die Studierenden wünschen sich also – wenn auch in unterschiedlicher Ausgestaltung – die Ergänzung von Beispielaufgaben (ggf. aus dem Testinstrument) sowie Musterlösungen zu Beispielaufgaben bzw. allen Testaufgaben (Kat. AF5i). Auf die Ergänzung der Rückmeldung um Aufgaben und Musterlösungen wird verzichtet, um einerseits die Nutzung des Reflexionstests weiterhin - auch mehrfach in Prä-Post-Szenarien - zu ermöglichen und den Umfang der Rückmeldung nicht zu erhöhen (es müssten 213 Antwortoptionen aufgeführt werden). Gleichzeitig ist die Bedeutung von Beispielaufgaben und Musterlösungen für die Studierenden als „Hilfe zur Erinnerung“ (A5, Z. 82) an das Instrument nachvollziehbar und wird durch das – zeitlich nach den hier diskutierten Interviews entwickelte und daher in deren Rahmen nicht berücksichtigte - Selbstlernmaterial aufgefangen. Im Rahmen des Materials sind zusätzlich zu allgemeinen Hinweisen zur Reflexion von Physikunterricht drei weitere Videovignetten aus dem Unterricht eingebettet. Diese Vignetten sollen von den Studierenden (angeleitet) reflektiert werden und sind mit Musterlösungen versehen, die entsprechend des Hinweises von A5 „weil ja vielleicht [...] nicht nur ein Weg der richtige ist“ (A5, Z. 84) „Beispiellösungen“ genannt werden. Mit diesem Selbstlernmaterial wird den Studierenden auf inhaltlicher Ebene ermöglicht, sich an die Situation der Unterrichtsstunde zu erinnern (obwohl hier andere Phasen gezeigt werden, als im Reflexionstest, sind die Rahmenbedingungen des Unterrichts erneut gegeben) sowie die Struktur des Reflexionstests in Bezug auf die Zweiteilung der Aufgaben zu jedem Aspekt in eine Aufgabe zur Bewertung und eine zum Vorschlagen von Alternativen nachzuvollziehen, die – wie oben beschrieben – auch im Rahmen der Rückmeldung expliziert wird.

In Bezug auf die Ergänzung von Literaturhinweisen in der Rückmeldung (Kat. AF5j) äußern sich die Studierenden eher vorsichtig. A1 sieht hier „für uns Studenten keinen großen Vorteil drin“ (A1, Z. 83), auch A5 vermutet, dass die Ergänzung von Literaturhinweisen „nicht unbedingt notwendig ist, weil man ja eh [in der Veranstaltung] mit den Thematiken konfrontiert wird“ (A5, Z. 86). A1 und A4 vermuten außerdem, dass solche Hinweise kaum genutzt würden (A1, Z. 85; A4, Z. 63). Demgegenüber halten A2, A3, A4 und A6 eine Ergänzung von Literaturhinweisen in unterschiedlicher Form für sinnvoll, laut A3 insbesondere „für die Leute, die das [Anm.: Vermutlich ist hier eine Verbesserung der Reflexionsfähigkeit gemeint] dann wirklich versuchen“ (A4, Z. 63). A2 findet Hinweise sinnvoll, „wenn man die Möglichkeit hat, an Stellen, die einen interessieren, nachzulesen“ (A2, Z. 66), würde diese aber nicht zu einem festen Bestandteil der Rückmeldung machen wollen (A2, Z. 66), und sieht ebenso wie A3 und A4 den Mehrwert insbesondere

bei der Verbindung von Literaturhinweisen mit Handlungsoptionen für die Studierenden (A2, Z. 66; A3, Z. 68, 73; A4, Z. 65). Außerdem hält A3 Literaturhinweise für den Bereich der Schülervorstellungen für geeignet, bevorzugt für den Bereich der Alternativen allerdings andere Hinweise: „[B]ei Alternativen [...] ist es vielleicht eher so, dass man einen Denkanstoß [...] geben muss, [...] wie man da ran geht, also vielleicht einfach mal, selbst wenn man eine gute Lösung gefunden hat, nochmal über eine Zweite nachdenken." (A3, Z. 40). A6 sieht einen Mehrwert bzw. eine „zusätzliche Hilfe" (A6, Z. 65) von Literaturhinweisen in der Rückmeldung, wenn sie als Nachschlageoption für das weitere Studium genutzt werden können (z. B. „vor einer wichtigen Prüfung oder ähm eine[r] wichtige Hausarbeit" (A6, Z. 63)).

Aufgrund der fehlenden Einigkeit auf Seiten der Studierenden in Bezug auf die Ergänzung von Literaturhinweisen wird diese nicht in der Rückmeldung, sondern im Rahmen des Selbstlernmaterials vorgenommen. Die ergänzten Literaturhinweise beziehen sich einerseits auf das Vorgehen bei der Reflexion von Unterricht und andererseits überwiegend auf die reflektierten Aspekte und bieten die Möglichkeit, fachdidaktische Inhalte nach individuellen Bedarfen nachzulesen. Dieser Bedarf kann entstehen, wenn die Studierenden durch das Selbstlernmaterial angeleitete Reflexionen durchführen und Abweichungen zwischen der eigenen Einschätzung und der Beispiellösung feststellen. So werden die Literaturhinweise, wie von A2, A3 und A4 gewünscht (A2, Z. 66; A3, Z. 68, 73; A4, Z. 65), zumindest mit Handlungen verknüpft (wenn auch nicht mit individuellen Handlungsempfehlungen für die Studierenden), die sich an die erhaltene Rückmeldung anschließen. Gleichzeitig bleibt die Rückmeldung, wie von A2 (Z. 66) gefordert, auch ohne Literaturhinweise verständlich. Da die Literaturhinweise auf die reflektierten Aspekte der im Selbstlernmaterial genutzten Vignetten aufbauen, werden so allerdings nicht alle Inhalte des Reflexionstests mit Literatur gestützt.

Wie bereits beschrieben wurde, haben die Studierenden Schwierigkeiten damit, aus der Rückmeldung konkrete Handlungen zur Umsetzung von identifizierten Verbesserungspotenzialen abzuleiten. Auf Nachfrage sehen auch alle Studierenden Bedarfe, die Entwicklung der Reflexionsfähigkeit z. B. durch ein Schulungsmaterial zu unterstützen (A1, Z. 87; A2, Z: 64; A3, Z. 36; A4, Z. 59; A5, Z. 88; A6, Z. 73ff.; Kat. AF5k), auch „um dann für das spätere Berufsleben gewappnet zu sein" (A5, Z. 88). A2 gibt zwar an, das Vorgehen einer Reflexion „bis zum gewissen Punkt verstanden" (A2, Z. 64) zu haben, bräuchte aber trotzdem eine Antwort auf die Frage: „[W]as mache ich jetzt mit dieser Erkenntnis, okay, in dem Bereich A habe ich halt noch Verbesserungspotenzial, wie setze ich das um" (A1, Z. 64). A3 und A4 wünschen sich auf das individuelle Ergebnis zugeschnittene Hinweise bzw. Handlungsempfehlungen (z. B. „bei diesem Beispielergebnis hier wäre, dass [...] dem ja (.) Teilnehmer empfohlen wird, dass er sich nochmal mit Schülervorstellungen auseinandersetzt" (A3, Z. 36; ähnlich auch A4, Z. 59)). A6 schränkt ein, dass Literaturhinweise an dieser Stelle nicht ausreichend hilfreich seien, sondern es besser wäre, „mit den Dozenten irgendwie ab[zu]sprechen, dass man das [...] [in]

einer Seminarstunde [...] trainiert“ (A6, Z. 77), um die Handlungen praktisch zu üben (A6, Z. 79). Außerdem sieht A6 den Vorteil, dass man nach dem Üben „den Test irgendwie in ein oder zwei Wochen [...] nochmal später durchführt [und] man dann eine [...] Verbesserung sieht“ (A6, Z. 73). A1 ist es wichtig, anzumerken, dass die Einbindung von Hinweisen zur Schulung der Reflexionsfähigkeit in die Rückmeldung nicht ausreichend ist, sondern „vielleicht die äh Lehre noch angepasst werden kann, um das besser zu schulen“ (A1, Z. 87). A1 betont, „da ist nicht nur der Student selbst in der Verantwortung“ (A1, Z. 87).

Um die Studierenden nach dem Erhalt der Rückmeldung dabei zu unterstützen, sich weiter mit der Reflexion von Unterricht auseinanderzusetzen, wird in der Rückmeldung das Selbstlernmaterial verlinkt. Dieses Material ermöglicht eine durch schriftliche Hinweise angeleitete Reflexion von drei Unterrichtsausschnitten und unterstützt auf methodischer Ebene so gezielt das Üben von Reflexion, wie es sich A6 (Z. 79) wünscht. Trotz dieser Unterstützungsmaßnahme kann die Beschäftigung mit und das Üben von Unterrichtsreflexion nicht als abgeschlossen betrachtet werden, sondern bedarf weiterer Unterstützung im Rahmen der Lehramtsausbildung (vgl. auch A1, Z. 87). So ist bspw. die Reflexion von eigenem Unterricht nicht Gegenstand des Materials und der Abgleich von eigener Einschätzung und Musterlösung muss selbständig erfolgen, sodass die Wirksamkeit bzw. der Mehrwert des Selbstlernmaterials abhängig von der Art der Auseinandersetzung mit diesem stark variieren kann.

Sonstige Anmerkungen zum Reflexionstest Zwei der Studierenden formulieren Ideen, wie der Reflexionstest gestaltet werden sollte (Kat. S1a), damit es ihnen einen Mehrwert bietet. A2 schlägt vor, dass „man einfach auch an der ein oder anderen Stelle einen relevanten Aspekt vorgibt, zu dem Rückmeldung gegeben werden soll“ (A2, Z. 38), da A2 nicht sicher ist, im Noticing die relevanten Aspekte identifiziert zu haben und dementsprechend auch die Reflexion zu vorgegebenen (relevanten) Aspekten für sinnvoll erachtet. A1 merkt an, dass sich die Anforderung durch die anwählbaren Antwortoptionen von einer offenen Situation unterscheiden und würde anschließend an das bearbeitete Testinstrument einen offeneren Test bearbeiten: „[E]s wäre für mich jetzt interessant, nachdem ich jetzt gesehen habe, dass ich das mit dem Auswählen vielleicht schon ganz gut kann, wie es denn dann wäre, wenn ich das selbst reproduzieren und anwenden muss, also vielleicht wäre dann so ein aufbauender Test irgendwie in die Richtung ähm was würde die Person dann selbst sagen, schreiben, vorschlagen oder so interessant“ (A1, Z. 55).

Passend zum Vorschlag von A2 werden die zu reflektierenden Aspekte unabhängig von der Bearbeitung der Aufgaben zum Noticing vorgegeben. Damit wird der Reflexionstest weiter standardisiert, sodass immer die gleichen Aufgaben bearbeitet werden. Die Überlegung von A1, anschließend an die Bearbeitung des geschlossenen Instruments ein offeneres Instrument bearbeiten zu wollen, um anstelle der

Auswahl gegebener Optionen aufgefordert zu sein, frei zu reflektieren (A1, Z. 55), kann im Rahmen der Validierung des Testinstruments nicht umfassend gelöst werden. Grundsätzlich bestünde z. B. die Möglichkeit, den offenen Performanztest zur Reflexion von Physikunterricht von Kempin, Kulgemeyer und Schecker (2018) als Instrument zur Testung von Handeln in authentischeren Situationen zu verwenden. Dieses Testinstrument ist allerdings mit einer vergleichsweise aufwendigen Auswertung verbunden, sodass ein ökonomischer Einsatz nicht möglich ist und beinhaltet die im geschlossenen Instrument verwendeten Videovignetten, sodass Studierende nach einer Bearbeitung des geschlossenen Instruments bereits bestimmte Inhalte reflektiert haben und ihre Einschätzung unter Umständen durch die gegebenen Antwortoptionen beeinflusst sind. Eine andere Möglichkeit, Reflexionen in einem offeneren, aber dennoch angeleiteten Setting durchzuführen, ist die Nutzung des neu entwickelten Selbstlernmaterials zur Reflexion von Physikunterricht (s. Kapitel 11.2), mit welchem Studierende angeleitet reflektieren und ihre Einschätzung mit Beispiellösungen abgleichen können. Hier erfolgten also keine Auswertung und Rückmeldung der Ergebnisse an die Studierenden, sondern ein selbstständiger Abgleich von eigenen Überlegungen und Lösungen.

Erhalt der Rückmeldung Einige Tage nach Abschluss der Testbearbeitung, konnten die Studierenden ihre Rückmeldung unter Eingabe ihres persönlichen Codes auf einer eigens dafür eingerichteten Website herunterladen. Drei der Studierenden sehen bezüglich dieses Vorgehens keinen Änderungsbedarf, solange den Studierenden der Link zur Website vorliegt (A4, Z. 72; A5, Z. 89; A6, Z. 80; Kat. S2a), A4 spricht lediglich die vage Überlegung an, die Rückmeldung mit e-Learning-Profilen zu verknüpfen (A4, Z. 72). Die übrigen Studierenden geben zwar auch an, dass sie ihre Rückmeldung ohne Schwierigkeiten erhalten haben (A1, Z. 91; A2, Z. 68; A3, Z. 77; Kat. S2a), würden aber den Erhalt der Rückmeldung sofort nach Abschluss des Testinstruments bzw. auf der letzten Testseite bevorzugen (A1, Z. 97; A2, Z: 72; A3, Z. 77, 79; Kat. S2b), denn „je länger das Ganze her ist, desto weniger interessiert man sich für die Rückmeldung“ (A2, Z. 72). A3 schlägt vor, zusätzlich die Option beizubehalten, die Rückmeldung herunterladen zu können, „wenn man das dann nicht gespeichert hat“ (A3, Z. 79; Kat. S2a). Außerdem sieht A1 bei einem hypothetischen Versand der Rückmeldung per E-Mail den Vorteil, dass „viele sie auch sehen und wahrnehmen und lesen, während wenn man die halt zum Runterladen anbietet, dann vielleicht Leute doch wieder nicht dazu kommen und es wieder vergessen und so und dann vielleicht ähm Bedarfe, die sie eigentlich haben, da noch weiter zu arbeiten, dann untergehen“ (A1, Z. 93; Kat. S2c).

Insgesamt zeigt sich, dass die Ausgabe der Rückmeldungen über die eingerichtete Website - sofern der Link zur Website vorliegt - problemlos funktioniert (u. a. A1, Z. 91; A2, Z. 68; A3, Z. 77) und somit kein unbedingter Änderungsbedarf besteht (A4, Z. 72; A5, Z. 89; A6, Z. 80; Kat. S2a). Gleichzeitig äußern einige

der Studierenden die Änderungswünsche, die Rückmeldung sofort nach Abschluss der Bearbeitung des Instruments zu erhalten (A1, Z. 97; A2, Z: 72; A3, Z. 77, 79; Kat. S2b) bzw. die Rückmeldung persönlich zu erhalten (z. B. per E-Mail, A1, Z. 93, oder über e-Learning-Systeme, A4, Z. 72; Kat S2c), sodass diese von mehr Studierenden wahrgenommen wird (A1, Z. 93). Letzteres ist bei Wahrung der Anonymität nicht möglich, daher erfolgt eine Zuordnung der Rückmeldung weiterhin anhand des individuellen persönlichen Codes.

Die Option, die Rückmeldung sofort nach Beendigung der Bearbeitung des Testinstruments zu erhalten (Kat. S2b) bedarf zumindest in der vorliegenden technischen Umsetzung des Testinstruments einer ausreichend großen Datenbasis, die den Rückmeldungen zugrunde gelegt werden kann. Diese Datenbasis liegt bislang nicht vor, sodass auch dieser Änderungsvorschlag vorläufig nicht umgesetzt wird. Perspektivisch erscheint eine solche vollautomatisierte Erstellung der Rückmeldungen sinnvoll, um einen flexiblen und individuellen Einsatz des Testinstruments zu ermöglichen.

Schlussfolgerungen für die Überarbeitung des Assessment-Feedbacks und dessen Evaluation

Insgesamt wird im Rahmen der Interviews deutlich, dass die Studierenden das erhaltene Assessment-Feedback überwiegend sinnvoll auf Grundlage der Einordnung der Testwerte in die Vergleichsgruppe interpretieren; Unklarheiten und Fehler erscheinen nicht gravierend, sodass das Assessment-Feedback mit den diskutierten Änderungen weiter eingesetzt wird. Die basierend auf den Erkenntnissen aus den Interviews umgesetzten Änderungen werden im nächsten Kapitel einleitend zur Evaluation des überarbeiteten Assessment-Feedbacks zusammengefasst. Die Bewertung des Indikators 4.1 *„Die Interpretationen der Ergebnisse durch die Teilnehmenden passen inhaltlich zum erhaltenen Feedback.“* erfolgt aufgrund großer Überschneidungen in den Ergebnissen anschließend an die Darstellung und Diskussion der Ergebnisse der Evaluation der überarbeiteten Version des Assessment-Feedbacks.

18.1.2 Evaluation des weiterentwickelten Assessment-Feedbacks

Die überarbeitete Version des Assessment-Feedback entspricht in seiner Struktur zu großen Teilen der vorigen Version: Auf die nun um die Angabe des Gesamtergebnisses ergänzte Einleitung folgt die Beschreibung des Begriffsverständnisses, ergänzt um eine ausführlichere Zuordnung der Aufgaben im Testinstrument zu den einzelnen Teilfacetten. Anschließend werden die individuellen Ergebnisse rückgemeldet, wobei der Boxplot um erläuternde Beschriftungen sowie eine Darstellung

der Effektstärke ergänzt wurde. Diese Einbindung der Effektstärke wird auf Basis informeller Rückmeldungen durch Dozierende (nicht auf Grundlage der mit den Studierenden geführten Interviews) vorgenommen. Zur besseren Einordnung der Ergebnisse in Bezug auf die Ergebnisse der Vergleichsgruppe wird im Boxplot zum Gesamtergebnis auf einer zweiten Achse die Effektstärke eingetragen (ein Beispiel für ein so überarbeitetes Diagramm ist auf Seite zwei der Beispielrückmeldung in Kapitel J.3.5 im Anhang ab Seite 503 enthalten). So wird der Unterschied zwischen dem eigenen Ergebnis und dem Mittelwert der Vergleichsgruppe in ein standardisiertes Maß überführt. Auch die bereits von den Studierenden geäußerte Bewertung der eigenen Ergebnisse als tendenziell positiv, sofern sie über dem Median liegen, und negativ, wenn sie unter dem Median liegen wird so gestützt, da Ergebnisse über bzw. unter dem Median mit einer positiven bzw. negativen Effektstärke verbunden sind. Anzumerken ist, dass die Effektstärke unabhängig von einer Signifikanzprüfung des Unterschieds dargestellt wird, da das Ergebnis einer Einzelperson mit dem der Vergleichsgruppe in Beziehung gesetzt wird. Zusätzlich wird das entwickelte Selbstlernmaterial eingebunden und als eine Möglichkeit zur weiterführenden Beschäftigung mit Unterrichtsreflexion beschrieben.

Durchführung der Interviews

Zum überarbeiteten Assessment-Feedback wurden im Sommersemester 2022 erneut $N = 6$ strukturierte leitfadengestütze Interviews mit Physik-Lehramtsstudierenden im Bachelorstudium der Universität Bremen geführt (nachfolgend insgesamt als Kohorte B bezeichnet; die Studierenden werden mit B1 bis B6 benannt), die inhaltlich auf den oben beschriebenen Interviews der Pilotierung aufbauen und sich in vergleichbare Phasen gliedern lassen: Im Einstieg wurden die Studierenden gebeten, wiederzugeben, welche Schlüsse sie aus ihrem individuellen Assessment-Feedback gezogen haben. Daran anschließend wurde anhand einer Beispielrückmeldung (s. Kapitel J.3.5 ab S. 503) die Interpretation und Verständlichkeit des Assessment-Feedbacks diskutiert. Weiterhin wurden das Vorgehen bei der Arbeit mit dem Selbstlernmaterial und die Perspektive der Studierenden auf das Material erfragt. Abschließend bestand die Möglichkeit, weitere Fragen oder Anmerkungen der Studierenden zu den eingesetzten Materialien zu klären.

Zusätzlich wurden im Wintersemester 2022/23 $N = 5$ Interviews mit Physik-Lehramtsstudierenden im Bachelorstudium der Universität Bremen geführt (nachfolgend als Kohorte C bezeichnet; die Studierenden werden mit C1 bis C5 benannt). Für vier der fünf Studierenden schloss sich das Interview an den Erhalt des Assessment-Feedbacks zur zweiten Bearbeitung des Reflexionstests an; mit zwei der Studierenden wurde zudem bereits im vorigen Semester ein Interview geführt.[181] Das Assessment-Feedback, das die Studierenden in dieser Kohorte

[181]Diesen Studierenden sind die Benennungen B4 und C2 bzw. B6 und C5 zugeordnet.

erhalten haben, unterscheidet sich in zwei Punkten von dem der Interviewkohorte B und ist in Kapitel J.5.7 (ab S. 511) einsehbar:

1. Die Darstellung der Effektstärke im Boxplot zum Gesamtergebnis wurde entfernt.
2. Der Verweis auf das Selbstlernmaterial wird - abhängig vom individuell niedrigeren Teilergebnis - um den konkreten Hinweis ergänzt, sich bei der Arbeit mit diesem auf das Bewerten von Unterricht oder Entwickeln von Alternativen zu fokussieren.

Aufgrund der inhaltlich ähnlichen Ausrichtung dieser Interviews zu denen der vorangegangenen Pilotierung weisen die Interviewleitfäden gewisse Übereinstimmung bezüglich der Wahrnehmung und Interpretation des Assessment-Feedback durch die Studierenden auf. Die Interviewleitfäden sind in Kapitel J.3 (ab S. 497) und J.5 (ab S. 504) einzusehen. So wird eine gewisse Vergleichbarkeit beider Interviewstudien ermöglicht und es können die vorgenommenen Änderungen und Ergänzungen evaluiert werden. Die Interviews wurden wieder vollständig audiografiert.

Auswertung der Interviews

Analog zu den Interviews der Pilotierung wurden die Aufnahmen transkribiert und kategoriengeleitet analysiert (vgl. Krüger und Riemeier, 2014).[182] Die Analyse basiert auf dem in Kapitel 18.1.1 beschriebenen Kategoriensystem zur Pilotierung des Assessment-Feedbacks, wobei zusätzlich entlang des Interviewleitfadens deduktiv die Kategorie AF6 zur Einbindung der Effektstärke im Assessment-Feedback sowie Kategorien zur Arbeit mit dem Selbstlernmaterial (SLM1 - SLM4) aufgenommen und induktiv um Subkategorien ergänzt wurden. Auch in den übrigen Kategorien wurden Subkategorien hinzugefügt. Das aktualisierte Kategoriensystem ist in Tabelle 18.2 dargestellt.

Ergebnisse und Schlussfolgerungen

Die Interviews weisen eine durchschnittliche Dauer von etwa 15 Minuten (neun bis 20 Minuten) auf. Analog zur Ergebnisdarstellung und -diskussion der ersten Interviewerhebung wird dieses Kapitel in inhaltliche Abschnitte gegliedert. Zunächst wird dabei wieder auf die Darstellung der Ergebnisse mit Boxplots, auf die aus den Ergebnissen gezogenen Schlussfolgerungen sowie auf die Bewertung (von Bestandteilen) des Assessment-Feedbacks fokussiert. Zusätzlich wird für die Studierenden der Kohorte C, die bereits zuvor ein Assessment-Feedback erhielten, auf hergestellte Bezüge zwischen beiden Ergebnissen eingegangen. Anschließend

[182] Auch für diese Interview sind die vollständigen, entlang der Sprechbeiträge segmentierten Transkripte der Arbeit angehängt (s. Kapitel P).

wird das Selbstlernmaterial in den Blick genommen. Hierbei wird betrachtet, ob (Interviewkohorte C) und wie (Interviewkohorte B) die Studierenden mit diesem arbeiten, welchen Mehrwert sie im Material und dessen Bestandteilen sehen und welche Änderungsbedarfe ggf. bestehen. Abschließend wird auf sonstige Kommentare der Studierenden zum Reflexionstest eingegangen.

Auf Entwicklungen in der Interpretation des Assessment-Feedbacks und im Umgang mit diesem wird nicht explizit eingegangen, da lediglich zwei Studierende Teil beider Interviewkohorten (B und C) sind.

Rückmeldung			
Kat. 1 Boxplots		**Kat. 4 Gewichtung der Bestandteile**	
AF1a	Nicht bekannt	AF4a	(Eigene) Ergebnisse
AF1b	Bekannt aus der Ausbildung	AF4b	Sonstige Bestandteile
AF1c	Bekannt aus der Schulzeit	AF4c	Text nicht gelesen
AF1d	Bekannt aus dem Studium	**Kat. 5 Änderungsbedarfe Rückmeldung**	
AF1e	Hilfestellungen zum „Lesen“ der Boxplots	AF5a	Keine
AF1f	Unklarheiten	AF5b	Umfang der Rückmeldung/-Überflüssige Inhalte
Kat. 2 Interpretation		AF5c	Formulierungen/inhaltliche Struktur
AF2a	Orientierung an der Vergleichsgruppe	AF5d	Ergänzen von Informationen zu Zielen des Projekts
AF2b	Orientierung an Prozentwerten bzw. am Maximalwert von 100%	AF5e	Ergänzen von Informationen zum Zielzustand der Reflexionsfähigkeit
AF2c	Vergleich der eigenen (Teil-)Ergebnisse	AF5f	Aufschlüsselung der Ergebnisse auf Aufgabenebene
AF2d	Fehler in der Interpretation	AF5g	Anpassung der Vergleichsgruppe
Kat. 3 Schlussfolgerungen		AF5h	Darstellung in Boxplots
AF3a	Bedeutung/Mehrwert des Ergebnisses	AF5i	Ergänzen einer Musterlösung und bzw. oder von Beispielaufgaben
AF3b	Bewertung des Ergebnisses (z. B. als positiv)	AF5j	Ergänzen von Literaturhinweisen
AF3c	Identifikation von Verbesserungspotenzial	AF5k	Informationen zur Förderung der Reflexionsfähigkeit

AF3d	Beschreibung möglicher Ursachen für die Ergebnisse	**Kat. 6 Effektstärke**	
AF3e	Unklarheiten bzgl. der (Interpretation der) Ergebnisse	AF6a	Nicht bekannt
		AF6b	Bekannt aus dem Studium
Selbstlernmaterial			
Kat. 1 Vorgehen		SLM3	**Kat. 3 Gewichtung**
SLM1a	(Entscheidung für die) Bearbeitung		
SLM1b	Betrachtung der Videos	**Kat. 4 Änderungsbedarfe Selbstlernmaterial**	
SLM1c	Eigene Überlegungen vor Abgleich mit Lösungen	SLM4a	Keine
SLM1d	Direkter Einbezug der Lösungen	SLM4b	Umfang des Materials
Kat. 2 Mehrwert		SLM4c	Wiederholungen
SLM2a	Selbstevaluation		
SLM2b	Informationen zum Vorgehen bei der Unterrichtsanalyse		
SLM2c	Sensibilisierung für konkrete Situationen		
SLM2d	„Nachschlagewerk“		
Sonstiges			
Kat. 1 zum Testinstrument		**Kat. 2 Erhalt der Rückmeldung**	
S1a	Gestaltung (Aufbau und Umfang) des Testinstruments	S2a	Problemlos
S1b	Interpretation als Beispiel für Prüfungsleistung	S2b	Instantan gewünscht
S1c	Gelungene Videos	S2c	Zusenden (z. B. per Mail) gewünscht
S1d	Interaktivität des Instruments/Materials		

Tabelle 18.2: Übersicht über die Kategorien zur Kodierung der Interviews zum überarbeiteten Assessment-Feedback inklusive Selbstlernmaterial (dargestellt ist das vollständige Kategoriensystem für alle Interviewstudien zum Assessment-Feedback, wobei die Kategorien, die im Rahmen der Pilotierung auftraten und hier nicht angesprochen wurden, grau gesetzt sind)

Ist die gewählte Darstellungsform der Ergebnisse in Boxplots für die Studierenden angemessen? Für sechs der neun Studierenden waren Boxplots bis zum Erhalt der Rückmeldung unbekannt (B1, Z. 16; B3, Z. 18; B4, Z. 19, 21, 25; B5, Z. 27, 29; C3, Z. 27; Kat. AF1a), bzw. erinnern sich zumindest nicht daran, mit ihnen gearbeitet zu haben (C4, Z. 30). B2 und B6 kennen Boxplots aus dem Mathematikunterricht der Schule (B2, Z. 10, 12; B6, Z. 26, 30; Kat AF1c); C1 gibt an, sie aus physikdidaktischen Veranstaltungen zu kennen (C1, Z. 13,15), und C2 und C5 sind sie lediglich aus dem zuvor erhaltenen Assessment-Feedback bekannt (C2, Z. 34ff.; C5, Z. 45ff.; beides Kat. AF1d). Fünf Studierende geben außerdem explizit an, die Boxplots (ggf. mit den Erläuterungen der Rückmeldung (B6, Z. 32) bzw. nach dem Lesen der Rückmeldung (B2, Z. 16)) zu verstehen (B1, Z. 18; B2, Z. 16; B4, Z. 27, B6, Z. 32; C3, Z. 27; Kat. AF1e), darunter auch beide Studierende (B2 und B6), denen Boxplots vorher schon bekannt waren (z. B. „also die Erklärung ist ja noch daneben, man kommt ja eigentlich gut damit klar“ (B6, Z. 32)). C4 beschreibt die Antennen des Boxplots fälschlicherweise als „Fehler“ (C4, Z. 54; Kat. AF1f), gibt aber nach einer Klärung des Missverständnisses an, die Rückmeldung „nicht genau genug durchgelesen“ (C4, Z. 84) zu haben, um die Informationen dazu aus der Rückmeldung zu entnehmen (Kat. AF4c).

Die Effektstärke, die im Assessment-Feedback für die Interviewkohorte B im Boxplot zum Gesamtergebnis dargestellt war, war fünf der sechs Studierenden dieser Kohorte zum Zeitpunkt des Erhalts der Rückmeldung unbekannt (B1, Z. 20; B2, Z. 18; B4, Z. 31; B5, Z. 31, 33; B6, Z. 34; Kat. AF6a). Keine Person spricht sie von sich aus bei der Interpretation der Rückmeldung an; B6 gibt an, die Effektstärke beim Erhalt der Rückmeldung wahrgenommen, aber nicht interpretiert zu haben (B6, Z. 105). Nur B2 interpretiert die Effektstärke auf Nachfrage korrekt als Maß für den Abstand vom eigenen Ergebnis zum Mittelwert der Vergleichsgruppe (B2, Z. 22). Die vier anderen Studierenden, denen die Effektstärke zuvor unbekannt war, können diese nicht interpretieren (B1, Z. 26; B4, Z. 59, 61; B5, Z. 39; B6, Z. 46, 105, 107). B3 ist die Effektstärke aus dem vorherigen Studium bekannt (B3, Z. 20; Kat. AF6b), ist bei der Interpretation der Effektstärke aber unsicher und vermutet, sie sei eine Angabe über die Abweichung des eigenen Ergebnisses vom Median: „Also die Effektstärke null ist ja hier bei dem Median, das heißt, wenn man schaut, okay, wie weit liegt man da drüber, dann (..) müsste das ja- Also null Komma fünf, hat äh eben Herr N [Anm: Professor für Physikdidaktik und Dozent der Veranstaltung in dessen Rahmen der Reflexionstest bearbeitet wurde] gesagt, ist eine hohe Effektstärke, also ab einer Effektstärke von null Komma fünf wird es als relevant oder als hoch eingeschätzt. Kann man schauen, okay gut, wie hoch ist der Unterschied vielleicht? [...] Nein, da bin ich mir nicht ganz sicher.“ (B3, Z. 26).

Bei der Einordnung der in der Rückmeldung dargestellten Ergebnisse (sowohl der eigenen Ergebnisse als auch die der Beispielrückmeldung) orientieren sich alle sechs Studierenden an der Vergleichsgruppe (u. a. dem Mittelwert, Median

und der Box des Boxplots sowie in Interviewkohorte C zusätzlich an der schriftlichen Einordnung des individuellen Gesamtergebnisses in Quartile zu Beginn des Assessment-Feedbacks (B1, Z. 22; B2, Z. 6, 20, 24; B3, Z. 22; B4, Z. 8, 49; B5, Z. 6, 15; B6, Z. 40, 42, 44; C1, Z. 10, 23, 29, 31; C2, Z. 42; C3, Z. 31; C4, Z. 32, 36; C5, Z. 53, 55, 60f.; Kat. AF2a)). Nur zwei Studierende orientieren sich explizit an den dargestellten Prozentwerten (B2, Z. 20; C5, Z. 8; Kat. AF2b).

Gleichzeitig werden bei der Interpretation der Boxplots fachliche Unklarheiten oder zumindest Ungenauigkeiten deutlich (Kat. AF2d): So wird beispielsweise der Median als „Mittelwert“ (z. B. B1, Z. 22) bzw. „Durchschnitt“ (z. B. B2, 6, 20; C1, Z. 23) oder die Box als Bereich „in dem die meisten Ergebnisse liegen“ (B4, Z. 49) bezeichnet. Unklar ist, ob das Einfluss auf die Interpretation der Ergebnisse hat und wie dieser ggf. aussieht. B5 missversteht außerdem den Median bzw. die Box, die die mittleren 50 % der Ergebnisse darstellt, als anzustrebenden Ergebnisbereich, korrigiert diese Vorstellung während der Arbeit mit der Rückmeldung aber selbst (B5, Z. 13, 23).

Insgesamt lässt sich festhalten, dass die Darstellung der Ergebnisse in Boxplots für die Studierenden zwar oft unbekannt (Kat. AF1a) war, allerdings trotzdem ein – wenn auch fachlich nicht immer umfänglich korrektes (Kat. AF2d) – Verständnis für die Ergebnisdarstellung aufgebracht werden kann. Die Darstellung der eigenen Ergebnisse gemeinsam mit denen der Vergleichsgruppe in Form eines Boxplots kann also für die Befragten als geeignet bewertet werden. Für die Studierenden der Kohorte B unbekannt und dementsprechend in der Interpretation meist unklar (Kat. AF6a) ist die Angabe der Effektstärke als Maß für die Abweichung des eigenen Ergebnisses vom Mittelwert der Vergleichsgruppe. Die Einbettung der Effektstärke in die Rückmeldung erscheint in dieser Form also nicht sinnvoll, da sie keinen inhaltlichen Beitrag zur Interpretation der Ergebnisse darstellt. Daher wird die Angabe der Effektstärke zwischen der Durchführung der Interviews mit den Kohorten B und C aus dem Assessment-Feedback für die Studierenden entfernt, um den Umfang der Rückmeldung (weitgehend) ohne Informationsverlust für die Befragten zu reduzieren und keine unbekannten Maße in die Rückmeldung einzubinden (vgl. Hosenfeld, 2010, S. 41). Alternativ dazu bedürfte die Darstellung einer einfach zugänglichen Erläuterung des Begriffs für die Studierenden. Bei der Einordnung der Ergebnisse orientieren sich alle Studierenden primär an den im Boxplot dargestellten Ergebnissen der Vergleichsgruppe (Kat. AF2a); ein Bezug zu den maximal erreichbaren 100 % der Punkte (Kat. AF2b) wird von zwei Personen hergestellt.

Welche Schlussfolgerungen ziehen die Studierenden aus der (erhaltenen bzw. Beispiel-) Rückmeldung? Die Studierenden äußern verschiedene Vorstellungen in Bezug darauf, was ihnen die Rückmeldung bietet und welches Gewicht diese Rückmeldung für sie hat (Kat. AF3a). B1 gibt an, dass die Rückmeldung das Ver-

besserungspotenzial in den Teilbereichen angibt (B1, Z. 10). Für B4 liefert sie eine „kurze [...] grobe Rückmeldung zu [B4s] didaktische[m] Verständnis“ (B4, Z. 33) in den Teilfähigkeiten, wobei zu vermuten ist, dass B4 zumindest das Ergebnis zum Erkennen und zum Umgang mit Schülervorstellungen nicht korrekt interpretiert (B4 beschreibt die Teilergebnisse als Informationen über „Schülervorstellungen, die ich selber noch habe“ und das „faire Bewerten“ (B4, Z. 37); das (eigene) Ergebnis zu den Alternativen kann B4 offenbar nicht einordnen, da B4 unklar ist, was im Kontext der Rückmeldung unter dem Begriff „Alternativen“ zu verstehen ist (B4, Z. 41)). Gleichzeitig sieht B4 ein schlechtes Ergebnis als unproblematisch an, da die Rückmeldung nur grob ist: „Ähm ich würde mir aber auch nicht zu viele Gedanken machen, wenn ich jetzt schlecht abschneiden würde im Test, ähm da das halt nur sehr grob ist.“ (B4, Z. 35). B3 sieht in Test und Rückmeldung (und Selbstlernmaterial) die Möglichkeit zur Selbsteinschätzung der ins Studium mitgebrachten Reflexionsfähigkeit im Vergleich zur Vergleichsgruppe („wir hatten Reflexion ja noch nicht in der Vorlesung, deswegen war es irgendwie eher etwas, okay, was bringe ich schon mit im Vergleich“ (B3, Z. 22)) sowie einen Einblick, wie Reflexionen in der Modulprüfung oder im Referendariat aufgegriffen werden können: „Also ich fand [...] den Reflexionstest insoweit gut, weil ich mir vorher für die das Modul auch gar nicht so sicher war, wie so eine Prüfungsleistung aussehen kann oder wie Fragen formuliert werden können. Und daran habe ich mich so ein bisschen orientiert jetzt“ (B3, Z. 12). C3 beurteilt die Relevanz des Ergebnisses vorsichtig und, da das Niveau der Reflexionsfähigkeit der Vergleichsgruppe nicht detailliert bekannt ist, als eher relevant „für die Forschung als für mich selber dann“ (C3, Z. 42).

Die Studierenden ziehen aus den (Teil-)Ergebnissen positive Schlussfolgerungen (B4, Z. 53; B5, Z. 35, B6, Z. 40-44; C1, Z. 31; C3, Z. 12, 34; C4, Z. 14; Kat. AF3b), wenn diese über dem Median liegen, wie z. B. „[D]iese Rückmeldung ist ja sehr positiv erstmal zum Bewerten[, w]eil das Ergebnis hier im oberen Mittel liegt.“ (B6, Z. 40, 42). B2 gibt an, auf Grundlage der Rückmeldung die Schlussfolgerung gezogen zu haben, auf das eigene Gefühl vertrauen zu können (B2, Z. 6).

In zehn der elf Interviews identifizieren die Studierenden aus der eigenen oder der Beispielrückmeldung Verbesserungspotenziale (Kat. AF3c) entweder allgemein in Bezug auf die Reflexionsfähigkeit (B6, Z. 14; C2, Z. 42; C5, Z. 53, 55) oder in Bezug auf die verschiedenen Teilfähigkeiten (B1, Z. 22; B4, Z. 53; B5, Z. 35; C1, Z. 10, 31; C3, Z. 34; C4, Z. 32, 36), wobei B5 dies unter Vorbehalt der Gültigkeit der eigenen Rückmeldung tut, die B5 als eingeschränkt vermutet, da B5 das Testinstrument (nach einem Absturz aufgrund langer Inaktivität) erneut und weniger intensiv bearbeitet hat (B5, Z. 6, 8; Kat. AF3d). Vergleichbares berichtet C5 (Z. 8). Auch B2 und B4 vermuten, dass sie bei einer intensiveren Auseinandersetzung mit dem Testinstrument besser abschneiden würden (B2 vermutet auch explizit die mangelnde Anstrengung bei der Bearbeitung des Tests als Ursache für das erzielte Ergebnis; B2, Z. 6, 28; B4, Z. 12, 14). C2 gibt zu bedenken, dass unklar ist, ob

die Personen in der Vergleichsgruppe „alle sehr gut abgeschnitten haben“ (C2, Z. 44) und das die Bewertung der Ergebnisse beeinflussen würde (C2, Z. 47f.; Kat. AF3e).

Konkrete Überlegungen zur Umsetzung des Verbesserungspotenzials werden nur in einem Interview geäußert: C3 würde „vielleicht auch nochmal bei Kollegen in [den] Unterricht gehen, wenn ich Zeit habe und die das dulden“ (C3, Z. 34), um die Bewertung von Unterricht zu üben und im Rahmen einer Selbstreflexion "vielleicht auch Muster [...] von Kollegen an mir selber wiederzuerkennen“ (C3, Z. 34). In den übrigen Interviews wird - vorrangig in Bezug auf die eigenen Ergebnisse (B2, Z. 6; B4, Z. 8; B5, Z. 10; B6, Z. 50, 54) - ein wenig aussagekräftiges Vorgehen formuliert (z. B. „nachbessern“ (B2, Z. 6) oder „versuchen, reflektierter zu sein“ (B6, Z. 54)). C5 gibt sogar explizit an, nicht zu wissen, „was ich jetzt genau zum Beispiel tun kann, um besser zu werden“ (C5, Z. 77). B5 schränkt die Umsetzung des Verbesserungspotenzials zusätzlich ein: „Schülervorstellungen würde ich mir dann wahrscheinlich nochmal angucken, hätte ich die Zeit.“ (B5, Z. 35). C4 gibt als Ziel der Verbesserung an, „mindestens mal im Durchschnitt alles anderen zu sein“ (C4, Z. 32). Nur B3 formuliert keinen expliziten Verbesserungsbedarf, allerdings auch keine explizite positive Bewertung der Ergebnisse.

Die globalen Interpretationen, die die Studierenden äußern sowie die Beschreibung des Mehrwerts, den ihnen das Assessment-Feedback bietet (Kat. AF2a), sind auf Basis der Konstruktion von Testinstrument und Rückmeldung überwiegend als angemessen zu bewerten. Dazu gehören z. B. die Einschätzung, die in der Rückmeldung dargestellten Ergebnisse zeigten Verbesserungspotenziale (B1, Z. 10) bzw. eine allgemeine, aber grobe, Einschätzung der Angemessenheit didaktischer Überlegungen („didaktische[s] Verständnis“ (B4, Z. 33, 35)) bzw. der ins Studium mitgebrachten Reflexionsfähigkeit (B3, Z. 24) sowie die Einschränkungen der Gültigkeit der Ergebnisse aufgrund des Multiple-Choice-Formats (B4, Z. 35), der unbekannten Vergleichsgruppe (C3, Z. 42; Kat. AF3a), eingeschränkter Motivation und Anstrengungsbereitschaft bei der Bearbeitung des Testinstruments (B2, Z. 6, 28; B4, Z. 12, 14; B5, Z. 6, 8; Kat. AF3d).

Im Gegensatz dazu ist die Einschätzung, Testinstrument und Rückmeldung würden einen Einblick in mögliche Prüfungsleistungen im Modul sowie die Durchführung von Reflexionen im Referendariat bieten (B3, Z. 12), nicht Anspruch des Testinstruments und insofern nicht überprüft und damit nicht bewertbar. Ebenfalls ist das vermutlich falsche Verständnis der Teilfähigkeiten, welches B4 beschreibt (B4, Z. 37, 41), unpassend. Diese beiden Studierenden äußern also auch problematische Interpretationen, die bei der Beurteilung der konsequentiellen Validität zu berücksichtigen sind.

Die Interpretation und Einordnung der einzelnen Ergebnisse – die positive Bewertung von Ergebnissen über dem Median von (Kat. AF3b) sowie die Identifikation von Verbesserungspotenzial (Kat. Af3c) bei Ergebnissen unter diesem – erscheint

umfassend angemessen. Ein mögliches Vorgehen zur Umsetzung des Verbesserungspotenzials wird mit nur einer Ausnahme (C3, Z. 34) von den Studierenden allerdings nicht konkretisiert. Dies wäre für den weiteren Professionalisierungsprozess zwar wünschenswert, gleichzeitig soll das entwickelte Selbstlernmaterial an dieser Stelle angreifen und eine Option zur Förderung der Reflexionsfähigkeit bieten.

Wie gewichten die Studierenden die Bestandteile der Rückmeldung? Als besonders hilfreich oder relevant stellen fast alle Studierenden, wenn auch in unterschiedlicher Form, die erzielten Ergebnisse heraus (Kat. AF4a). B1, B3, B5 und C4 geben hier die Boxplots im Allgemeinen an (z. B. „[D]as fand ich auch vor allen Dingen durch die Graphen ziemlich anschaulich dargestellt." (B1, Z. 12; ähnlich auch B1, Z. 30; B3, Z. 28; B5, Z. 41)), C4 begründet dies: „weil am Ende kommt es ja auf jeden der Aspekte an" (C4, Z. 48). B1, B2, B3, B4, C2 und C3 benennen die Aufschlüsselung der Ergebnisse in die drei Teilergebnisse (z. B. „Das fand ich sehr hilfreich, um zu sehen, in welchem Bereich das nochmal einzeln gewertet wurde oder wo ich mich in einem einzelnen Bereich nochmal verbessern musste." (B2, Z. 28; ähnlich auch B1, Z. 30; B3, Z. 28; B4, Z. 66)). B2, B4 und C5 betonen explizit den Vergleich zur Vergleichsgruppe als positiv (B2, Z. 24; B4, Z. 66; C5, Z. 63, 65). B4 stellt die „maximal erreichten Punktzahlen" (B4, Z. 66) und B6 das persönliche Ergebnis zum Umgang mit Schülervorstellungen heraus (B6, Z. 50). B5 und C5 bewerten die überblicksartige Darstellung des Gesamtergebnisses in Textform zu Beginn der Rückmeldung als besonders hilfreich oder relevant: „[U]nd da ich diese Rückmeldung wirklich einfach nur überflogen habe und das jetzt auch gerade wieder gemacht habe, diese prägnanten Aussagen, wie ich bin im oberen Bereich" (B5, Z. 41; ähnlich auch C5, Z. 73).

Darüber hinaus werden von einzelnen Studierenden unterschiedliche Aspekte hervorgehoben (Kat. AF4b). C1 stellt die Information heraus, „dass ähm Reflexion äh wichtig ist, um den eigenen Unterricht zu verbessern" (C1, Z. 34). Außerdem bewerten B1, C1, C3 und C4 die Erläuterungen zu den Teilfähigkeiten der Reflexion bzw. deren Abbildung im Testinstrument als hilfreich (B1, Z. 30; C1, Z. 36; C3, Z. 42; C4, Z. 48), B2 die Erläuterungen zum Verständnis der Boxplots (B2, Z. 16) und B5 die Erläuterungen zur Effektstärke als besonders hilfreich oder relevant (B5, Z. 41), B2 erwähnt in Bezug auf die Rückmeldung die Abfrage der Motivation bei der Bearbeitung am Ende des Testinstruments zur besseren Einordnung der erzielten Ergebnisse („Also ich fand das auch gut, dass dann am Ende des Tests nochmal gefragt wurde, wie intensiv man sich auf diesen Test dann nochmal eingelassen hat, ähm wie doll man sich damit auseinander gesetzt hat" (B2, Z. 28)) und B6 (bei einem falschen Verständnis dessen als Musterlösung für das Testinstrument (B6, Z. 14, 18, 69)) sowie C4 den Link zum Selbstlernmaterial, um sich mit alternativen Handlungsoptionen auseinanderzusetzen: „[W]as ich gut fand war eben jetzt zum Beispiel dieser Link, wenn man jetzt jetzt (.) sein Ergebnis hat

und dann auch nicht damit zufrieden ist, dass man noch eben Alternativen hat, wo man nochmal (.) sich informieren kann, was man eben auch besser machen kann, oder wo man etwas finden kann, wie man es besser machen kann.“ (B6, Z. 36; ähnlich auch C4, Z. 48).

Fünf der Studierenden geben an, die in der Rückmeldung enthaltenen Texte nicht (aufmerksam bzw. vollständig) gelesen zu haben (Kat. AF4c), da diese als nicht relevant bzw. nicht auf das eigene Ergebnis zugeschnitten eingeschätzt wurden (B3, Z. 28; B4, Z. 70; B5, Z. 41, 43, C3, Z. 12; C4, Z. 6, 8, 14).

Für alle befragten Studierenden stehen also die in der Rückmeldung enthalten Ergebnisse zur (eigenen) Bearbeitung des Reflexionstests als besonders hilfreich oder relevant im Fokus (Kat. AF4a). Die übrigen Bestandteile der Rückmeldung (u. a. Erläuterungen zum Reflexionsbegriff, zur Aufschlüsselung der Teilergebnisse) werden von fünf Studierenden beim Erhalt der Rückmeldung gar nicht gelesen (Kat. AF4c). Die Fokussierung der Studierenden auf die (eigenen) Ergebnisse entspricht dem Ziel der Rückmeldung, den Studierenden eine Einschätzung ihrer Fähigkeiten auf Grundlage der Bearbeitung des Testinstruments zu geben. Die fehlende Beschäftigung mit den Erläuterungstexten (Kat. AF4c) ist nicht intendiert und könnte Einfluss auf das Verständnis oder die Interpretation der Rückmeldung nehmen, wenn bspw. das Ergebnis zu Schülervorstellungen fälschlicherweise als Rückmeldung über eigene Schülervorstellungen verstanden wird und nicht als Rückmeldung dazu, wie gut Schülervorstellungen bei Lernenden im Reflexionstest erkannt wurden (vgl. B4, Z. 37, 41).

Die übrigen als besonders bedeutsam oder hilfreich hervorgehobenen Aspekte sind individuell sehr unterschiedlich, beziehen sich aber überwiegend auf Aspekte, die die Interpretation der Ergebnisse beeinflussen, z. B. die Berücksichtigung der (eigenen) Testmotivation (B2, Z. 28), bzw. unterstützen, wie die Erläuterungen zu den dargestellten Teilfähigkeiten (B1, Z. 30; B2, Z. 16; B5, Z. 41).

Insgesamt stützen auch diese Aussagen über die Gewichtung der Bestandteile des Assessment-Feedbacks den Schluss, dass die Studierenden sich stark auf die (eigenen) Ergebnisse als Rückmeldung des Ist-Zustands (Feed-Back) fokussieren. Der Ausblick auf den Umgang mit diesen Ergebnissen, die Arbeit an der eigenen Reflexionsfähigkeit, steht – ebenso, wie bei der Interpretation der Rückmeldung durch die Studierenden – offenbar nicht im Fokus (nur B6 stellt den Verweis auf das Selbstlernmaterial als positiv heraus (B6, Z. 36), missversteht dieses dabei allerdings als Musterlösung zum Testinstrument (B6, Z. 14, 18, 69)). Auf Grundlage dieser Ergebnisse kann also weiterhin nicht davon ausgegangen werden, dass das Assessment-Feedback allein ausreichend Anstoß für die Arbeit an der Reflexionsfähigkeit oder damit verbundenen Aspekten liefert.

Wo bestehen aus der Perspektive der Studierenden Änderungsbedarfe in Bezug auf die Rückmeldung? Welche Missverständnisse auf Seiten der Studierenden werden deutlich, die ggf. einen Änderungsbedarf indizieren? Auf Nachfrage sehen die Studierenden in den meisten Interviews keinen Änderungsbedarf in Bezug auf die Rückmeldung (B1, Z. 32, 34, 36; B2, Z. 6, 28; B3, Z. 32, 34, 36; B6, Z. 52, 54, 59; C1, Z. 46, 55f.; C2, Z. 44, 56; C3, Z. 48; Kat. AF5a). B3 gibt ergänzend an, dass die Informationen zu Schülervorstellungen, die in der Rückmeldung gegeben sind, zwar bekannt sind, deren Erwähnung aber trotzdem gut ist. Zwei Studierende können keine Aussage dazu treffen, ob Inhalte der Rückmeldung überflüssig sind, da sie angeben, diese nicht vollständig gelesen zu haben (B4, Z. 70; B5, Z. 43).

Dennoch wird der Umfang des Assessment-Feedbacks in den Fokus gerückt (Kat. AF5b). So bewertet B6 es als „umfangreich“ (Z. 36) und B5 stellt die Frage, ob davon ausgegangen werden muss, „dass alle nur dreißig Sekunden Zeit haben, um sich ihre Ergebnisse anzugucken“ (B5 Z. 45) und die Rückmeldung dementsprechend angepasst werden müsste.

B4 wünscht sich zusätzlich zur Rückmeldung eine Musterlösung zu den Aufgaben des Testinstruments (Kat. AF5i), um feststellen zu können, wo bei der Bearbeitung des Tests Fehler gemacht wurden: „[Ich hätte gerne die Information gehabt,] was die (.) bestmöglichen Lösungen da wären, weil die weiß man ja jetzt ja nicht. Man weiß nur was, also wo man also prozentuell steht, aber nicht, was man genau falsch gemacht hat.“ (B4, Z. 75; ähnlich auch B4, Z. 126, 134; B5, Z. 69; C5, Z. 81). Zusätzlich wird eine ausdifferenzierte Rückmeldung auf Aufgabenebene (bzw. auf Ebene der Antwortoptionen) vorgeschlagen (B5, Z. 45; C5, Z. 78f.); für C5 könnte dies eine Möglichkeit sein, die Frage zu beantworten, „was ich jetzt genau zum Beispiel tun kann, um besser zu werden“ (C5, Z. 77; Kat. AF5k). B5 sieht hingegen keine Notwendigkeit, diese Aspekte zu ergänzen, da das Selbstlernmaterial vergleichbare Hinweise liefere (Kat. SLM2d): „Naja also hier ist ja jetzt keine Rückmeldung dazu, wie ich die einzelnen Fragen richtig oder falsch beantwortet habe, aber das ist ja im Onlinematerial möglich, die richtigen oder mögliche richtige Antworten (.) äh sich äh zu den einzelnen Fragen anzuschauen, deswegen da eigentlich nicht.“ (B5, Z. 45; ähnlich auch B5, Z. 51). Auch C3 kommt zu dem Schluss, so eine Aufschlüsselung sei „eigentlich auch gar nicht so (.) wichtig“ (C3, Z. 48).

Darüber hinaus schlägt B5 vor, zur Darstellung der Ergebnisse Balkendiagramme anstelle von Boxplots zu werden (Kat. AF5h), da diese niedrigschwelliger seien (B5, Z. 45).

Außerdem werden in den Interviews zwei (potenzielle) Missverständnisse deutlich, die einen Änderungsbedarf indizieren:

- B2 bewertet die Fragen zur Testmotivation bzw. Intensität der Bearbeitung am Ende des Testinstruments positiv zur besseren Einordnung der erzielten

Ergebnisse („denn dann wurde ja auch anhand dessen das Ergebnis festgemacht“ (B2, Z. 28)). Unklar ist, ob B2 damit die eigene Interpretation der Ergebnisse meint (wie aus der sich anschließenden Aussage über ein besseres Abschneiden bei höherer Anstrengung vermutet werden kann: „man [konnte] dann auch sehen, okay, selbst wenn ich da nicht so viel aufgepasst habe (unv.) in dem Bereich, hätte ich mehr aufgepasst oder mitgemacht, dann hätte ich oder wäre ich in dem Bereich gewesen“ (B2, Z. 28)) oder (fälschlicherweise) davon ausgeht, die Rückmeldung würde unter Berücksichtigung dieser Selbsteinschätzung erstellt. In beiden Fällen betrifft diese Überlegung mögliche Ursachen für die dargestellten Ergebnisse (Kat. AF3d).

- B6 geht offenbar davon aus, dass die Videos im Testinstrument und im Selbstlernmaterial identisch sind und das Selbstlernmaterial somit die Musterlösung zum Test bietet (B6, Z. 14, 18, 69).

Im Rahmen der Interviews werden keine von allen bzw. dem Großteil der Studierenden geteilten Änderungsbedarfe in Bezug auf die Rückmeldung deutlich (Kat. AF5a). Insofern kann davon ausgegangen werden, dass die genannten Aspekte keine grundsätzlichen bzw. verbreiteten Schwierigkeiten in Bezug auf das Verständnis bzw. die Interpretation der Rückmeldung darstellen. Auffällig ist allerdings, dass B6 die Rückmeldung als „umfangreich“ (B6, Z. 36; Kat. AF5b) bezeichnet und die Rückmeldung in fünf weiteren Fällen nicht umfassend gelesen wird (Kat. AF4c), was das korrekte Verständnis des Assessment-Feedbacks möglicherweise erschwert. Die Inhalte der Erläuterungen im Assessment-Feedback sind auch zur Interpretation der Ergebnisse relevant. Es ist also anzustreben, dass diese von den Studierenden wahrgenommen werden. Da das Assessment-Feedback allerdings selbstständig von den Studierenden und grundsätzlich auch unabhängig von universitären Veranstaltungen heruntergeladen werden kann, muss es ein in sich schlüssiges, abgeschlossenes Dokument darstellen. Eine Reduktion des Umfangs erscheint daher nur unter Einbußen hinsichtlich der Verständlichkeit realisierbar. Möglicherweise könnte beim Herunterladen der Rückmeldung auf die Relevanz der Texte für die Interpretation des Ergebnisses aufmerksam gemacht werden. Da einige Studierende die Rückmeldung nicht (vollständig) lesen, kann allerdings nicht davon ausgegangen werden, dass die Ergänzung schriftlicher Hinweise – wenn auch an anderer Stelle – ausreicht. Zentral erscheint hier, dass die Studierenden ausreichend Zeit für die Beschäftigung mit der Rückmeldung einplanen (vgl. B5, Z. 45). Um dies sicherzustellen, könnte die Interpretation der Rückmeldung auch im Rahmen von Veranstaltungen eingeplant und ggf. sogar unter Zuhilfenahme der Erläuterungen besprochen werden.

Auch um den (potenziellen) Missverständnissen vorzubeugen, die in den Interviews mit B2 und B6 deutlich werden (bezüglich der Interpretation der Rückmeldung auf Grundlage der Testmotivation bzw. Anstrengungsbereitschaft (B2, Z. 28; Kat. AF3d) sowie dem falschen Verständnis des Selbstlernmaterials als Musterlö-

sung (B6, Z. 14, 18, 69; Kat. AF5i)), könnten entsprechende Aspekte entweder schriftlich im Rahmen der Rückmeldung (bzw. im Selbstlernmaterial) oder ggf. im Veranstaltungskontext explizit thematisiert werden. Diese Missverständnisse scheinen allerdings individueller Natur zu sein, sodass auch hier die Ergänzung weiterer (schriftlicher) Hinweise im Assessment-Feedback gegenüber einer weiteren Erhöhung des Umfangs abzuwägen ist.

Die Ergänzung einer Musterlösung (Kat. AF5i), wie von drei Studierenden gewünscht, erscheint - wie bereits zu den Interviews der Kohorte A diskutiert - nicht sinnvoll: Einerseits würde diese den Umfang des Assessment-Feedback stark erhöhen, was insbesondere vor dem Hintergrund, dass die Rückmeldung bereits jetzt von den Studierenden als umfangreich wahrgenommen und teilweise - auch von den drei Studierenden, die sich eine solche Musterlösung wünschen - nicht umfassend gelesen wird. Andererseits soll das Testinstrument auch mehrfach eingesetzt werden können (z. B. zur Untersuchung der Fähigkeiten zu Beginn und zum Ende einer Veranstaltung); vorliegende und genutzte Musterlösungen könnten dann zukünftige Testergebnis verfälschen.

Der Vorschlag, anstelle der Boxplots Balkendiagramme zur Darstellung der Ergebnisse zu verwenden (B5, Z. 45; Kat. AF5h) wird auch verworfen, da die Boxplots einen deutlich differenzierteren Vergleich von eigenem Ergebnis und den Ergebnissen der Vergleichsgruppe ermöglichen und von den Studierenden offenbar inhaltlich angemessen interpretiert werden können. Reine Balkendiagramme erscheinen überdies möglicherweise auch aus motivationaler Perspektive nachteilig: Ein Ergebnis im Bereich der unteren 25 % - 50 % der Vergleichsgruppe liegt im Boxplot noch innerhalb der Box, es wird also deutlich, dass auch diverse andere Studierende Ergebnisse in diesem Bereich erzielen; im Balkendiagramm würde nur deutlich, dass das Ergebnis unter dem Median liegt, also schlechter als das mittlere Ergebnis der Vergleichsgruppe ist.

Welchen Bezug stellen die Studierenden zum zuvor erhaltenen Assessment-Feedback her? Die Studierenden C2, C3, C4 und C5 haben den Reflexionstest zweifach bearbeitet und zu beiden Bearbeitungen ein Assessment-Feedback erhalten. Alle Studierenden geben - teilweise auf Nachfrage - an, sich an das letzte Ergebnis erinnern zu können (C2, Z. 14; C4, Z. 22; C5, Z. 24) bzw. es zum Vergleich erneut angesehen zu haben (C3, Z. 18; Kat. AF2c). C2 (Z. 14), C3 (Z, 18) und C4 (Z. 22) geben an, sich verbessert zu haben und „froh“ (C2, Z. 14) darüber zu sein bzw. dies erwartet zu haben, da „man sich noch gar nicht mit dem Thema beschäftigt hatte“ (C4, Z. 22) als die erste Bearbeitung stattfand. C5 hingegen zieht die Schlussfolgerung, dass C5 trotz der (Besprechung der) letzten Bearbeitung des Reflexionstests in einer universitären Veranstaltung „noch nicht so genau weiß, worauf es jetzt [bei der Reflexion von Unterricht] ankommt“ (C5, Z. 16).

Auch wenn hier strukturell gleiches Assessment-Feedback zum selben Reflexionstest gegeben wird, kann der von den Studierenden (z. T. auf Rückfrage) hergestellte Rückbezug zum ersten Feedback als Beitrag für wiederholtes Feedback verstanden werden, was zur Effektivität von Feedback beiträgt (Henderson et al., 2019). Die gezogenen Schlussfolgerungen der Studierenden, eine Verbesserung als positiv einzustufen, ist grundsätzlich angemessen. Auch die Schlussfolgerung von C5, allein durch die Besprechung der vorigen Bearbeitung des Reflexionstests kein als ausreichend zu bewertendes Fähigkeitsniveau erlangt zu haben, erscheint nachvollziehbar. Grundsätzlich kann vermutet werden, dass durch die Bearbeitung der Lernumgebung sowie zwischenzeitlich besuchte universitäre Veranstaltungen Lerngelegenheiten zur Unterrichtsreflexion geboten werden (können); inwiefern diese durch die Studierenden genutzt wurden und mit welchen Effekten die Nutzung einhergeht, ist allerdings unklar.

Wie gehen die Studierenden bei der Bearbeitung des Selbstlernmaterials vor? Für die Studierenden der Interviewkohorte B wurde die Bearbeitung des Selbstlernmaterials in eine Seminarsitzung integriert, sodass alle Studierenden in diesem Rahmen etwa 20 Minuten mit diesem gearbeitet haben. Vier der sechs Studierenden (B1, Z. 38; B2, Z. 36; B3, Z. 44; B6, Z. 64) haben das Selbstlernmaterial allein bearbeitet; B4 und B5 haben gemeinsam am Material gearbeitet (B5, Z. 53; Kat. SLM1a). Die Studierenden der Interviewkohorte C wurden nicht explizit zur Arbeit mit dem Selbstlernmaterial aufgefordert. In dieser Gruppe hat nur eine Person sich überhaupt mit dem Beginn des Materials beschäftigt (C1, Z. 59-68; Kat. SLM1a). Die übrigen Studierenden haben das Material nicht bearbeitet (C2, Z. 60; C3, Z. 52; C4, Z. 88; C5, Z. 84f.), wobei zwei Personen angeben, dies zukünftig tun zu wollen (C2, Z. 52) oder es sich zumindest vorstellen können (C4, Z. 88). Als ursächlich dafür, dass die Studierenden (bislang) nicht mit dem Selbstlernmaterial gearbeitet haben, geben sie mangelnde zeitliche Ressourcen (C2, Z. 68, 75-78) und die Fülle an erhaltenem Material im Lehramtsstudium an (C4, Z. 88). C5 hat das Selbstlernmaterial nicht wahrgenommen, da das Assessment-Feedback nicht vollständig gelesen wurde (C5, Z. 87-91). C4 sieht darüber hinaus keine Möglichkeit, die Nutzung des Selbstlernmaterials für die Studierenden attraktiver zu gestalten: „[D]a ist ein Link, ich glaube, [...] mehr kann man nicht machen an eurer Stelle“ (C4, Z. 97).

Bei der Beschreibung des eigenen Vorgehens bei der Bearbeitung des Selbstlernmaterials sprechen vier der sechs Studierenden explizit an, die Videos angesehen zu haben (B1, Z. 40; B2, Z. 38; B4, Z. 38; B5, Z. 53, Kat. SLM1b). B6 gibt an, die Videos übersprungen zu haben, da B6 sich noch an diese erinnert: „Dann habe ich die Videosequenzen übersprungen, weil das erst zwei Wochen her war und ich das für nötig, sage ich mal, gehalten habe, weil das noch im Hinterkopf war.“ (B6, Z. 64). Hier liegt allerdings die falsche Annahme zugrunde, die Szenen

in Test und Selbstlernmaterial wären identisch und das Selbstlernmaterial böte die Musterlösung zum Testinstrument.

In Bezug auf das Vorgehen bei der Bearbeitung des Materials nach dem Ansehen der Unterrichtsausschnitte lassen sich die Studierenden in zwei Gruppen einteilen: B1, B3, B4 und B5 geben an, nach dem Schauen der Videos zunächst selbstständig (B4 und B5 allerdings gemeinsam) Überlegungen in Bezug auf die im Selbstlernmaterial angegebenen Fragestellungen angestellt zu haben (B1, Z. 40; B3, Z. 44; B5, Z. 53; Kat. SLM1c (B4 erinnert sich im Interview nicht mehr an das konkrete Vorgehen bei der Analyse der Unterrichtsausschnitte (B4, Z. 86), daher wird hier auf die Schilderung von B5 zurückgegriffen)). Erst im Anschluss lesen sie die Musterlösung und gleichen die eigenen Überlegungen mit der Musterlösung ab (z. B. „[I]ch habe mir die einzelnen Videos angeguckt, habe dann erstmal versucht, die Fragestellung selbst zu beantworten. Nachdem ich dann mir gedacht habe, gut, jetzt habe ich die Antwort gefunden, habe ich mir dann die Lösung angeguckt und habe dann das abgeglichen“ (B1, Z. 40; ähnlich auch B3, Z. 44; B5, Z. 53)). Nur B5 erwähnt dabei explizit, in der Bearbeitung des Materials mit B4 die Tipps genutzt zu haben, um die eigenen Überlegungen abzugleichen, bevor die Musterlösung betrachtet wird: „Wir hatten dann oft auch unterschiedliche Standpunkte und haben uns dann Tipps die Tipps dazu angeguckt und daraufhin nochmal weiter diskutiert“ (B5, Z. 53).

B2 und B6 beschreiben ihr Vorgehen so, dass sie direkt nach dem Schauen der Videos bzw. nach dem Übersprungen der Videos (da diese B6 vermeintlich bekannt waren) die Lösungen gelesen haben (B2, Z. 36, 38; B6, Z. 69), also die Situation nicht zunächst selbstständig analysiert haben (Kat. SLM1d): „[D]ie [Videos] habe ich mir dann nochmal einzeln angeschaut. Ähm habe da versucht, da so viel mitzunehmen, eigentlich aus der Situation, [...] als ob ich so im Unterricht säße und das bewerten müsste. [...] und danach eben bin ich auf die Kommentare [...] da unten (da?) dann nochmal gegangen und ähm habe die mir durchgelesen, ob ich das hätte genauso gemacht, ob ich das nachvollziehen konnte“ (B2, Z. 38). B2 gibt an, beim Lesen der Lösungen trotzdem abgeglichen zu haben, inwiefern diese zur eigenen Einschätzung passt und eine „Selbstreflexion“ (B2, Z. 36, 38) in Bezug auf die angemessene Interpretation der Situation durchgeführt zu haben (B2, Z. 36, 38). Aufgrund der falschen Annahme, die Szenen im Selbstlernmaterial entsprächen denen im Reflexionstest, geht B6 davon aus, dass ein Abgleich der Lösung mit den eigenen Überlegungen nicht möglich war, da B6 sich nicht an die zu den einzelnen Unterrichtsausschnitten im Test gegebenen Antworten erinnert (B6, Z. 69).

Bei einer erneuten Bearbeitung des Selbstlernmaterials würde B2 nicht erneut allein mit dem Material arbeiten (Kat. SLM1a), sondern hätte „Notizen nebenbei gemacht und äh wäre gerne ins Gespräch gekommen mit [den] Kommilitonen“ (B2, Z. 44).

Mit Ausnahme des Vorgehens von B6, welches weiterhin auf der falschen Annahme beruht, die Unterrichtsausschnitte in Test und Selbstlernmaterial wären identisch (B6, Z. 14, 18, 69), beschreiben die Studierenden grundsätzlich sinnvolle Vorgehensweisen bei der Arbeit mit dem Material: Nach dem Schauen der Videos werden entweder in Einzel- oder Partnerarbeit zunächst eigene Überlegungen in Bezug auf den Unterrichtsausschnitt angestellt und diese anschließend mit der Lösung verglichen (Kat. SLM1a, SLM1b, SLM1c) oder es werden direkt die Lösungen betrachtet und nachvollzogen bzw. abgelehnt (SLM1d). Es wird also stets die Musterlösung zur Reflexion des entsprechenden Unterrichtsausschnitts betrachtet und mit der eigenen Position abgeglichen. Es ist zu vermuten, dass die selbständig bzw. im Gespräch durchgeführte Reflexion vor der Betrachtung der Lösungen gegenüber dem reinen Lesen der Lösungen die aktive Auseinandersetzung mit dem Material unterstützt und damit die nachhaltigere Variante ist. Ob für die Arbeit mit dem Material Einzel- oder Partnerarbeit bevorzugt werden sollte, kann an dieser Stelle nicht entschieden werden; auch hier ist vorstellbar, dass durch das Gespräch die aktive Auseinandersetzung mit den Unterrichtsinhalten unterstützt wird.

Welchen Mehrwert bietet das Material den Studierenden? Auf explizite Nachfrage sprechen alle Studierenden der Interviewkohorte B dem Material einen Mehrwert zu (B1, Z. 46; B2, Z. 42; B3, Z. 48; B4, Z. 92; B5, Z. 61; B6, Z. 79), fokussieren dabei aber unterschiedliche Aspekte.

B1, B2 und B5 stellen - wenn auch unterschiedlich konkret - die Selbstevaluation (Kat. SLM2a) in den Vordergrund; so findet B5 das Selbstlernmaterial hilfreich, um das eigene „Reflektieren zu reflektieren“ (B5, Z. 59), B2 gibt an, überprüfen zu können, ob die Situation korrekt interpretiert wurde und die prioritären Inhalte der Situation richtig erkannt wurden (z. B. „[I]ch fand, wie gesagt, das Material hilfreich, um überhaupt die Situation einzuschätzen“ (B2, Z. 48; ähnlich auch B2, Z. 42)), B1 gibt allgemeiner an, einen Überblick über das eigene Abschneiden zu erhalten (B1, Z. 42, 44, 46) und feststellen zu können, „wo [B1] [s]ich noch verbessern kann und wo [B1] auch schon gut [ist]“ (B1, Z. 44). Und B2 zieht die Schlussfolgerung, auf das eigene Gefühl vertrauen zu können, wenn es darum geht, die Schüler*innenperspektive einzunehmen bzw. nachvollziehen zu können (B2, Z. 40).

Weiterhin wird die Anleitung zur Unterrichtsreflexion als „Hilfestellung“ (C3, Z. 52; auch C2, Z. 74) bzw. Möglichkeit zur Übung (C3, Z. 52) und demnach mit Mehrwert beschrieben (Kat. SLM2b): Auch B3 nimmt aus der Arbeit mit dem Material – wie schon in Bezug auf die Rückmeldung und das Testinstrument – mit, am Beispiel der Musterlösung sehen zu können, wie auf die Fragen zur Analyse des Unterrichts geantwortet werden kann bzw. wie an die Analyse der Unterrichtssituationen herangegangen werden kann (B3, Z. 46).

Weiterhin geben zwei Studierende an, das Selbstlernmaterial diene zur Sensibilisierung (Kat. SLM2c), da sie in den Unterrichtsausschnitten kritische Situationen erkennen, für die man aus ihrer Perspektive als Lehrkraft sensibilisiert sein sollte („[E]s ist halt sehr hilfreich, [...] wenn man [...] ins Auge gesetzt bekommen möchte, wo man eher aufpassen soll“ (B2, Z. 42; ähnlich auch B2, Z. 40)) bzw. die im eigenen Unterricht möglicherweise gar nicht auffallen, sondern nur durch die Perspektive von außen sichtbar werden („[W]enn man die Lehrperson ist, die die Fehler macht, ist es schwer, die Fehler zu erkennen.“ (B4, Z. 88)). Darüber hinaus beschreibt B4, dass die Arbeit mit dem (Förder-)Material dazu geführt hat, dass B4 sich selbst in Bezug auf den Umgang mit Schüler*innen hinterfragt hat („Aber vielleicht hätte ich das auch so sagen können und dann bringt mich so ein Video doch auch zum Nachdenken, ob das überhaupt das richtig ist. Auch, wenn es überhaupt nicht böse gemeint ist oder so.“ (B4, Z. 92; die Szene, auf die B4 sich hier bezieht, ist allerdings Teil des Testinstruments, nicht des Selbstlernmaterials)). B1 gib an, aus der Arbeit mit dem Selbstlernmaterial die Musterlösung als Entwicklungsziel abzuleiten: „Also es gibt mir einen guten Überblick, [...] wo ich hinmöchte.“ (B1, Z. 44).

B6, gibt – unter der Annahme, das Selbstlernmaterial beinhaltete die gleichen Szenen, wie das Testinstrument – an, dass B6, um konkrete Schlüsse aus der Arbeit mit dem Instrument ziehen zu können, auch die selbst gegebenen Antworten im Abgleich benötigen würde (B6, Z. 81), äußert aber die Absicht, das Material möglicherweise zukünftig als eine Art Nachschlagewerk (Kat. SLM2d) für geeignete Reaktionen auf die dargestellten Situationen zu nutzen, bevor B6 ersten eigenen Unterricht hält.

Die von den Studierenden beschriebenen Mehrwerte, die das Selbstlernmaterial bietet, stehen überwiegend in Einklang mit dem damit verfolgten, übergeordneten Entwicklungsziel: Die Förderung der Professionalisierung in Bezug auf die Reflexionsfähigkeit.

Die Evaluation des eigenen Handelns bzw. der eigenen Überlegungen (Kat. SLM2a), die von drei Studierenden in den Fokus gerückt wird (B5, Z. 59; B2, Z. 42, 48; B1, Z. 42, 44, 46), gibt zwar keinen Aufschluss darüber, inwiefern hier eine Entwicklung vorhanden bzw. zu erwarten ist, als Ausgangspunkt für eine Entwicklung kann diese Selbstevaluation aber – ebenso, wie die auf Grundlage der Rückmeldung durchgeführte Einschätzung – geeignet sein, Entwicklungsbedarfe aufzuzeigen und zu motivieren.

Das Kennenlernen eines möglichen Vorgehens zur Analyse von Unterricht (Kat. SLM2b) liefert einen Beitrag zur Förderung der Reflexionsfähigkeit auf methodischer Ebene. Außerdem sensibilisiert das Material einige der Studierenden für bestimmte (unterschiedliche) Situationen (Kat. SLM2c) bzw. führt dazu, dass sie ihr eigenes Handeln in Bezug auf die beobachteten bzw. vergleichbare Situationen hinterfragen (B1, Z. 40, 42; B4, Z. 88), oder das Selbstlernmaterial

zukünftig als Nachschlagewerk nutzen wollen (B6, Z. 75; Kat. SLM2d; diese Aussage ist potenziell unabhängig von der zugrundeliegenden falschen Annahme, das Selbstlernmaterial beinhaltete die gleichen Videos wie das Testinstrument). Beide Aspekte können zu einer Entwicklung der Reflexionsfähigkeit auf methodischer Ebene bzw. in Bezug auf konkrete Inhalte beitragen. Diese angestoßenen Prozesse für die Professionalisierung dürfen damit aber nicht als abgeschlossen verstanden werden.

Wie gewichten die Studierenden die Bestandteile des Selbstlernmaterials?
Als besonders hilfreiche oder interessante Bestandteile des Selbstlernmaterials bzw. aus der Arbeit mit diesem benennen die Studierenden individuell unterschiedliche Aspekte (Kat. SLM3). So verweist B5 auf die Tipps, die im Material zu allen Reflexionspunkten aufgeführt sind: „Ich fand es gut, dass man einen Tipp bekommen hat, bevor man die Lösung hatte, weil dann konnte man selber nochmal seine eigentlichen Gedanken reflektieren und gucken okay, passt das auch mit dem Tipp." (B5, Z. 63). Abgesehen davon werden von den Studierenden keine Inhalte des Materials explizit herausgestellt (z. B. B1, Z. 48; B2, Z. 44; B3, Z. 50).

B4 gibt als zentrale Erkenntnis die angestellten Überlegungen und Selbstreflexion zum Umgang mit den Schüler*innen an, die sich allerdings auf Unterrichtsausschnitte aus dem Testinstrument beziehen („Dass das so kleine Dinge sind, äh die aber einen großen Einfluss haben können" (B4, Z. 98)). B6 betont die dargestellten Musterlösungen zu den Unterrichtsausschnitten als besonders hilfreich (B6, Z. 83) und geht davon aus, nach einer Bearbeitung des Selbstlernmaterials im Test besser abzuschneiden, allerdings unter der falschen Annahme, das Selbstlernmaterial böte die Musterlösung zur Testbearbeitung: „[W]enn ich mir jetzt [...] nochmal Gedanken machen würde und das Material nochmal durcharbeiten würde und den Test nochmal machen würde und es wäre vielleicht immernoch schlecht, dann würde ich mir doch Fragen stellen, warum das nicht besser geworden ist" (B6, Z. 54).

Auf Grundlage der Interviews lassen sich keine klaren Aussagen über die Gewichtung der Bestandteile des Selbstlernmaterial in Bezug auf deren Nützlichkeit für die Studierenden ziehen. Dem Selbstlernmaterial wird eher insgesamt ein Mehrwert zugesprochen (z. B. B1, Z. 48; B2, Z. 44; Kat. SLM3). Möglicherweise wird eine solche Gewichtung dadurch erschwert, dass die Studierenden das Material üblicherweise nicht (vollständig) bearbeitet haben. Auffällig ist, dass keine*r der Studierenden der Interviewkohorte B, die das Selbstlernmaterial bearbeitet haben, die weiterführenden Literaturhinweise zu spezifischen Unterrichtssituationen im Selbstlernmaterial thematisiert. Dies lässt vermuten, dass die Studierenden ihnen - zumindest bei der Arbeit mit dem Selbstlernmaterial in diesem Setting - wenig Relevanz zusprechen.

Wo bestehen aus der Perspektive der Studierenden Änderungsbedarfe in Bezug auf das Selbstlernmaterial? Welche Missverständnisse auf Seiten der Studierenden werden deutlich, die ggf. einen Änderungsbedarf indizieren?
In den meisten Interviews wird kein Änderungsbedarf in Bezug auf das Selbstlernmaterial formuliert (B1, Z. 28, 50, 52; B3, Z. 52, 54; B4, Z. 105, 107; B5, Z. 65, 67; C1, Z. 78; Kat. SLM4a). B1 und B6 schätzen das Material allerdings als umfangreich („Es war ziemlich viel“ (B1, Z. 38)) bzw. die Bearbeitung als zeitaufwendig ein („wenn man viel Zeit darein steckt und [...] in dem Fall haben wir das ja gemacht, dann kommen wir auch zu den richtigen Ergebnissen“ (B5, Z. 53); Kat. SLM4b), was als Bedarf für Kürzungen interpretiert werden könnte.

Kritisiert werden darüber hinaus inhaltliche Wiederholungen im Selbstlernmaterial (Kat. SLM4c). B2 bemängelt die Dopplung von schriftlichen und in Videos von Robert mündlich dargestellten Situationsbeschreibungen und schlägt vor, die schriftlichen Beschreibungen auszublenden und somit optional sichtbar zu machen: „vielleicht ist es auch ganz gut, wenn man [...] das Geschriebene erstmal nicht gesehen hätte, (.) weil (das?) dann das irgendwie so doppelt ist“ (B2, Z. 46). Unklar ist, ob sich B2 hierbei auf die Situationsbeschreibung zu Beginn des Testinstruments oder die Einleitungen in die unterschiedlichen Unterrichtsausschnitte bezieht. B6 gibt an, die sich wiederholenden (Anstoß-)Fragen zur Reflexion der Unterrichtsausschnitte als überflüssig zu bewerten (B6, Z. 69, 85), sieht aber keinen unbedingten Änderungsbedarf, da diese standardmäßig eingeklappt und somit nicht störend sind: „Weil es halt eingeklappt war, (hat?) es halt nicht gestört dann.“ (B6, Z. 87).

Das bei B6 aufgetretene falsche Verständnis des Selbstlernmaterials als Musterlösung für das Testinstrument (B6, Z. 14, 18, 69) zeigt auch einen grundlegenden Klärungsbedarf auf, wurde aber bereits in Bezug auf die Rückmeldung diskutiert und wird daher hier nicht erneut explizit aufgegriffen. Der teilweise als groß wahrgenommene Umfang des Selbstlernmaterials (Kat. SLM4b) wird nicht als umzusetzender Änderungsbedarf interpretiert, da auf Grundlage der Interviews kaum Einschätzungen zum gesamten Material vorliegen.

Insgesamt wird in den Aussagen der Studierenden also nur ein weiterer umzusetzender Änderungsbedarf in Bezug auf das Selbstlernmaterial deutlich: Die Reduktion der Redundanz in schriftlichen und von Robert erzählten Situationsbeschreibungen (vgl. B2, Z. 46; Kat. SLM4c). Den drei in das Selbstlernmaterial eingebundenen Unterrichtsausschnitten ist jeweils ein kurzes, schriftliches Statement von Robert sowie eine Einordnung in das Stundenraster vorangestellt. Diese dienen zur Einordnung des nachfolgenden Ausschnitts in die Unterrichtsstunde. Aufgrund der Kürze dieser Einleitungen (die Einordnung in das Stundenraster ist standardmäßig ausgeblendet und kann bei Bedarf angezeigt werden) und der geringen tatsächlichen Redundanz wird hier keine Änderung vorgenommen. Eine umfangreiche Überschneidung von geschriebenen und gesprochenen Situationsbe-

schreibungen liegt allerdings zu Beginn des Selbstlernmaterials (zur Lerngruppe, dem Vorwissen der Schüler*innen, den Unterrichtszielen und dem Stundenraster) vor: Hier werden die Beschreibungen sowohl in Form kurzer, zusammenfassender Texte als auch in Form von Videos, in denen Robert die Situation beschreibt, gegeben. Im Selbstlernmaterial wird zwar darauf hingewiesen, dass hier dieselben Informationen in Text- und Videoform redundant vorliegen, wie sie auch im Testinstrument geschildert werden, um es den Studierenden zu ermöglichen, auszusuchen, welche der Darstellungsformen sie nutzen wollen. Dies wird offenbar nicht allen Studierenden deutlich. Da durch die Reduktion auf eine Darstellungsform keine Informationen verloren gehen und die schriftliche Version (optisch) kürzer ist und gezieltes Nachlesen ermöglicht, bleiben die schriftlichen Beschreibungen stehen und die Videos werden entfernt.

Sonstige Anmerkungen zum Reflexionstest Im Rahmen der Interviews wird zusätzlich deutlich, dass fünf Studierende das Testinstrument als umfangreich wahrnehmen oder zumindest ein kürzeres Testinstrument erwartet haben (B3, Z. 36; B4, Z. 14; B5, Z. 53, 79; B6, Z. 14; C4, Z. 108; Kat. S1a), was zu einer Abnahme der Motivation im Verlauf des Instruments führen könnte (C4, Z. 108). Um die Bearbeitung des Testinstruments ohne Kürzung für die Studierenden zeitlich flexibler zu gestalten, wird der Vorschlag von B5, das Testinstrument in mehreren Teilen durchführen zu können (B5, Z. 81), umgesetzt.

Die von B3 geäußerte Wahrnehmung des Testinstruments bzw. der gesamten Lernumgebung als Orientierung für eine Prüfungsleistung im Modul und das Vorgehen bei Unterrichtsanalysen im Referendariat (B3, Z. 12, 48; Kat. S1b) ist insofern problematisch, als dass dies nicht das Entwicklungsziel der Lernumgebung ist und diese daher nicht in Bezug auf diese Interpretation validiert ist. Hier kann – wie bereits oben erwähnt – eine Thematisierung im Veranstaltungskontext Klarheit schaffen.

Positiv erwähnt wurden außerdem die Videos an sich („[I]ch hatte in meiner Antwort da geschrieben, dass ich die Videos cool fand. Also cool gemacht.“ (B4, Z. 111)) bzw. deren Eignung, trotz übertriebener Darstellungen zu sensibilisieren (B2, Z. 40) und einen Einblick in problematischen Unterricht zu bieten (B3, Z. 12; alles Kat. S1c), sowie die Interaktivität (Kat. S1d) von Test und bzw. oder Material (der Bezug wird nicht klar geäußert): „Also es ist eine coole [...] Sache gewesen, (da?) man ein bisschen interaktiv auch ein bisschen mitarbeiten konnte. Das war ganz gut.“ (B3, Z. 56; auch C1, Z. 50).

18.1.3 Ergänzende Erkenntnisse aus dem Akzeptanzfragebogen

Zur Einschätzung des Indikators 4.2 *„Die Teilnehmenden sprechen dem Assessment-Feedback (inklusive Selbstlernmaterial) einen Mehrwert zu.“* wurde ergänzend ein Fragebogen zur Einschätzung der Akzeptanz der Teilnehmenden eingesetzt (s. Kapitel 13.3). Um Einschätzungen zum Assessment-Feedback (inklusive Selbstlernmaterial) zu erhalten, wurden Teile des Fragebogens nach Erhalt der Rückmeldung (Teil 2 des Akzeptanzfragebogens) bzw. Bearbeitung des Selbstlernmaterials (Teil 3 des Akzeptanzfragebogens) bearbeitet. Da das Selbstlernmaterial kaum (selbstständig) bearbeitet wurde, liegen lediglich von fünf Studierenden Einschätzung zum Assessment-Feedbacks[183] und von acht Studierenden[184] Einschätzungen zum Selbstlernmaterial vor. Eine Übersicht über die Einschätzungen der jeweiligen Items durch die Studierenden findet sich im Anhang in Tabelle K.1 (S. 514).

Dabei zeigt sich eine insgesamt neutrale, tendenziell positive Einschätzung der Lernumgebung: Weder beim Herunterladen des Assessment-Feedbacks noch bei der Bearbeitung des Selbstlernmaterials traten (technische) Schwierigkeiten auf. Für alle übrigen Einschätzungen u. a. zur Relevanz, Qualität und Nutzungsabsicht geben jeweils mehr Studierende Zustimmung als Ablehnung zu den entsprechenden Aussagen an, insbesondere das Selbstlernmaterial wird von den Studierenden tendenziell positiv bewertet. Aufgrund der geringen Anzahl an Einschätzungen können diese Ergebnisse zwar nicht als repräsentativ für die gesamte Stichprobe eingestuft werden, ergänzen aber die aus den Interviews gezogenen Schlussfolgerungen.

18.1.4 Bewertung der Indikatoren 4.1 und 4.2

Zur Beurteilung des Indikators 4.1 *„Die Interpretationen der Ergebnisse durch die Teilnehmenden passen inhaltlich zum erhaltenen Feedback.“* werden Ergebnisse der drei Interviewkohorten zusammengeführt. Dabei zeigt sich, dass die Studierenden das Assessment-Feedback überwiegend angemessen interpretieren. Über alle drei Interviewkohorten hinweg waren Boxplots den meisten Studierenden zwar zuvor unbekannt, erweisen sich dennoch als geeignete Darstellungsform, da die Studierenden die Ergebnisse sinnvoll und überwiegend unter Einbezug der Vergleichsgruppe (insb. des Medians als Markierung einer Grenze für gute Ergebnisse bzw. Verbesserungsbedarfe) einordnen. Auch Einschränkungen der Gültigkeit der Ergebnisse (z. B. aufgrund der eigenen Motivation während der Testbearbeitung) werden von den Studierenden sinnvoll ergänzt.

Die im Anschluss an die Interviews zur Pilotierung des Assessment-Feedback vorgenommenen Änderungen (z. B. die Erläuterung des Boxplots im Diagramm

[183] Dies sind die fünf Studierenden der Interviewkohorte C. Ausgefüllt haben die Studierenden den Fragebogen jeweils vor der Durchführung des Interviews.

[184] Darunter die sechs Studierenden der Interviewkohorte B.

oder die schriftliche Darstellung des individuellen Gesamtergebnisses zu Beginn des Assessment-Feedback) erscheinen angemessen, da sie in den nachfolgenden Interviews mit den Kohorten B und C offenbar nicht zu Unklarheiten führen, sondern für die Interpretation des Assessment-Feedback genutzt werden.

Als zentrale Ergänzung im Assessment-Feedback ist für die Studierenden der Kohorten B und C gegenüber der Kohorte A ist die Einbindung des Selbstlernmaterials zu berücksichtigen, das den Studierenden im Sinne eines Feed-Forward eine erste Möglichkeit zur selbstständigen Arbeit an ihrer Reflexionsfähigkeit bieten soll. Auffallend ist, dass die Studierenden das Material kaum bearbeiten, wenn es nicht in eine universitäre Veranstaltung integriert ist. In diesem Fall kommt das Selbstlernmaterial also nicht zum Einsatz, und die Schwierigkeit der Studierenden, auf Grundlage des Assessment-Feedbacks konkrete Schlussfolgerungen zur persönlichen Weiterentwicklung zu ziehen, bleibt bestehen. Die Studierenden der Kohorte B, die das Material im Rahmen einer Veranstaltung bearbeiteten, beschreiben unterschiedliche, aber grundsätzlich sinnvolle Vorgehensweisen. Dabei lassen sich Hinweise darauf finden, dass die Arbeit mit dem Selbstlernmaterial einen Beitrag zur Professionalisierung liefern kann: einerseits im Rahmen der Selbstevaluation als möglicher Anstoß für Entwicklung, andererseits auf methodischer und inhaltlicher Ebene, indem das Vorgehen bei der Reflexion greifbar gemacht wird und konkrete Situationen reflektiert werden. Diese Perspektive der Studierenden auf das Selbstlernmaterial stellt Evidenz für die Erfüllung des Indikators 4.2 *„Die Teilnehmenden sprechen dem Assessment-Feedback (inklusive Selbstlernmaterial) einen Mehrwert zu.“* dar. Auch die Einschätzungen über die Bedeutung des Assessment-Feedbacks lassen Rückschlüsse auf durch die Studierenden erkannte Mehrwerte zu: das Erhalten einer Rückmeldung zu den eigenen Fähigkeiten (B3, Z. 24; B4, Z. 33, 35) und Verbesserungspotenzialen (B1, Z. 10) sowie die Nutzung des Assessment-Feedbacks als Anlass für Selbstreflexion (A2, Z. 24). Insgesamt kann der Indikator 4.2 daher als erfüllt angesehen werden. Die mit dem Akzeptanzfragebogen erhobenen insgesamt tendenziell positiven Beurteilungen des Assessment-Feedbacks u. a. als nützlich oder hilfreich zur Einschätzung der eigenen Reflexionsfähigkeit sowie des Selbstlernmaterials als nützlich zur Verbesserung der Reflexionsfähigkeit und inhaltlich relevant für das Lehramtsstudium, sind aufgrund der sehr kleinen Stichprobenzahl vorsichtig zu interpretieren, widersprechen dieser Einschätzung aber nicht.

Einschränkend muss berücksichtigt werden, dass die Zuschreibungen einzelner Studierender, das Assessment-Feedback gäbe Aufschluss über die Eignung für den Beruf als Lehrkraft (A1, Z. 24) oder einen exemplarischen Einblick in Prüfungsleistungen in Studium oder Referendariat (B3, Z. 12), als problematische (Über-)Interpretationen einzustufen sind, da das Material für diese Aussagen nicht validiert wurde. Zusätzlich wird das Assessment-Feedback von verschiedenen Studierenden nicht vollständig gelesen und das Selbstlernmaterial von den befragten Studierenden offenbar kaum aus eigenem Antrieb bearbeitet. Daher kann

vermutet werden, dass die Studierenden zwar einen Mehrwert im Assessment-Feedback (inklusive Selbstlernmaterial) erkennen, dieser aber nicht ausreicht, um die Studierenden zu einer weiterführenden Beschäftigung mit Unterrichtsreflexion anzuregen. Aus diesem Grund erscheint bspw. die Einbettung der Lernumgebung in eine Lehrveranstaltung sinnvoll, da hier einerseits Relevanz für die Bearbeitung des Materials geschaffen werden kann und andererseits ggf. individuell auftretende, problematische Interpretationen direkt adressiert und ausgeräumt werden können.

Als methodische Limitation der Auswertung ist festzuhalten, dass die Auswertung der Interviews ausschließlich durch eine Person vorgenommen wurde. Zur Sicherstellung der Qualität der Auswertung und Nachvollziehbarkeit der entwickelten Kategorien wäre ergänzend eine unabhängige Kodierung (von Teilen) des Materials zu empfehlen (Kuckartz, 2018, S. 205).

18.2 Beurteilung des Assessment-Feedbacks vor dem Hintergrund von Kriterien für effektives Feedback (Indikator 4.3)

Bereits bei der Entwicklung und Überarbeitung des Assessment-Feedbacks wurden Erkenntnisse zur Gestaltung effektiven Feedbacks berücksichtigt. Das Assessment-Feedback wird als Angebot an Teilnehmende verstanden, auf Grundlage des Testbearbeitung Informationen über die eigene Reflexionsfähigkeit zu erhalten sowie einen Ausgangspunkt für die weitere Professionalisierung zu bilden. Die fundierte Gestaltung des Feedbacks soll dabei die Wahrscheinlichkeit erhöhen, ein relevantes Feedback darzustellen, das von den Teilnehmenden angenommen wird. Mit dem Indikator 4.3 *„Das Assessment-Feedback stellt inhaltlich relevante Testergebnisse unter Berücksichtigung von Qualitätskriterien für effektives Feedback dar.“* wird dessen Evaluation daher um eine theoretische Perspektive ergänzt. Dazu wird zunächst die inhaltliche Relevanz der im Assessment-Feedback dargestellten Ergebnisse begründet, bevor auf die Umsetzung weiterer Gestaltungsmerkmale für effektives Feedback eingegangen wird.

Im Assessment-Feedback werden neben dem Testwert des Reflexionstests zur Reflexionsfähigkeit (dem Gesamtergebnis) Testwerte zu drei Teilfähigkeiten (zum Bewerten der Aspekte, Vorschlagen von Alternativen und Erkennen von und Umgang mit Schülervorstellungen) rückgemeldet. Die inhaltliche Relevanz dieser Ergebnisse - insbesondere des Gesamtergebnisses - liegt grundsätzlich in der Bedeutung von Unterrichtsreflexion für die Professionalisierung von Lehrkräften begründet (s. Kapitel 7.2), die gleichermaßen die Relevanz der entwickelten Lernumgebung insgesamt begründet. Um ein stärker ausdifferenziertes Feedback zu ermöglichen, wird auf das Reflexionsmodell von Nowak, Kempin, Kulgemeyer und Borowski (2019) zurückgegriffen, das bereits bei der Entwicklung des Reflexions-

tests zugrunde gelegt wurde: Da Bewertungen und Alternativen (inklusive weniger Konsequenzen) zu jedem reflektierten Aspekt jeweils in getrennten Aufgaben erfasst werden, können auch diese beiden Teilfähigkeiten explizit rückgemeldet werden. Als drittes Teilergebnis wird rückgemeldet, wie die Teilnehmenden bei der Diagnose von und dem Umgang mit Schülervorstellungen im Reflexionstest abgeschnitten haben. Damit wird hier also nicht auf ein Reflexionselement eingegangen, sondern es wird ein inhaltlich relevanter Aspekt der Physikdidaktik gegenüber den anderen reflektierten Aspekten im Reflexionstest besonders hervorgehoben, was mit der Bedeutung von Schülervorstellungen für die zu den Lernvoraussetzungen der Schüler*innen passende Gestaltung von Physikunterricht (Schecker & Duit, 2018) begründet werden kann. Da die Aufgaben zu diesem Aspekt nicht entlang der Reflexionselemente Bewertungen und Alternativen formuliert sind, sondern entsprechend auf die Diagnose von und den Umgang mit Schülervorstellungen (s. Kapitel 10.1), ermöglicht dieses Vorgehen den Einbezug aller Aufgaben des Reflexionstests in je eine Teilfähigkeit.

Die inhaltliche Gestaltung des Assessment-Feedbacks wird entlang der in den Kapiteln 8.1 und 8.2 dargestellten Erkenntnisse diskutiert. Im Sinne des Feedback-Modells von Hattie und Timperley (2007) sollte das Assessment-Feedback zum Reflexionstest also zunächst die drei Bestandteile Feed Up, Feed Back und Feed Forward inkludieren. Diese Elemente finden sich grundsätzlich in der Struktur des Assessment-Feedbacks wieder. Auf die Einleitung folgt die Beschreibung der getesteten Fähigkeit im Abschnitt *Reflektieren - Was ist das genau & wie kommt das im Test vor?* Hier wird die für Lehrkräfte relevante und daher zu entwickelnde Reflexionsfähigkeit allgemein beschrieben sowie deren Abbildung im Reflexionstest entlang der Teilfähigkeiten des Bewertens und Aufzeigens von Alternativen, was als Teil des Feed Up nach Hattie und Timperley (2007) eingeordnet werden kann. Diese Erläuterungen zu den Elementen der Unterrichtsreflexion sprechen im Prinzip die Prozess- bzw. Selbstregulations-Ebene für Unterrichtsreflexion an, allerdings auf einer vom Reflexionstest abstrahierten Ebene, da eigene Reflexionen in Bezug auf das Vorkommen dieser Elemente evaluiert werden können. Aufgrund der engen Führung durch die MC-Aufgaben im Reflexionstest kann nicht davon ausgegangen werden, dass hier Unterstützung explizit zum Vorgehen bei der Bearbeitung des Reflexionstests geboten wird, sondern Hinweise durch Durchführung von Unterrichtsreflexionen allgemein. Außerdem ist eine grundlegende Einschränkung bei der Bereitstellung von Feedback zur Reflexionsfähigkeit zu berücksichtigen: Feedback dient grundsätzlich dazu, eine Diskrepanz zwischen Ist- und Sollzustand zunächst transparent zu machen und anschließend zu überwinden (z. B. Hattie und Timperley, 2007). Eine normative Zielvorgabe für die Fähigkeit zur Unterrichtsreflexion kann auf Grundlage bisheriger Befunde u. a. zum Zusammenhang zwischen der Reflexionsfähigkeit und Unterrichtsqualität bzw. Professionalität allerdings nicht abgeleitet (Häcker, 2022) und auch auf Grundlage Standards für die Lehrkräftebildung kaum formuliert werden (vgl. von Aufschnai-

ter, Hofmann, Geisler und Kirschner, 2019). Das Feed Up (die Zielvorgabe) kann hier also nur eingeschränkt adressiert werden. Konkret auf den Reflexionstest bezogen kann das möglichst gute Abschneiden als Zielvorgabe ausgewiesen werden, da das Ergebnis als Maß für die Ausprägung der Reflexionsfähigkeit interpretiert wird. Auch aus diesem Grund ist die Ergebnisdarstellung im Assessment-Feedback an einer sozialen Bezugsnorm orientiert, was dem grundsätzlich für Feedback empfohlenen Sachbezug widerspricht (Sippel, 2009). Den Teilnehmenden wird durch die kurze Beschreibung der Vergleichsgruppe (Anzahl der Studierenden mit Angabe des durchschnittlichen Fachsemesters und Referendar*innen) ermöglicht, sich bezüglich des Studienfortschritts in die Vergleichsgruppe einzuordnen. Durch die wiederholte Bearbeitung des Reflexionstests können außerdem Entwicklungen über intraindividuelle Vergleiche festgestellt werden.

Anschließend werden die Testwerte, d. h. Einschätzungen der individuellen Fähigkeiten auf Grundlage der Testbearbeitungen, dargestellt, was durch die Angabe der Testwerte vorrangig einem Feed Back (Darstellung des Ist-Zustands) auf Aufgaben-Ebene im Sinne des Feedback-Modells von Hattie und Timperley (2007) entspricht. Vor dem Hintergrund, dass Abweichungen zwischen Ist- und Sollzustand möglichst präzise rückgemeldet und ggf. auch um Musterlösungen ergänzt werden sollten, könnte ein höherer Auflösungsgrad in der Ergebnisdarstellung (z. B. auf Aufgabenebene) angestrebt werden. Gleichzeitig würde eine solche Aufschlüsselung der Ergebnisse auf Aufgabenebene den Umfang der Rückmeldung stark erhöhen und die Option, die Lernumgebung wiederholt einzusetzen, ausschließen. Beispiele für gelungene Reflexionen sind allerdings im Selbstlernmaterial eingebunden. Damit wird der Empfehlung, im Rahmen einer Rückmeldung zu handlungsnahen Fähigkeiten optimales Verhalten statt bzw- als Analogon zu einer Musterlösung zu beschreiben (Hosenfeld, 2010, S. 42) im Selbstlernmaterial für weitere Unterrichtsvignetten nachgekommen.

Auf dieses Selbstlernmaterial werden die Teilnehmenden abschließend als eine Möglichkeit zur praktischen Übung der Reflexionsfähigkeit verwiesen, wobei auf Grundlage der individuellen Ergebnisse eine Fokussierung auf die Entwicklung von Bewertungen bzw. Alternativen (sowie bei Testwerten von unter 50 % in den entsprechenden Aufgaben zusätzlich auf Schülervorstellungen) empfohlen wird. Damit sind im Sinne eines Feed Forward sinnvolle nächste Schritte zur Arbeit an der eigenen Reflexionsfähigkeit eingebunden.

Shute (2008) beschreibt, dass zu umfangreiches Assessment-Feedback nicht umfassend wahrgenommen wird; unklar bleibt allerdings, welcher Umfang als angemessen angesehen werden kann, sodass hier auf die Erkenntnisse aus den Interviews (s. Kapitel 18.1) zurückgegriffen werden muss. Dabei wurde deutlich, dass das Assessment-Feedback von einigen Studierenden als (zu) umfangreich wahrgenommen wird. Gleichzeitig können aus den Interviews keine Rückschlüsse auf sinnvolle weitere Kürzung des Assessment-Feedbacks gezogen werden. Die strukturelle

Trennung zwischen der Rückmeldung der Testwerte im PDF-Dokument sowie dem verlinkten Selbstlernmaterial in einer *LimeSurvey*-Umgebung kann dabei einerseits eine Hürde für die Bearbeitung des Selbstlernmaterials darstellen, da es explizit aufgerufen werden muss. Andererseits würde die Zusammenführung beider Bestandteile das Assessment-Feedback deutlich umfangreicher gestalten.

Neben Kriterien der inhaltlichen Gestaltung kommt außerdem dem Verfahren zur Bereitstellung des Assessment-Feedbacks Bedeutung zu. Die Teilnehmenden können ihr Assessment-Feedback üblicherweise wenige Stunden bis Tage nach Bearbeitung des Reflexionstests anonym herunterladen. Dies ermöglicht den Studierenden, sich zu einem individuell geeigneten Zeitpunkt mit dem eigenen Assessment-Feedback auseinanderzusetzen, was nach Parkin et al. (2012) von Studierenden wertschätzend aufgenommen wird. Die Bereitstellung des Assessment-Feedbacks als PDF-Dokument ermöglicht es den Teilnehmenden außerdem, ein erhaltenes Assessment-Feedback zu sichern und zukünftig darauf zurückzugreifen, sodass bspw. nach einer erneuten Bearbeitung des Reflexionstests das erhaltene Assessment-Feedback mit dem vorigen verglichen werden und eine mögliche Entwicklung festgestellt werden kann.

Die Teilnehmenden erhalten hier keine Rückmeldung zu den konkreten Aufgaben, sondern zur übergeordneten bzw. daraus abgeleiteten Fähigkeit zur Unterrichtsreflexion. Aufgrund der Komplexität der rückgemeldeten Fähigkeit könnte allerdings gerade für wenig fortgeschrittene Lernende ein unmittelbares Feedback lohnenswert sein (vgl. Shute, 2008). Da die Vergleichsgruppe, die im Assessment-Feedback als soziale Bezugsnorm zugrunde gelegt wird, sich zum Zeitpunkt der Erhebungen im Aufbau befand, konnte dies bislang nicht umgesetzt werden; perspektivisch wäre ein vollständig automatisiertes Assessment-Feedback realisierbar, das unmittelbar im Anschluss an die Testbearbeitung ausgegeben wird.

18.3 Beurteilung des Teils 4 des INA: Übersetzung Testwerte in Assessment-Feedback

Die Prüfung des vierten und letzten Teils des INA soll Aufschluss über die konsequentielle Validität der Testwertinterpretation liefern und sicherstellen, dass das Assessment-Feedback eine geeignete Form der Rückmeldung der Testwerte darstellt. Dazu wurde vorrangig die Perspektive der Studierenden auf das Assessment-Feedback erhoben. Insgesamt zeigt sich, dass die zentralen Elemente (insbesondere die Ergebnisdarstellung in Boxplots) trotz einzelner Unsicherheiten und fachlicher Ungenauigkeiten für die Studierenden verständlich sind und die Studierenden größtenteils in der Lage sind, ihre Ergebnisse angemessen zu interpretieren (Indikator 4.1). Im Rahmen der Interviews (und tendenziell auch im Akzeptanzfragebogen) sprechen die Studierenden außerdem sowohl der Er-

gebnisdarstellung im Assessment-Feedback als auch dem Selbstlernmaterial einen Mehrwert zu (Indikator 4.2). In der Bewertung dieser Einschätzungen ist allerdings zu berücksichtigen, dass hieraus kaum Erkenntnisse über einen tatsächlichen Mehrwert in Bezug auf die Entwicklung der Reflexionsfähigkeit abgeleitet werden können. Im Rahmen der Interviews werden lediglich Selbstauskünfte bezüglich des Vorgehens bei der Bearbeitung des Selbstlernmaterials bzw. der Intention, das Material zukünftig zu bearbeiten, erfragt. Wie bereits zu anderen Indikatoren diskutiert, wurden die Interviews nur mit einer Gelegenheitsstichprobe durchgeführt, die sich hier lediglich aus Bachelorstudierenden zusammensetzt, sodass limitierend festzuhalten ist, dass keine empirisch fundierten Argumente für die Generalisierbarkeit der Ergebnisse vorliegen. Insbesondere die Erkenntnisse aus dem eingesetzten Akzeptanzfragebogen nach Erhalt des Assessment-Feedbacks und Bearbeitung des Selbstlernmaterials sind aufgrund der sehr geringen Stichprobengröße als vorläufig einzuschätzen. Dass die Studierenden im Rahmen der Interviews vorrangig nicht ihr eigenes, sondern eine Beispiel-Rückmeldung diskutieren, ist außerdem als limitierend festzuhalten, da mögliche emotionale und motivationale Einflüsse (z. B. aufgrund von (fehlendem) Selbstvertrauen (Butler & Winne, 1995)) auf die Bewertung des individuellen Assessment-Feedbacks aufgrund der persönlichen Betroffenheit nicht berücksichtigt werden (können). Diese Einschränkung wurde in Kauf genommen, um die Hürde zur Teilnahme an den Interviews zu senken, da so die Anonymität der Testergebnisse weiterhin gewährleistet werden konnte.

Mit dem zusätzlichen Abgleich der Gestaltung des Assessment-Feedback mit evidenzbasierten bzw. theoretisch begründeten Kriterien für effektives Feedback wurde sichergestellt, dass sich das Assessment-Feedback auf inhaltlich relevante Aspekte der Reflexionsfähigkeit bezieht und Kriterien für effektives Feedback, wie z. B. das Adressieren von Feed Up, Back und Forward nach Hattie und Timperley (2007), erfüllen (Indikator 4.3). Ergänzend zum theoretischen Abgleich mit Kriterien zur Effektivität des Assessment-Feedbacks könnten vergleichend verschiedene Formen des Feedbacks, die z. B. in ihrem Umfang oder Detailgrad variieren oder das Assessment-Feedback und Selbstlernmaterial gemeinsam online darstellen, eingesetzt werden, um einen detaillierteren Einblick in die Wahrnehmung des Feedbacks (inklusive Selbstlernmaterial) durch die Teilnehmenden zu erhalten.

Dass die drei untersuchten Indikatoren als erfüllt angesehen werden können, stützt den hier untersuchten Aspekt der konsequentiellen Validität der Interpretation der Testwerte: Es wird, wie intendiert, den Studierenden ein Assessment-Feedback zur Verfügung gestellt, das diese weitgehend angemessen vor dem Hintergrund der Ergebnisse der Vergleichsgruppe interpretieren. Aufgetretene Fehlinterpretationen bezüglich der Aussagekraft des Assessment-Feedbacks für künftige Prüfungen können im Rahmen eines Einsatzes in Lehrveranstaltungen explizit adressiert und so präventiv ausgeräumt werden und stellen somit keine besondere Gefährdung der validen Testwertinterpretation dar. Einschränkend ist festzuhalten, dass die Evaluation des Selbstlernmaterials ausschließlich auf Selbstauskünften der Studierenden

basiert. Inwiefern also tatsächlich eine Auseinandersetzung mit dem Material stattgefunden hat bzw. inwiefern Absichten, das Material zu nutzen, bestehen oder andere konkrete Maßnahmen zur individuellen Weiterentwicklung ergriffen werden, kann nicht beurteilt werden. Das hier untersuchte Assessment-Feedback stellt lediglich eine Voraussetzung dafür dar.

19 Validitätsargumentation zum Reflexionstest (Beantwortung der Forschungsfrage 1)

Die Entwicklung der Lernumgebung bestehend aus Reflexionstest und Assessment-Feedback (inklusive Selbstlernmaterial) war mit dem Ziel verbunden, Teilnehmenden eine fundierte Selbsteinschätzung ihrer Reflexionsfähigkeit zu ermöglichen und einen Ausgangspunkt für mögliche Weiterentwicklung zu bieten. Die Bewertung der Eignung des Testinstruments für diesen intendierten Zweck erfolgt auf Grundlage der in Kapitel 12 formulierten Forschungsfrage 1, die hier auf Grundlage der durchgeführten Untersuchungen beantwortet wird:

Forschungsfrage 1: Inwieweit kann die folgende Testwertinterpretation gerechtfertigt werden?

Die Testwerte des Reflexionstests können (a) als Maß für die Reflexionsfähigkeit und (b) als Ausgangspunkt für einen weiteren Professionalisierungsprozess aufgefasst werden.

Das ebenfalls in Kapitel 12 formulierte INA bildet den Bezugsrahmen für die Bewertung der Forschungsfrage. Es wird um übergeordnete Qualitätsanforderungen ergänzt, die an die vier durchlaufenen und evaluierten Übersetzungsschritte gestellt werden:

Anforderung an Teil 1 des INA: Die Aufgaben im Reflexionstest bilden relevante und repräsentative Inhalte der Fähigkeit zur Unterrichtsreflexion präzise ab.

Anforderung an Teil 2 des INA: Die Proband*innen zeigen bei der Bearbeitung des Reflexionstests konstruktrelevantes Verhalten.

Anforderung an Teil 3 des INA: Das (Ankreuz-)Verhalten der Proband*innen während der Testbearbeitung wird adäquat in Testwerte übersetzt.

Anforderung an Teil 4 des INA: Das Assessment-Feedback stellt eine geeignete Form der Rückmeldung der Testwerte dar.

Zu diesen Qualitätsanforderungen wurden jeweils Indikatoren formuliert (s. Kapitel 12), deren empirische oder theoretische Prüfung in den Kapiteln 15 bis 18 beschrieben wurde. Eine Übersicht über die zu prüfenden Anforderungen und untersuchten Indikatoren sowie dazu gewählte Methoden für alle Teile des INA befindet sich in den Abbildungen 14.1 und 14.2 auf den Seiten 186 und 188. Diese Untersuchungen bilden die Grundlage für die Beurteilung der Gültigkeit der vier Teile des INA. Zur abschließenden Beurteilung der Gültigkeit sind insbesondere auch Limitationen der Untersuchungen als Einschränkungen bzw. Gefährdungen für die valide Interpretation der Testwerte zu berücksichtigen. In Anlehnung an Dickmann (2016, S. 141f.) und Bartels (2018, S. 158) kann in der Bewertung folgendermaßen zwischen bedrohenden und nicht-bedrohenden Einschränkungen unterschieden werden, die in der nachfolgenden zusammenfassenden Bewertung der Bestandteile des INA berücksichtigt wird:

- *Bedrohende Einschränkungen* schwächen die Validitätsargumentation oder schränken diese ein (wenn z. B. erwartete Zusammenhänge im Rahmen der externen Validierung nicht bestätigt werden können).
- *Nicht-bedrohende Einschränkungen* ergeben sich „aufgrund einer eingeschränkten Datenbasis“ (Dickmann, 2016, S. 142) und können ohne hohen Aufwand und die Durchführung weiterer Studien nicht vermieden werden. Sie sprechen allerdings nicht pauschal gegen das Beibehalten der formulierten Indikatoren.

Ergänzend können für die umfassende Evaluation der Lernumgebung weitere (klassische) Gütekriterien betrachtet werden, die wie in Kapitel 6.5 angekündigt teilweise explizit Teil der Indikatoren zur Überprüfung des INA sind, teilweise implizit enthalten bzw. vorausgesetzt werden. Diese werden ergänzend in Kapitel 19.5 diskutiert.

19.1 Zusammenfassende Bewertung des Teils 1 des INA: Übersetzung Reflexionsfähigkeit in Testinstrument

Die Bewertung des ersten Teils des INA *„Die Aufgaben im Reflexionstest bilden relevante und repräsentative Inhalte der Fähigkeit zur Unterrichtsreflexion präzise ab.“* fokussiert auf die inhaltliche Validität der Testwertinterpretation. Dazu wurde zunächst diskutiert, inwieweit *die Auswahl der Szenen und Aufgaben ein umfassendes Abbild der Fähigkeit zur Unterrichtsreflexion in Bezug auf einen inhaltlich bedeutsamen Fachinhalt von Physikunterricht darstellen* (Indikator 1.1). Hier wurde in Bezug auf die inhaltsvalide Abbildung von Physikunterricht als Reflexionsgrundlage argumentiert, dass die Vignetten im Reflexionstest geeignet sind, Physikunterricht zwar eingeschränkt auf einen Fachinhalt der Mechanik,

aber dennoch in gewisser Breite über Unterrichtsphasen (durch das Nachzeichnen einer Doppelstunde) und -methoden abzubilden. Bezüglich der Abbildung der einzelnen Dimensionen von Unterrichtsreflexion gemäß dem Modell von Nowak, Kempin, Kulgemeyer und Borowski (2019) wurde argumentiert, dass erstens trotz der Fokussierung auf Situationen, die für den *Physik*unterricht spezifisch sind, also vorrangig auf Grundlage von fachlichem und fachdidaktischem Wissen reflektiert werden, alle Wissensbasen (fachliches, fachdidaktisches und pädagogisches Wissen) im Reflexionstest abgebildet werden, zweitens die Fokussierung auf Aufgaben zu Bewertungen und Alternativen (inklusive Konsequenzen) die Reflexionselemente (Beschreibungen, Bewertungen, Alternativen und Konsequenzen) nur in reduzierter Form, aber angemessen vor dem Hintergrund der realen Reflexionen von Studierenden zu den Videovignetten, die Grundlage für die entwickelten MC-Aufgaben sind, abbildet. Zusätzlich wurde die Klassifikation der Antwortoptionen auf Grundlage des Indikators 1.2 *„Die als korrekt definierten Antwortoptionen repräsentieren die zentralen Erkenntnisse der Unterrichtsreflexion zu den dargestellten Szenen.“* diskutiert. Grundlage für die Einschätzung, den Indikator als bestätigt anzunehmen, ist die empirische Fundierung durch die Expert*innenbefragung des zur Bewertung zugrundegelegten Kodiermanuals zum Reflexionsperformanztest von Kempin, Kulgemeyer und Schecker (2018). Ein Überblick über die Limitationen zu beiden Indikatoren ist in Tabelle 19.1 dargestellt, wobei die inhaltliche Einschränkung auf einen Teilbereich der Mechanik sowie die reduzierte Authentizität und Komplexität der Reflexionssituation, die mit der eingeschränkten Handlungsfreiheit durch die Verwendung von MC-Aufgaben in der Reflexion einhergeht, Einschränkungen der Testwertinterpretation darstellen. Hier kann eine Verschiebung von der Erfassung der Reflexionsfähigkeit mit einem Performanztest (mit dem Test von Kempin, Kulgemeyer und Schecker (2018)) hin zum Kompetenztest nach Miller (1990) unterstellt werden. Vor dem Hintergrund der theoretisch fundierten Abbildung von Unterrichtsreflexion im Reflexionstest sowie der festgestellten mittleren Korrelation, die an der Schwelle zur hohen Korrelation ist, zwischen dem Reflexionstest und dem der Entwicklung zugrundeliegenden Reflexionsperformanztest von Kempin, Kulgemeyer und Schecker (2018), die im Rahmen der Untersuchung von Indiktor 3.4 festgestellt wurde (s. Kapitel 17.4), erscheint die Validität der Testwertinterpretation nicht grundsätzlich gefährdet.

Insgesamt liegen mit den begründeten Indikatoren also Argumente vor, die für die Angemessenheit der Übersetzung des Konstrukts Unterrichtsreflexion in Testaufgaben sprechen, sodass die Anforderung an den ersten Teil des INA als bestätigt angesehen werden kann. Zusammenfassend kann unter Berücksichtigung der zu den Indikatoren diskutierten Limitationen festgehalten werden:

Teil 1 des INA: Die Aufgaben im Reflexionstest bilden relevante und repräsentative Inhalte der Fähigkeit zur Unterrichtsreflexion präzise ab.		
Indikator	Bestätigt?	Limitationen
1.1 Die Auswahl der Szenen und Aufgaben stellt ein umfassendes Abbild der Fähigkeit zur Unterrichtsreflexion in Bezug auf einen inhaltlich bedeutsamen Fachinhalt von Physikunterricht dar.	(✓)	• Einschränkung auf ein konkretes Unterrichtsszenario (insb. Fachinhalt, Lerngruppe) • (Empirisch fundierte) Fokussierung auf Bewertungen und Alternativen als Teile von Unterrichtsreflexion • **Beschränkte Handlungsfreiheit durch Einsatz von MC-Aufgaben**
1.2 Die als korrekt definierten Antwortoptionen repräsentieren die zentralen Erkenntnisse der Unterrichtsreflexion zu den dargestellten Szenen.	✓	• Auswahl der Expert*innen zur Erstellung der Bewertungsgrundlage (Kodiermanual zum Reflexionsperformanztest)

Tabelle 19.1: Übersicht über Bewertung der Indikatoren und Limitationen zum Teil 1 des INA (fett gedruckt sind Limitationen, die als die Testwertinterpretation einschränkend zu bewerten sind)

Die Aufgaben im Reflexionstest bilden relevante und repräsentative Inhalte der Fähigkeit zur Unterrichtsreflexion am Beispiel eines Fachinhalts der Mechanik und unter eingeschränkter Handlungsfreiheit im Sinne eines Kompetenztests nach Miller (1990) präzise ab.

19.2 Zusammenfassende Bewertung des Teils 2 des INA: Übersetzung Testinstrument in Verhalten

Der zweite Teil des INA fokussiert auf das Verhalten der Teilnehmenden während der Testbearbeitung und damit auf die kognitive Validität; es ist sicherzustellen, dass *die Proband*innen bei der Bearbeitung des Reflexionstests konstruktrelevantes Verhalten zeigen.* Hierzu wurden im Rahmen der Prüfung des Indikators 2.1 *„Die Überlegungen der Teilnehmenden (a) beziehen sich auf die Qualität des Unterrichts bzw. die Professionalisierung und (b) passen zur Aufgabenstellung.“* im Rahmen einer Think-Aloud-Studie die Überlegungen von $N = 7$ Studierenden untersucht,

wobei deutlich wurde, dass der größte Teil der Überlegungen Elemente von Reflexionen darstellte oder sonstigen direkten Bezug zum Unterricht aufwies. Eine darüberhinausgehende Passung zwischen den Überlegungen der Teilnehmenden und den konkreten Aufgaben in Bezug auf die Reflexionselemente Bewertungen und Alternativen konnte allerdings nicht festgestellt werden. Zusätzlich wurden zur Überprüfung des Indikators 2.2 *„Die Teilnehmenden akzeptieren das Testverfahren.“* Fragebögen zur Akzeptanz und Motivation eingesetzt sowie Freitextantworten im Kommentarfeld am Ende des Reflexionstests ausgewertet. Mit Ausnahme weniger technischer Schwierigkeiten bei der Bearbeitung des Reflexionstests, die teilweise behoben werden konnten, und dem Eindruck, der Reflexionstest sei eher umfangreich, scheint er insgesamt auf breite Akzeptanz zu treffen.

Weitere Limitationen dieser Untersuchungen liegen bezüglich der Stichprobengröße und -auswahl sowie des Detailgrades der Auswertung in den Think-Aloud-Interviews sowie der Erhebungsmethode zur Akzeptanz des Testverfahrens vor (s. Tabelle 19.2): Eine umfangreichere Datenbasis würde möglicherweise differenziertere Ergebnisse zu beiden Indikatoren liefern. Diese Limitationen sind allerdings nicht als bedrohlich für die Beurteilung der Validität der Testwertinterpretation einzuschätzen.

Insgesamt liegt also empirische Evidenz für die beide Indikatoren zur Anforderung an den zweiten Teil des INA vor, sodass dieser uneingeschränkt als bestätigt eingeschätzt und festgehalten werden kann:

*Die Proband*innen zeigen bei der Bearbeitung des Reflexionstests konstruktrelevantes Verhalten.*

19.3 Zusammenfassende Bewertung des Teils 3 des INA: Übersetzung Verhalten in Testwerte

Die Untersuchungen zum dritten Teil des INA dienen der Prüfung der *adäquaten Übersetzung des (Ankreuz-)Verhaltens der Proband*innen während der Testbearbeitung in Testwerte.* Dazu wurden zunächst als Voraussetzung für die und im Sinne der strukturellen Validität Skalen- und Itemanalysen durchgeführt, die zum Indikator 3.1 *„Der Reflexionstest ist dazu geeignet, verschiedene Ausprägungen der Reflexionsfähigkeit auch innerhalb der verschiedenen Ausbildungsphasen (Bachelor- und Masterstudium) differenziert darzustellen.“* Analysen der Verteilung der Testwerte sowieso der Höhe der Lösungshäufigkeiten beinhalten und zum Indikator 3.2 *„Die zur Bestimmung der Testwerte verwendeten Skalen sind reliabel auch über die verschiedenen Ausbildungsphasen (Bachelor- und Masterstudium) hinweg.“* um Betrachtungen der internen Konsistenz und Aufgabentrennschärfen ergänzt

Teil 2 des INA: Die Proband*innen zeigen bei der Bearbeitung des Reflexionstests konstruktrelevantes Verhalten.		
Indikator	Bestätigt?	Limitationen
2.1 Die Überlegungen der Teilnehmenden (a) beziehen sich auf die Qualität des Unterrichts bzw. die Professionalisierung und (b) passen zur Aufgabenstellung.	✓	• Kleine Gelegenheitsstichprobe • Detailgrad der Auswertung (Fokussierung auf Verbalisiertes) • Passung zwischen verbalisierten Reflexionselementen (Bewertungen bzw. Alternativen) und Aufgabe
2.2 Die Teilnehmenden akzeptieren das Testverfahren.	✓	• Keine etablierten Instrumente • Freiwillige Nutzung des Kommentarfelds

Tabelle 19.2: Übersicht über Bewertung der Indikatoren und Limitationen zum Teil 2 des INA

wurden. Obwohl die Testwerte insgesamt eher hoch ausfallen, konnten die Indikatoren jeweils für die Gesamtstichprobe und beide Studienphasen bestätigt werden. Zum Indikator 3.3 *„Die Struktur der Testwerte ist theoretisch erklärbar.“* wurde eine explorative Faktorenanalyse durchgeführt, die auf die Eindimensionalität des erfassten Konstrukts hinweist. Als limitierend ist hier die Zusammensetzung der Stichprobe anzuführen, die als Gelegenheitsstichprobe zwar relativ heterogen über den Studienverlauf und verschiedene Ausbildungsstandorte ausfällt, allerdings nicht als repräsentativ z. B. in Bezug auf den Lehramtstypen angenommen werden kann.

Zusätzlich wurden Untersuchungen zur externen Validierung durchgeführt, deren Ergebnisse nachfolgend zusammengefasst werden:

- Indikator 3.4 *„Die Testwerte des Reflexionstests weisen Zusammenhänge zur Qualität von Reflexionen in authentischeren Settings (repräsentiert durch die Testwerte des Reflexionsperformanztests) auf.“* konnte für eine kleine Stichprobe von $N = 41$ Studierenden mit beinahe hoher Effektstärke ($r_{\text{Pearson}} = 0,49$) bestätigt werden. Hier kann allerdings ein Einfluss der großen inhaltlichen Überschneidungen beider Testinstrumente nicht ausgeschlossen werden.
- Zum Zusammenhang der Testwerte mit Lerngelegenheiten und sonstigen demografischen Variablen liegen ambivalente Befunde vor: Während z. B. Zusammenhänge zwischen dem Studienfortschritt und im Rahmen von Pflichtpraktika gesammelter Praxiserfahrung stützende Evidenz für Indikator 3.5 *„Unterschiedliche Lerngelegenheiten der Teilnehmenden führen zu nach-*

vollziehbaren Unterschieden in den Testwerten.“ liefern, könnten Zusammenhänge zu einzelnen demografischen Variablen oder der Abiturnote auf konstruktirrelevante Varianz (Messick, 1989) in den Testwerten hinweisen. Dies betrifft beobachtete Mittelwertsunterschiede abhängig vom Geschlecht und Zusammenhänge zur Abitur- bzw. letzten Deutschnote, die tendenziell zu Ergebnissen anderer Studien passen, die Zusammenhänge zwischen Schulnoten und z. B. dem Studienerfolg untersuchen (z. B. Trapmann, Hell, Weigand und Schuler, 2007). Da der Großteil der (nicht erwartungskonformen) Analysen nicht auf realen Längsschnittdaten basiert, besteht weiterer Untersuchungsbedarf; der Indikator kann auf dieser Grundlage aber nur eingeschränkt als bestätigt angesehen werden.

- Der Indikator 3.6 *„Die Testwerte des Reflexionstests weisen einen Zusammenhang zum fachdidaktischen Wissen auf.“* konnte empirisch nicht bestätigt werden. Die ausbleibende, aber tendenziell sogar negative Korrelation zum fachdidaktischen Wissen ist als potenziell gefährdend für die valide Testwertinterpretation einzuschätzen, da gemäß dem Reflexionsmodell von Nowak, Kempin, Kulgemeyer und Borowski (2019) u. a. das fachdidaktische Wissen als Grundlage für fundierte Unterrichtsreflexionen verstanden wird. Vor dem Hintergrund der geringen Stichprobengröße sollten weitere Evidenzen zur Bewertung dieses Zusammenhangs generiert werden. Dabei sollte zusätzlich auf eine heterogenere Zusammensetzung der Stichprobe geachtet werden, um (zumindest über die Studienphasen hinweg) generalisierbare Schlüsse zu ermöglichen.

Insgesamt liegen mit den bestätigten Indikatoren 3.1 bis 3.3 also Hinweise auf die statistisch angemessene Skalenbildung für den Reflexionstest insgesamt und für die Phasen des Bachelor- und Masterstudiums vor. Ergänzend stützen die Zusammenhänge zum Reflexionsperformanztest (Indikator 3.4) sowie zu (stellvertretenden Maßen für) Lerngelegenheiten (Indikator 3.5), wie dem Studienfortschritt repräsentiert durch die Studienphase oder das Fachsemester, diese Interpretation. Als limitierend ist insbesondere der ausbleibende, Zusammenhang zum fachdidaktischen Wissen (Indikator 3.6) aufzuführen. Bezüglich der Anforderung an den dritten Teil des INA wird daher festgehalten:

*Das (Ankreuz-)Verhalten der Proband*innen während der Testbearbeitung wird vorbehaltlich nicht beobachteter Zusammenhänge zum fachdidaktischen Wissen adäquat in Testwerte übersetzt.*

Teil 3 des INA: Das (Ankreuz-)Verhalten der Proband*innen während der Testbearbeitung wird adäquat in Testwerte übersetzt.		
Indikator	Bestätigt?	Limitationen
3.1 Der Reflexionstest ist dazu geeignet, verschiedene Ausprägungen der Reflexionsfähigkeit auch innerhalb der verschiedenen Ausbildungsphasen (Bachelor- und Masterstudium) differenziert darzustellen.	✓	• Gelegenheitsstichprobe
3.2 Die zur Bestimmung der Testwerte verwendeten Skalen sind reliabel auch über die verschiedenen Ausbildungsphasen (Bachelor- und Masterstudium) hinweg.	✓	• Gelegenheitsstichprobe
3.3 Die Struktur der Testwerte ist theoretisch erklärbar.	✓	• Gelegenheitsstichprobe • andere Extraktionsmethoden deuten auf abweichende Anzahl an Faktoren hin
3.4 Die Testwerte des Reflexionstests weisen Zusammenhänge zur Qualität von Reflexionen in authentischeren Settings (repräsentiert durch die Testwerte des Reflexionsperformanztests) auf.	✓	• Stichprobengröße und -zusammensetzung • Inhaltliche Ähnlichkeit der Instrumente
3.5 Unterschiedliche Lerngelegenheiten der Teilnehmenden führen zu nachvollziehbaren Unterschieden in den Testwerten.	(✓)	• Überwiegend nur Quasi-Längsschnitt • Hinweise auf konstruktirrelevante Varianz
3.6 Die Testwerte des Reflexionstests weisen einen Zusammenhang zum fachdidaktischen Wissen auf.	✗	• Stichprobengröße und - zusammensetzung • **Keine Korrelation zum fachdidaktischen Wissen**

Tabelle 19.3: Übersicht über Bewertung der Indikatoren und Limitationen zum Teil 3 des INA (fett gedruckt sind Limitationen, die als die Testwertinterpretation einschränkend zu bewerten sind)

19.4 Zusammenfassende Bewertung des Teils 4 des INA: Übersetzung Testwerte in Assessment-Feedback

Der vierte und letzte Teil des INA berücksichtigt die konsequentielle Validität indem die Angemessenheit der Interpretation des Assessment-Feedbacks und daraus gezogener Konsequenzen durch die Teilnehmenden evaluiert wird. Überprüft wird die Anforderung: *Das Assessment-Feedback stellt eine geeignete Form der Rückmeldung der Testwerte dar.* Dazu wurden mit einer Interviewstudie die Indikatoren 4.1 *„Die Interpretationen der Ergebnisse durch die Teilnehmenden passen inhaltlich zum erhaltenen Feedback.“* und 4.2 *„Die Teilnehmenden sprechen dem Assessment-Feedback (inklusive Selbstlernmaterial) einen Mehrwert zu.“* untersucht. Dabei zeigte sich insgesamt, dass die Teilnehmenden trotz einiger auftretender fachlicher Ungenauigkeiten (z. B. in der Verwechslung von Median und Durchschnitt) inhaltlich sinnvolle Interpretationen und Bewertungen des erhaltenen Assessment-Feedbacks vornehmen und dem Feedback (inklusive Selbstlernmaterial) einen Mehrwert zusprechen, was die Validitätsargumentation stützt. Einschränkend ist festzuhalten, dass die Interviewstichprobe eher homogen aus Bachelorstudierenden zusammengesetzt ist; eine heterogenere Stichprobe könnte ein differenzierteres Bild ggf. auch über die Entwicklung der Interpretationen und Einschätzungen im Verlauf des Studiums bieten. Außerdem liegen den Interviews überwiegend Beispielrückmeldungen zugrunde, sodass möglicherweise auftretende Einflüsse der persönlichen Betroffenheit (z. B. (fehlendes) Selbstvertrauen (Butler & Winne, 1995)) kaum berücksichtigt werden können. Aus methodischer Perspektive wäre die zusätzliche Kodierung des Materials durch mehrere unabhängige Kodierende anzustreben (Kuckartz, 2018, S. 205). Zusätzlich wurde zu Indikator 4.3 *„Das Assessment-Feedback stellt inhaltlich relevante Testergebnisse unter Berücksichtigung von Qualitätskriterien für effektives Feedback dar.“* theoriebasiert argumentiert, welche Qualitätskriterien für effektives Feedback in der Gestaltung des Assessment-Feedback erfüllt sind, wobei vorrangig das Feedback-Modell von Hattie und Timperley (2007) zugrunde gelegt wurde.

Vor dem Hintergrund der angestrebten Interpretation der Testwerte im Reflexionstest als (a) Maß für die Reflexionsfähigkeit und (b) als Ausgangspunkt für einen weiteren Professionalisierungsprozess muss als potenziell gefährdende Einschränkung für die Validität der Testwertinterpretation berücksichtigt werden, dass im Rahmen der Interviews lediglich Selbstauskünfte und Absichten der Studierenden bezüglich der individuellen Weiterentwicklung geäußert wurden; konkrete Maßnahmen wurden kaum genannt und das Selbstlernmaterial kaum außerhalb einer Veranstaltung bearbeitet.

Insgesamt liegt also Evidenz dafür vor, dass die Anforderung an den vierten Teil des INA - vorrangig für den ersten Teil der Testwertinterpretation - als erfüllt angesehen werden kann:

Das Assessment-Feedback stellt eine geeignete Form der Rückmeldung der Testwerte zur fundierten Selbsteinschätzung dar.

Um eine fundierte Bewertung eines etwaigen Beitrags des Assessment-Feedback für die individuelle Weiterentwicklung zu ermöglichen, bedarf es weiterer Untersuchungen, die die Entwicklung der Studierenden längerfristig begleiten.

19.5 Bewertung sonstiger Gütekriterien

Wie bereits in Kapitel 6.5 beschrieben, liegt der Fokus dieser Arbeit auf der Entwicklung einer Validitätsargumentation für die fundierte Interpretation der Testwerte des Reflexionstests. Unter dem holistischen Verständnis der Validität werden dabei auch weitere Gütekriterien subsumiert. Die vorliegende Evidenz zu den jeweiligen Gütekriterien wird im Folgenden überblicksartig zusammengefasst, um vorliegende Bezüge zur Validitätsargumentation zu explizieren. Entsprechend der klassischen Evaluation steht dabei die Einschätzung des (Einsatzes des) Reflexionstests im Mittelpunkt, nicht die Interpretation der Testwerte.

Die vollständige Einbindung des Reflexionstests in *LimeSurvey* sowie der Information auf der Startseite des Instruments, dass keine weiteren Hilfsmittel zur Bearbeitung herangezogen werden sollten, trägt grundsätzlich zur Standardisierung des Testeinsatzes und damit der Durchführungsobjektivität bei. Darüber hinaus kann für den Einsatz des Reflexionstests als individuell und nach Bedarf einsetzbares Testinstrument kaum echte Durchführungsobjektivität erreicht werden, da die Teilnehmenden das Testinstrument selbständig z. B. zu Hause in Vorbereitung auf eine Veranstaltung durchführen können. Da die Teilnehmenden ihre Ergebnisse nach Erhalt der Rückmeldung selbständig interpretieren, können sie z. B. den Einfluss etwaiger störender Bearbeitungsumstände auf die Konzentration während der Bearbeitung ggf. selbstständig in die Beurteilung ihrer Ergebnisse einbeziehen. Nach der Bearbeitung des Testinstruments gewährleistet die automatisierte Bepunktung der MC-Aufgaben die Auswertungsobjektivität zum Reflexionstest. Hinweise auf Interpretationsobjektivität liegen auf Grundlage der zur konsequentiellen Validierung durchgeführten Interviews vor, da die Studierenden ähnliche Kriterien zur Beurteilung ihrer Ergebnisse heranziehen (v. a. die positive Bewertung von Ergebnissen über dem Median und das Ableiten von Verbesserungsbedarfen bei Ergebnissen darunter).

Zur Beurteilung der Reliabilität liegen auf Grundlage der durchführten Untersuchungen lediglich Informationen zur internen Konsistenz (bestimmt wurde

Teil 4 des INA: Das Assessment-Feedback stellt eine geeignete Form der Rückmeldung der Testwerte dar.		
Indikator	Bestätigt?	Limitationen
4.1 Die Interpretationen der Ergebnisse durch die Teilnehmenden passen inhaltlich zum erhaltenen Feedback.	✓	• (Auch) fachliche Ungenauigkeiten • Vorwiegend Betrachtung von Beispielrückmeldungen • Zusammensetzung der Stichprobe
4.2 Die Teilnehmenden sprechen dem Assessment-Feedback (inklusive Selbstlernmaterial) einen Mehrwert zu.	✓	• **Ausschließlich Selbstauskünfte (keine Untersuchung der Arbeit mit Selbstlernmaterial und Entwicklung der Reflexionsfähigkeit)** • Keine Vergleich der Kodierung durch unabhängige Kodierer*innen • Zusammensetzung der Stichprobe
4.3 Das Assessment-Feedback stellt inhaltlich relevante Testergebnisse unter Berücksichtigung von Qualitätskriterien für effektives Feedback dar.	✓	• Ausschließlich theoretische Einschätzung

Tabelle 19.4: Übersicht über Bewertung der Indikatoren und Limitationen zum Teil 4 des INA (fett gedruckt sind Limitationen, die als die Testwertinterpretation einschränkend zu bewerten sind)

α_{Cronbach}) vor. Für den Reflexionstest konnte auf Grundlage der Gesamtstichprobe ($N = 192$) eine sehr hohe interne Konsistenz von $\alpha_{\text{Cronbach}} = 0,95$ festgestellt werden. Da die Reflexionsfähigkeit kein stabiles Merkmal, sondern eine zu entwickelnde Fähigkeit darstellt (wie die Zusammenhänge der Testwerte zu Lerngelegenheiten nahelegen; s. Kapitel 17.5), kann keine Retest-Reliabilität bestimmt werden und der Aufwand für die Entwicklung und Validierung eines Paralleltests zur Untersuchung der Paralleltest-Reliabilität ist im Rahmen dieser Arbeit nicht leistbar.

Bezüglich des Nebengütekriteriums der Normierung muss berücksichtigt werden, dass keine externen bzw. normativ gesetzten Anforderungen an die Ausprägung der Reflexionsfähigkeit vorliegen, sodass lediglich eine Einordnung der Ergebnisse in eine Referenzstichprobe möglich ist (und im Rahmen des Assessment-Feedback zur Einordnung eingesetzt wird). Diese Stichprobe wurde im Rahmen der durchgeführten Studien sukzessive aufgebaut, weist mit $N = 192$ Teilnehmenden allerdings weiterhin eine eher geringe Größe auf (vgl. Pospeschill, 2010, S. 28f.) und kann nicht als repräsentativ angenommen werden.

Dem Gütekriterium der Vergleichbarkeit kann - wie bereits zur Paralleltest-Reliabilität beschrieben - aufgrund des hohen Aufwands der Testentwicklung nicht Rechnung getragen werden. Es liegt zwar mit dem Reflexionsperformanztest von Kempin, Kulgemeyer und Schecker (2018) ein Instrument vor, das Reflexion zu denselben Unterrichtssituationen erfasst, allerdings aufgrund des abweichenden Erhebungsformats nicht als Parallelform verstanden werden kann. Zusätzlich liegen verschiedene andere Verfahren vor, die zur Erfassung von Unterrichtsreflexionen entwickelt wurden (z. B. die Bewertung schriftlicher Unterrichtsreflexionen von u. a. Mientus (2023, S. 31ff.); Nowak (2023, S. 44ff.) und Kobl (2021, S. 109ff.)), denen zusätzlich zur abweichenden Erhebungsmethode andere theoretische Modellierungen zugrunde liegen und daher ebenfalls nicht als Paralleltestformen verstanden werden können. Instrumente zur Erfassung der Reflexionsfähigkeit mit geschlossenem Antwortformat sind der Autorin nicht bekannt.

Die Gütekriterien der Ökonomie und Nützlichkeit sind als erfolgreich berücksichtigt anzusehen, da sie Vorbedingungen für die Entwicklung des Instruments darstellten: Ziel war die Entwicklung eines Testinstruments, das - auch im Vergleich zum Reflexionsperformanztest von Kempin, Kulgemeyer und Schecker (2018) - einen ökonomischen Einsatz sowie insbesondere eine ökonomische Auswertung ermöglicht, um Teilnehmenden eine fundierte Selbsteinschätzung auf Grundlage des zugehörigen Assessment-Feedbacks zu ermöglichen. Auch, wenn einige Teilnehmende den Reflexionstest als (zu) umfangreich bewerteten (s. Kapitel 18.1.2), ist die Zumutbarkeit des Testinstruments als annehmbar einzuschätzen: Mit einer durchschnittlichen Bearbeitungsdauer von etwa einer Stunde (etwa 65 min mit einer Standardabweichung von etwa 28 min) erscheint der zeitliche Aufwand ähnlich zum zuvor eingesetzten, auf 70 Minuten Bearbeitungszeit begrenzten

Reflexionsperformanztest; möglicherweise ist er sogar weiter reduzierbar, da der Test bislang lediglich mit einer zeitlichen Empfehlung, aber ohne Zeitbegrenzung eingesetzt wurde. Zu prüfen bliebe, ob eine im Umfang reduzierte Variante des Reflexionstests die Reflexionsfähigkeit weiterhin mit ausreichender Genauigkeit und in ausreichender Breite erfassen würde, um die zeitliche Belastung der Teilnehmenden durch die Testdurchführung zu reduzieren.

Das Kriterium der Unverfälschbarkeit soll sicherstellen, dass die Testwerte nicht bewusst beeinflusst werden. Die eingesetzte Bepunktungsvorschrift ermöglicht grundsätzlich die (bewusste) Verfälschung der Testwerte: Die Auswahl jeweils aller Antwortoptionen führt zum maximalen Testwert, was zunächst problematisch erscheinen mag. Aufgrund des intendierten Einsatzes des Reflexionstests mit dem Ziel einer Rückmeldung über die individuellen Fähigkeiten und der anonymen Durchführung des Tests besteht für die Teilnehmenden allerdings kein Bedarf für die bewusste Verfälschung des eigenen Ergebnisses, sodass diese Einschränkung des Gütekriteriums den sinnhaften Einsatz des Reflexionstests nicht grundsätzlich infrage stellt.

Vier weitere klassische Nebengütekriterien werden unmittelbar im Rahmen der Validierung adressiert:

- Das Kriterium der Transparenz, d. h. der Nachvollziehbarkeit der Instruktionen, ist Teil der kognitiven Validierung im Rahmen des zweiten Teils des INA und aufgrund dessen Bestätigung (und nicht auftretender Unklarheiten bezüglich der Testdurchführung) als erfüllt anzunehmen; die mit dem Einsatz des Reflexionstests verbundenen Ziele sind für die Teilnehmenden auf der Startseite des Instruments dargestellt.
- Auch die Akzeptanz des Testinstruments und -verfahrens wird vorrangig im Rahmen der kognitiven Validität (Teil 2 des INA) adressiert und im Rahmen des vierten Teils des INA um die Perspektiven von Studierenden auf das Assessment-Feedback ergänzt, wobei die Ergebnisse insgesamt positive Schlussfolgerungen bzw. der Akzeptanz zulassen.
- Die Skalierung umfasst die angemessene Bildung von Testwerten und ist Gegenstand der Untersuchungen zum dritten Teil des INA. Im Rahmen der statistischen Auswertungen zum Indikator 3.5 *„Unterschiedliche Lerngelegenheiten der Teilnehmenden führen zu nachvollziehbaren Unterschieden in den Testwerten.“* werden auch Analysen durchgeführt, die Aufschluss über die (möglicherweise mangelnde) Fairness des Instruments liefern können, wenn konstrukt(ir)relevante Varianz auftritt, wie z. B. in Bezug auf die festgestellten signifikanten Unterschiede in den Testwerten nach Geschlecht der Teilnehmenden. Hier besteht Bedarf für weitere Untersuchungen.
- Die Generalisierbarkeit wird als einer der von Messick (1995) formulierten Validitätsaspekte an verschiedenen Stellen für unterschiedliche Stichproben

diskutiert; u. a. im Rahmen der Itemanalyse im dritten Teil des INA (Kapitel 17.1 und 17.2) für Bachelor- und Masterstudierende sowie in Kapitel 17.4 für Studierende des Lehramts an Grund- bzw. weiterführenden Schulen.

19.6 Fazit

Im Rahmen der vorliegenden Arbeit wurde das Vorgehen der Entwicklung des Reflexionstests und des zugehörigen Assessment-Feedbacks systematisch evaluiert, wobei die Entwicklung einer Validitätsargumentation für die Interpretation der Testwerte des Reflexionstests (a) als Maß für die Reflexionsfähigkeit und (b) als Ausgangspunkt für einen weiteren Professionalisierungsprozess im Fokus stand. Die zur Evaluation der Bestandteile des INA durchgeführten Untersuchungen bieten eine vergleichsweise ausführliche Basis für die transparente Beurteilung der Gültigkeit dieser Testwertinterpretation auf Grundlage des Prozesses von der Testentwicklung bis hin zur Interpretation des Assessment-Feedbacks durch die Teilnehmenden; insbesondere sind alle Validitätsaspekte nach Messick (1995) in die Validitätsargumentation einbezogen. Auf Grundlage der in den Kapiteln 19.1 bis 19.4 zusammengefassten Validitätsargumentation wird das INA mit einigen Einschränkungen als gültig angenommen. Eine zusammenfassende Übersicht über die Bewertung der Indikatoren sowie darauf aufbauend die eingeschränkte Formulierung des INA enthält Abbildung 19.1. Auch aus den übrigen diskutierten Gütekriterien (s. Kapitel 19.5) sind keine unmittelbaren Einschränkungen für die Testwertinterpretation bzw. den Einsatz der Lernumgebung zu berücksichtigen. Für die Beantwortung der Forschungsfrage 1 kann daher abschließend für die beiden Teile der Testwertinterpretation festgehalten werden:

- Für die valide Interpretation der Testwerte des Reflexionstests als (a) Maß für die Reflexionsfähigkeit liegt im Wesentlichen keine Gefährdung vor. Durch die Beschränkung der Handlungsfreiheit durch MC-Aufgaben ist der Reflexionstest allerdings als Kompetenztest nach Miller (1990) einzustufen. Näherer Untersuchungsbedarf besteht bezüglich des Zusammenhangs zum Professionswissen, insbesondere dem fachdidaktischen Wissen, da hier (s. Kapitel 17.6) im Rahmen der externen Validierung keine erwarteten Zusammenhänge festgestellt werden konnten. Übrige Limitationen liegen in der Aussagekraft einzelner Studien (z. B. aufgrund einer geringen Stichprobengröße) begründet, stellen aber keine Gefahr für die valide Testwertinterpretation dar und werden vor dem Hintergrund der insgesamt breiten Validitätsargumentation für diese Arbeit akzeptiert.
- Die Interpretation der Testwerte des Reflexionstests als (b) Ausgangspunkt für einen weiteren Professionalisierungsprozess ist mit einer zentralen Limitation verbunden: Zwar konnte im Rahmen der Untersuchungen zum vierten Teil des INA Evidenz dafür gesammelt werden, dass das Assessment-

Feedback eine geeignete Form der Rückmeldung der Testwerte in Bezug auf die angemessene Interpretation und Bewertung der erzielten Ergebnisse durch die Teilnehmenden darstellt. Damit ermöglichen die erhaltenen Testwerte z. B. die Einschätzung in Bezug auf welche Teilfähigkeiten (verstärktes) Entwicklungspotenzial besteht. Gleichzeitig ist nicht davon auszugehen, dass basierend auf dem Assessment-Feedback zur Testbearbeitung tatsächlich individuelle Professionalisierungsprozesse initiiert werden, da weder davon ausgegangen werden kann, dass alle Teilnehmenden das Assessment-Feedback lesen bzw. das inkludierte Selbstlernmaterial bearbeiten, noch, dass selbständig andere Maßnahmen zur Entwicklung ergriffen werden (vgl. Kapitel 18.1). Dies könnte auf einen Bedarf zur stärker angeleiteten Förderung der Reflexionsfähigkeit hinweisen.

19.7 Ausblick

In dieser Arbeit wurde - auf Grundlage der jeweils vorliegenden und vorbehaltlich der z. T. eher schwachen Datenbasis - argumentiert, dass die Testwerte des Reflexionstests valide als Maß für Reflexionsfähigkeit verstanden werden können und so auch durch die Teilnehmenden verstanden werden. Gleichzeitig liegen gewisse, in den Kapiteln 19.1 bis 19.4 diskutierte, Limitationen zu den durchgeführten Studien vor, die zu gewissen Einschränkungen in der Testwertinterpretation führen. Im Sinne der nach Kane (1992) nie abschließbaren Validitätsuntersuchungen bieten sich als ergänzende Untersuchungen die folgenden zur externen Validierung und Untersuchung der Generalisierbarkeit an:

- Die Durchführung von Doppelstudien zu Zusammenhängen zwischen Testwerten im Reflexionstest und den Bestandteilen der professionellen Handlungskompetenz (den Facetten des Professionswissens sowie der affektiv-motivationalen Bestandteile wie z. B. die Reflexionsbereitschaft), hierbei insbesondere auch zum fachdidaktischen Wissen, da die in dieser Arbeit vorgestellten Befunde möglicherweise durch die kleine Stichprobengröße beeinflusst sind.[185]
- Die Durchführung umfangreicherer Untersuchungen bezüglich der Zusammenhänge zwischen den Testwerten im Reflexionstest und demografischen Variablen, um einen präziseren Einblick auch in (möglicherweise ursächliche) Zusammenhänge z. B. zwischen den Testwerten und der Erfahrung als Vertretungslehrkraft zu erhalten.

[185] Bedeutsam erscheint außerdem die Untersuchung von Zusammenhängen zwischen der Reflexionsfähigkeit und realem Handeln von Lehrkräften (z. B. erfasst über die Unterrichtsqualität), da hierzu bislang kaum empirische Befunde vorliegen. Gleichzeitig kann solch eine Untersuchung aus ebendiesem Grund nicht als Beitrag zur Validierung bewertet werden, sondern dient der Gewinnung empirischer Erkenntnisse zum vielfach postulierten positiven Beitrag der Reflexionsfähigkeit auf die Unterrichtsqualität.

- Die Durchführung von Doppelstudien zur Erweiterung des Vergleichs unterschiedlicher Instrumente zur Erfassung der Reflexionsfähigkeit über den Vergleich zum inhaltlich ähnlichen Performanztest hinaus; ggf. können hier authentischere Formate für Reflexionen (z. B. (schriftliche) Reflexionen von (hospitiertem) Unterricht im Schulpraktikum als Handlungstests nach Miller (1990)) herangezogen werden. Aufgrund der Vielzahl der Modellierungen des Reflexionsbegriffs kommt hier der Auswahl des Bewertungsschemas besondere Bedeutung zu. Zusätzlich könnten Einblicke in ggf. variierende kognitive Prozesse für eine differenzierte Einschätzung der erfassten Fähigkeit generiert werden (vgl. Eßling, Sunder, Steffensky und Meschede, 2024).
- Die Untersuchung von (expliziten) Lerngelegenheiten für Unterrichtsreflexion und deren Einfluss auf die Testwerte im Reflexionstest im Längsschnitt: Liegen z. B. qualitative Unterschiede für theoretische (z. B. universitäre Veranstaltungen) und praktische Lerngelegenheiten (z. B. begleitete Praktika) vor? Dazu bliebe allerdings vorab zu prüfen, ob der Reflexionstest Unterschiede in der Reflexionsfähigkeit ausreichend empfindlich abbildet; erste Hinweise darauf liefern die beobachteten Entwicklungen in den Testwerten im Verlaufe eines Semesters (s. Kapitel 17.5).
- Die Entwicklung weiterer Testszenarien zu weiteren Fachinhalten ggf. auch außerhalb der Mechanik zur Untersuchung der Übertragbarkeit der Ergebnisse.
- Die Erweiterung der Stichprobe um Referendar*innen und ggf. erfahrene Lehrkräfte zur Untersuchung der Generalisierbarkeit der Ergebnisse auf die verschiedenen Ausbildungsphasen und -hintergründe (z. B. Zugangswege zum Referendariat).
- Die umfangreichere Untersuchung der Arbeit mit dem Selbstlernmaterial über Selbstauskünfte hinaus, um weitere Erkenntnisse bezüglich der Eignung des Materials für die Teilnehmenden und als Feed Forward zu erhalten. Hier könnten ggf. weiterführende Erkenntnisse bezüglich der Interpretation der Testwerte als Ausgangspunkt für einen weiteren Professionalisierungsprozess der Teilnehmenden erlangt werden.

Weiterhin erscheint die Frage relevant, welcher Zusammenhang zwischen den Testwerten im Reflexionstest und der Fähigkeit zur Selbstreflexion besteht. Diese ist notwendig für die individuelle Weiterentwicklung von (angehenden) Lehrkräften auf Grundlage ihres eigenen Unterrichts. Dieser Frage könnte im Rahmen von Untersuchungen zur prädiktiven Validität nachgegangen werden, betrifft aber darüber hinaus auch den bislang nicht abschließend geklärten Zusammenhangs zwischen den Fähigkeiten zur Selbst- und Fremdreflexion.

Über weiterführende Untersuchungen zur validen Testwertinterpretation hinaus wird die fortlaufende Nutzung der entwickelten Lernumgebung in den Blick genom-

men. Auch wenn diese grundsätzlich selbständig und unabhängig von (universitären) Veranstaltungen bearbeitet werden kann, ist auf Grundlage der Erfahrungen während der Erprobungsphase sowie der Beobachtung, dass bereits das Assessment-Feedback nach Bearbeitung des Reflexionstests nicht selbstständig (vollständig) wahrgenommen und bearbeitet wird, davon auszugehen, dass extern initiierte Bearbeitungsanlässe z. B. gekoppelt an Veranstaltungen sowohl für die Durchführung des Reflexionstests als auch die Auseinandersetzung mit dem Selbstlernmaterial geschaffen werden sollten, um angehende Lehrkräfte an die Lernumgebung heranzuführen. Möglicherweise könnte für die zeitlich effizientere Erfassung der Reflexionsfähigkeit eine Kurzversion des Reflexionstests zum Einsatz kommen. Hier ist allerdings der Einfluss von Kürzungen auf die inhaltliche Breite des dargestellten Unterrichts und somit die inhaltliche Validität zu berücksichtigen.

Perspektivisch könnte außerdem die Weiterentwicklung der Rückmeldung zur präziseren Einordnung der individuellen Testwerte im Sinne des Feed Back (Hattie & Timperley, 2007) angestrebt werden: Während ein normativer Bewertungsmaßstab für Unterrichtsreflexion weiterhin offen bleibt und auf breiterem Konsens als dem hier vorstellten Reflexionstest aufbauen sollte, könnte z. B. eine Erweiterung des Datensatzes angestrebt werden, die die Verortung nicht nur in der Gesamtstichprobe, sondern in einer Stichprobe der aktuellen Ausbildungsphase (Bachelor-, Masterstudium oder Referendariat) ermöglicht oder sogar (empirisch fundierte) Entwicklungsstufen der Reflexionsfähigkeit aufgreift. Erste Untersuchungen zu Profilen der Reflexionsfähigkeiten werden nachfolgend im Kapitel 21 vorgestellt.

Abschließend sei an dieser Stelle darauf hingewiesen, dass die hier durchgeführte Validierung eng mit der diskutierten Testwertinterpretation verbunden ist. Der Einsatz des Reflexionstests zu anderen Zwecken bedarf einer jeweils angepassten Validitätsargumentation, für die die hier vorgestellten Untersuchungen zwar eine Grundlage bilden (können), aber nicht pauschal als ausreichend angenommen werden dürfen. Dies gilt insbesondere dann, wenn auf Grundlage der erreichten Testwerte Schlüsse gezogen werden, die unmittelbare Konsequenzen für die Teilnehmenden haben (z. B. in Form von Benotungen von Studierenden). Hierfür bedarf es einer zusätzlichen Absicherung der Ergebnisse mit weiteren Studien bzw. größeren Stichproben mit einem Fokus auf die konsequentielle Validität.

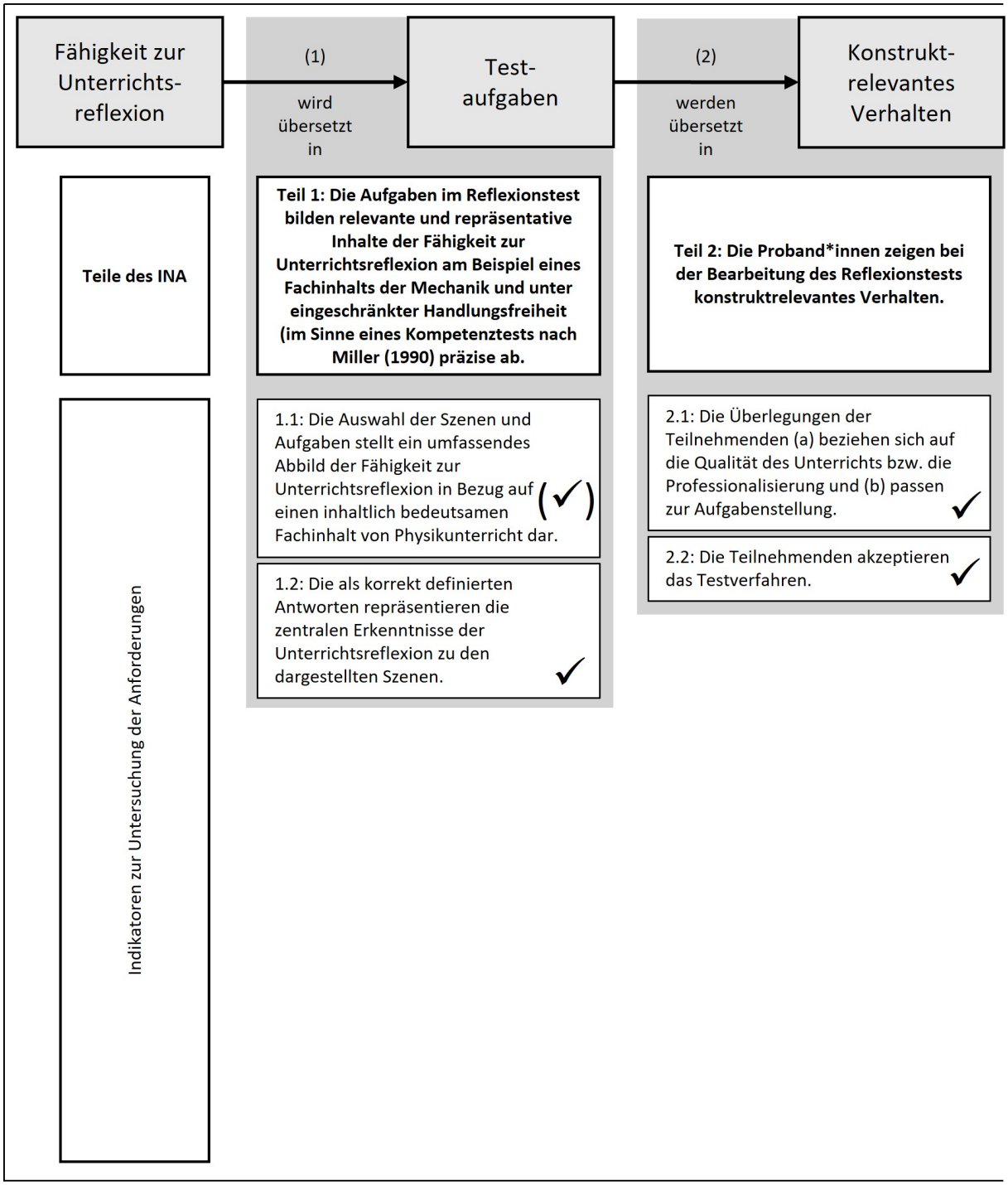
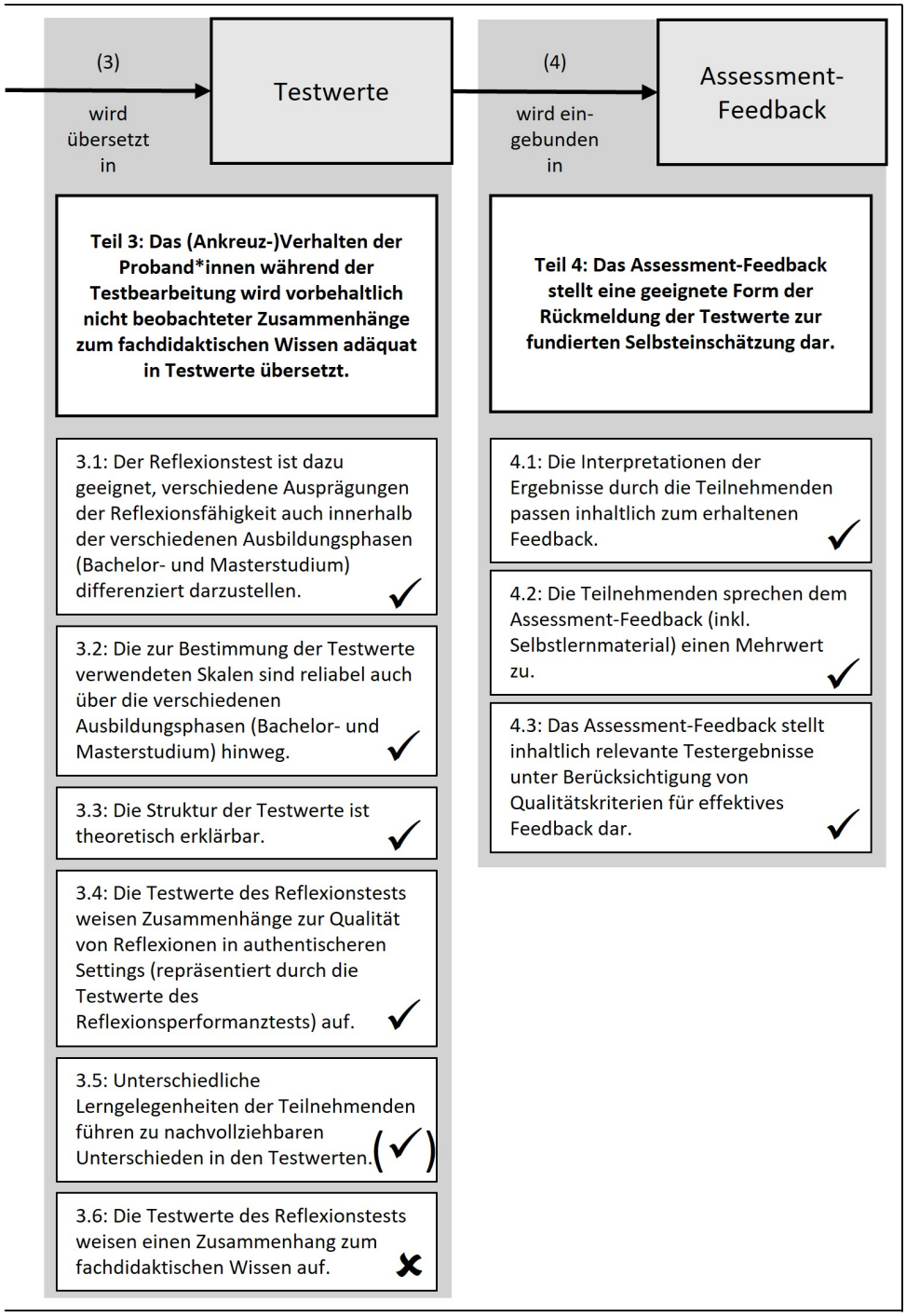


Abbildung 19.1: Darstellung des INA für die Interpretation der Testwerte des Reflexionstests mit Assessment-Feedback, ergänzt um die Bewertung der jeweils überprüften Indikatoren auf Grundlage der durchgeführten Validierungsstudien


(3)
wird übersetzt in
Testwerte
(4)
wird eingebunden in
Assessment-Feedback
Teil 3: Das (Ankreuz-)Verhalten der Proband*innen während der Testbearbeitung wird vorbehaltlich nicht beobachteter Zusammenhänge zum fachdidaktischen Wissen adäquat in Testwerte übersetzt.
Teil 4: Das Assessment-Feedback stellt eine geeignete Form der Rückmeldung der Testwerte zur fundierten Selbsteinschätzung dar.
3.1: Der Reflexionstest ist dazu geeignet, verschiedene Ausprägungen der Reflexionsfähigkeit auch innerhalb der verschiedenen Ausbildungsphasen (Bachelor- und Masterstudium) differenziert darzustellen. ✓
3.2: Die zur Bestimmung der Testwerte verwendeten Skalen sind reliabel auch über die verschiedenen Ausbildungsphasen (Bachelor- und Masterstudium) hinweg. ✓
3.3: Die Struktur der Testwerte ist theoretisch erklärbar. ✓
3.4: Die Testwerte des Reflexionstests weisen Zusammenhänge zur Qualität von Reflexionen in authentischeren Settings (repräsentiert durch die Testwerte des Reflexionsperformanztests) auf. ✓
3.5: Unterschiedliche Lerngelegenheiten der Teilnehmenden führen zu nachvollziehbaren Unterschieden in den Testwerten. (✓)
3.6: Die Testwerte des Reflexionstests weisen einen Zusammenhang zum fachdidaktischen Wissen auf. ✗
4.1: Die Interpretationen der Ergebnisse durch die Teilnehmenden passen inhaltlich zum erhaltenen Feedback. ✓
4.2: Die Teilnehmenden sprechen dem Assessment-Feedback (inkl. Selbstlernmaterial) einen Mehrwert zu. ✓
4.3: Das Assessment-Feedback stellt inhaltlich relevante Testergebnisse unter Berücksichtigung von Qualitätskriterien für effektives Feedback dar. ✓

Teil IV

Untersuchungen zur Reflexionsfähigkeit und ihrer Entwicklung

20 Ziele und Forschungsfragen der Untersuchungen

Wie in Kapitel 7 ausführlich dargelegt, wird Unterrichtsreflexion einhellig als zentral für die Lehrkräfte, ihren Unterricht und ihre persönliche Entwicklung angesehen. Gleichzeitig ist die Fähigkeit zur Unterrichtsreflexion bei (angehenden) Lehrkräften oft nur gering ausgeprägt. Unklar ist insbesondere, wie sich die Reflexionsfähigkeit auch im Verlauf des Studiums entwickelt und - obwohl verschiedene Ansätze zur Förderung berichtet und teilweise evaluiert wurden - weitgehend auch, welche Unterstützungsmaßnahmen dazu besonders geeignet sind (z. B. Reintjes und Kunze, 2022). Auch der Zusammenhang zwischen der Unterrichtsreflexion und ihrer Entwicklung und der - zumindest theoretisch eng mit ihr verknüpften - Fähigkeit zum Noticing (s. Kapitel 7.6) ist kaum untersucht.

Ziel dieses Abschnitts ist daher die weitere Exploration der Struktur der Reflexionsfähigkeit, ihrer Entwicklung sowie von Zusammenhängen zum Noticing auf Grundlage der mit dem Reflexionstest gewonnenen Daten. Zugrunde liegen diesen Überlegungen folgende Annahmen zum Zusammenhang zwischen Professionswissen, Noticing und Unterrichtsreflexion:

- Noticing, die „wissensgesteuerte Identifikation relevanter Situationen und Ereignisse im Unterrichtsgeschehen“ (Seidel, Blomberg & Stürmer, 2010, S. 297), wird - wie bereits in Kapitel 7.6 beschrieben - aus theoretischer Perspektive als eine Voraussetzung für (bedeutsame) Unterrichtsreflexion verstanden: Nur, wenn (auch) die relevanten Aspekte des Unterrichts erkannt (und anschließend reflektiert) werden, kann davon ausgegangen werden, dass die Umsetzung darauf basierender Handlungen zu einer relevanten Verbesserung der Unterrichtsqualität bzw. einer individuellen Weiterentwicklung führt.
- (Beobachtbares) Noticing und (beobachtbare) Reflexion sind Ausdruck der zugrundeliegenden Fähigkeiten (Blömeke, Gustafsson & Shavelson, 2015).
- (Beobachtbares) Noticing und (beobachtbare) Reflexion bzw. die zugrundeliegenden Fähigkeiten sind durch gewisse Dispositionen - wie das Professi-

onswissen - geprägt (z. B. Blömeke, Gustafsson und Shavelson, 2015; Schäfer und Seidel, 2015; Weber, Neuber und Prilop, 2023).[186]

Zusammengefasst sind diese Zusammenhänge als heuristisches Modell in Abbildung 20.1[187] dargestellt. Die in den folgenden Kapiteln beschriebenen Untersuchungen orientieren sich an den Bestandteilen des Modells. Während in Abbildung 20.1 die Zusammenhänge lediglich unidirektional dargestellt sind (Pfeile 1, 2 und 3), können evidenzbasiert weitere Einflüsse von Reflexionsfähigkeit auf Noticingfähigkeit (ggf. mediiert über das Professionswissen) sowie von beiden Fähigkeiten auf das Professionswissen angenommen werden[188], die hier allerdings gemäß dem Fokus auf die Förderung der Reflexionsfähigkeit nicht weiter untersucht werden.

Auf Grundlage der in diesem Teil der Arbeit berichteten Analysen soll das Modell zum Zusammenhang von Noticing und Unterrichtsreflexion explorativ erweitert werden, um zur Fortentwicklung der Theorie in diesem Bereich beizutragen. Dazu wird das vorliegende Modell in den folgenden Kapiteln schrittweise auf Grundlage der Ergebnisse ergänzt, sodass in Kapitel 24 mit einem Modell abgeschlossen wird, das Hypothesen über die Zusammenhänge und zugrundeliegenden Mechanismen enthält, die einen Ausgangspunkt für eine empirisch fundierte, zielgerichtete Förderung der Reflexionsfähigkeit bilden können. Die entwickelten Hypothesen und vermuteten Zusammenhänge können dann im Rahmen von Folgestudien untersucht und validiert werden.

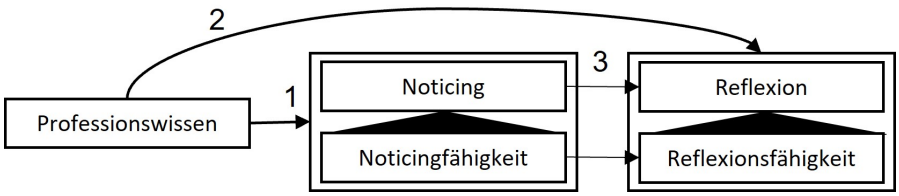


Abbildung 20.1: Basismodell zum Zusammenhang von Noticing und Unterrichtsreflexion

[186] Vogelsang, Kulgemeyer und Riese (2022) stellen zumindest Zusammenhänge zwischen Facetten des professionellen Wissens und der Reflexionsfähigkeit fest.

[187] Optisch weist das Modell Ähnlichkeit zum Kompetenzmodell von Blömeke, Gustafsson und Shavelson (2015) auf, in dem die Fähigkeit zum Noticing als situationsspezifische Fähigkeit und als zwischen den Dispositionen und beobachtbarer Performanz mediierend verstanden wird; hier wird sie hier allerdings als eigenständige Fähigkeit betrachtet.

[188] Passend dazu liegen Befunde dazu vor, dass die Reflexionsfähigkeit Einfluss auf die Entwicklung des fachlichen und fachdidaktischen Wissens im Verlauf des Praxissemesters nimmt (Kulgemeyer et al., 2021).

Konkret wird zunächst untersucht, ob sich latente Profile der Reflexionsfähigkeit identifizieren lassen (Kapitel 21). Die Identifikation solcher Profile könnte einerseits Rückschlüsse auf verschiedene Ausprägungen der Reflexionsfähigkeit zulassen und andererseits insbesondere einen Ausgangspunkt für eine individuell besser zugeschnittene Förderung der Reflexionsfähigkeit ermöglichen, wenn z. B. zwischen den Profilen variierende Förderbedarfe festgestellt werden. Hierzu dient Forschungsfrage 2:

> **Forschungsfrage 2:** Welche latenten Profile der Reflexionsfähigkeit lassen sich auf Grundlage der Bearbeitungen des Reflexionstests identifizieren?

Anschließend werden Zusammenhänge der Fähigkeiten zur Reflexion und zum Noticing sowie der Entwicklung beider Fähigkeiten analysiert, um das theoriebasiert angenommene enge Verhältnis beider Fähigkeiten empirisch zu untersuchen.

Die Untersuchung der Zusammenhänge zwischen beiden Fähigkeiten erfolgt in Kapitel 22 auf Grundlage der (theoretisch begründeten) Annahme, dass Noticing eine Voraussetzung für bedeutsame Reflexion darstellt (s. Kapitel 7.6) anhand Forschungsfrage 3:

> **Forschungsfrage 3:** Wie hängen die Fähigkeiten zum Noticing und zur Unterrichtsreflexion zusammen?

Zur darauf aufbauenden Untersuchung von Entwicklungszusammenhängen zwischen beiden Fähigkeiten wird in Kapitel 23 Forschungsfrage 4 betrachtet:

> **Forschungsfrage 4:** Wie beeinflussen die Fähigkeiten zum Noticing und zur Unterrichtsreflexion wechselseitig ihre Entwicklung?

Hypothesen über vorliegende Entwicklungszusammenhänge lassen sich nicht eindeutig aus theoretischen Überlegungen oder bisherigen Befunden ableiten. Vorstellbar ist ein Einfluss der Noticingfähigkeit auf die Entwicklung der Reflexionsfähigkeit da das Noticing als Voraussetzung für Reflexion anzunehmen ist und Gelegenheiten zur Reflexion auf gelungenes Noticing aufbauen. Andersherum wird Reflexion als Mechanismus zur Entwicklung von Wissen aus Erfahrung verstanden (s. Kapitel 7.2; empirische Hinweise auf diesen Zusammenhang beschreiben Kulgemeyer et al. (2021)), das wiederum als Grundlage für gelingendes Noticing fungieren kann.

Für die folgenden Analysen wird auf die vorliegenden Daten aus den in Teil III berichteten Validierungsstudien zurückgegriffen. Als Gesamtstichprobe liegen $N = 189$ Bearbeitungen des Reflexionstests von (angehenden) Physik-Lehrkräften

vor (s. Kapitel 14.1; drei Bearbeitungen wurden aufgrund des unklaren Ausbildungshintergrunds (Angabe „Sonstiges“) ausgeschlossen). Zusätzlich wird die vorliegende Teilstichprobe von $N = 41$ Physik-Lehramtsstudierenden betrachtet, die bereits im Rahmen der Untersuchung von Indikator 3.5 herangezogen wurde (s. Abbildung 14.2 auf Seite 188). Sie setzt sich zusammen aus $N = 15$ Bachelorstudierenden, die eine physikdidaktische Grundlagenveranstaltung besuchten, und $N = 26$ Masterstudierenden im Praxissemester.

21 Analyse latenter Profile von Reflexionsfähigkeit

In diesem Kapitel werden die zur Unterrichtsreflexion gewonnenen Daten näher betrachtet. Die Untersuchungen sind dem Bereich der Reflexion bzw. Reflexionsfähigkeit im Modell zum Zusammenhang von Noticing und Unterrichtsreflexion zuzuordnen.

Für die Förderung der Reflexionsfähigkeit ist es erstrebenswert, den Teilnehmenden ein möglichst präzises Assessment-Feedback zur Verfügung zu stellen. Eine Möglichkeit zur näheren inhaltlichen Beschreibung der (individuellen) Fähigkeiten ist die Zuordnung der Personen zu Fähigkeitsprofilen, die qualitative und bzw. oder quantitative Fähigkeitsunterschiede zwischen den Profilen aufzeigen können (Blömeke, Kaiser, König & Jentsch, 2020). Auf Grundlage solcher Profile könnten Erkenntnisse über die Ausbildung der Reflexionsfähigkeit gewonnen und in einem nächsten Schritt z. B. gezielte Fördermaßnahmen entwickelt werden, um ein ausführlicheres Feed Up und Feed Forward (Hattie & Timperley, 2007) zu ermöglichen. Auch für die folgenden Untersuchungen von Zusammenhängen, z. B. zur Fähigkeit zum Noticing, könnte eine Betrachtung entlang der Profile detailliertere Einblicke ermöglichen. Dazu wird Forschungsfrage 2 untersucht:

Forschungsfrage 2: Welche latenten Profile der Reflexionsfähigkeit lassen sich auf Grundlage der Bearbeitungen des Reflexionstests identifizieren?

21.1 Latente Profilanalyse

Die latente Profilanalyse (LPA) wird als probabilistisches Verfahren zur Identifikation latenter Gruppen in vorliegenden Datensätzen eingesetzt (Bauer, 2022). Der LPA liegt die Annahme zugrunde, dass Personen den verschiedenen Gruppen mit gewisser Wahrscheinlichkeit zugeordnet werden können und die Zuordnung durch Muster in den erreichten Aufgabenscores erklärt werden kann (Spurk et al., 2020).

Grundlage für die hier durchgeführte LPA sind die $N = 189$ vorliegenden Bearbeitungen des Reflexionstests von (angehenden) Physik- bzw. Sachunterrichtslehrkräften mit Wahlfach Physik. In die Modellberechnung gehen als Variablen die 32

Aufgaben des Reflexionstests ein, die sich auf jeweils 16 Aufgaben zu Bewertungen und zu Alternativen aufteilen. Durchgeführt wurde die LPA mit dem *R*-Paket *tidyLPA* (Rosenberg et al., 2018).

21.2 Ergebnisse der latenten Profilanalyse

Es wurden Modelle mit bis zu acht Profilen berechnet.[189] Die Obergrenze von acht Profilen wurde zunächst gewählt, um eine Zerfaserung in zu kleine Profile zu vermeiden (bei einer Gleichverteilung wären auch bei acht Profilen immer noch über zwanzig Personen in einem Profil); klare Empfehlungen zur minimalen Stichprobengröße liegen nicht vor (Weller, Bowen & Faubert, 2020).

Für alle Modelle mit fünf oder mehr Profilen zeigt sich allerdings eine Zuordnung zu den Profilen, in der jeweils mindestens ein Profil mit nur einer Person besetzt ist. Nachfolgend werden also lediglich die Modelle mit maximal vier Profilen berücksichtigt. Zur Bestimmung des am besten passenden Modells werden die Modellgüteparameter AIC, BIC, sample-size adjusted BIC (SABIC), Entropy und BLRT p-Wert berücksichtigt, die die Güte des Fits vor dem Hintergrund der Stichprobengröße angeben (AIC, BIC, SABIC), ein Maß für die Trennschärfe zwischen den Profilen darstellen (Entropy) sowie Auskunft darüber geben, ob das jeweilige Modell mit k Profilen signifikant besser ist als das mit $k - 1$ Profilen (BLRT p-Wert) (Spurk et al., 2020). Tabelle 21.1 bietet eine Übersicht über die Parameter für die Modelle mit ein bis vier Profilen. Die Auswahl dieser Parameter, insbesondere die parallele Betrachtung von AIC und BIC, beruht auf Empfehlungen für Stichproben mit unter 300 Fällen (Spurk et al., 2020).[190] Mit Ausnahme der Entropy, die im Vergleich zu den anderen Modellen einen eher niedrigen Wert aufweist (der allerdings trotzdem deutlich über der Grenze für akzeptable Werte von $0,8$ liegt (Weller, Bowen & Faubert, 2020)), weisen alle Parameter darauf hin, dass das beste Modell jenes mit vier Profilen ist. Basierend auf den Empfehlungen von Akogul und Erisoglu (2017), die eine analytische Hierarchisierung verschiedener Modellgüteparameter (darunter auch AIB und BIC) vornehmen, um die optimale Zahl an Profilen festzustellen, würde das Modell mit zwei Profilen als das beste eingeschätzt werden. Dafür würden auch die bestimmte Werte der Entropy sprechen. Das Modell mit zwei Profilen differenziert zwischen einem eher kleinen unteren und einem großen oberen Profil. Das obere Profil zeichnet sich dadurch aus, dass ihm eher weiter fortgeschrittene angehende Lehrkräfte sowie

[189] Das *R*-Skript zur latenten Profilanalyse und den nachfolgen statistischen Analysen befindet sich im Anhang in Kapitel S ab Seite 923.

[190] Einige der Parameter erscheinen nicht robust bezüglich nicht normalverteilter Variablen, die in der LPA berücksichtigt werden (Spurk et al., 2020). Vor dem Hintergrund des Einsatzes der LPA für die Hypothesengenerierung werden trotzdem alle 32 Aufgaben des Reflexionstests in die LPA einbezogen.

Anzahl der Profile	AIC	BIC	SABIC	Entropy	BLRT p-Wert
	Möglichst niedrig	*Möglichst niedrig*	*Möglichst niedrig*	*Möglichst dicht an 1*	*< 0,01 (besseres Modell als vorherige Zeile)*
1	686,86	894,32	691,60	1	0,01
2	-866,50	-552,05	-859,30	0,99	0,01
3	-1472,20	-1050,77	-1462,55	0,96	0,01
4	-1613,20	-1084,80	-1601,11	0,93	0,01

Tabelle 21.1: Übersicht über die Modellgüteparameter und ihre Interpretation für die verschiedenen Modelle der LPA

alle Referendar*innen angehören. Insgesamt gibt es große Übereinstimmung zwischen den unteren Profilen in den Modellen mit zwei und vier Profilen; das obere Profil der Zwei-Profil-Lösung entspricht überwiegend den oberen drei Profilen der Vier-Profil-Lösung. Damit ermöglicht das Modell mit zwei Profilen insgesamt weniger differenzierte Einblicke, was vor dem Ziel der Hypothesengenerierung ungünstig erscheint.[191] Nachfolgend wird daher das Modell mit vier Profilen näher untersucht.

21.3 Beschreibung der Profile

Eine grafische Übersicht über die vier Profile ist in Abbildung 21.1 dargestellt. Auffällig ist, dass drei der vier Profile näherungsweise parallel verlaufen und auch das vierte kein grundsätzlich entgegengesetztes Verhalten zeigt. Die Profile können global, aber auch bezogen auf alle Aufgaben, als unterschiedlich erfolgreich im Reflexionstest beschrieben werden, die durchschnittlichen Lösungshäufigkeiten steigen dementsprechend an (Profil 1: $0,31$; Profil 2: $0,61$; Profil 3: $0,69$; Profil 4: $0,79$), wobei die Lösungshäufigkeiten im unteren Profil 1 in allen Aufgaben unter der durchschnittlichen Lösungshäufigkeit über alle Profile liegen, die der Profile 3 und 4 jeweils darüber und die Lösungshäufigkeiten im mittleren Profil 2 in etwa der Hälfte der Aufgaben über (17 von 32 Aufgaben) bzw. unter (15 von 32 Aufgaben) der durchschnittlichen Lösungshäufigkeit aller Teilnehmenden. Sowohl im Testwert zur Reflexionsfähigkeit also auch in den Testwerten zu Bewertungen und zu Alternativen ergeben sich gemäß Kruskal-Wallis-Test und Post-Hoc-Analyse mit paarweisen Wilcoxon-Tests mit Bonferroni-Korrektur große signifikante Unterschiede zwischen allen vier Profilen (für die Reflexionsfähigkeit: $H(3) = 167,47$; $p < 0,001$, für die Teilfähigkeit Bewerten: $H(3) = 155,35$;

[191] Ergänzende inhaltlichen Argumentationen zur Auswahl des Modells bzw. Fragen der Interpretierbarkeit werden auch von anderen Autor*innen herangezogen (z. B. Radkowitsch et al., 2023; Weller, Bowen und Faubert, 2020; Spurk et al., 2020).

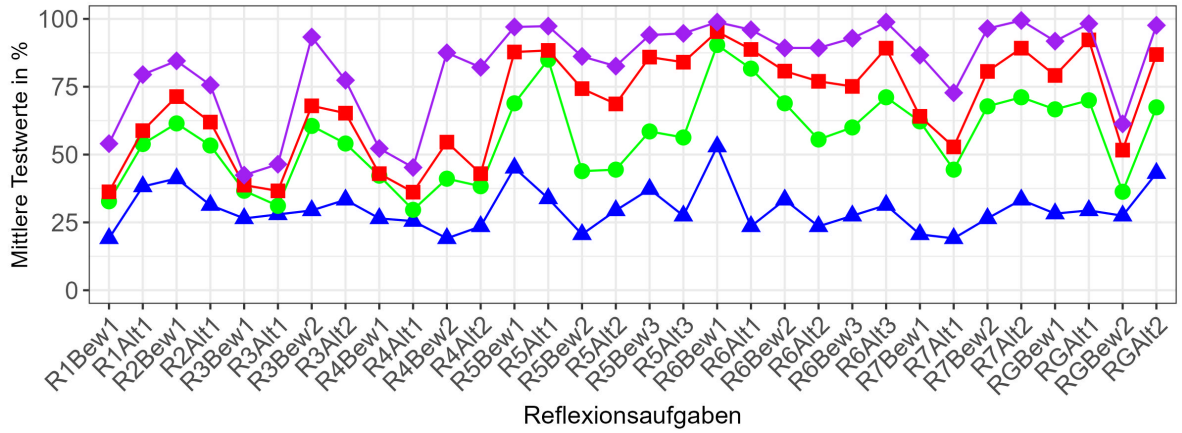


Abbildung 21.1: Latente Profile der Reflexionsfähigkeit (in aufsteigender Reihenfolge: 1 *(blau ▲)*, 2 *(grün ●)*, 3 *(rot ■)*, 4 *(lila ◆)*)

$p < 0,001$ und für die Teilfähigkeit Vorschlagen von Alternativen: $H(3) = 162,42$; $p < 0,001$).[192] Die vollständige Übersicht über die in diesem Kapitel berichteten Tests und Effektstärken befindet sich im Anhang in Kapitel L ab Seite 515.

Auf Aufgabenebene lässt sich ein etwas differenzierteres Bild bezüglich der Unterschiede zwischen den Profilen zeichnen. Während die Profile in den ersten Aufgaben des Testinstruments überwiegend näher beieinander liegen, zeigt sich ab Aufgabe R5Bew1 eine überwiegend größere Spannbreite über die Profile (s. Abbildung 21.1). Lediglich die Profile 1 (das untere Profil) und 4 (das obere Profil) unterscheiden sich in allen Aufgaben signifikant. Die mittleren Profile 2 und 3 unterscheiden sich hingegen nur in etwa der Hälfte der Aufgaben (15 von 32) signifikant voneinander. Auffällig ist die augenscheinlich besonders starke Differenzierung zwischen dem unteren Profil 1 und den oberen drei Profilen in den Aufgaben R5Alt1 und R6Alt1. Diese Aufgaben beziehen sich auf Alternativen zum stereotypischen Verhalten von Robert gegenüber den Schüler*innen. Hier scheinen die Personen im unteren Profil 1 zwar Probleme zu erkennen, wie die etwas höheren Scores in den zugehörigen Bewertungsaufgaben (R5Bew1 und R6Bew1) vermuten lassen, sie sind aber nicht in gleichem Maße in der Lage, angemessene alternative Vorgehensweisen auszuwählen. Für die drei oberen Profile ist dieses Verhältnis zwischen Bewertungen und Alternativen tendenziell umgekehrt (bezüglich der Aufgaben R5Bew1 und R5Alt1) bzw. deutlich abgeschwächt (bezüglich der Aufgaben R6Bew1 und R6Alt1). In den Aufgaben zur Verwendung der Fachsprache bei der Einführung des Impulsbegriffs (R3Bew1 und R3Alt1) sowie zur Wahl

[192] Im Rahmen der Untersuchung von Unterschieden zwischen den Profilen wird in diesem Kapitel 21 durchweg auf nicht parametrische Verfahren zurückgegriffen, da zunächst lediglich explorative Untersuchungen zur Charakterisierung der Profile durchgeführt werden und Voraussetzungen zur Verwendung parametrischer Testverfahren nicht umfassend gegeben sind. Zur Bewertung von Effekten werden weiterhin die Richtwerte nach Cohen (1988; 1992) zur Orientierung herangezogen.

des Freihandexperiments zur Veranschaulichung der Abhängigkeit des Impulses von Masse und Geschwindigkeit (R4Bew1 und R4Alt1) liegen die vier Profile hingegen augenscheinlich besonders dicht beieinander. Alle vier Profile erreichen hier eher niedrige Scores. Diese Aufgaben sind außerdem jene mit den niedrigsten Trennschärfen im Testinstrument (zwischen $0,29$ (R3Bew1) und $0,35$ (R4Bew1)), was insgesamt in der vorliegenden Stichprobe auf einen Bodeneffekt für diese Items hinweisen könnte.

21.4 Interpretation der Profile als Fähigkeitsstufen

Wie im vorigen Abschnitt beschrieben, lassen sich also vier Profile ausmachen, die einheitlich über alle Aufgaben hinweg unterschiedlich hohe Aufgabenscores aufweisen. Sie werden daher als vier Stufen der Reflexionsfähigkeit interpretiert.[193]

Die Interpretation der Profile als Fähigkeitsstufen wird darüber hinaus durch die Beobachtung gestützt, dass Studierende, die den Reflexionstest mehrfach bearbeiteten, tendenziell auf derselben Fähigkeitsstufe verbleiben bzw. auf eine höhere Stufe wechseln: Der Reflexionstest wurde in einigen Kohorten zu zwei MZPs eingesetzt. Die Stichprobe, die das Instrument zweifach bearbeitet hat (insgesamt $N = 41$ Studierende), setzt sich zusammen aus $N = 25$ Studierenden im Praxissemester (Testeinsatz vor bzw. zu Beginn und nach bzw. am Ende des Praxissemesters) und Studierenden in zwei Bachelorveranstaltungen zu Grundlagen der Physikdidaktik, die laut Studienverlaufsplan im dritten bzw. vierten Semester ($N = 12$) bzw. vierten und fünften Semester ($N = 4$) verankert sind. Es ist davon auszugehen, dass die universitären Veranstaltungen bzw. Erfahrungen im Praxissemester, insbesondere die Hospitation von (Physik-)Unterricht sowie gemeinsames Reflektieren mit Betreuenden der Universität (vgl. Kulgemeyer, Kempin und Weißbach, 2021), Lerngelegenheiten zur Entwicklung der Reflexionsfähigkeit darstellen. Da allerdings das tatsächliche Ausmaß der Lerngelegenheiten unbekannt ist und die Förderung der Reflexionsfähigkeit kein explizites Ziel der Veranstaltungen war, sind keine starken Weiterentwicklungen zu erwarten.

Abbildung 21.2 stellt eine Übersicht über die Entwicklung der Zuordnungen zu den Fähigkeitsstufen über beide MZPs dar. Insgesamt sind nur zwei Wechsel in niedrigere Fähigkeitsstufen (innerhalb der mittleren Profile von Profil 3 in 2) beobachtbar, während 13 Personen zum zweiten MZP einer höheren Fähigkeitsstufe zugeordnet werden. Keine Person wechselt von einer anderen in die niedrigste Fähigkeitsstufe 1 und keine Person wechselt aus der höchsten Fähigkeitsstufe 4 in eine niedrigere.

[193] Ähnlich nutzen auch Watermann und Klieme (2006) latente Klassenmodelle zur Prüfung des Kompetenzstufenmodells für den Untertest zur euklidischen Geometrie der TIMSS (*Trends in International Mathematics and Science Study*)/III-Untersuchung.

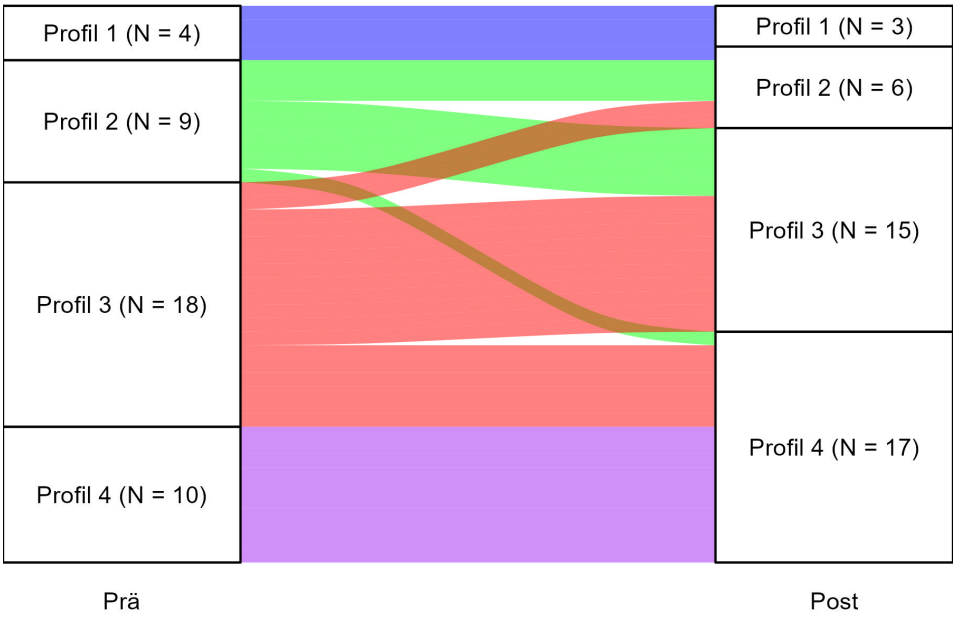


Abbildung 21.2: Entwicklung der Zuweisung der Personen zu den Fähigkeitsstufen über beide Messzeitpunkte

Nachfolgend wird untersucht, inwieweit sich die Profile auf Grundlage inhaltlicher Aufgabenmerkmale sowie demografischer Daten näher charakterisieren lassen.

21.5 Charakterisierung der Fähigkeitsstufen über inhaltliche Merkmale der Aufgaben

Basierend auf dem Reflexionsmodell von Nowak, Kempin, Kulgemeyer und Borowski (2019) kann für verschiedene inhaltliche Merkmale, die den einzelnen Aufgaben zugeordnet werden können, ein Einfluss auf die Zuordnung zu den Fähigkeitsstufen vermutet werden, die nachfolgend betrachtet werden:

1. Das dem Testinstrument zugrundeliegende Reflexionsmodell legt aufgrund der Stufung der Reflexionselemente nahe, dass diese auch als Schwierigkeitsstufen interpretiert werden könnten. Hier könnte vermutet werden, dass Aufgaben zu Bewertungen, da sie das Abbild der niedrigeren Stufe darstellen, grundsätzlich besser gelöst werden als die zu Alternativen.
2. Die in den Aufgaben reflektierten Aspekte können den Wissensbasen des Reflexionsmodells (fachliches, fachdidaktisches und pädagogisches Wissen) zugeordnet werden. Vorstellbar wäre, dass Profile sich besonders in den Aufgaben unterscheiden, die Bezug zu unterschiedlichen Wissensbasen aufweisen.

Fähigkeitsstufen	R1Bew1	R1Alt1	R2Bew1	R2Alt2	R3Bew1	R3Alt1	R3Bew2	R3Alt2	R4Bew1	R4Alt1	R4Bew2	R4Alt2	R5Bew1	R5BAlt1	R5Bew2	R5Alt2
	f/fd		fd		fd		f		f/fd/p		fd/f		fd		f	
1,2				X			**X**	**X**	**X**		X			**X**	X	
2,3															**X**	X
3,4	X	X		X			**X**				**X**	**X**		X	X	X

Fähigkeitsstufen	R5Bew3	R5Alt3	R6Bew1	R6Alt1	R6Bew2	R6Alt2	R6Bew3	R6Alt3	R7Bew1	R7Alt1	R7Bew2	R7Alt2	RGBew1	RGAlt1	RGBew2	RGAlt2
	p		fd		f		f		fd		p		fd/p		f/fd	
1,2	X	X	**X**	**X**	X	X	**X**	**X**	**X**	**X**	**X**	X	**X**	**X**		
2,3	**X**	X				X		X			X	X		**X**	X	X
3,4							X		X	X	X	X	X			X

Tabelle 21.2: Übersicht über vorliegende signifikante Unterschiede mit mittlerer („X“) bzw. großer Effektstärke („**X**“) in den Aufgabenscores benachbarter Fähigkeitsstufen mit Zuordnung der Aufgaben zu den Wissensbasen (f - fachliches, fd - fachdidaktisches und p - pädagogisches Wissen).

Bezüglich des Verhältnisses zwischen den Aufgaben zu Bewertungen und Alternativen sind allerdings weder über alle Fähigkeitsstufen hinweg noch zwischen den einzelnen Fähigkeitsstufen Muster identifizierbar (es kann z. B. nicht festgestellt werden, dass Aufgaben zu Bewertungen oder Alternativen grundsätzlich höhere oder niedrigere Lösungshäufigkeiten aufwiesen oder dass die Fähigkeitsstufen für eine der beiden Aufgabengruppen stärkere Unterschiede aufweisen). Die Interpretation der Reflexionselemente als hierarchische Schwierigkeitsstufen der Unterrichtsreflexion wird also durch die LPA nicht gestützt.

In einem zweiten Schritt wird untersucht, ob eine Charakterisierung der Fähigkeitsstufen entlang der reflektierten Inhalte möglich wird. Da sich die Fähigkeitsstufen in den einzelnen Aufgaben zum Großteil signifikant voneinander unterscheiden, werden zur Annäherung an eine inhaltliche Differenzierung der Fähigkeitsstufen nachfolgend ausschließlich die Unterschiede zwischen benachbarten Fähigkeitsstufen näher betrachtet, die eine mindestens mittlere Effektstärke aufweisen. Eine Übersicht über diese Aufgaben sowie die Zuordnung der Aufgaben zu den drei

Wissensbasen fachliches, fachdidaktisches und pädagogisches Wissen[194] bietet Tabelle 21.2.

Hier zeigt sich, dass sich die unteren beiden Fähigkeitsstufen in etwa einem Drittel (10) der Aufgaben signifikant mit großem Effekt unterscheiden, und die mittleren beiden sowie oberen beiden Fähigkeitsstufen in jeweils nur drei (unterschiedlichen) Aufgaben. Eine Zuordnung dieser stark differenzierenden Aufgaben zu den Wissensbasen offenbart ebenfalls keine deutlichen Muster - möglicherweise auch, da nicht alle Aufgaben eindeutig einer einzelnen Wissensbasis zugeordnet werden können, sondern die enthaltenen Antwortoptionen gemäß dem Kodiermanual für den Reflexionsperformanztest, auf dem diese Zuordnung beruht, verschiedene Wissensbasen ansprechen. Dies führt dazu, dass die in Tabelle 21.2 gekennzeichneten differenzierenden Aufgaben für alle drei Vergleiche der Fähigkeitsstufen eine Auswahl an Antwortoptionen enthalten, die insgesamt alle drei Wissensbasen ansprechen. Möglicherweise sind die Differenzen also auf noch feinere Unterschiede mit Zusammenhang zu den Antwortoptionen in den Aufgaben zurückzuführen, die auf Aufgabenebene nicht rekonstruiert werden können.

21.6 Charakterisierung der Fähigkeitsstufen über demografische Variablen

Zur näheren Charakterisierung der Fähigkeitsstufen werden die erhobenen demografischen Variablen herangezogen. Dabei werden vorrangig die Variablen herangezogen, die sich im Sinne der Validierung als bedeutsam herausgestellt haben, also der Studienfortschritt repräsentiert durch die Ausbildungsphase und das Physik-Fachsemester sowie der Umfang absolvierter Pflichtpraktika (vgl. Kapitel 17.5).

Die Betrachtung der Häufigkeiten der in den Fähigkeitsstufen vertretenen Ausbildungsphasen (s. Tabelle 21.3) verdeutlicht, dass der Anteil der Masterstudierenden gegenüber den Bachelorstudierenden jeweils mit aufsteigender Fähigkeitsstufe zunimmt. Die elf im Datensatz enthaltenen Referendar*innen verteilen sich auf die oberen drei Fähigkeitsstufen. Denkbar wäre, dass hier Unterschiede in Bezug auf die vorige Ausbildung (grundständiges Lehramtsstudium oder Quer- oder Seiteneinstieg) auftreten, dies kann aufgrund fehlender Daten allerdings nicht überprüft werden.[195] Auf Grundlage eines Kruskal-Wallis-Tests zur Bestimmung der Gruppenunterschiede lassen sich signifikante Unterschiede in Bezug auf die Aus-

[194] Die Zuordnung der Aufgaben zu den Wissensbasen basiert auf der Zuordnung der reflektierten Aspekte zu den Wissensbasen im Kodiermanual zum Reflexionsperformanztest von Kempin, Kulgemeyer und Schecker (2018).

[195] Es ist zwar bekannt, dass die Stichprobe der Referendar*innen sich sowohl aus grundständig ausgebildeten Referendar*innen sowie Referendar*innen im Quer- oder Seiteneinstieg zusammensetzt, die genaue Stichprobenzusammensetzung wurde allerdings nicht erfasst.

Ausbildungsphase	Fähigkeits-stufe 1		Fähigkeits-stufe 2		Fähigkeits-stufe 3		Fähigkeits-stufe 4	
Gesamtzahl	17		45		71		56	
davon Anteilig in den jeweiligen Ausbildungsphasen:								
	in %	N	in %	N	in %	N	in %	N
Bachelor	94 %	16	56 %	25	41 %	29	36 %	20
Master	6 %	1	36 %	16	49 %	35	61 %	34
Staatsexamen	0 %	0	3 %	0	3 %	2	0 %	0
Referendariat	0 %	0	7 %	4	7 %	5	4 %	2

Tabelle 21.3: Übersicht über die Zusammensetzung der Fähigkeitsstufen nach Ausbildungsphasen in Prozent- und Absolutwerten

bildungsphase der zugeordneten Personen ausmachen ($H(3) = 18,10$; $p < 0,01$), allerdings nur zwischen der unteren Fähigkeitsstufen 1 und jeweils den drei übrigen Fähigkeitsstufen 2 ($p < 0,05$), 3 ($p < 0,01$) und 4 ($p < 0,01$) Auf Grundlage paarweiser Vergleiche mit Bonferroni-korrigierten Wilcoxon-Tests lassen sich die Effektstärken als moderat einordnen (Cohen, 1988, S. 79ff.). Die oberen drei Fähigkeitsstufen weisen untereinander keine signifikanten Unterschiede in ihrer Zusammensetzung hinsichtlich der Ausbildungsphase auf.

In Bezug auf das Physik-Fachsemester lassen sich keine signifikanten Unterschiede zwischen den Fähigkeitsstufen ausmachen ($H(3) = 3,66$; $p = 0,30$). Unklar ist der Einfluss falscher Angaben, wenn z. B. lediglich die Anzahl der Fachsemester im Masterstudium und nicht die Gesamtzahl an Fachsemestern angegeben wurde. Als weitere demographische Variable mit Zusammenhang zum Studienfortschritt wurde die Anzahl der Wochen in Pflichtpraktika erhoben. Der Großteil der Studierenden gibt an, keinerlei Praxiserfahrung gesammelt zu haben. Die Berechnung des Kruskal-Wallis-Tests weist dennoch auf Unterschiede zwischen den Gruppen hin ($H(3) = 13,06$; $p < 0,05$). Die Post-hoc-Analyse ergibt signifikante Unterschiede zwischen der oberen Fähigkeitsstufe 4 (mit durchschnittlich $M = 11,52$ ($SD = 10,54$) absolvierten Wochen in Pflichtpraktika) und den unteren beiden Fähigkeitsstufen mit (knapp noch) kleiner Effektstärke von $r_{\text{Pearson}} = 0,29$ für die Fähigkeitsstufen 2 und 4 bzw. mittlerer Effektstärke $r_{\text{Pearson}} = 0,33$ für die Fähigkeitsstufen 1 und 4 (jeweils $p < 0,05$ mit $M = 3,71$ ($SD = 4,21$) absolvierten Wochen in Pflichtpraktika in Fähigkeitsstufe 1 und $M = 6,62$ ($SD = 10,54$) absolvierten Wochen in Pflichtpraktika in Fähigkeitsstufe 2).

Schulische Erfahrungen können die Studierenden in der Stichprobe auch außerhalb des Studiums sammeln, indem sie als Vertretungslehrkraft arbeiten. Diese Erfahrung wurde auf Grundlage zweier Fragen im Fragebogen zur Demografie quantifiziert, die erfassen, wie viele Stunden die Studierenden in einer Woche als

Vertretungslehrkraft arbeiten und wie viele Semester sie dies bisher tun. Einzeln betrachtet geben die Variablen Aufschluss über Intensität (Stundenzahl je Woche) und Dauer der Erfahrung (Anzahl der Semester). In allen vier Fähigkeitsstufen gibt der überwiegende Anteil der Personen an, nicht als Vertretungslehrkraft zu arbeiten bzw. gearbeitet zu haben (zwischen etwa 65% der Personen in Profil 2 und etwa 82% in Profil 3), sodass sich diese Variablen nicht zur Differenzierung der Fähigkeitsstufen eignen.

21.7 Untersuchung von Zusammenhängen zwischen Testmotivation sowie Bearbeitungsdauer und der Zuordnung zu Fähigkeitsstufen

Zusätzlich wurden die angegebene Motivation während der Testbearbeitung (s. Kapitel 13.2) sowie die Bearbeitungsdauer als zwei Variablen erfasst, die sich möglicherweise über die Fähigkeitsstufen unterscheiden. Insgesamt liegt die durchschnittlich angegebene Motivation in allen vier Fähigkeitsstufen deutlich über dem theoretischen Mittelwert von 9 (s. Tabelle 21.4)[196], sodass über alle Fähigkeitsstufen hinweg grundsätzlich von einer engagierten Bearbeitung des Reflexionstests auszugehen ist. Dennoch scheint auch die Motivation zwischen den Fähigkeitsstufen unterschiedlich stark ausgeprägt. Auf Grundlage des Kruskal-Wallis-Tests ($H(3) = 15,74$; $p < 0,01$) und einer Post-hoc-Analyse zum paarweisen Vergleich der Fähigkeitsstufen ist von signifikanten Unterschieden kleiner bis mittlerer Effektstärke ($p < 0,05$) jeweils zwischen der oberen Fähigkeitsstufe 4 und den drei übrigen Fähigkeitsstufen auszugehen. Insgesamt entspricht die aufsteigende Reihenfolge der angegebenen Motivation auch den aufsteigenden Testwerten (Fähigkeitsstufen 1, 2, 3, 4).

Auch die Bearbeitungsdauer des Reflexionstests nimmt über die Profile zu (s. Tabelle 21.4). Der Kruskal-Wallis-Test weist auf signifikante Unterschiede zwischen den Gruppen hin ($H(3) = 23,65$; $p < 0,01$). Gemäß Post-hoc-Analyse liegen signifikante Unterschiede zwischen der Fähigkeitsstufe 1 und den übrigen Fähigkeitsstufen mit jeweils mittlerer Effektstärke vor (jeweils mit $p < 0,01$), wobei wiederum die höheren Fähigkeitsstufen jeweils eine höhere durchschnittliche Bearbeitungsdauer und geringere Streuung als Fähigkeitsstufe 1 aufweisen.

Eine geringere Testmotivation oder Bearbeitungsdauer sind allerdings nicht zwangsläufig als kausal für das schlechtere Abschneiden im Reflexionstest anzunehmen. Denkbar wäre auch ein umgekehrter Zusammenhang, indem z. B. Unsicherheit oder Überforderung während der Bearbeitung zu geringerer Motivation führen

[196] Die Angaben zur Zustimmung zu den drei Items entlang der jeweils fünf-stufigen Likert-Skala wurden aufsummiert, sodass sich für die Testmotivation Werte von 3 bis 15 ergeben können.

	Fähigkeitsstufe 1		Fähigkeitsstufe 2		Fähigkeitsstufe 3		Fähigkeitsstufe 4	
	M	SD	M	SD	M	SD	M	SD
Mittlere Testmotivation (Skala von 1 bis 15)	$10,3$	$2,2$	$10,5$	$2,2$	$10,8$	$3,0$	$12,2$	$2,7$
Mittlere Bearbeitungsdauer gerundet auf ganze Sekunden	2431	2541	3706	1532	4326	1505	3840	1373

Tabelle 21.4: Durchschnittliche Motivation und Bearbeitungsdauer des Reflexionstests (mit Standardabweicheungen) in den einzelnen Fähigkeitsstufen

oder weniger umfangreiche Reflexionen mit einer kürzeren Bearbeitungsdauer einhergehen.

21.8 Erste Erweiterung des Modells zum Zusammenhang von Noticing und Unterrichtsreflexion (Beantwortung der Forschungsfrage 2)

Ziel der durchgeführten Untersuchungen war die Beantwortung der Forschungsfrage 2:

> **Forschungsfrage 2:** Welche latenten Profile der Reflexionsfähigkeit lassen sich auf Grundlage der Bearbeitungen des Reflexionstests identifizieren?

Im Rahmen der Profilanalyse konnten vier Profile der Unterrichtsreflexion identifiziert werden, die in allen Aufgaben in gleicher Reihenfolge aufsteigende durchschnittliche Lösungshäufigkeiten aufweisen und daher als Fähigkeitsstufen interpretiert wurden. Gestützt wird diese Interpretation durch die Beobachtung der Entwicklung der Profilzuordnungen in einem Prä-Post-Design: Von den $N = 41$ Studierenden verbleibt der Großteil zum zweiten Messzeitpunkt im selben Profil, während 13 Personen in ein höheres und nur zwei Personen in ein niedrigeres Profil eingruppiert werden. Da davon ausgegangen werden kann, dass zwischen beiden Messzeitpunkten zwar Lerngelegenheiten für Reflexion (z. B. im Praxissemester oder universitären Veranstaltungen) bestanden, die Entwicklung der Reflexionsfähigkeit auf dieser Grundlage aber nicht als selbstverständlich angenommen werden

kann (Kulgemeyer et al., 2021), erscheint die Interpretation als Fähigkeitsstufen plausibel. Gleichzeitig stützt die Beobachtung des zunehmenden Anteils an Masterstudierenden in aufsteigender Reihenfolge der Profile die Interpretation, dass umfangreichere Lerngelegenheiten, die hier durch den Studienfortschritt repräsentiert werden, tendenziell mit einer höher ausgeprägten Reflexionsfähigkeit und damit der Zuordnung zu einer höheren Fähigkeitsstufe einher gehen.

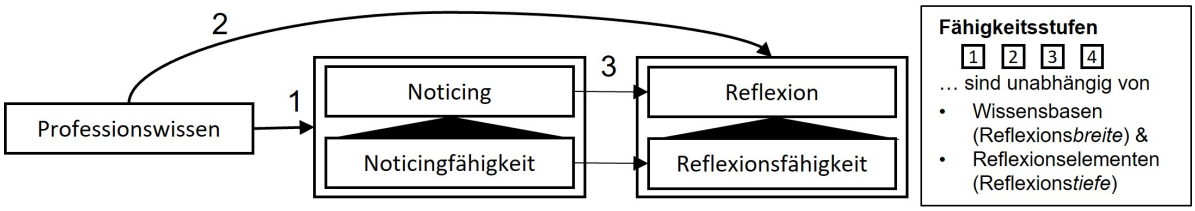


Abbildung 21.3: Modell zum Zusammenhang von Noticing und Unterrichtsreflexion (erste Erweiterung)

Zwischen den Profilen bzw. Fähigkeitsstufen und den Reflexionselementen (der Reflexionstiefe) oder den der Reflexion zugrundeliegenden Wissensbasen (die Reflexionsbreite) ließen sich keine Zusammenhänge finden; insbesondere scheinen beide Merkmale nicht die schwierigkeitsbestimmenden Parameter des Testinstruments zu sein (zusammenfassend dargestellt in Abbildung 21.3). Schlussfolgernd könnte vermutet werden, dass zur Interpretation der Fähigkeitsstufen weitere Fähigkeiten herangezogen werden müssen, die bei der Bewältigung der Testaufgaben helfen. Denkbar wäre, dass die Teilnehmenden unterschiedlich gut in der Lage sind, die zu reflektierenden Aspekte in den Videovignetten zu erkennen und die Aufgaben zur Reflexion daher unterschiedlich gut bearbeiten können, also ein Zusammenhang zum Noticing besteht, oder Unterschiede im Professionswissen z. B. Einfluss auf die Wahl adäquater Alternativen nehmen. Zusammenhänge zwischen den Fähigkeiten zur Reflexion und zum Noticing werden nachfolgend untersucht.

22 Zum Zusammenhang von Noticing und Reflexion

Im entwickelten Reflexionstest müssen die Teilnehmenden nicht selbst entscheiden, welche Aspekte sie reflektieren, stattdessen werden die Reflexionen durch den Einsatz der Multiple-Choice-Aufgaben zu einzelnen Aspekten der Videovignetten inhaltlich stark gelenkt. Das Noticing ist also kein inhärenter Bestandteil der Aufgabenbearbeitung. Gleichzeitig bearbeiten die Teilnehmenden nach jeder Vignette eine Aufgabe zum Noticing, in der sie angeben, welche Aspekte der gesehenen Vignette aus ihrer Perspektive besonders relevant für die anschließende Reflexion scheinen (s. Kapitel 10.1). Diese Aufgaben tragen einerseits zur Authentizität der Reflexionssituation bei (s. Kapitel 15.1) und bieten andererseits einen ersten Zugang zur Erfassung der Fähigkeit zum Noticing. So können nachfolgend beide mit dem Reflexionstest erfassten Fähigkeiten (zur Unterrichtsreflexion und zum Noticing) miteinander in Beziehung gesetzt werden, um die Forschungsfrage 3 zu beantworten.

Forschungsfrage 3: Wie hängen die Fähigkeiten zum Noticing und zur Unterrichtsreflexion zusammen?

Noticing wird - wie bereits in Kapitel 7.6 beschrieben - aus theoretischer Perspektive als eine Voraussetzung für bedeutsame Unterrichtsreflexion verstanden: Nur, wenn (auch) die relevanten Aspekte des Unterrichts erkannt und anschließend reflektiert werden, kann von einer relevanten Verbesserung der Unterrichtsqualität bzw. einer individuellen Weiterentwicklung ausgegangen werden.

Zur empirischen Bestätigung dieses theoretisch unterstellten Zusammenhangs werden die Hypothesen 1 bis 3 geprüft:

Hypothese 1: Zwischen den Testwerten zur Noticing- und Reflexionsfähigkeit besteht ein positiver Zusammenhang.

Hypothese 2: Zwischen den Fähigkeitsstufen zur Unterrichtsreflexion bestehen Unterschiede in den Testwerten zur Noticingfähigkeit.

Hypothese 3: Die Teilnehmenden schneiden in Aufgaben zur Reflexion besser ab, wenn sie die reflektierten Aspekte zuvor in den Noticing-Aufgaben als bedeutsam einschätzen.

22.1 Ergebnisse

Nachfolgend werden einige Voruntersuchungen zur Verwendung der Noticing-Aufgaben als Skala durchgeführt, bevor die Analysen zu den einzelnen Hypothesen berichtet werden. Dabei werden hier lediglich die Ergebnisse der zentralen statistischen Tests berichtet; eine Übersicht über die jeweilige Prüfung der Voraussetzung für die Verwendung der Tests (z. B. die Normalverteilung der betrachteten Testwerte) befindet sich im Anhang in Kapitel M ab Seite 525.[197]

22.1.1 Voruntersuchungen zu den Testwerten zur Noticingfähigkeit

Vor der Prüfung der aufgestellten Hypothesen bezüglich des Zusammenhangs zwischen Noticing- und Reflexionsfähigkeit wird geprüft, inwiefern die statistischen Voraussetzungen zur Verwendung der Noticing-Aufgaben als Skala erfüllt sind. Dazu werden die Normalverteilung und Reliabilität der Skala sowie die Trennschärfen und Lösungshäufigkeiten als Aufgabenparameter geprüft (Döring & Bortz, 2016, S. 476).

Der Reflexionstest enthält insgesamt acht Aufgaben zum Noticing (eine im Anschluss an jede Videovignette sowie vor der globalen Rückmeldung zur Unterrichtsstunde). Die $N = 189$ Teilnehmenden erreichen Noticing-Testwerte von etwa 19 % bis 100 %. Auf Grundlage des nicht signifikanten Ergebnisses des Kolmogorow-Smirnow-Tests von $D(189) = 0,052; p = 0,68$ kann von einer Normalverteilung der Daten mit Mittelwert von $M = 65$ % und Standardabweichung von $SD = 18$ % ausgegangen werden. Die Skala erreicht eine ausreichende Reliabilität von $\alpha_{\text{Cronbach}} = 0,86$ (Döring & Bortz, 2016, S. 443); alle Aufgaben erreichen hohe Trennschärfen von $> 0,5$ (durchschnittlich $0,62$) (Döring & Bortz, 2016, S. 478). Auch die Lösungshäufigkeiten liegen mit durchschnittlich $0,65$ im bevorzugten Bereich zwischen $0,20$ und $0,80$ (Döring & Bortz, 2016, S. 477). Lediglich die Lösungshäufigkeit zur letzten Noticing-Aufgabe vor der Reflexion des Unterrichts insgesamt liegt mit $0,82$ geringfügig über diesem Intervall.

22.1.2 Prüfung der Hypothese 1

Zur Prüfung der Hypothese 1 *„Zwischen den Testwerten zur Noticing- und Reflexionsfähigkeit besteht ein positiver Zusammenhang.“* wird die Korrelation zwischen dem Testwert zur Reflexion mit dem zum Noticing im Reflexionstest bestimmt. Dabei ergibt sich ein hochsignifikanter Zusammenhang von $r_{\text{Pearson}} = 0,61$ ($p < 0,01$) mit großem Effekt (Cohen, 1988, S. 79ff.).

[197] Das zugehörige *R*-Skript zu den Analysen befindet sich im Anhang in Kapitel T ab Seite 945.

Damit kann Hypothese 1 als bestätigt angesehen werden: die Fähigkeiten zum Noticing und zur Unterrichtsreflexion scheinen also zusammenzuhängen. Da gleichzeitig die Korrelation beider Fähigkeiten deutlich nicht vollständig miteinander korrelieren, ist nicht davon auszugehen, dass beide Skalen dieselbe Fähigkeit erfassen. Darauf aufbauend werden die Fähigkeiten zum Noticing und zur Unterrichtsreflexion als unterscheidbar interpretiert.[198]

22.1.3 Prüfung der Hypothese 2

Zur Untersuchung der Hypothese 2 *„Zwischen den Fähigkeitsstufen zur Unterrichtsreflexion bestehen Unterschiede in den Testwerten zur Noticingfähigkeit.“* wird für die Gesamtstichprobe der $N = 189$ berücksichtigten Bearbeitungen des Reflexionstests von (angehenden) Physik-Lehrkräften eine ANOVA durchgeführt, die auf hochsignifikante Unterschiede in den Testwerten zum Noticing zwischen den Fähigkeitsstufen hindeutet $F(3; 185) = 30,04; p < 0,01$. Zusätzlich werden direkte Vergleiche der Fähigkeitsstufen mit paarweisen t-Tests mit Bonferroni-Korrektur angestellt und d_{Cohen} als Maß für die Effektstärke der signifikanten Unterschiede bestimmt. Hier zeigt sich, dass die obere Fähigkeitsstufe 4 signifikant bessere Testwerte im Noticing erreicht, als alle darunterliegenden ($p < 0,01$ mit jeweils großer Effektstärke zwischen $d_{\text{Cohen}} = 1,00$ und $d_{\text{Cohen}} = 2,00$ (Döring & Bortz, 2016, S. 820)). Zusätzlich weist die untere Fähigkeitsstufe 1 signifikant niedrigere Testwerte im Noticing als die beiden mittleren Fähigkeitsstufen auf ($p < 0,05$ mit knapp noch moderater Effektstärke von $d_{\text{Cohen}} = 0,77$ zu Fähigkeitsstufe 2 und $p < 0,01$ mit großer Effektstärke von $d_{\text{Cohen}} \approx 1,11$ zu Fähigkeitsstufe 3). Zwischen den mittleren Fähigkeitsstufen 2 und 3 liegt kein signifikanter Unterschied in den Testwerten zum Noticing vor ($p = 0,15$).

Insgesamt konnten also drei Gruppen ausgemacht werden, die sich signifikant voneinander unterscheiden: die untere Fähigkeitsstufe, die beiden mittleren Fähigkeitsstufen sowie die obere Fähigkeitsstufe, sodass Hypothese 2 als bestätigt anzusehen ist.

22.1.4 Prüfung der Hypothese 3

Zur Prüfung der Hypothese 3 *„Die Teilnehmenden schneiden in Aufgaben zur Reflexion besser ab, wenn sie die reflektierten Aspekte zuvor in der Noticing-Aufgabe als bedeutsam einschätzten.“* werden zwei Testwerte zur Reflexionsfähigkeit

[198] Dies passt zu den Ergebnissen von Bastian et al. (2022), die die empirische Trennbarkeit von Bestandteilen der professionellen Unterrichtswahrnehmung untersuchten (von ihnen allerdings als „Noticing“ bezeichnet) und die Wahrnehmung von zentralen Ereignissen („Perception“ entsprechend der hier als Noticing bezeichneten Fähigkeit) sowie die beiden analytischen Bestandteilen Interpretation und Entwicklung von Handlungsoptionen differenzieren konnten.

bestimmt, die jeweils die durchschnittliche, in einer Aufgabe erreichte Punktzahl für zuvor im Noticing angewählte bzw. nicht angewählte Aspekte angeben:

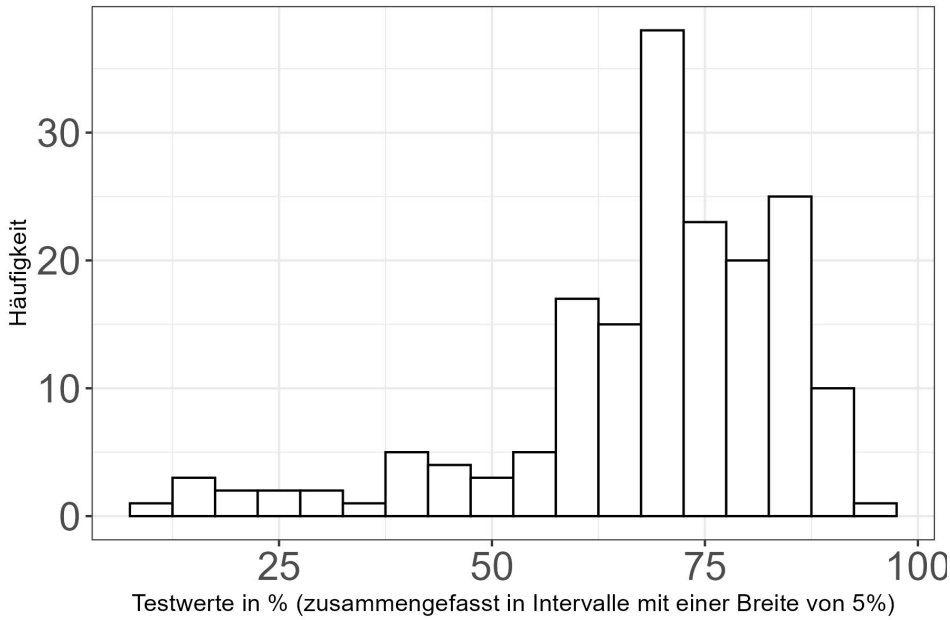
$$\text{Reflexionsscore}_{\text{angewählt im Noticing}} = \frac{\text{Summe der Aufgabenscores in Aufgaben zu angewählten Aspekten}}{\text{Anzahl der angewählten Aspekte, die anschließend reflektiert werden}}$$

$$\text{Reflexionsscore}_{\textit{nicht} \text{ angewählt im Noticing}} = \frac{\text{Summe der Aufgabenscores in Aufgaben zu } \textit{nicht} \text{ angewählten Aspekten}}{\text{Anzahl der } \textit{nicht} \text{ angewählten Aspekte, die anschließend reflektiert werden}}$$

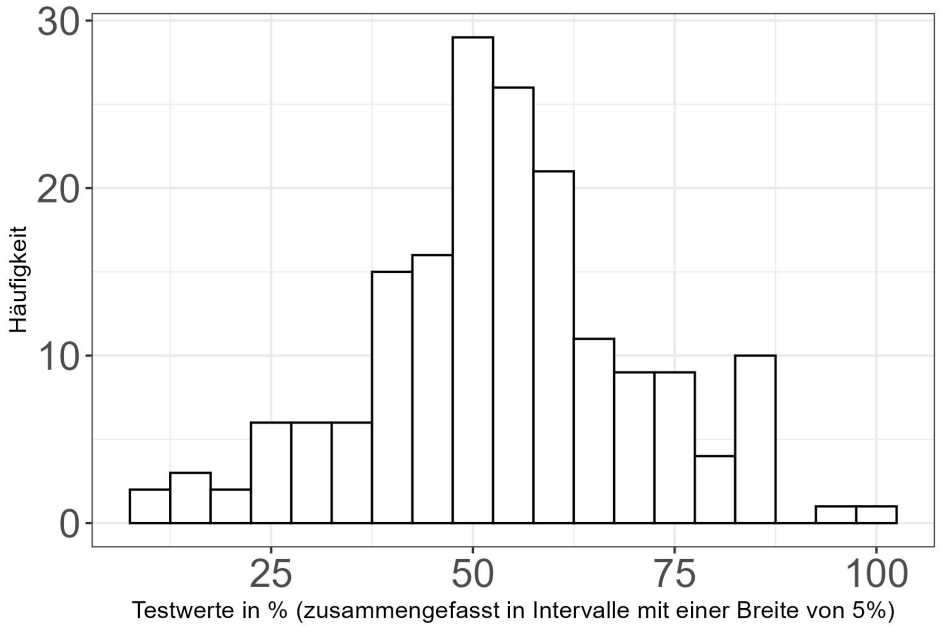
Bereits die Häufigkeitsverteilungen dieser beiden mittleren Aufgabenscores (s. Abbildung 22.1) deuten darauf hin, dass die Teilnehmenden in den Aufgaben zu Aspekten, die sie zuvor als relevant identifiziert haben, tendenziell bessere Aufgabenscores erreichen. Für den $\text{Reflexionsscore}_{\textit{nicht} \text{ angewählt im Noticing}}$ kann - unter Ausschluss von zwölf Fällen für die dieser Score nicht bestimmt werden konnte, da in den Noticing-Aufgaben alle Aspekte als relevant angewählt wurden, die anschließend in den Aufgaben zu Bewertungen und Alternativen reflektiert werden - nach dem Kolmogorow-Smirnow-Test von einer Normalverteilung der Daten ausgegangen werden ($D(177) = 0,059$; $p = 0,57$). Für den $\text{Reflexionsscore}_{\text{angewählt im Noticing}}$ wird der Kolmogorow-Smirnow-Test allerdings signifikant ($D(177) = 0,15$; $p < 0,01$), sodass nicht von normalverteilten Daten ausgegangen werden kann.

Zur ergänzenden statistischen Untersuchung des Unterschieds zwischen beiden Testwerten wird ein gepaarter t-Test berechnet. Voraussetzung hierfür ist die Normalverteilung der Differenz beider Reflexionsscores, die nach Kolmogorow-Smirnow-Test gegeben ist ($D(177) = 0,07$; $p = 0,44$). Auf Grundlage des gepaarten t-Test kann von einem signifikanten Unterschied zwischen beiden Reflexionsscores ausgegangen werden ($t(177) = 12,0$; $p < 0,01$): Die Teilnehmenden erreichen mit großer Effektstärke ($d_{\text{Cohen},z} = 0,90$) durchschnittlich dann bessere Testwerte, wenn sie Aufgaben zu Aspekten bearbeiten, die sie zuvor (in den Noticing-Aufgaben) als bedeutsam eingeschätzt haben: Sie erreichen dann durchschnittlich $M = 69$ % ($SD = 17$ %) gegenüber $M = 54$ % ($SD = 17$ %) für nicht als relevant eingeschätzte Aspekte.

Entsprechende Unterschiede in den beiden bestimmten Reflexionsscores lassen sich analog für die einzelnen Profile untersuchen. Für die oberen drei Profile ergeben sich dabei ebenfalls signifikante Unterschiede mit großer Effektstärke zwischen den erreichten Punktzahlen abhängig davon, ob die jeweils reflektierten



(a)



(b)

Abbildung 22.1: Histogramme zu den jeweils erreichten Testwerten (in %) in den Aufgaben zur Reflexion von in Rahmen der Aufgaben zum Noticing (a) als bedeutsam bzw. (b) nicht als bedeutsam eingeschätzten Aspekten

Aspekte im Rahmen des Noticing als relevant eingeschätzt wurden.[199] Für die untere Fähigkeitsstufe lässt sich dieser Zusammenhang nicht feststellen ($t(16) = 0,07$; $p = 0,94$). Über die Ursache des ausbleibenden Zusammenhangs kann auf Grundlage der vorliegenden Daten nur spekuliert werden. Möglicherweise ist die Reflexionsfähigkeit in dieser Stufe (noch) nicht ausreichend ausgeprägt, um mit dem Reflexionstest unterschiedlich bedeutsame Reflexionen zu differenzieren.[200] Auch ist denkbar, dass die Fähigkeit zum Noticing in dieser Fähigkeitsstufe (noch) nicht ausreichend ausgebildet ist, um einen feststellbaren Einfluss auf die Reflexionen zu nehmen.

Gelingendes Noticing scheint also - zumindest für die oberen Fähigkeitsstufen - Voraussetzung für bedeutsame Reflexion zu sein. Daher ist Hypothese 3 grundsätzlich als bestätigt anzusehen.

22.2 Zweite Erweiterung des Modells zum Zusammenhang von Noticing und Unterrichtsreflexion (Beantwortung der Forschungsfrage 3)

Die berichteten Untersuchungen zu Forschungsfrage 3 adressieren den Zusammenhang der Fähigkeiten zum Noticing und zur Unterrichtsreflexion:

Forschungsfrage 3: Wie hängen die Fähigkeiten zum Noticing und zur Unterrichtsreflexion zusammen?

Auf Grundlage der durchgeführten Untersuchungen können zwei zentrale Erkenntnisse bezüglich des Zusammenhangs zwischen beiden Fähigkeiten festgehalten und im Modell zum Zusammenhang von Noticing und Unterrichtsreflexion ergänzt werden (s. Abbildung 22.2):

1. Die Bestätigung der Hypothesen 1 und 2 weist darauf hin, dass die Fähigkeiten zum Noticing und zur Unterrichtsreflexion unterscheidbar sind, aber zusammenhängen und eine höhere Reflexionsfähigkeit in der Regel mit einer höheren Fähigkeit zum Noticing einhergeht.

[199] Fähigkeitsstufe 2: $t(44) = 7,32; p < 0,01$ mit Effektstärke $d_{\text{Cohen}} \approx 1,09$; Fähigkeitsstufe 3: $t(69) = 7,06; p < 0,01$ mit Effektstärke $d_{\text{Cohen}} \approx 0,84$; Fähigkeitsstufe 4: $t(44) = 8,92; p < 0,01$ mit Effektstärke $d_{\text{Cohen}} \approx 1,33$ (die jeweils erreichten mittleren Testwerte sind ebenfalls im Anhang in Kapitel M.3 ab Seite 526 aufgeführt)

[200] Ein Einfluss des Auflösungsgrads des Reflexionstests ist dabei nicht auszuschließen. Für die Modellierung und Messung von Kompetenzen stellt Neumann (2013) allgemein einen Bezug zwischen der Entwicklung einer Kompetenz und der „Feinkörnigkeit“ (Neumann, 2013, S. 36) her, in dem die Kompetenz praktisch gemessen werden kann.

2. Dass die Teilnehmenden außerdem in den Aufgaben zur Reflexion dann besser abschneiden, wenn sie die reflektierten Aspekte zuvor in den Noticing-Aufgaben als bedeutsam einschätzten (Hypothese 3), spricht dafür, dass die Fähigkeiten zum Noticing und zur Unterrichtsreflexion nicht nur verbunden sind, sondern Noticing Voraussetzung für (bedeutsame) Reflexion ist.

Die Bestätigung dieser Hypothesen kann außerdem als Hinweis darauf interpretiert werden, dass Noticing- und Reflexionsfähigkeit möglicherweise von derselben bzw. denselben Disposition(en) profitieren (vgl. Pfeile 1 und 3 in Abbildung 22.2). Die im Modell dargestellte Annahme, solch eine Disposition sei das Professionswissen, liegt darin begründet, dass Noticing und Reflexion in ihren Definitionen als wissensgesteuerte bzw. theoriebasierte Prozesse verstanden werden. So ist bspw. Fachwissen Voraussetzung dafür, fachliche Fehler zu erkennen, und Wissen über den Umgang mit Schülervorstellungen ist Grundlage für die Entwicklung fundierter Handlungsalternativen. Im Rahmen der Validierungsstudien konnte der Zusammenhang zwischen den Fähigkeiten zum Noticing und zur Reflexion und dem fachdidaktischen Wissen als Bestandteil des Professionswissens allerdings nicht bestätigt werden (s. auch Kapitel 17.6). Da lediglich ein Bestandteil des Professionswissens in einer eher kleinen Stichprobe ($N = 33$ Fälle) erhoben wurde, besteht weiterer Bedarf zur Untersuchung des etwaigen Zusammenhangs.

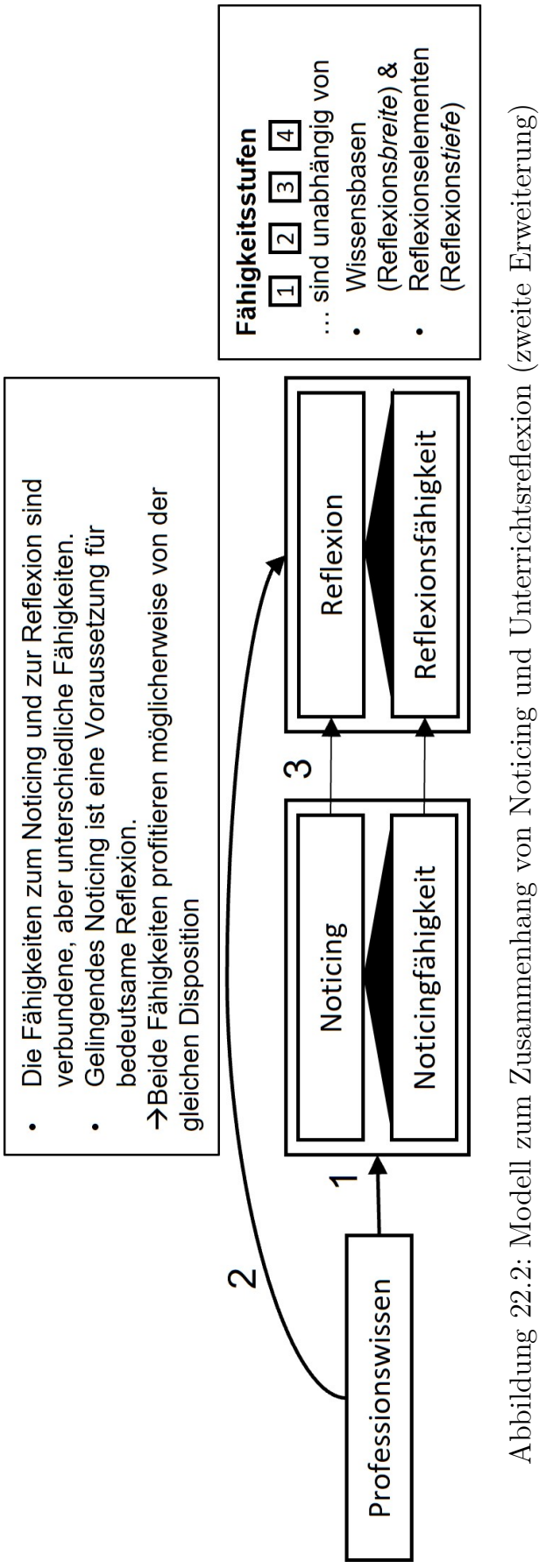


Abbildung 22.2: Modell zum Zusammenhang von Noticing und Unterrichtsreflexion (zweite Erweiterung)

23 Zum Entwicklungszusammenhang der Fähigkeiten zum Noticing und zur Reflexion

In diesem Kapitel wird die Untersuchung des modellierten Zusammenhangs zwischen Noticingfähigkeit und Reflexionsfähigkeit um die Untersuchung möglicher kausaler Entwicklungszusammenhänge zwischen beiden Fähigkeiten erweitert. Die hier berichteten Analysen ergänzen damit die Untersuchung des Zusammenhangs zwischen beiden Fähigkeiten um eine Betrachtung der Entwicklung. Explizit wird Forschungsfrage 4 untersucht:

> **Forschungsfrage 4:** Wie beeinflussen die Fähigkeiten zum Noticing und zur Unterrichtsreflexion wechselseitig ihre Entwicklung?

Gestützt durch die festgestellten Zusammenhänge zwischen Noticing und Unterrichtsreflexion in Kapitel 22 kann dabei ein Einfluss der Noticingfähigkeit auf die Entwicklung der Reflexionsfähigkeit nach folgendem Mechanismus vermutet werden: Eine stärker ausgeprägte Noticingfähigkeit ermöglicht es, zentrale Elemente in Unterrichtssituationen zu identifizieren und anschließend zu reflektieren. Die Reflexion dieser Elemente stellt jeweils eine Lerngelegenheit für bedeutsame Unterrichtsreflexionen dar.[201]

Analog wäre allerdings auch ein umgekehrter Einfluss der Reflexionsfähigkeit auf die Entwicklung der Fähigkeit zum Noticing wie folgt denkbar: Ist die Reflexionsfähigkeit stärker ausgeprägt, finden sinnvolle Reflexionen zu mehr (zentralen) Aspekten des Unterrichts statt, was einen Beitrag zur Entwicklung des Professionswissens liefern kann (Kulgemeyer et al., 2021). So können die Aspekte, die zuvor inhaltlich sinnvoll reflektiert wurden, nachfolgend besser als zentrale Elemente des Unterrichts erkannt werden.

[201] Die Fokussierung auf die Reflexion bedeutsamer Aspekte des Unterrichts ist hier relevant, da auch im Reflexionstest auf bedeutsame Aspekte des (Physik-)Unterrichts fokussiert wird und z. B. die Reflexion der Körpersprache aufgrund der untergeordneten Bedeutung nicht enthalten ist.

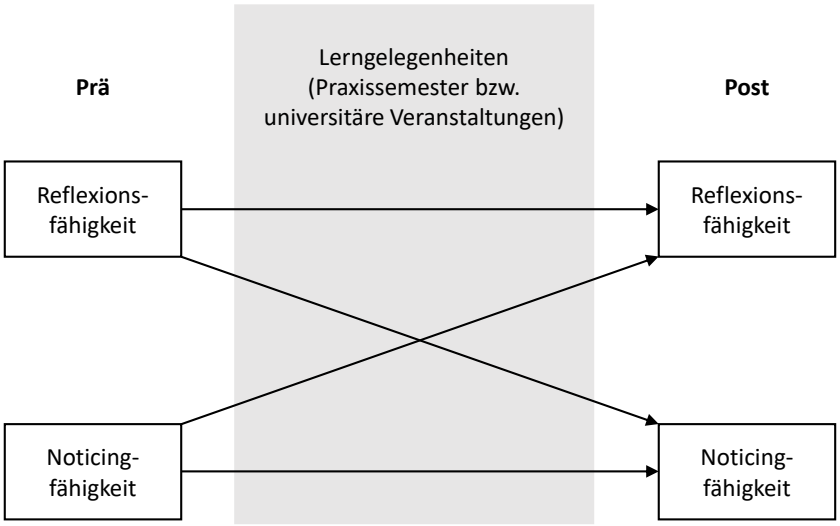


Abbildung 23.1: CLP-Design für die Untersuchung von Entwicklungszusammenhängen zwischen den Fähigkeiten zum Noticing und zur Unterrichtsreflexion

Da ein erwarteter Einfluss also nicht eindeutig theoretisch begründet werden kann, erfolgt die Untersuchung möglicher Kausalzusammenhänge explorativ.

23.1 Statistische Analysen im Cross-Lagged-Panel-Design

Für die Untersuchung von Entwicklungszusammenhängen zwischen der Noticingfähigkeit und der Reflexionsfähigkeit wird wiederum die Stichprobe herangezogen, die den Reflexionstest zweifach bearbeitet hat ($N = 41$; s. Kapitel 21.4). Um mögliche Kausalzusammenhänge in der Entwicklung von Noticing- und Reflexionsfähigkeit zu untersuchen, werden Analysen in einem Cross-Lagged-Panel-Design (CLP-Design) durchgeführt (s. Abbildung 23.1). Dazu wird ein Partial Least Square-Strukturgleichungsmodell (PLS-SEM) berechnet, das im Gegensatz zur schrittweisen Bestimmung partieller Korrelationen, die zur konservativeren Abschätzung ergänzend bestimmt wurden, die gleichzeitige Schätzung der Kreuzpfade ermöglicht (Reinders, 2006) und relativ voraussetzungsarm angewendet werden kann (u. a. für kleine Stichproben und ohne Anforderung an die Verteilung der vorliegenden Daten (Hair et al., 2021, S. 11f.)).

Der vorliegende Datensatz ist mit $N = 41$ Personen zwar relativ klein, erfüllt aber eine oft herangezogene Faustregel für die minimale Stichprobengröße: Er umfasst zehnmal so viele Personen wie das Modell geschätzte Zusammenhänge

enthält (Hair et al., 2021, S. 16). Für die explorative und hypothesengenerierende Untersuchung soll dieses Modell in Anerkennung der Kritik an dieser Abschätzung der minimalen Stichprobengröße (z. B. Hair et al., 2021, S. 16ff.)[202] und unter der Einschränkung, dass lediglich große Effekte untersucht werden können, genutzt werden.

23.2 Ergebnisse

Die nachfolgend berichteten Modelle sind saturiert, sodass keine Prüfungen der Modellgüte vorgenommen werden kann. Die Interpretation der Ergebnisse erfolgt daher unter der angenommenen Angemessenheit des Modells.[203]

23.2.1 PLS-SEM für alle vorliegenden Prä-Post-Datensätze

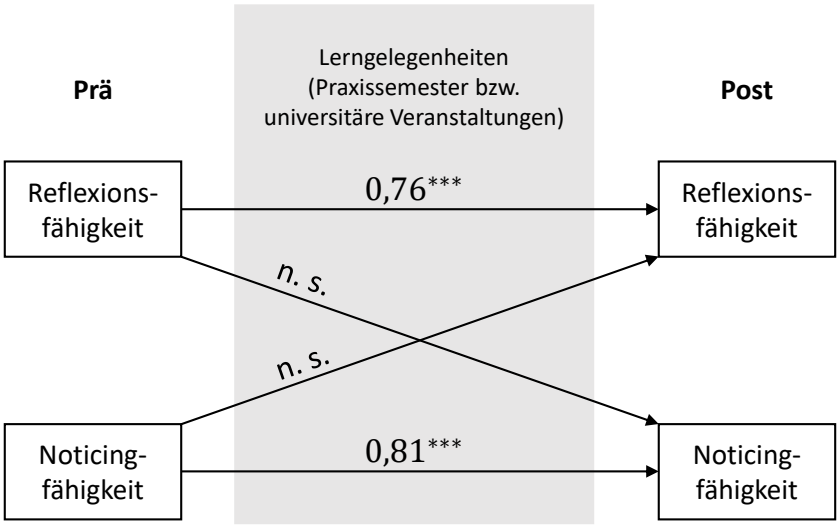
Im PLS-SEM (s. Abbildung 23.2a) wird deutlich, dass für die Fähigkeiten zum Noticing und zur Reflexion jeweils hochsignifikante Zusammenhänge zwischen den beiden MZPs bestehen ($\beta_{\text{Reflexion}} = 0,76$; $p < 0,001$ und $\beta_{\text{Noticing}} = 0,81$; $p < 0,001$). Die Kreuzpfade werden allerdings nicht signifikant; es kann also nicht davon ausgegangen werden, dass global kausale Zusammenhänge zwischen den Fähigkeiten bestehen.

23.2.2 PLS-SEM für die Teilstichprobe der Personen in den oberen drei Fähigkeitsstufen

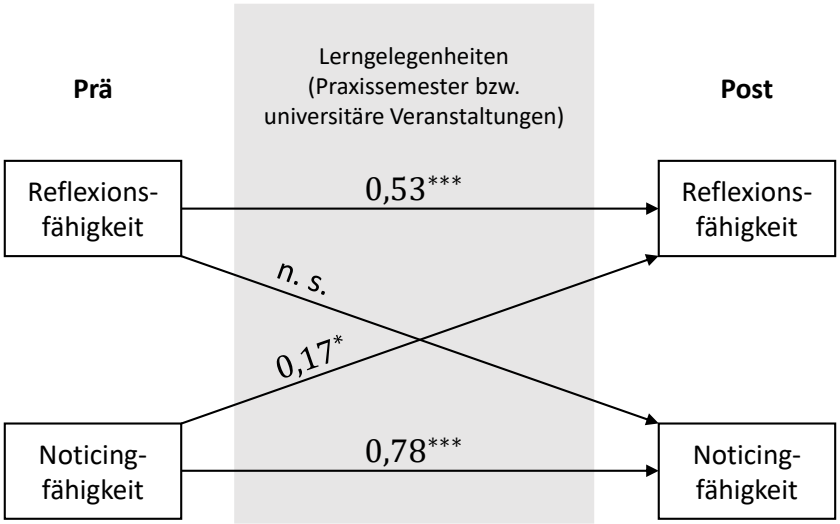
Wie in Kapitel 22.1.4 berichtet, konnte die Hypothese 3 *„Die Teilnehmenden schneiden in Aufgaben zur Reflexion besser ab, wenn sie die reflektierten Aspekte zuvor in der Noticing-Aufgabe als bedeutsam einschätzten.“* zwar global über alle Teilnehmenden bestätigt werden und gilt jeweils für die oberen drei Fähigkeitsstufen, allerdings nicht für die untere Fähigkeitsstufe. Damit scheint die untere Fähigkeitsstufe bezüglich des Zusammenhangs zwischen Noticing und Reflexion eine Sonderstellung einzunehmen. Deren Ursache ist allerdings unklar; möglicherweise sind Zusammenhänge zwischen Noticing, Reflexion und ggf. auch dem Professionswissen noch nicht ausreichend ausgebildet und stellen für diese Gruppe einen Förderbedarf dar. Daher wird nachfolgend erneut ein PLS-SEM unter Ausschluss dieser Fähigkeitsstufe berechnet. Die Stichprobe der oberen drei Fähigkeitsstufen besteht aus $N = 37$ Studierenden und verfehlt damit knapp die zuvor beschriebene untere Grenze für die Stichprobengröße (das Zehnfache der

[202] Andere Methoden zur Bestimmung der minimalen Stichprobengröße lieferten mit $N > 150$ vielfache dieser Schätzung (vgl. Hair et al., 2021, S. 16ff.).

[203] Das *R*-Skript zu den Analysen befindet sich im Anhang in Kapitel U ab Seite 955.



(a)



(b)

Abbildung 23.2: Ergebnisse des PLS-SEM zu Zusammenhängen zwischen den Fähigkeiten zur Reflexion und zum Noticing im CLP-Design für (a) alle vorliegenden Prä-Post-Datensätze und (b) die Teilstichprobe der Personen in den oberen drei Fähigkeitsstufen ($^{*}p < 0,05$; $^{***}p < 0,001$)

geschätzten Zusammenhänge). Um die Vergleichbarkeit zum vorigen PLS-SEM zu ermöglichen, wird unter Verweis auf den explorativen Charakter dieser Analysen weiterhin ein PLS-SEM bestimmt.

Auch für diese Teilstichprobe sind die Pfadkoeffizienten zwischen der Reflexionsfähigkeit zu beiden MZPs sowie der Noticingfähigkeit zu beiden MZPs hochsignifikant ($\beta_{\text{Reflexion}} = 0,53$; $p < 0,001$ und $\beta_{\text{Noticing}} = 0,78$; $p < 0,001$; s. Abbildung 23.2b). Im Gegensatz zum PLS-SEM für die Gesamtstichprobe ist hier zusätzlich der Pfad knapp signifikant, der den Einfluss der Notcingfähigkeit$_{\text{Prä}}$ auf die Reflexionsfähigkeit$_{\text{Post}}$ beschreibt. Der Pfad, der andersherum den Einfluss der Reflexionsfähigkeit$_{\text{Prä}}$ auf die Noticingfähigkeit$_{\text{Post}}$ beschreibt, ist auch hier nicht signifikant.

Für die oberen drei Fähigkeitsstufen kann im PLS-SEM zum CLP-Design also ein Einfluss der Noticingfähigkeit$_{\text{Prä}}$ auf die Reflexionsfähigkeit$_{\text{Post}}$, also auf die Entwicklung der Reflexionsfähigkeit, festgestellt werden. Einschränkend ist festzuhalten, dass dieser Einfluss aufgrund seiner geringen Größe im Vergleich zum Einfluss der Reflexionsfähigkeit$_{\text{Prä}}$ nicht als bestimmend für die Entwicklung der Reflexionsfähigkeit angenommen werden kann.

23.2.3 Cross-Lagged-Panel-Analysen mit partiellen Korrelationen

Als konservativere Methode können alternativ die partiellen Kreuzkorrelationen zur Analyse dieser Pfade im CLP-Design bestimmt werden (Riese et al., 2022), wobei nach Reinders (2006) der Nachweis für kausale Zusammenhänge in der längsschnittlichen Erhebung gegeben ist, wenn (1) die Ursache (hier die vorhandene Noticing- bzw. Reflexionsfähigkeit zum ersten Messzeitpunkt) der Wirkung (d. h. Veränderung der jeweils anderen Fähigkeit) vorgelagert ist, (2) Ursache und Wirkung kovariieren sowie (3) die Ursache eine zentrale (oder sogar alleinige) Erklärung für die Wirkung darstellt. Die zeitlich aufeinanderfolgende Messung ist durch das längsschnittliche Design erfüllt. Die zweite und dritte Bedingung sind dann erfüllt, wenn einerseits die Noticingfähigkeit zum ersten Messzeitpunkt mit der Entwicklung der Reflexionsfähigkeit korreliert (bzw. die Reflexionsfähigkeit mit der Entwicklung der Noticingfähigkeit) und andererseits - unter der Voraussetzung, dass die zentralen potenziellen Einflussfaktoren erfasst wurden - die (partiellen) Kreuzkorrelationen zwischen Noticing- und Reflexionsfähigkeit größer als die jeweilige Autokorrelation der Noticing- bzw. Reflexionsfähigkeit zwischen den verschiedenen Messzeitpunkten sind (vgl. Riese et al., 2022).

Für die vorliegende Stichprobe sind letztere Bedingungen nicht erfüllt. Sowohl für die Gesamtstichprobe als auch für die Teilstichprobe der oberen drei Fähigkeitsstufen können die Testwerte zur Reflexionsfähigkeit und zur Noticingfähigkeit zu beiden MZPs nach dem Kolmogorow-Smirnow-Test als normalverteilt angenommen

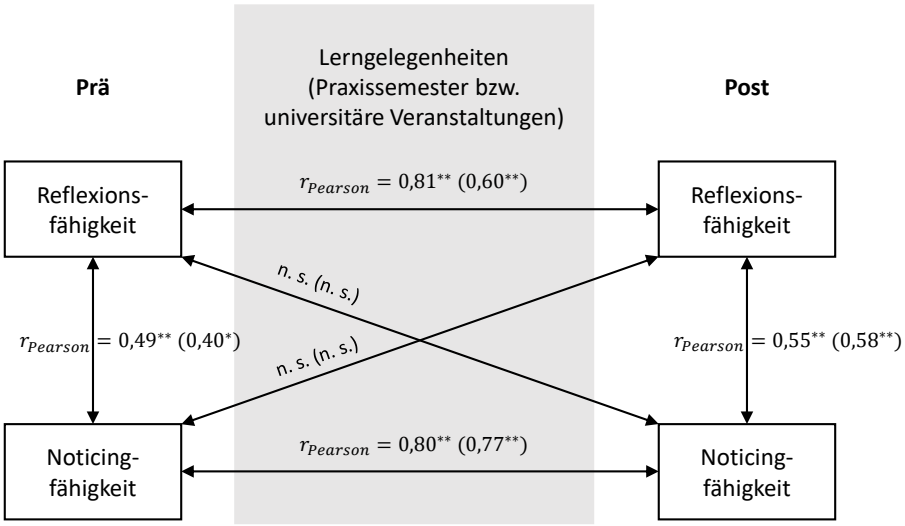


Abbildung 23.3: Übersicht über die Analysen zum CLP-Design für die Gesamtstichprobe (bzw. die Teilstichprobe der Personen in den oberen drei Fähigkeitsstufen); bestimmt wurden jeweils die bivariaten Korrelationen zwischen den Fähigkeiten zum Noticing und zur Unterrichtsreflexion zu den beiden MZP sowie für beide Fähigkeiten zwischen den MZP und die partiellen Kreuzkorrelationen für die Kreuzpfade ($^{*}p < 0,05$; $^{**}p < 0,01$)

werden.[204] Auch liegen zu beiden MZPs sowohl für die Gesamtstichprobe als auch die Teilstichprobe der oberen drei Fähigkeitsstufen signifikante Zusammenhänge zwischen Noticing- und Reflexionsfähigkeit sowie Korrelationen der Reflexions- sowie Noticingfähigkeit zu den beiden MZPs vor (s. Abbildung 23.3). Die partiellen Kreuzkorrelationen von $\text{Noticingfähigkeit}_{\text{Prä}}$ und $\text{Reflexionsfähigkeit}_{\text{Post}}$ sowie $\text{Reflexionsfähigkeit}_{\text{Prä}}$ und $\text{Noticingfähigkeit}_{\text{Post}}$ sind allerdings weder für die Gesamtstichprobe noch die zuletzt betrachtete Teilstichprobe der oberen drei Fähigkeitsstufen signifikant, sodass auf Grundlage dieser Ergebnisse keine wechselseitige Beeinflussung der Entwicklung der Fähigkeiten zur Unterrichtsreflexion und zum Noticing unterstellt werden kann.

[204] Die Ergebnisse der Kolmogorow-Smirnow-Tests sind im Anhang in Kapitel N ab Seite 529 berichtet.

23.3 Dritte Erweiterung des Modells zum Zusammenhang von Noticing und Unterrichtsreflexion (Beantwortung der Forschungsfrage 4)

Ziel der Analysen im CLP-Design war die explorative Untersuchung möglicher Zusammenhänge in der Entwicklung der Noticing- und Reflexionsfähigkeit über ein Semester mit fachdidaktischen Lerngelegenheiten gemäß Forschungsfrage 4:

> **Forschungsfrage 4:** Wie beeinflussen die Fähigkeiten zum Noticing und zur Unterrichtsreflexion wechselseitig ihre Entwicklung?

Die Analysen zu Entwicklungszusammenhängen der Noticing- und Reflexionsfähigkeit liefern kein eindeutiges Bild. Während mit PLS-SEMs nur für die Teilstichprobe der Studierenden in den oberen drei Fähigkeitsstufen ein Einfluss der Noticingfähigkeit auf die Entwicklung der Reflexionsfähigkeit festgestellt werden konnte, wurde dieser in der konservativeren Analyse mit partiellen Kreuzkorrelationen weder für die Gesamt- noch diese Teilstichprobe beobachtet. Insgesamt können diese Ergebnisse daher höchstens als erster Hinweis darauf interpretiert werden, dass die Fähigkeit zum Noticing möglicherweise Einfluss auf die Entwicklung der Reflexionsfähigkeit nimmt - aber nur wenn ein gewisses Maß an Reflexion- und bzw. oder Noticingfähigkeit vorhanden ist.[205] Einen Erklärungsansatz stellt der folgende Mechanismus dar:

Gelingendes Noticing, d. h. das Erkennen der zentralen Elemente von Unterricht, ermöglicht jeweils die anschließende Reflexion. Somit schafft Noticing Lerngelegenheiten für Unterrichtsreflexion.

Gleichzeitig ist zu vermuten, dass ein gewisses Maß an Noticingfähigkeit vorhanden sein muss, um jene Aspekte zur Reflexion auszuwählen, die adäquate Lerngelegenheiten darstellen. Dies ergänzt das Modell zum Zusammenhang von Noticing und Reflexion um Vermutungen bezüglich der Entwicklungszusammenhänge zwischen beiden Fähigkeiten (s. Abbildung 23.4). Es besteht weiterhin Bedarf zur weiteren Untersuchung der Ausgestaltung des Entwicklungszusammenhangs beider Fähigkeiten, insbesondere bezüglich möglicher weiterer (äußerer) Einflussfaktoren.

[205] Der Ausschluss der unteren Fähigkeitsstufe beruht zwar auf der Zuweisung zu Fähigkeitsstufen auf Grundlage der Reflexionsfähigkeit, da allerdings zwischen den Fähigkeitsstufen signifikante Unterschiede in der Noticingfähigkeit bestehen (s. Kapitel 22.1.3), kann diese Voraussetzung prinzipiell für beide Fähigkeiten vermutet werden.

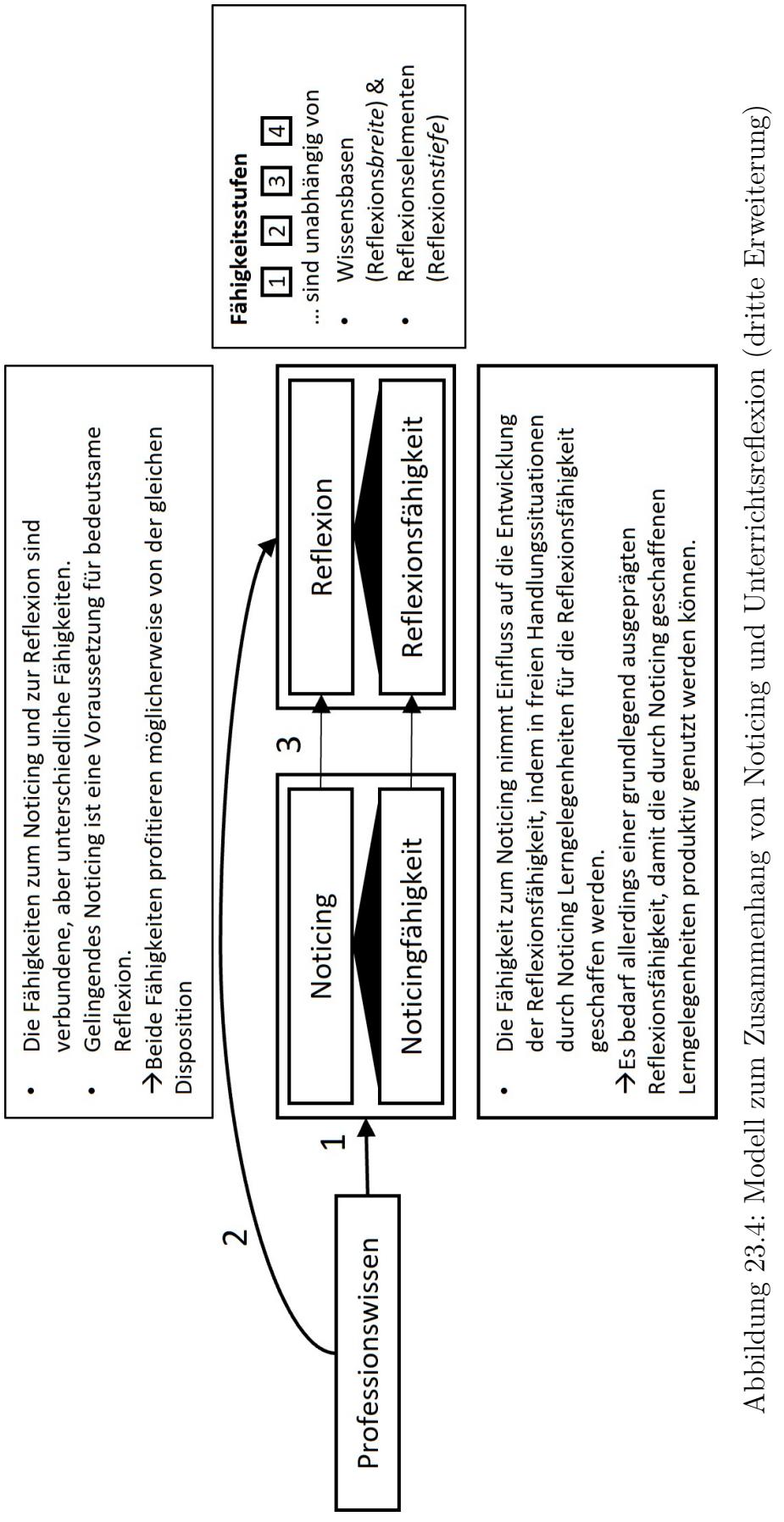


Abbildung 23.4: Modell zum Zusammenhang von Noticing und Unterrichtsreflexion (dritte Erweiterung)

24 Zusammenfassung und Fazit zu den explorativen Untersuchungen zur Reflexionsfähigkeit

Ziel dieses IV. Teils war die Untersuchung der Reflexionsfähigkeit sowie von (Entwicklungs-)Zusammenhängen zur Noticingfähigkeit auf Grundlage der mit dem Reflexionstest gewonnenen Daten. Entlang des Modells zum Zusammenhang von Noticing und Unterrichtsreflexion wurden Analysen zu diesen drei Aspekten durchgeführt. Die Ergebnisse dieser Analysen werden nachfolgend einzeln zusammengefasst (Kapitel 24.1 bis 24.3) und im erweiterten Modell zum Zusammenhang von Noticing und Unterrichtsreflexion zusammengeführt (Kapitel 24.4) bevor abschließend auf Limitationen der Untersuchungen eingegangen wird (Kapitel 24.5).

24.1 Stufen der Reflexionsfähigkeit

Zur explorativen Untersuchung der Struktur der Reflexionsfähigkeit wurde eine latente Profilanalyse durchgeführt. Hierbei erwies sich eine Lösung mit vier Profilen als geeignet. Die Profile weisen über alle in die Profilanalyse einbezogenen Aufgaben hinweg quantitative Unterschiede in der Reflexionsfähigkeit auf und werden daher als vier unterscheidbare Fähigkeitsstufen interpretiert. Diese Interpretation wird durch die tendenziell aufsteigenden Entwicklungen mit zunehmenden Lerngelegenheiten (im Längsschnitt im Rahmen der zwischenzeitlich besuchten Veranstaltungen bzw. des Praxissemesters und im Querschnitt gröber repräsentiert durch die Studienphase) gestützt. Die vorliegenden Daten lassen keine präzise inhaltliche Charakterisierung der Fähigkeitsstufen auf Grundlage der Dimensionen Reflexionselemente und Wissensbasen des Reflexionsmodells nach Nowak, Kempin, Kulgemeyer und Borowski (2019) zu, das der Testentwicklung zugrunde gelegt wurde. Vielmehr deuten sie auf eine globale Stufung der Reflexionsfähigkeit hin, deren Entwicklung möglicherweise durch zusätzliche Fähigkeiten beeinflusst sein könnte.

24.2 Zum Zusammenhang von Noticing und Reflexion

Auf Grundlage der vorliegenden Daten konnte ein enger Zusammenhang von Noticing- und Reflexionsfähigkeit bestätigt werden. Für alle Teilnehmenden am Reflexionstest ($N = 192$) ergab sich eine hochsignifikante Korrelation von $r_{\text{Pearson}} = 0,62$ ($p < 0,01$). Auch konnten signifikante Unterschiede in der Noticingfähigkeit zwischen Fähigkeitsstufen im Reflektieren gefunden werden (mit Ausnahme der beiden mittleren Fähigkeitsstufen). Zusätzlich zeigte sich, dass die Teilnehmenden in den Reflexionsaufgaben dann höhere Scores erreichen, wenn sie die reflektierten Aspekte zuvor im Rahmen der Noticing-Aufgaben als relevant einschätzten ($t(179) = 12,2$; $p < 0,01$). Dies stützt die Interpretation, dass gelingendes Noticing eine Voraussetzung für bedeutsame Reflexion ist und liefert einen Beitrag zur Klärung des Verhältnisses von Noticing und weiteren Aspekten professionellen Handelns (Amador & Weston, 2024).

Insgesamt kann also davon ausgegangen werden, dass die Fähigkeiten zu Noticing und Reflexion zwar eng verbunden sind, aber aufeinander aufbauen und möglicherweise von derselben Disposition, dem Professionswissen, profitieren, da dieses sowohl Grundlage für das Erkennen relevanter Elemente des Unterrichts im Noticing sowie für die fundierte Entscheidung über Bewertungen und Handlungsalternativen in der Reflexion bilden kann.

24.3 Noticingfähigkeit als Voraussetzung für die Entwicklung der Reflexionsfähigkeit

Zuletzt wurden Entwicklungszusammenhänge zwischen den Fähigkeiten zum Noticing und zur Unterrichtsreflexion in einem CLP-Design untersucht. Während für die Gesamtstichprobe der Personen, die den Test zu zwei MZPs bearbeiteten ($N = 41$), keine Entwicklungszusammenhänge beider Fähigkeiten gefunden wurden, weisen Analysen mit einer Teilstichprobe unter Ausschluss der Studierenden, die zum ersten MZP der untersten Fähigkeitsstufe zugeordnet wurden ($N = 37$), zumindest im PLS-SEM einen signifikanten Einfluss der Noticingfähigkeit zum ersten MZP auf die Entwicklung der Reflexionsfähigkeit auf. Als ursächlich für diesen Zusammenhang kann vermutet werden, dass eine ausgeprägtere Noticingfähigkeit Lerngelegenheiten für die Reflexionsfähigkeit schafft, indem (zwischen beiden MZPs) die Elemente in beobachtetem oder durchgeführtem Unterricht identifiziert werden, deren Reflexion bedeutsam für die Unterrichtsqualität bzw. Lehrkräfteprofessionalität ist.

Dass dieser Zusammenhang nur für die oberen drei Fähigkeitsstufen beobachtbar ist, führt zu folgender Hypothese:

Es bedarf einer grundlegend ausgeprägten Reflexionsfähigkeit, damit (durch Noticing geschaffene) Lerngelegenheiten für Unterrichtsreflexion produktiv genutzt werden können.

Damit einher geht die Voraussetzung einer grundlegend ausgeprägten Fähigkeit zum Noticing, damit zentrale Situationen im Unterricht erkannt und somit die Lerngelegenheiten für Reflexion überhaupt geschaffen werden können. Offen bleibt allerdings, wie dieses Mindestmaß beider Fähigkeiten ausgebildet wird. Hier könnte ein Einfluss gewisser Dispositionen, wie dem hier berücksichtigten Professionswissen, vermutet werden: Ist Wissen vorhanden, kann dieses genutzt werden, um z. B. relevante Aspekte des Unterrichts zu erkennen oder fundierte alternative Handlungsoptionen zu entwickeln.

24.4 Erweitertes Modell zum Zusammenhang von Noticing und Unterrichtsreflexion

In Abbildung 24.1 sind die Ergebnisse der Analysen zur Struktur und Entwicklung der Reflexionsfähigkeit zusammenfassend entlang des Modells zum Zusammenhang von Noticing und Unterrichtsreflexion dargestellt. Vor dem Hintergrund des Entwicklungsziels der entwickelten Lernumgebung zur Unterrichtsreflexion, einen Ausgangspunkt für die Weiterentwicklung der Reflexionsfähigkeit zu schaffen, wird abschließend auf die Bedeutung dieser Ergebnisse für die Förderung der Reflexionsfähigkeit eingegangen. Aus der Untersuchung der Entwicklungszusammenhänge im CLP-Design lässt sich ableiten, dass nicht nur die unmittelbare Förderung der Reflexionsfähigkeit angestrebt werden sollte, sondern zusätzlich die Förderung der Noticingfähigkeit bedeutsame Reflexionen sowie die Entwicklung der Reflexionsfähigkeit unterstützen kann. So können drei Handlungsbereiche in der Förderung berücksichtigt werden:

- die (bewusste) Nutzung von Professionswissen, um relevante Aspekte des Unterrichts zu erkennen (Pfeil 1; Üben von Noticing),
- die Reflexion vorgegebener Situationen, um Bewertungen auf Grundlage das Professionswissen vorzunehmen und fundierte Handlungsalternativen und Konsequenzen abzuleiten (Pfeil 2; Üben von Reflexion)
- sowie darauf aufbauend die Reflexion selbst bestimmter bzw. erkannter Situationen (beides auf Grundlage des Professionswissens; Pfeile 1 und 2), um reale Reflexionssituationen abzubilden (Pfeil 3; gemeinsames Üben von Noticing und Reflexion).

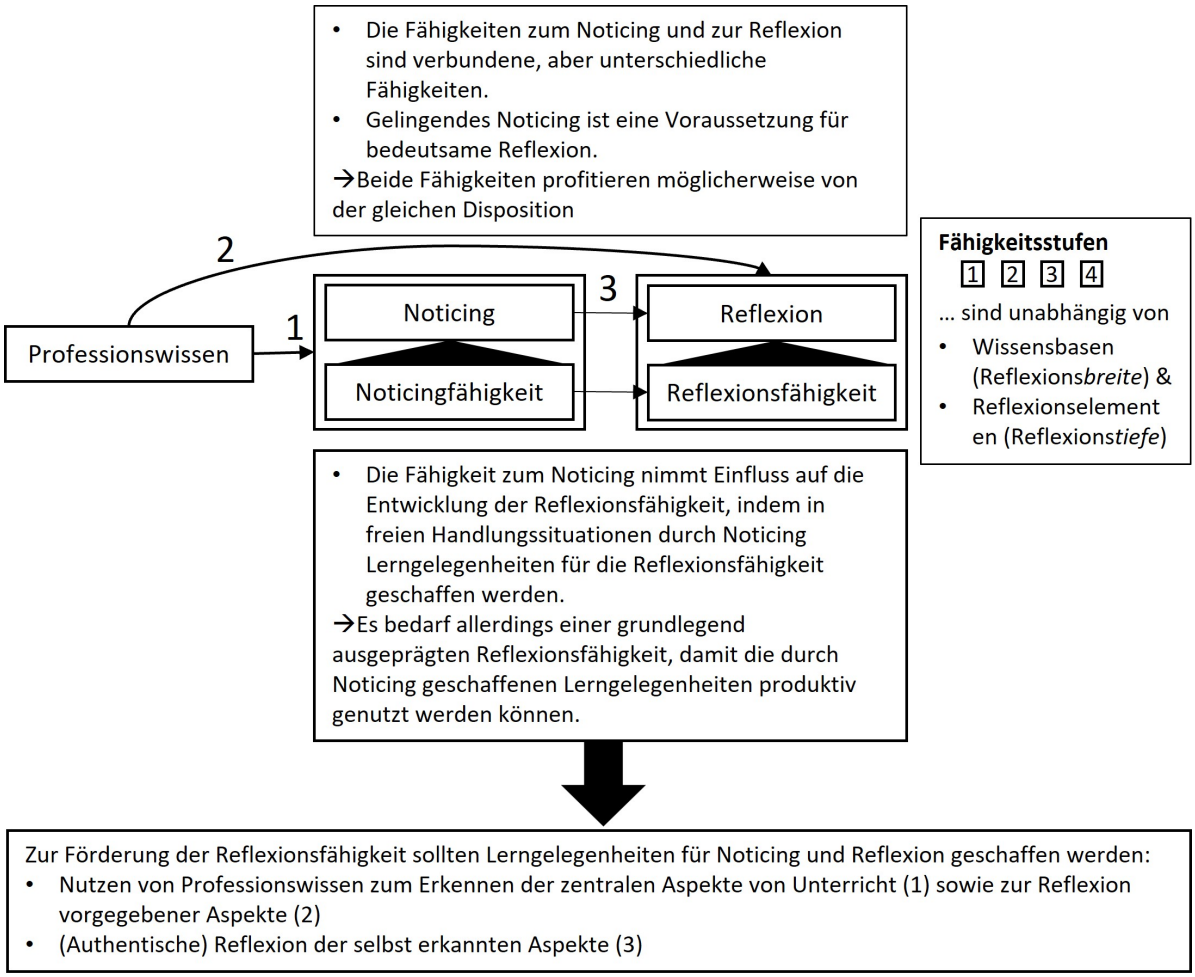


Abbildung 24.1: Erweitertes Modell zum Zusammenhang von Noticing und Unterrichtsreflexion

Die Bearbeitung des Reflexionstests spricht dabei insbesondere die Pfeile 1 und 2 an, wobei auch die Fokussierung auf die in den Videovignetten handelnde Lehrkraft eher der Perspektive von Lehramtsstudierenden entspricht (Stahnke & Blömeke, 2021). Für die Förderung gerade der Fähigkeit zum Noticing könnte darüber hinaus die Fokussierung auf das Handeln der Schüler*innen, wie sie bei erfahrenen Lehrkräften beobachtet werden kann, angeregt werden.

24.5 Limitationen der explorativen Untersuchungen

Die hier vorgestellten Ergebnisse der Untersuchungen zur Reflexionsfähigkeit und ihrer Entwicklung unter Berücksichtigung von Zusammenhängen zum Noticing sind mit gewissen Limitationen bezüglich der Erfassung der Noticingfähigkeit sowie den eingesetzten statistischen Methoden verbunden.

In den berichteten Analysen wurden die Fähigkeiten zum Noticing und zur Reflexion fokussiert; das Modell zum Zusammenhang von Noticing und Unterrichtsreflexion umfasst zusätzlich das Professionswissen als (angenommene) Disposition für beide Fähigkeiten. Für diese Zusammenhänge liegen allerdings keine belastbaren empirischen Daten vor. Für eine Teilstichprobe von $N = 33$ Studierenden im Praxissemester wurde zu mindestens einem der beiden MZPs ein Test zum fachdidaktischen Wissen bearbeitet (s. Kapitel 13.4); Bearbeitungen zu beiden MZPs liegen nur von $N = 14$ Personen vor.

Dass die Untersuchung von Zusammenhängen zwischen Reflexionsfähigkeit und fachdidaktischem Wissen für diese Teilstichprobe keine signifikanten Zusammenhänge lieferte (s. Kapitel 17.6), stützt das angenommene Modell zunächst nicht. Da allerdings die Stichprobe insbesondere für die Untersuchung von Entwicklungszusammenhängen zu gering ist und mit dem Test zum fachdidaktischen Wissen nur ein Teil des Professionswissens erfasst wird, besteht Bedarf für weitere Untersuchungen zur Klärung etwaiger Zusammenhänge zwischen dem Professionswissen und den Fähigkeiten zu Noticing und Reflexion.

Die Erfassung der Fähigkeit zum Noticing erfolgt im Rahmen des Reflexionstests mit Multiple-Choice-Aufgaben. Hier kann - analog zur Erfassung der Reflexion mit Multiple-Choice-Aufgaben - nicht ausgeschlossen werden, dass die Bereitstellung möglicher Antwortoptionen das Antwortverhalten gegenüber offenen Aufgabenformaten beeinflusst. Im Vergleich zu anderen Instrumenten zum Noticing oder der professionellen Unterrichtswahrnehmung sind die Antwortoptionen außerdem kurz und stichwortartig (vgl. Weyers et al., 2023), sodass ohne weiterführende Untersuchungen (z. B. Think-Aloud-Studien) nicht gesichert davon ausgegangen werden kann, dass die Antwortoptionen jeweils wie intendiert auf die Videovignette bezogen werden und darauf aufbauend die Testwerte valide als Maß der Noticingfähigkeit interpretiert werden können.

Zusätzlich ist einschränkend festzuhalten, dass die Fähigkeit zum Noticing in dieser Arbeit als eine der Reflexion vorausgehende Fähigkeit verstanden wird, die Erfassung aber nicht vollständig getrennt von Reflexionsprozessen erfolgt, da einerseits die Noticing-Aufgaben jeweils erst im Anschluss an die Videovignetten bearbeitet werden (und nicht wie bspw. im Instrument von Wöhlke (2020, S. 58f.) eine unmittelbare Kennzeichnung der bedeutsamen Aspekte der Vignette erfolgt) und andererseits eine möglicherweise lenkende Frage gestellt wird: „Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung (Lob und/oder Kritik) erhalten, um etwas für sich bzw. seinen Unterricht zu lernen?“. Das hier vorgegebene Ziel, etwas für den Unterricht zu lernen, dient zur Selektion der von den Teilnehmenden als bedeutsam empfundenen Aspekte, erfordert aber die (bewusste) Abwägung der Relevanz der einzelnen Aspekte.

Weiterhin sind die vorliegenden Stichprobengrößen für einige der verwendeten statistischen Methoden (die LPA und insbesondere die Cross-Lagged-Panel-Analysen)

als sehr gering einzustufen. Da die hier vorgestellten Analysen insgesamt explorativ angelegt sind, ist dies vorrangig mit der Einschränkung verbunden, dass lediglich große Effekte festgestellt werden können, dennoch sollte zur Bestätigung der Ergebnisse eine Erweiterung der Stichprobe angestrebt werden. Auch ein Einfluss der Zusammensetzung der Stichproben kann nicht ausgeschlossen werden; für alle Analysen wurden Gelegenheitsstichproben herangezogen. Während die Gesamtstichprobe (unter Ausschluss dreier Personen, die sich weder in der ersten noch zweiten Ausbildungsphase zur Physik-Lehrkraft befinden) von $N = 189$ Studierenden sich zwar aus Personen aus verschiedenen Ausbildungsphasen und von verschiedenen Ausbildungs- bzw. Universitätsstandorten zusammensetzt, besteht die Teilstichprobe für die Cross-Lagged-Panel-Analysen aus zwei möglicherweise sehr unterschiedlichen Gruppen: $N = 25$ Masterstudierenden im Praxissemester sowie $N = 16$ Bachelorstudierenden. Die gemeinsame Betrachtung dieser Gruppen erscheint einerseits sinnvoll, da so eher über den Studienfortschritt generalisierbare Ergebnisse erzielt werden können, andererseits können Spezifika der Gruppen wie insbesondere die deutlich unterschiedlichen Lerngelegenheiten für beide Gruppen zwischen den Messzeitpunkten (Praxissemester gegenüber rein universitärer Lehrveranstaltungen) so nicht berücksichtigt werden; eine getrennte Betrachtung beider Gruppen ist aufgrund der geringen Stichprobengröße nicht möglich.

Aus methodischer Perspektive bleibt außerdem sicherzustellen, dass die vorliegenden Aufgaben trotz der inhaltlich geschachtelten Struktur aufgrund der Bezüge zu den Videovignetten sowie insbesondere den reflektierten Aspekten tatsächlich - wie hier umgesetzt - als lokal stochastisch unabhängig angenommen werden können (vgl. Meschede, Steffensky, Wolters und Möller, 2015). Als strukturelle Voraussetzung dafür ist erfüllt, dass stets alle Aufgaben des Reflexionstests bearbeitet werden und z. B. keine Abhängigkeit der reflektierten Aspekte vom Noticing besteht. Auch die Beobachtung aus dem Reflexionsperformanztest von Kempin, Kulgemeyer und Schecker (2018), dass Bewertungen und Alternativen im freien Antwortformat weder stets in dieser Reihenfolge noch überhaupt gemeinsam auftreten, lässt die Vermutung zu, dass sowohl Bewertungen als auch Alternativen unabhängig voneinander benannt werden können. Unklar bleibt dennoch der Einfluss der inhaltlich engen Verknüpfung insbesondere zwischen den Aufgaben zu Bewertungen und Alternativen zu den reflektierten Aspekten.

Teil V

Schlussbemerkungen

25 Beitrag der Arbeit für die Erfassung und Förderung der Fähigkeit zur Unterrichtsreflexion

Reflexion gilt als essentieller Bestandteil des professionellen Handelns von Lehrkräften. Sie dient zur Entwicklung der Unterrichtsqualität und allgemein zur Entwicklung der Professionalität von Lehrkräften. Auch für Lehramtsstudierende ist die Reflexion bereits relevant: unmittelbar in Praxisphasen oder im Rahmen von, häufig ebenfalls im Kontext von Praxisphasen, bearbeiteten Reflexionsaufträgen, aber auch darüber hinaus in ihrer Funktion zur Relationierung von (im Studium erlernter) Theorie und Praxis. (Angehende) Lehrkräfte reflektieren allerdings überwiegend nicht auf hohen Qualitätsniveaus, sodass ein Bedarf zur Förderung der Fähigkeit auch schon im Studium unterstellt werden kann. Ausgangspunkt für die Förderung der Reflexionsfähigkeit kann die Einschätzung des individuellen Fähigkeitsstands sein. So kann vorliegendes Entwicklungspotential einerseits erkannt und andererseits anschließend zielgerichtet gefördert werden.

Mit der entwickelten Lernumgebung bestehend aus Reflexionstest und Assessment-Feedback mit Selbstlernmaterial liegt nun ein Verfahren vor, dass es (angehenden) Physiklehrkräften ermöglicht, eine fundierte Einschätzung ihrer Reflexionsfähigkeit vorzunehmen. Die Verwendung von MC-Aufgaben geht zwangsläufig mit einer Einschränkung der Handlungsfreiheit der Teilnehmenden während der Unterrichtsreflexion einher, die sich insbesondere darin niederschlägt, dass lediglich vorgegebene Aspekte reflektiert werden und z. B. Bewertungen und Alternativen nicht losgelöst von Begründungen erfasst werden können. Vorteile des entwickelten Testinstruments gegenüber anderen Formaten sind die Möglichkeit zum selbstständigen und anonymen Einsatz sowie zur automatisierten Auswertung und zur Einbettung der Testwerte in eine offenbar angemessene, d. h. für die Teilnehmenden verständliche und von ihnen akzeptierte Form des Assessment-Feedbacks, das der fundierten Selbsteinschätzung der Reflexionsfähigkeit dient.

Ergänzend zeigt sich auf Grundlage der explorativen Untersuchungen in Teil IV, dass mit dem evaluierten Reflexionstest verschiedene Stufen der Reflexionsfähigkeit detektiert werden können. Somit könnte eine Eignung des Reflexionstests auch für die Generierung von Erkenntnissen zur Entwicklung der Reflexionsfähigkeit, z. B. im Rahmen der Untersuchung von Lerngelegenheiten oder Einflussfaktoren auf

die Reflexionsfähigkeit, sowie zur Untersuchung von Zusammenhängen zwischen Reflexionsfähigkeit und Unterrichtsqualität angenommen werden. Außerdem scheinen Zusammenhänge zwischen den Fähigkeiten zur Unterrichtsreflexion und zum Noticing vorzuliegen. Dazu gehört insbesondere, dass gelingendes Noticing als Voraussetzung für bedeutsame Reflexion verstanden werden und zur Entwicklung der Reflexionsfähigkeit beitragen kann - zumindest, wenn die Reflexionsfähigkeit bereits grundlegend ausgeprägt ist. Als zugrundeliegender Mechanismus wird vermutet, dass durch Noticing Lerngelegenheiten für Reflexion geschaffen werden können, indem zentrale Elemente des Unterrichts für die Reflexion herangezogen werden. Bei der Förderung der Reflexionsfähigkeit erscheint daher ebenfalls die Berücksichtigung der Fähigkeit zum Noticing angeraten - entweder durch die (zusätzliche) explizite Förderung dieser Fähigkeit oder aber durch die Vorgabe des Noticing zur Fokussierung auf die Reflexion. Mit dem um die Ergebnisse der explorativen Untersuchungen und Hypothesen zur Förderung der Reflexionsfähigkeit erweiterten Modell zur Untersuchung des Zusammenhangs von Noticing und Unterrichtsreflexion liegt ein heuristisches Modell vor, dessen Annahmen in Folgestudien zur fundierten Weiterentwicklung der Theorie zur Entwicklung von Unterrichtsreflexion näher untersucht werden können.

Literaturverzeichnis

Abels, S. (2011). *LehrerInnen als "Reflective Practitioner": Reflexionskompetenz für einen demokratieförderlichen Naturwissenschaftsunterricht: Zugl.: Hamburg, Universität, FB Erziehungswissenschaften, Dissertation, 2010*. VS Verlag für Sozialwissenschaften.

Abou Baker El-Dib, M. (2007). Levels of reflection in action research. An overview and an assessment tool. *Teaching and Teacher Education*, *23*(1), 24–35. https://doi.org/10.1016/j.tate.2006.04.002

Adl-Amini, K., Jehle, M., McLean, P., Müller, S., & Seifert, A. (2022). Reflektierte Normativität, normierte Reflexivität? Möglichkeiten und Grenzen empirischer Zugänge. In C. Reintjes & I. Kunze (Hrsg.), *Reflexion und Reflexivität in Unterricht, Schule und Lehrer:innenbildung* (S. 21–38). Verlag Julius Klinkhardt.

Aeppli, J., & Lötscher, H. (2016). EDAMA - Ein Rahmenmodell für Reflexion. *Beiträge zur Lehrerinnen- und Lehrerbildung*, *34*(1), 78–97. https://doi.org/10.25656/01:13921

Ajzen, I. (1985). From Intentions to Actions: A Theory of Planned Behavior. In J. Kuhl & J. Beckmann (Hrsg.), *Action Control* (S. 11–39). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-69746-3_2

Ajzen, I. (2002 (überarbeitet 2006)). Constructing a TPB questionnaire: Conceptual and methodological considerations. https://people.umass.edu/aizen/pdf/tpb.measurement.pdf

Ajzen, I., & Madden, T. J. (1986). Prediction of goal-directed behavior: Attitudes, intentions, and perceived behavioral control. *Journal of Experimental Social Psychology*, *22*(5), 453–474. https://doi.org/10.1016/0022-1031(86)90045-4

Akogul, S., & Erisoglu, M. (2017). An Approach for Determining the Number of Clusters in a Model-Based Cluster Analysis. *Entropy*, *19*(9), 452. https://doi.org/10.3390/e19090452

Albu, C., & Lindmeier, A. (2023). Performance assessment in teacher education research—A scoping review of characteristics of assessment instruments in the DACH region. *Zeitschrift für Erziehungswissenschaft*, *26*(3), 751–778. https://doi.org/10.1007/s11618-023-01167-7

Aliakbari, M., Khany, R., & Adibpour, M. (2020). EFL teachers' reflective practice, job satisfaction, and school context variables: Exploring possible relationships. *TESOL Journal*, *11*(1), 1–20. https://doi.org/10.1002/tesj.461

Amador, J. M., & Weston, T. L. (2024). A Review of Analytic Frameworks for Noticing in Mathematics and Science: Comparing Noticing Frameworks Across Disciplines and over Time. *International Journal of Science and Mathematics Education*. https://doi.org/10.1007/s10763-024-10452-8

American Educational Research Association, American Psychological Association & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*.

Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J., & Wittrock, M. C. (Hrsg.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives* (Complete ed.). Longman.

Andrade, H., & Valtcheva, A. (2009). Promoting Learning and Achievement Through Self-Assessment. *Theory Into Practice*, *48*(1), 12–19. https://doi.org/10.1080/00405840802577544

Baer, M., Dörr, G., Fraefel, U., Kocher, M., Küster, O., Larcher, S., Müller, P., Sempert, W., & Wyss, C. (2007). Werden angehende Lehrpersonen durch das Studium kompetenter? Kompetenzaufbau und Standarderreichung in der berufswissenschaftlichen Ausbildung an drei Pädagogischen Hochschulen in der Schweiz und in Deutschland: Unterrichtswissenschaft 35 (2007) 1, S. 15-47. *Unterrichtswissenschaft*, *35*. https://doi.org/10.25656/01:5485

Bain, J. D., Ballantyne, R., Packer, J., & Mills, C. (1999). Using Journal Writing to Enhance Student Teachers' Reflectivity During Field Experience Placements. *Teachers and Teaching*, *5*(1), 51–73. https://doi.org/10.1080/1354060990050104

Ballard, K. K. (2006). *Using von Manen's Model to Assess Levels of Reflectivity Among Preservice Physical Education Teachers* [Dissertation]. Texas A&M University. https://oaktrust.library.tamu.edu/bitstream/handle/1969.1/4373/etd-tamu-2006B-KINE-Ballard.pdf?sequence=1

Bartels, H. (2018). *Entwicklung und Bewertung eines performanznahen Videovignettentests zur Messung der Erklärfähigkeit von Physiklehrkräften: Dissertation* (Bd. 264). Logos.

Bastian, A., Kaiser, G., Meyer, D., Schwarz, B., & König, J. (2022). Teacher noticing and its growth toward expertise: an expert–novice comparison with pre-service and in-service secondary mathematics teachers. *Educational Studies in Mathematics*, *110*(2), 205–232. https://doi.org/10.1007/s10649-021-10128-y

Bauer, J. (2022). A Primer to Latent Profile and Latent Class Analysis. In M. Goller, E. Kyndt, S. Paloniemi & C. Damşa (Hrsg.), *Methods for Researching Professional Learning and Development* (S. 243–268, Bd. 33). Springer International Publishing. https://doi.org/10.1007/978-3-031-08518-5_11

Baumert, J., & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, *9*(4), 469–520. https://doi.org/10.1007/s11618-006-0165-2

Baumert, J., & Kunter, M. (2011). Das Kompetenzmodell von COACTIV. In M. Kunter, J. Baumert & W. Blum (Hrsg.), *Professionelle Kompetenz von Lehrkräften* (S. 29–53). Waxmann.

Beauchamp, C. (2006). *Understanding reflection in teaching: A framework for analysing the literature*. Library and Archives Canada.

Behnke, K. (2016). *Umgang mit Feedback im Kontext Schule*. Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-10223-4

Berger, R., Kulgemeyer, C., & Lensing, P. (2019). Ein Multiple-Choice-Test zum konzeptuellen Verständnis der Kraftwirkung auf Ladungsträger in statischen elektrischen und magnetischen Feldern. *Zeitschrift für Didaktik der Naturwissenschaften*, *25*(1), 197–210. https://doi.org/10.1007/s40573-019-00099-2

Bilandzic, H. (2017). Lautes Denken. In L. Mikos & C. Wegener (Hrsg.), *Qualitative Medienforschung* (S. 406–413). UVK Verlagsgesellschaft mbH; UVK Lucius.

Bischoff, S., Brühwiler, C., & Baer, M. (2005). Videotest zur Erfassung «adaptiver Lehrkompetenz». *Beiträge zur Lehrerbildung*, *23*(3), 382–397. https://doi.org/10.25656/01:13584

Black, P., & Wiliam, D. (2018). Classroom assessment and pedagogy. *Assessment in Education: Principles, Policy & Practice*, *25*(6), 551–575. https://doi.org/10.1080/0969594X.2018.1441807

Blömeke, S., Gustafsson, J.-E., & Shavelson, R. J. (2015). Beyond Dichotomies. *Zeitschrift für Psychologie*, *223*(1), 3–13. https://doi.org/10.1027/2151-2604/a000194

Blömeke, S., Kaiser, G., König, J., & Jentsch, A. (2020). Profiles of mathematics teachers' competence and their relation to instructional quality. *ZDM*, *52*(2), 329–342. https://doi.org/10.1007/s11858-020-01128-y

Blömeke, S., Kaiser, G., & Lehmann, R. (Hrsg.). (2008). *Professionelle Kompetenz angehender Lehrerinnen und Lehrer: Wissen, Überzeugungen und Lerngelegenheiten deutscher Mathematikstudierender und -referendare ; erste Ergebnisse zur Wirksamkeit der Lehrerausbildung*. Waxmann.

Blömeke, S., Kaiser, G., & Lehmann, R. (Hrsg.). (2010). *TEDS-M 2008: Professionelle Kompetenz und Lerngelegenheiten angehender Mathematiklehrkräfte für die Sekundarstufe I im internationalen Vergleich*. Waxmann.

Blunk, O. (2020). *Unterstützung von Aspekten kollaborativer Reflexion durch Prompts in Communities of Practice* [Dissertation]. Technische Universität Clausthal. Fakultät für Mathematik/Informatik und Maschinenbau. https://dokumente.ub.tu-clausthal.de/servlets/MCRFileNodeServlet/clausthal_derivate_00001164/Db114426.pdf

Bogner, A., Littig, B., & Menz, W. (Hrsg.). (2014). *Interviews mit Experten*. Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-531-19416-5

Bohndick, C., Kohlmeyer, S., & Buhl, H. M. (2019). Anstoß und Unterstützung der professionellen Weiterentwicklung im Lehramtsstudium durch weiter-

entwicklungsorientiertes Online-Self-Assessment. *Lehrerbildung auf dem Prüfstand*, *12*(1), 19–36.

Bohnsack, R. (2020). *Professionalisierung in praxeologischer Perspektive: Zur Eigenlogik der Praxis in Lehramt, Sozialer Arbeit und Frühpädagogik* (Bd. 5355). Verlag Barbara Budrich.

Borowski, A., Neuhaus, B. J., Tepner, O., Wirth, J., Fischer, H. E., Leutner, D., Sandmann, A., & Sumfleth, E. (2010). Professionswissen von Lehrkräften in den Naturwissenschaften (ProwiN) – Kurzdarstellung des BMBF-Projekts. *Zeitschrift für Didaktik der Naturwissenschaften*, *16*, 341–349.

Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2004). The concept of validity. *Psychological review*, *111*(4), 1061–1071. https://doi.org/10.1037/0033-295X.111.4.1061

Boud, D., Keogh, R., & Walker, D. (1985). Promoting Reflection in Learning: a Model. In D. Boud, R. Keogh & D. Walker (Hrsg.), *Reflection* (S. 18–40). Kogan Page.

Brazeau, C., Boyd, L., & Crosson, J. (2002). Changing an Existing OSCE to a Teaching Tool: The Making of a Teaching OSCE. *Academic Medicine*, *77*(9), 932.

Brendel, N. (2017). *Reflexives Denken im Geographieunterricht: Eine empirische Studie zur Bestimmung von Schülerreflexion mithilfe von Weblogs im Kontext Globalen Lernens: Dissertation* (1. Aufl., Bd. 10). Waxmann.

Bromme, R. (1992). *Der Lehrer als Experte: Zur Psychologie des professionellen Wissens* (1. Aufl.). Huber.

Bromme, R. (1997). Kompetenzen, Funktionen und unterrichtliches Handeln des Lehrers. In F. E. Weinert (Hrsg.), *Psychologie des Unterrichts und der Schule* (S. 177–212). Hogrefe Verl. für Psychologie.

Brovelli, D., Bölsterli, K., Rehm, M., & Wilhelm, M. (2013). Erfassen professioneller Kompetenzen für den naturwissenschaftlichen Unterricht: Ein Vignettentest mit authentisch komplexen Unterrichtssituationen und offenem Antwortformat. *Unterrichtswissenschaft*, *41*(4), 306–329.

Brückmann, M., & Duit, R. (2014). Videobasierte Analyse unterrichtlicher Sachstrukturen. In D. Krüger, I. Parchmann & H. Schecker (Hrsg.), *Methoden in der naturwissenschaftsdidaktischen Forschung* (S. 189–201). Springer Berlin Heidelberg.

Brühwiler, C., Hollenstein, L., Affolter, B., Biedermann, H., & Oser, F. (2017). Welches Wissen ist unterrichtsrelevant? *Zeitschrift für Bildungsforschung*, *7*(3), 209–228. https://doi.org/10.1007/s35834-017-0196-1

Buber, R. (2009). Denke-Laut-Protokolle. In R. Buber & H. H. Holzmüller (Hrsg.), *Qualitative Marktforschung* (S. 555–568). Gabler. https://doi.org/10.1007/978-3-8349-9441-7_35

Buhl, H. M., Sagolla, N., & Bohndick, C. (2022). Die Förderung von Reflexionskompetenz und Reflexion durch die angeleitete Bearbeitung eines weiterentwicklungsorientierten Online-Self-Assessments. In N. Schaper & C. Vogelsang

(Hrsg.), *Themenheft: Erfassung und Förderung von Reflexionskompetenz in der Lehrerbildung* (S. 117–136). Verlag Empirische Pädagogik.

Bühner, M. (2011). *Einführung in die Test- und Fragebogenkonstruktion* (3., aktualisierte Auflage). Pearson Studium.

Buholzer, A., Baer, M., Zulliger, S., Torchetti, L., Ruelmann, M., Häfliger, A., & Lötscher, H. (2020). Formatives Assessment im alltäglichen Mathematikunterricht von Primarlehrpersonen: Häufigkeit, Dauer und Qualität. *Unterrichtswissenschaft*, *48*(4), 629–661. https://doi.org/10.1007/s42010-020-00083-7

Buholzer, A., & Brovelli, D. (2023). Einleitung. In A. Buholzer & D. Brovelli (Hrsg.), *Formatives Assessment* (S. 7–15). Waxmann Verlag GmbH.

Busker, M. (2014). Entwicklung eines Fragebogens zur Untersuchungen des Fachinteresses. In D. Krüger, I. Parchmann & H. Schecker (Hrsg.), *Methoden in der naturwissenschaftsdidaktischen Forschung* (S. 269–281). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-37827-0_22

Butler, D. L., & Winne, P. H. (1995). Feedback and Self-Regulated Learning: A Theoretical Synthesis. *Review of Educational Research*, *65*(3), 245–281. https://doi.org/10.3102/00346543065003245

Carlson, J., Daehler, K. R., Alonzo, A. C., Barendsen, E., Berry, A., Borowski, A., Carpendale, J., Kam Ho Chan, K., Cooper, R., Friedrichsen, P., Gess-Newsome, J., Henze-Rietveld, I., Hume, A., Kirschner, S., Liepertz, S., Loughran, J., Mavhunga, E., Neumann, K., Nilsson, P., … Wilson, C. D. (2019). The Refined Consensus Model of Pedagogical Content Knowledge in Science Education. In A. Hume, R. Cooper & A. Borowski (Hrsg.), *Repositioning Pedagogical Content Knowledge in Teachers' Knowledge for Teaching Science* (S. 77–94). Springer Singapore. https://doi.org/10.1007/978-981-13-5898-2_2

Casanova, D., Alsop, G., & Huet, I. (2021). Giving away some of their powers! Towards learner agency in digital assessment and feedback. *Research and practice in technology enhanced learning*, *16*(1), 20. https://doi.org/10.1186/s41039-021-00168-6

Cauet, E. (2016). *Testen wir relevantes Wissen? Zusammenhang zwischen dem Professionswissen von Physiklehrkräften und gutem und erfolgreichem Unterrichten: Dissertation* (Bd. 204). Logos.

Cauet, E., Liepertz, S., Borowski, A., & Fischer, H. E. (2015). Does it Matter What We Measure? Domain-specific Professional Knowledge of Physics Teachers. *Revue suisse des sciences de l'éducation*, *37*(3), 462–479.

Chan, K. K. H., & Hume, A. (2019). Towards a Consensus Model: Literature Review of How Science Teachers' Pedagogical Content Knowledge Is Investigated in Empirical Studies. In A. Hume, R. Cooper & A. Borowski (Hrsg.), *Repositioning Pedagogical Content Knowledge in Teachers' Knowledge for Teaching Science* (S. 3–76). Springer Singapore. https://doi.org/10.1007/978-981-13-5898-2_1

Chan, K. K. H., Xu, L., Cooper, R., Berry, A., & van Driel, J. H. (2021). Teacher noticing in science education: do you see what I see? *Studies in Science Education*, *57*(1), 1–44. https://doi.org/10.1080/03057267.2020.1755803

Cheng, E. W. L. (2019). Choosing between the theory of planned behavior (TPB) and the technology acceptance model (TAM). *Educational Technology Research and Development*, *67*(1), 21–37. https://doi.org/10.1007/s11423-018-9598-6

Cizek, G. J. (2010). An Introduction to Formative Assessment: History, Characteristics, and Challenges. In G. J. Cizek (Hrsg.), *Handbook of formative assessment* (S. 3–17). Routledge.

Clarà, M. (2015). What Is Reflection? Looking for Clarity in an Ambiguous Notion. *Journal of Teacher Education*, *66*(3), 261–271. https://doi.org/10.1177/0022487114552028

Coe, R. (1998). Can feedback improve teaching? A review of the social science literature with a view to identifying the conditions under which giving feedback to teachers will result in improved performance. *Research Papers in Education*, *13*(1), 43–66. https://doi.org/10.1080/0267152980130104

Coffey, A. M. (2014). Using Video to Develop Skills in Reflection in Teacher Education Students. *Australian Journal of Teacher Education*, *39*(9). https://doi.org/10.14221/ajte.2014v39n9.7

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159. https://doi.org/10.1037/0033-2909.112.1.155

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2. Aufl.). Routledge.

Combe, A., & Helsper, W. (2002). Professionalität. In H.-U. Otto, T. Rauschenbach & P. Vogel (Hrsg.), *Erziehungswissenschaft: Professionalität und Kompetenz* (S. 29–47). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-663-08029-9_3

Combe, A., & Kolbe, F.-U. (2008). Lehrerprofessionalität: Wissen, Können, Handeln. In W. Helsper & J. Böhme (Hrsg.), *Handbuch der Schulforschung* (S. 857–875). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-91095-6_35

Conner, M., & Norman, P. (2022). Understanding the intention-behavior gap: The role of intention strength. *Frontiers in psychology*, *13*, 1–16. https://doi.org/10.3389/fpsyg.2022.923464

Copeland, W. D., Birmingham, C., de La Cruz, E., & Lewin, B. (1993). The reflective practitioner in teaching: Toward a research agenda. *Teaching and Teacher Education*, *9*(4), 347–359. https://doi.org/10.1016/0742-051X(93)90002-X

Cramer, C., Harant, M., Merk, S., Drahmann, M., & Emmerich, M. (2019). Meta-Reflexivität und Professionalität im Lehrerinnnen- und Lehrerberuf. *Zeitschrift für Pädagogik*, *65*(3), 401–423. https://doi.org/10.25656/01:23949

Crasborn, F., Hennissen, P., Brouwer, N., Korthagen, F., & Bergen, T. (2008). Promoting versatility in mentor teachers' use of supervisory skills. *Teaching and Teacher Education*, *24*(3), 499–514. https://doi.org/10.1016/j.tate.2007.05.001

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281–302. https://doi.org/10.1037/h0040957

Crooks, T. J., Kane, M. T., & Cohen, A. S. (1996). Threats to the Valid Use of Assessments. *Assessment in Education*, *3*(3), 265–285.

Damköhler, J., Elsholz, M., & Trefzger, T. (2023). Förderung der Reflexionskompetenz im Lehr-Lern-Labor: Unterstützung der Entwicklung der Reflexionskompetenz Studierender in einem Lehr-Lern-Labor-Seminar. *PhyDid B - Didaktik Der Physik - Beiträge Zur DPG-Frühjahrstagung*, 195–198.

Dauber, H. (2006). Selbstreflexion im Zentrum pädagogischer Praxis. In H. Dauber & R. Zwiebel (Hrsg.), *Professionelle Selbstreflexion aus pädagogischer und psychoanalytischer Sicht* (S. 11–39). Klinkhardt.

Davis, E. A. (2003). Prompting Middle School Science Students for Productive Reflection: Generic and Directed Prompts. *Journal of the Learning Sciences*, *12*(1), 91–142. https://doi.org/10.1207/S15327809JLS1201_4

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, *13*(3), 319–340. https://doi.org/10.2307/249008

Day, C. (1993). Reflection: a necessary but not sufficient condition for professional development. *British Educational Research Journal*, *19*(1), 83–93.

DeLone, W. H., & McLean, E. R. (2003). The DeLone and McLean Model of Information Systems Success: A Ten-Year Update. *Journal of Management Information Systems*, *19*(4), 9–30. https://doi.org/10.1080/07421222.2003.11045748

Dewey, J. (1933). *How We Think: A Restatement of the Relation of Reflective Thinking to the Educative Process*. D.C. Heath & Co Publishers.

Dewey, J. (1910/2002). *Wie wir denken: Mit einem Nachwort neu herausgegeben von Rebekka Horlacher und Jürgen Oelkers* (Bd. 2). Verlag Pestalozzianum.

Dickmann, M. (2016). *Messung von Experimentierfähigkeiten: Validierungsstudien zur Qualität eines computerbasierten Testverfahrens: Dissertation* (Bd. 210). Logos.

Diez, M. E. (2010). It Is Complicated: Unpacking the Flow of Teacher Education's Impact on Student Learning. *Journal of Teacher Education*, *61*(5), 441–450. https://doi.org/10.1177/0022487110372927

Dittrich, A.-K. (2020). *Pädagogisches Wissen im LehrerInnenberuf: Rekonstruktive Befunde aus der schulischen Praxis: Dissertation*. Julius Klinkhardt.

Dollny, S. (2011). *Entwicklung und Evaluation eines Testinstruments zur Erfassung des fachspezifischen Professionswissens von Chemielehrkräften: Dissertation* (Bd. 127). Logos Verlag.

Döring, N., & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften* (5. vollständig überarbeitete, aktualisierte und erweiterte Auflage). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-41089-5

Dyment, J. E., & O'Connell, T. S. (2010). The Quality of Reflection in Student Journals: A Review of Limiting and Enabling Factors. *Innovative Higher Education*, *35*(4), 233–244. https://doi.org/10.1007/s10755-010-9143-y

Enkrott, P. (2021). *Entwicklung des fachlichen Wissens angehender Physiklehrkräfte: Dissertation, Universität Potsdam*. https://doi.org/10.25932/publishup-50040

Ericsson, K. A., & Simon, H. A. (1993). *Protocol Analysis*. The MIT Press. https://doi.org/10.7551/mitpress/5657.001.0001

Eßling, I., Sunder, C., Steffensky, M., & Meschede, N. (2024). Videobasierte Erfassung professioneller Unterrichtswahrnehmung: Vergleich offener und geschlossener Formate im Kontext von Lernunterstützung. *Lehrerbildung auf dem Prüfstand*, *17*(1), 127–149. https://doi.org/10.62350/LTIK9621

Felske, C., König, J., Kaiser, G., Klemenz, S., Ross, N., & Blömeke, S. (2020). Pädagogisches Wissen von berufstätigen Mathematiklehrkräften: Validierung der Konstruktrepräsentation im TEDS-M-Test. *Diagnostica*, *66*(2), 110–122. https://doi.org/10.1026/0012-1924/a000250

Fengler, J. (2010). Feedback als Interventions-Methode. *Gruppendynamik und Organisationsberatung*, *41*(1), 5–20. https://doi.org/10.1007/s11612-010-0100-0

Ferguson, P. (2011). Student perceptions of quality feedback in teacher education. *Assessment & Evaluation in Higher Education*, *36*(1), 51–62. https://doi.org/10.1080/02602930903197883

Field, A. (2009). *Discovering statistics using SPSS: (and sex and drugs and rock 'n' roll)* (3. ed.). Sage Publ. http://www.uk.sagepub.com/field3e/main.htm

Fischer, H. E., Neumann, K., Labudde, P., & Viiri, J. (2014). Theoretical Framework. In H. E. Fischer, P. Labudde, K. Neumann & J. Viiri (Hrsg.), *Quality of instruction in physics* (S. 13–30). Waxmann.

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention and behaviour: An introduction to theory and research*. Addison-Wesley.

Fisseni, H.-J. (2004). *Lehrbuch der psychologischen Diagnostik: Mit Hinweisen zur Intervention* (3., überarbeitete und erweiterte Auflage). Hogrefe.

Forster, E. (2014). Reflexivität. In C. Wulf & J. Zirfas (Hrsg.), *Handbuch Pädagogische Anthropologie* (S. 589–597). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-531-18970-3_54

Fraij, A. (2018). Skalendokumentation der Gießener Offensive Lehrerbildung zur Reflexionsbereitschaft. https://doi.org/10.22029/jlupub-17142

Fraij, A., & Kirschner, S. (2017). Die Messung reflexionsbezogenen Wissens von Studierenden des Lehramts und der Erziehungswissenschaft. https://aepf2017.de/data/abstracts.pdf

Frank, K. (2022). Empirischer Forschungsstand zur Validierung der CORA-Aufgaben. In K. Frank (Hrsg.), *Critical Online Reasoning* (S. 29–39). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-39698-5_4

Frey, A. (2006). Methoden und Instrumente zur Diagnose beruflicher Kompetenzen von Lehrkräften. Eine erste Standortbestimmung zu bereits publizierten Instrumenten. *Kompetenzen und Kompetenzentwicklung von Lehrerinnen und Lehrern*, *Zeitschrift für Pädagogik, Beiheft*(51), 30–46. https://doi.org/10.25656/01:7369

Friege, G., & Lind, G. (2004). Leistungsmessung im Leistungskurs. *Der Mathematische und Naturwissenschaftliche Unterricht - MNU*, *57*(5), 259–265.

Friesen, M., Kuntze, S., & Vogel, M. (2018). Videos, Texte oder Comics? Die Rolle des Vignettenformats bei der Erhebung fachdidaktischer Analysekompetenz zum Umgang mit Darstellungen im Mathematikunterricht. In J. Rutsch, M. Rehm, M. Vogel, M. Seidenfuß & T. Dörfler (Hrsg.), *Effektive Kompetenzdiagnose in der Lehrerbildung* (S. 153–177). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-20121-0_8

Futter, K. (2017). *Lernwirksame Unterrichtsbesprechungen im Praktikum. Nutzung von Lerngelegenheiten durch Lehramtsstudierende und Unterstützungsverhalten der Praxislehrpersonen.* Verlag Julius Klinkhardt : Bad Heilbrunn. https://doi.org/10.25656/01:13048

Gebhardt, A., & Jenert, T. (2011). Besseres Feedback, mehr Reflexion? Fertigkeiten und Einstellungen Studierender zum Bloggen in Praxisprojekten. *Wissensgemeinschaften. Digitale Medien – Öffnung und Offenheit in Forschung und Lehre*, *Medien in der Wissenschaft*. https://doi.org/10.25656/01:11669

Gess-Newsome, J. (1999). Pedagogical Content Knowledge: An Introduction and Orientation. In J. Gess-Newsome & N. G. Lederman (Hrsg.), *Examining pedagogical content knowledge* (S. 3–17). Kluwer Academic Publ.

Gibbs, G. (1988). *Learning by Doing: A Guide to Teaching and Learning Methods.* Oxford Brooks University.

Gilles Raiche & David Magis. (2022). nFactors: Parallel Analysis and Other Non Graphical Solutions to the Cattell Scree Test. https://CRAN.R-project.org/package=nFactors

Göbel, K., & Gösch, A. (2019). Die Nutzung kollegialer Reflexion von Unterrichtsvideos im Praxissemester. In M. Degeling, N. Franken, S. Freund, S. Greiten, D. Neuhaus & J. Schellenbach-Zell (Hrsg.), *Herausforderung Kohärenz: Praxisphasen in der universitären Lehrerbildung* (S. 277–288). Verlag Julius Klinkhardt. https://doi.org/10.25656/01:17281

Goffree, F., & Oonk, W. (1999). Educating Primary School Mathematics Teachers in the Netherlands: Back to the Classroom. *Journal of Mathematics Teacher Education*, *2*, 207–214.

Graf, D. (2007). Die Theorie des geplanten Verhaltens. In D. Krüger & H. Vogt (Hrsg.), *Theorien in der biologiedidaktischen Forschung* (S. 33–43). Springer.

Gramzow, Y. (2015). *Fachdidaktisches Wissen von Lehramtsstudierenden im Fach Physik: Modellierung und Testkonstruktion: Dissertation* (Bd. 181). Logos.

Gramzow, Y., Riese, J., & Reinhold, P. (2013). Modellierung fachdidaktikschen Wissens angehender Physiklehrkräfte. *Zeitschrift für Didaktik der Naturwissenschaften*, *19*, 7–30.

Gräsel, C., & Göbel, K. (2011). Unterrichtsqualität. In H. Reinders, H. Ditton, C. Gräsel & B. Gniewosz (Hrsg.), *Empirische Bildungsforschung* (S. 87–98). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-93021-3_8

Große, A., Szogs, M., Lamprecht, J., & Korneck, F. (2021). Rating der Qualität kollegialer Reflexionen im Prä-/Post Vergleich. In S. Habig (Hrsg.), *Naturwissenschaftlicher Unterricht und Lehrerbildung im Umbruch* (S. 677–680). Universität Duisburg-Essen.

Gruber, H., Heid, H., & Bauer, J. (2005). *Forschungsprogramm Kompetenzentwicklung im Arbeitsprozess: Die Rolle von Arbeitsplatz-Veränderungen und Reflexion für Workplace Learning* (Bd. 17). Universität Regensburg, Lehrstuhl für Lehr-Lern-Forschung.

Gruber, H., & Rehrl, M. (2005). *Praktikum statt Theorie? Eine Analyse relevanten Wissens zum Aufbau pädagogischer Handlungskompetenz* (Bd. 15). Universität Regensburg, Lehrstuhl für Lehr-Lern-Forschung.

Häcker, T. (2017). Grundlagen und Implikationen der Forderung nach Förderung von Reflexivität in der Lehrerinnen- und Lehrerbildung. In C. Berndt, T. H. Häcker & T. Leonhard (Hrsg.), *Reflexive Lehrerbildung revisited* (S. 21–45). Verlag Julius Klinkhardt.

Häcker, T. (2019). Reflexive Professionalisierung. Anmerkungen zu dem ambitionierten Anspruch, die Reflexionskompetenz angehender Lehrkräfte umfassend zu fördern. In M. Degeling, N. Franken, S. Freund, S. Greiten, D. Neuhaus & J. Schellenbach-Zell (Hrsg.), *Herausforderung Kohärenz: Praxisphasen in der universitären Lehrerbildung* (S. 81–96). Verlag Julius Klinkhardt. https://doi.org/10.25656/01:17267

Häcker, T. (2022). Reflexive Lehrer*innenbildung. Versuch einer Lokalisierung in pragmatischer Absicht. In C. Reintjes & I. Kunze (Hrsg.), *Reflexion und Reflexivität in Unterricht, Schule und Lehrer:innenbildung* (S. 94–114). Verlag Julius Klinkhardt. https://doi.org/10.35468/5969-06

Hadenfeldt, J. C., & Neumann, K. (2014). Die kognitive Validität von Ordered Multiple Choice Aufgaben zur Erfassung des Verständnisses von Materie. *Zeitschrift für Didaktik der Naturwissenschaften*, *20*(1), 57–68. https://doi.org/10.1007/s40573-014-0003-7

Hair, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M., Danks, N. P., & Ray, S. (2021). *Partial Least Squares Structural Equation Modeling (PLS-SEM) Using R*. Springer International Publishing. https://doi.org/10.1007/978-3-030-80519-7

Hall, L. A. (2018). Using Blogs to Support Reflection in Teacher Education. *Literacy Research and Instruction*, *57*(1), 26–43. https://doi.org/10.1080/19388071.2017.1367055

Hamel, C., & Viau-Guay, A. (2019). Using video to support teachers' reflective practice : A literature review. *Cogent Education*, *6*(1), 1–14. https://doi.org/10.1080/2331186X.2019.1673689

Harden, R. M. (1988). What is an OSCE? *Medical teacher*, *10*(1), 19–22. https://doi.org/10.3109/01421598809019321

Hartig, J., Frey, A., & Jude, N. (2012). Validität. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 143–171). Springer. https://doi.org/10.1007/978-3-642-20072-4_7

Hashweh, M. Z. (1987). Effects of subject-matter knowledge in the teaching of biology and physics. *Teaching and Teacher Education*, *3*(2), 109–120. https://doi.org/10.1016/0742-051X(87)90012-6

Hashweh, M. Z. (2005). Teacher pedagogical constructions: a reconfiguration of pedagogical content knowledge. *Teachers and Teaching*, *11*(3), 273–292. https://doi.org/10.1080/13450600500105502

Hattie, J. A. C. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement* (Reprinted.). Routledge.

Hattie, J. A. C., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, *77*(1), 81–112. https://doi.org/10.3102/003465430298487

Hatton, N., & Smith, D. (1995). Reflection in teacher education: Towards definition and implementation. *Teaching and Teacher Education*, *11*(1), 33–49. https://doi.org/10.1016/0742-051X(94)00012-U

Helmke, A. (2015). *Unterrichtsqualität und Lehrerprofessionalität: Diagnose, Evaluation und Verbesserung des Unterrichts ; Franz Emanuel Weinert gewidmet* (6. Aufl.). Klett; Kallmeyer.

Henderson, M., Phillips, M., Ryan, T., Boud, D., Dawson, P., Molloy, E., & Mahoney, P. (2019). Conditions that enable effective feedback. *Higher Education Research & Development*, *38*(7), 1401–1416. https://doi.org/10.1080/07294360.2019.1657807

Herzig, B., Grafe, S., & Reinhold, P. (2005). Reflexives Lernen mit digitalen Videos - zur Auseinandersetzung mit dem Theorie-Praxisverhältnis der Lehrerausbildung. In M. Welzel & H. Stadler (Hrsg.), *Nimm doch mal die Kamera!* (S. 45–64). Waxmann.

Herzog, W. (1995). Reflexive Praktika in der Lehrerinnen- und Lehrerbildung: Beiträge zur Lehrerbildung 13 (1995) 3, S. 253-273. *Beiträge zur Lehrerbildung*, *13*. https://doi.org/10.25656/01:13309

Hiebert, J., Morris, A. K., Berk, D., & Jansen, A. (2007). Preparing Teachers to Learn from Teaching. *Journal of Teacher Education*, *58*(1), 47–61.

Hillje, M. (2012). *Fachdidaktisches Wissen von Lehrerinnen und Lehrern und die didaktische Strukturierung von Mathematikunterricht: Fallanalysen zur kognitiven Aktivierung in Unterrichtsplanungen und realisiertem Unterricht*

[Dissertation]. Carl von Ossietzky Universität Oldenburg. Fakultät für Mathematik und Naturwissenschaften.

Hilzensauer, W. (2017). *Wie kommt die Reflexion in den Lehrberuf? Dissertation* (Bd. 644). Waxmann.

Himme, A. (2007). Gütekriterien der Messung: Reliabilität, Validität und Generalisierbarkeit. In S. Albers, D. Klapper, U. Konradt, A. Walter & J. Wolf (Hrsg.), *Methodik der empirischen Forschung: 2., überarbeitete und erweiterte Auflage* (S. 375–390). Gabler Verlag.

Hofmann, B. (2011). *Analyse von Blickbewegungen von Schülern beim Lesen von physikbezogenen Texten mit Bildern: Eye Tracking als Methodenwerkzeug in der physikdidaktischen Forschung: Dissertation* (Bd. 119). Logos.

Holtz, P., & Gnambs, T. (2017). The improvement of student teachers' instructional quality during a 15-week field experience: a latent multimethod change analysis. *Higher Education*, *74*(4), 669–685. https://doi.org/10.1007/s10734-016-0071-3

Hosenfeld, A. (2010). *Führt Unterrichtsrückmeldung zu Unterrichtsentwicklung? Die Wirkung von videographischer und schriftlicher Rückmeldung bei Lehrkräften der vierten Jahrgangsstufe: Dissertation* (Bd. 22). Waxmann.

Hubley, A. M., & Zumbo, B. D. (2011). Validity and the Consequences of Test Interpretation and Use. *Social Indicators Research*, *103*(2), 219–230. https://doi.org/10.1007/s11205-011-9843-4

Hullfish, H. G., & Philip, G. S. (1963). *Reflective Thinking: The Method of Education* (2. Auflage). Dodd, Mead & Company.

Jahncke, H., Berding, F., Porath, J., & Magh, K. (2018). Einfluss von Feedback auf die (Selbst-) Reflexion von Lehramtsstudierenden. *die hochschullehre*, *4*, 505–530. https://doi.org/10.3278/HSL1830W

Janík, T., Seidel, T., & Najvar, P. (2009). Introduction: On the Power of Video Studies in Investigating Teaching and Learning. In T. Janík & T. Seidel (Hrsg.), *The power of video studies in investigating teaching and learning in the classroom* (S. 7–19). Waxmann.

Jeffries, C., & Maeder, D. W. (2005). Using Vignettes To Build and Assess Teacher Understanding of Instructional Strategies. *The Professional Educator*, *27*(1 & 2), 17–28.

Jeffries, C., & Maeder, D. W. (2011). Comparing Vignette Instruction and Assessment Tasks to Classroom Observations and Reflections. *The Teacher Educator*, *46*(2), 161–175. https://doi.org/10.1080/08878730.2011.552667

Jeschke, C., Lindmeier, A., & Heinze, A. (2021). Vom Wissen zum Handeln: Vermittelt die Kompetenz zur Unterrichtsreflexion zwischen mathematischem Professionswissen und der Kompetenz zum Handeln im Mathematikunterricht? Eine Mediationsanalyse. *Journal für Mathematik-Didaktik*, *42*(1), 159–186. https://doi.org/10.1007/s13138-020-00171-2

Jonkisz, E., Moosbrugger, H., & Brandt, H. (2012). Planung und Entwicklung von Tests und Fragebogen. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie*

und Fragebogenkonstruktion (S. 27–74). Springer. https://doi.org/10.1007/978-3-642-20072-4_3

Jordans, M., & Riese, J. (2023). Unterrichtsplanung mit sinnvoller Einbettung digitaler Medien im PU. In H. van Vorst (Hrsg.), *Lernen, Lehren und Forschen in einer digital geprägten Welt* (S. 794–797). Universität Duisburg-Essen.

Jordans, M., Zeller, J., Große-Heilmann, R., & Riese, J. (2022). Weiterentwicklung eines physikdidaktischen Tests zum Online-Assessment. In S. Habig & H. van Vorst (Hrsg.), *Unsicherheit als Element von naturwissenschaftsbezogenen Bildungsprozessen* (S. 764–767). Universität Duisburg-Essen.

Kadluba, A., & Obersteiner, A. (2024). How to Assess Mathematics Teachers' TPACK? A Comparison Between Self-Reports and Knowledge Tests. *International Journal of Science and Mathematics Education*. https://doi.org/10.1007/s10763-024-10490-2

Kaiser, G., Busse, A., Hoth, J., König, J., & Blömeke, S. (2015). About the Complexities of Video-Based Assessments: Theoretical and Methodological Approaches to Overcoming Shortcomings of Research on Teachers' Competence. *International Journal of Science and Mathematics Education*, *13*(2), 369–387. https://doi.org/10.1007/s10763-015-9616-7

Kane, M. T. (1992). An Argument-Based Approach to Validity. *Psychological Bulletin*, *112*(3), 527–535.

Kane, M. T. (2001). Current Concerns in Validity Theory. *Journal of Educational Measurement*, *38*(4), 319–342. https://www.jstor.org/stable/1435453

Kane, M. T. (2009). Validating the Interpretations and Uses of Test Scores. In Robert W. Lissitz (Hrsg.), *The concept of validity* (S. 39–64). Information Age Pub.

Kane, M. T. (2012). Validating score interpretations and uses. *Language Testing*, *29*(1), 3–17. https://doi.org/10.1177/0265532211417210

Kane, M. T. (2013). Validating the Interpretations and Uses of Test Scores: Educational Testint Service. *Journal of Educational Measurement*, *50*(1), 1–73.

Kempin, M., Kulgemeyer, C., & Schecker, H. (2018). Reflexion von Physikunterricht: Ein Performanztest. In C. Maurer (Hrsg.), *Qualitätsvoller Chemie- und Physikunterricht - normative und empirische Dimensionen* (S. 867–870). Universität Regensburg.

Kempin, M., Kulgemeyer, C., & Schecker, H. (2019). Erste Einblicke in die Entwicklung der Reflexionsfähigkeit von Physiklehramtsstudierenden im Praxissemester. In C. Maurer (Hrsg.), *Naturwissenschaftliche Bildung als Grundlage für berufliche und gesellschaftliche Teilhabe* (S. 357–360). Universität Regensburg.

Kempin, M., Kulgemeyer, C., & Schecker, H. (2020). Wirkung von Professionswissen und Praxisphasen auf die Reflexionsfähigkeit von Physiklehramtsstu-

dierenden. In S. Habig (Hrsg.), *Naturwissenschaftliche Kompetenzen in der Gesellschaft von morgen* (S. 439–442). Universität Duisburg-Essen.

Kennedy, M. M. (1999). Approximations to Indicators of Student Outcomes. *Educational Evaluation and Policy Analysis*, *21*(4), 345–363. https://doi.org/10.2307/1164482

Kennedy, M. M. (2010). Attribution Error and the Quest for Teacher Quality. *Educational Researcher*, *39*(8), 591–598. https://doi.org/10.3102/0013189X10390804

Kersting, M. (2008). Zur Akzeptanz von Intelligenz- und Leistungstests. *Report Psychologie*, *33*, 420–433.

Kinnear, B., Kelleher, M., May, B., Sall, D., Schauer, D. P., Schumacher, D. J., & Warm, E. J. (2021). Constructing a Validity Map for a Workplace-Based Assessment System: Cross-Walking Messick and Kane. *Academic medicine : journal of the Association of American Medical Colleges*, *96*(7S), S64–S69. https://doi.org/10.1097/ACM.0000000000004112

Kirschner, S. (2013). *Modellierung und Analyse des Professionswissens von Physiklehrkräften: Dissertation* (Bd. 161). Logos.

Kirschner, S., Borowski, A., Fischer, H. E., Gess-Newsome, J., & von Aufschnaiter, C. (2016). Developing and evaluating a paper-and-pencil test to assess components of physics teachers' pedagogical content knowledge. *International Journal of Science Education*, *38*(8), 1343–1372. https://doi.org/10.1080/09500693.2016.1190479

Klewin, G., Köker, A., & Störtländer, J. C. (2020). Veranlasste und benotete Reflexion: ein unmögliches Prüfungsformat? Professionalisierung durch (Praxis-)Reflexion in der Lehrer_innenbildung? *Herausforderung Lehrer*innenbildung - Zeitschrift zur Konzeption, Gestaltung und Diskussion*, *3*(2), 108–121. https://doi.org/10.4119/hlz-2530

Klieme, E., & Hartig, J. (2007). Kompetenzkonzepte in den Sozialwissenschaften und im erziehungswissenschaftlichen Diskurs. In M. Prenzel, I. Gogolin & H.-H. Krüger (Hrsg.), *Kompetenzdiagnostik* (S. 11–29). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-90865-6_2

Klieme, E., & Leutner, D. (2006). Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen. Beschreibung eines neu eingerichteten Schwerpunktprogramms der DFG. *Zeitschrift für Pädagogik*, *52*(6), 876–903. https://doi.org/10.25656/01:4493

Klosa, O. (2016). *Online-Sehen: Qualität und Akzeptanz von Web-TV: Dissertation.* Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-15182-9

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, *119*(2), 254–284. https://doi.org/10.1037/0033-2909.119.2.254

Knievel, I., Lindmeier, A. M., & Heinze, A. (2015). Beyond Knowledge: Measuring Primary Teachers' Subject-Specific Competences in and for Teaching Mathematics with Items Based on Video Vignettes. *International Journal of Science and Mathematics Education*, *13*(2), 309–329. https://doi.org/10.1007/s10763-014-9608-z

Kobl, C. (2021). *Förderung und Erfassung der Reflexionskompetenz im Fach Chemie: Dissertation* (Bd. 312). Logos Verlag.

Kohler, B., & Wacker, A. (2013). Das Angebots-Nutzungs-Modell. Überlegungen zu Chancen und Grenzen des derzeit prominentesten Wirkmodells der Schul- und Unterrichtsforschung. *Die deutsche Schule*, *105*(3), 241–257. https://doi.org/10.25656/01:25755

Kollmann, T. (1998). *Akzeptanz innovativer Nutzungsgüter und -systeme*. Gabler Verlag. https://doi.org/10.1007/978-3-663-09235-3

König, J., & Blömeke, S. (2009). Pädagogisches Wissen von angehenden Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, *12*(3), 499–527. https://doi.org/10.1007/s11618-009-0085-z

König, J., Darge, K., Klemenz, S., & Seifert, A. (2018). Pädagogisches Wissen von Lehramtsstudierenden im Praxissemester: Ziel schulpraktischen Lernens? In J. König, M. Rothland & N. Schaper (Hrsg.), *Learning to Practice, Learning to Reflect?* (S. 287–323). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-19536-6_11

König, J., Kaiser, G., & Felbrich, A. (2012). Spiegelt sich pädagogisches Wissen in den Kompetenzselbsteinschätzungen angehender Lehrkräfte? Zum Zusammenhang von Wissen und Überzeugungen am Ende der Lehrerausbildung. *Zeitschrift für Pädagogik*, *58*(4), 476–491. https://doi.org/10.25656/01:10390

König, J., & Klemenz, S. (2015). Der Erwerb von pädagogischem Wissen bei angehenden Lehrkräften in unterschiedlichen Ausbildungskontexten: Zur Wirksamkeit der Lehrerausbildung in Deutschland und Österreich. *Zeitschrift für Erziehungswissenschaft*, *18*(2), 247–277. https://doi.org/10.1007/s11618-015-0623-9

König, J., & Kramer, C. (2016). Teacher professional knowledge and classroom management: on the relation of general pedagogical knowledge (GPK) and classroom management expertise (CME). *ZDM*, *48*(1-2), 139–151. https://doi.org/10.1007/s11858-015-0705-4

König, J., & Pflanzl, B. (2016). Is teacher knowledge associated with performance? On the relationship between teachers' general pedagogical knowledge and instructional quality. *European Journal of Teacher Education*, *39*(4), 419–436. https://doi.org/10.1080/02619768.2016.1214128

König, J., & Rothland, M. (2018). Das Praxissemester in der Lehrerbildung: Stand der Forschung und zentrale Ergebnisse des Projekts Learning to Practice. In J. König, M. Rothland & N. Schaper (Hrsg.), *Learning to*

Practice, Learning to Reflect? (S. 1–62). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-19536-6_1

Konrad, K. (2010). Lautes Denken. In G. Mey & K. Mruck (Hrsg.), *Handbuch qualitative Forschung in der Psychologie* (S. 476–490). VS Verlag. https://doi.org/10.1007/978-3-531-92052-8_34

Körkkö, M. (2021). Towards Meaningful Reflection and a Holistic Approach: Creating a Reflection Framework in Teacher Education. *Scandinavian Journal of Educational Research*, *65*(2), 258–275. https://doi.org/10.1080/00313831.2019.1676306

Körkkö, M., Kyrö-Ämmälä, O., & Turunen, T. (2016). Professional development through reflection in teacher education. *Teaching and Teacher Education*, *55*, 198–206. https://doi.org/10.1016/j.tate.2016.01.014

Korneck, F., Krüger, M., & Szogs, M. (2017). Professionswissen, Lehrerüberzeugungen und Unterrichtsqualität angehender Physiklehrkräfte unterschiedlicher Schulformen. In H. E. Fischer & E. Sumfleth (Hrsg.), *Professionelle Kompetenz von Lehrkräften der Chemie und Physik* (S. 113–134). Logos.

Korneck, F., Oettinghaus, L., Kunter, M., & Redinger, R. (2016). Überzeugungen und Handlungen von Lehrpersonen: Messung von Unterrichtsqualität in komplexitätsreduzierten Settings des Physikunterrichts. In U. Rauin, M. Herrle & T. Engartner (Hrsg.), *Videoanalysen in der Unterrichtsforschung* (S. 174–197). Beltz Juventa.

Korthagen, F., & Vasalos, A. (2005). Levels in reflection: core reflection as a means to enhance professional growth. *Teachers and Teaching*, *11*(1), 47–71. https://doi.org/10.1080/1354060042000337093

Korthagen, F. A. J. (1985). Reflective Teaching and Preservice Teacher Education in the Netherlands. *Journal of Teacher Education*, *36*(5), 11–15. https://doi.org/10.1177/002248718503600502

Kost, D. (2019). *Reflexionsprozesse von Studierenden des Physiklehramtes* [Dissertation]. Justus-Liebig-Universität. Institut für Didaktik der Physik. http://geb.uni-giessen.de/geb/volltexte/2020/15006/

Kost, D., Kirschner, S., & von Aufschnaiter, C. (2016). Reflexion im Schulpraktikum - Pilotstudie. In C. Maurer (Hrsg.), *Authentizität und Lernen - das Fach in der Fachdidaktik* (S. 461–463). Universität Regensburg.

Krauss, S., Neubrand, M., Blum, W., Baumert, J., Brunner, M., Kunter, M., & Jordan, A. (2008). Die Untersuchung des professionellen Wissens deutscher Mathematik-Lehrerinnen und -Lehrer im Rahmen der COACTIV-Studie. *Journal für Mathematik-Didaktik*, *29*(3/4), 223–258.

Krüger, D., & Riemeier, T. (2014). Die qualitative Inhaltsanalyse – eine Methode zur Auswertung von Interviews. In D. Krüger, I. Parchmann & H. Schecker (Hrsg.), *Methoden in der naturwissenschaftsdidaktischen Forschung* (S. 133–145). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-37827-0_11

Krumphals, I., & Haagen-Schützenhöfer, C. (2020). Förderung von Diagnosekompetenz im Ph-LA Studium - ein DBR-Projekt. In S. Habig (Hrsg.), *Naturwissenschaftliche Kompetenzen in der Gesellschaft von morgen* (S. 266–269). Universität Duisburg-Essen.

Kuckartz, U. (2018). *Qualitative Inhaltsanalyse: Methoden, Praxis, Computerunterstützung* (4. Auflage). Beltz Juventa.

Kulgemeyer, C. (2016). Explaning science: Using a "performance test" to assess physics teachers' explaining skills. In I. Eilks, S. Markic & B. Ralle (Hrsg.), *Science education research and practical work* (S. 205–214). Shaker Verlag.

Kulgemeyer, C. (2020). Forschungsverbund ProfiLe-P+ Professionskompetenz im Lehramtsstudium Physik – Entwicklungsverläufe und Performanz in unterrichtlichen Anforderungssituationen: Teilprojekt Auswirkungen fachlicher und fachdidaktischer Kompetenz auf die Analyse von Physikunterricht und Veränderung der Analyseperformanz durch Praxisphasen: Schlussbericht für ein Forschungsvorhaben im Förderprogramm dees Bundesministeriums für Bildung und Forschung Themenfeld "Kompetenzmodelle und Instrumente der Kompetenzerfassung im Hochschulsektor - Validierung und methodische Innovatiotnn (KoKoHs)".

Kulgemeyer, C., Borowski, A., Fischer, H. E., Gramzow, Y., Reinhold, P., Riese, J., Schecker, H., Tomczyszyn, E., & Walzer, M. (2012). ProfiLe-P – Professionswissen in der Lehramtsausbildung Physik: Vorstellung eines Forschungsverbundes. *PhyDid B - Didaktik Der Physik - Beiträge Zur DPG-Frühjahrstagung*, 1–6.

Kulgemeyer, C., Kempin, M., & Weißbach, A. (2021). Entwicklung von Professionswissen und Reflexionsfähigkeit im Praxissemester. In S. Habig (Hrsg.), *Naturwissenschaftlicher Unterricht und Lehrerbildung im Umbruch* (S. 262–265). Universität Duisburg-Essen.

Kulgemeyer, C., Kempin, M., Weißbach, A., Borowski, A., Buschhüter, D., Enkrott, P., Reinhold, P., Riese, J., Schecker, H., Schröder, J., & Vogelsang, C. (2021). Exploring the impact of pre-service science teachers' reflection skills on the development of professional knowledge during a field experience. *International Journal of Science Education*, *43*(18), 3035–3057. https://doi.org/10.1080/09500693.2021.2006820

Kulgemeyer, C., & Riese, J. (2018). From professional knowledge to professional performance: The impact of CK and PCK on teaching quality in explaining situations. *Journal of Research in Science Teaching*, *55*(10), 1393–1418. https://doi.org/10.1002/tea.21457

Kulgemeyer, C., Riese, J., Vogelsang, C., Buschhüter, D., Borowski, A., Weißbach, A., Jordans, M., Reinhold, P., & Schecker, H. (2023). How authenticity impacts validity: Developing a model of teacher education assessment and exploring the effects of the digitisation of assessment methods. *Zeitschrift für Erziehungswissenschaft*, 601–625. https://doi.org/10.1007/s11618-023-01154-y

Kulgemeyer, C., & Tomczyszyn, E. (2015). Physik erklären – Messung der Erklärensfähigkeit angehender Physiklehrkräfte in einer simulierten Unterrichtssituation. *Zeitschrift für Didaktik der Naturwissenschaften*, *21*(1), 111–126. https://doi.org/10.1007/s40573-015-0029-5

Kultusministerkonferenz. (2020). Bildungsstandards im Fach Physik für die Allgemeine Hochschulreife: Beschluss der Kultusministerkonferenz vom 18.06.2020. https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2020/2020_06_18-BildungsstandardsAHR_Physik.pdf

Kunter, M., Baumert, J., & Blum, W. (Hrsg.). (2011). *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV*. Waxmann.

Kunter, M., & Klusmann, U. (2010). Kompetenzmessung bei Lehrkräften - Methodische Herausforderungen. *Unterrichtswissenschaft*, *38*(1), 68–86.

Kunter, M., Klusmann, U., & Baumert, J. (2009). Professionelle Kompetenz von Mathematiklehrkräften: Das COACTIV-Modell. In O. Zlatkin-Troitschanskaia, K. Beck, D. Sembill, R. Nickolaus & R. H. Mulder (Hrsg.), *Lehrprofessionalität* (S. 153–165). Beltz.

Kunter, M., & Voss, T. (2011). Das Modell der Unterrichtsqualität in COACTIV: Eine multikriteriale Analyse. In M. Kunter, J. Baumert & W. Blum (Hrsg.), *Professionelle Kompetenz von Lehrkräften* (S. 85–113). Waxmann.

Labott, D., & Reintjes, C. (2022). Unvereinbarkeit von Bewertung und Reflexionsaufgaben in der Lehrer*innenbildung. In C. Reintjes & I. Kunze (Hrsg.), *Reflexion und Reflexivität in Unterricht, Schule und Lehrer:innenbildung* (S. 170–184). Verlag Julius Klinkhardt. https://doi.org/10.35468/5969-10

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in psychology*, *4*, 863. https://doi.org/10.3389/fpsyg.2013.00863

Lam, D. S. H., & Chan, K. K. H. (2020). Characterising pre-service secondary science teachers' noticing of different forms of evidence of student thinking. *International Journal of Science Education*, *42*(4), 576–597. https://doi.org/10.1080/09500693.2020.1717672

Larrivee, B. (2008). Development of a tool to assess teachers' level of reflective practice. *Reflective Practice*, *9*(3), 341–360. https://doi.org/10.1080/14623940802207451

Lee, H.-J. (2005). Understanding and assessing preservice teachers' reflective thinking. *Teaching and Teacher Education*, *21*(6), 699–715. https://doi.org/10.1016/j.tate.2005.05.007

Leijen, Ä., Valtna, K., Leijen, D. A., & Pedaste, M. (2012). How to determine the quality of students' reflections? *Studies in Higher Education*, *37*(2), 203–217. https://doi.org/10.1080/03075079.2010.504814

Lenske, G., & Lohse-Bossenz, H. (2023). Stichwort: Reflexion im Pädagogischen Kontext. *Zeitschrift für Erziehungswissenschaft*, *26*(5), 1133–1164. https://doi.org/10.1007/s11618-023-01196-2

Lenske, G., Thillmann, H., Wirth, J., Dicke, T., & Leutner, D. (2015). Pädagogisch-psychologisches Professionswissen von Lehrkräften: Evaluation des ProwiN-Tests. *Zeitschrift für Erziehungswissenschaft*, *18*(2), 225–245. https://doi.org/10.1007/s11618-015-0627-5

Leonhard, T., & Abels, S. (2017). Der „reflective practitioner". Leitfigur oder Kategorienfehler einer reflexiven Lehrerinnen- und Lehrerbildung? In C. Berndt, T. H. Häcker & T. Leonhard (Hrsg.), *Reflexive Lehrerbildung revisited* (S. 46–55). Verlag Julius Klinkhardt.

Leonhard, T., & Rihm, T. (2011). Erhöhung der Reflexionskompetenz durch Begleitveranstaltungen zum Schulpraktikum? Konzeption und Ergebnisse eines Pilotprojekts mit Lehramtsstudierenden. *Lehrerbildung auf dem Prüfstand*, *4*, 240–270. https://doi.org/10.25656/01:14722

Leuders, T. (2014). Modellierungen mathematischer Kompetenzen – Kriterien für eine Validitätsprüfung aus fachdidaktischer Sicht. *Journal für Mathematik-Didaktik*, *35*(1), 7–48. https://doi.org/10.1007/s13138-013-0060-3

Li, J., & De Luca, R. (2014). Review of assessment feedback. *Studies in Higher Education*, *39*(2), 378–393. https://doi.org/10.1080/03075079.2012.709494

Lienert, G. A., & Raatz, U. (1998). *Testaufbau und Testanalyse: 6. Auflage*. Beltz.

Liepertz, S., & Borowski, A. (2019). Testing the Consensus Model: relationships among physics teachers' professional knowledge, interconnectedness of content structure and student achievement. *International Journal of Science Education*, *41*(7), 890–910. https://doi.org/10.1080/09500693.2018.1478165

LimeSurvey GmbH. (n. d.). LimeSurvey: An Open Source survey tool.

Lipowsky, F. (2006). Auf den Lehrer kommt es an. Empirische Evidenzen für Zusammenhänge zwischen Lehrerkompetenzen, Lehrerhandeln und dem Lernen der Schüler. In C. Allemann-Ghionda & E. Terhart (Hrsg.), *Kompetenzen und Kompetenzentwicklung von Lehrerinnen und Lehrern. Ausbildung und Beruf: Zeitschrift für Pädagogik, Beiheft 51* (S. 47–70). Beltz. https://doi.org/10.25656/01:7370

Lissitz, R. W., & Samuelsen, K. (2007). A Suggested Change in Terminology and Emphasis Regarding Validity and Education. *Educational Researcher*, *36*(8), 437–448. https://doi.org/10.3102/0013189X07311286

Lohse-Bossenz, H., Holzberger, D., Kunina-Habenicht, O., Seidel, T., & Kunter, M. (2018). Wie fach(un)abhängig ist bildungswissenschaftliches Wissen? *Zeitschrift für Erziehungswissenschaft*, *21*(5), 991–1019. https://doi.org/10.1007/s11618-018-0817-z

Lohse-Bossenz, H., Schmitt, M., Lenske, G., & Gold, B. (2023). „The same or different?" – Effekte von Unterrichtsanalyse und Unterrichtsreflexion auf die Veränderung kognitiver und motivationaler Merkmale professioneller Lehrkompetenz. *Zeitschrift für Erziehungswissenschaft*, *26*(5), 1281–1300. https://doi.org/10.1007/s11618-023-01197-1

Lucke, D. (1995). *Akzeptanz*. VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-663-09234-6

Luft, J. A., Huang, Y., Singh, H., Ozen-Tasdemir, H., DeLuca, J., Watson, S., Ayano, E., & Whitworth, B. A. (2024). Exploring the noticing of science teachers: What teachers' notice and using video to capture teacher knowledge. *School Science and Mathematics*, *124*(3), 159–170. https://doi.org/10.1111/ssm.12623

Ma, L. (1999). *Knowing and teaching elementary mathematics: Teachers, understanding of fundamental mathematics in China and the United States*. Erlbaum.

Maddox, B., Bayliss, A. P., Fleming, P., Engelhardt, P. E., Edwards, S. G., & Borgonovi, F. (2018). Observing response processes with eye tracking in international large-scale assessments: evidence from the OECD PIAAC assessment. *European Journal of Psychology of Education*, *33*(3), 543–558. https://doi.org/10.1007/s10212-018-0380-2

Magnusson, S., Krajcik, J., & Borko, H. (1999). Nature, Sources, and Development of Pedagogical Content Knowledge for Science Teaching. In J. Gess-Newsome & N. G. Lederman (Hrsg.), *Examining pedagogical content knowledge* (S. 95–132, Bd. 6). Kluwer Academic Publ. https://doi.org/10.1007/0-306-47217-1_4

Marangunić, N., & Granić, A. (2015). Technology acceptance model: a literature review from 1986 to 2013. *Universal Access in the Information Society*, *14*(1), 81–95. https://doi.org/10.1007/s10209-014-0348-1

Marcos, J. M., Sanchez, E., & Tillema, H. H. (2011). Promoting teacher reflection: what is said to be done. *Journal of Education for Teaching*, *37*(1), 21–36. https://doi.org/10.1080/02607476.2011.538269

Mathieson, K. (1991). Predicting User Intentions: Comparing the Technology Acceptance Model with the Theory of Planned Behavior. *Information Systems Research*, *2*(3), 173–191. https://www.jstor.org/stable/23010882

Mayr, J., & Nieskens, B. (2015). Self-Assessments für angehende Lehrpersonen: Was sie bezwecken, was sie leisten und was man von ihnen nicht erwarten darf. *Das Hochschulwesen*, *63*(3+4), 81–86.

McAlpine, L., Weston, C., Beauchamp, C., Wiseman, C., & Beauchamp, J. (1999). Building a metacognitive model of reflection. *Higher Education*, *37*(2), 105–131. https://doi.org/10.1023/A:1003548425626

McCray, J. S., & Chen, J.-Q. (2012). Pedagogical Content Knowledge for Preschool Mathematics: Construct Validity of a New Teacher Interview. *Journal of Research in Childhood Education*, *26*(3), 291–307. https://doi.org/10.1080/02568543.2012.685123

Meier, J. (2022). *„LauRa – In der Lehramtsausbildung Reflexionskompetenz analysieren" Untersuchungen zum Verhältnis von Theorie, Praxis und Reflexion im Rahmen der Professionalisierung angehender Lehrpersonen* [Manteltext zur Dissertation zur Erlangung des akademischen Grades Doktor der Philosophie (Dr. phil.) im Fach Erziehungswissenschaft der Universität Paderborn]. Universität Paderborn. Fakultät für Kulturwissenschaften.

Meier, J., Vogelsang, C., Watson, C., & Schaper, N. (2022). Reflexion ist erzwungenes Nachdenken: Zusammenhänge zwischen dem Reflexionsverständnis Lehramtsstudierender und Facetten ihrer Reflexionsperformanz. In N. Schaper & C. Vogelsang (Hrsg.), *Themenheft: Erfassung und Förderung von Reflexionskompetenz in der Lehrerbildung* (S. 39–58). Verlag Empirische Pädagogik.

Meinhardt, C., Rabe, T., & Krey, O. (2018). Formulierung eines evidenzbasierten Validitätsarguments am Beispiel der Erfassung physikdidaktischer Selbstwirksamkeitserwartungen mit einem neu entwickelten Instrument. *Zeitschrift für Didaktik der Naturwissenschaften*, *24*(1), 131–150. https://doi.org/10.1007/s40573-018-0079-6

Melo, L., Cañada-Cañada, F., González-Gómez, D., & Jeong, J. S. (2020). Exploring Pedagogical Content Knowledge (PCK) of Physics Teachers in a Colombian Secondary School. *Education Sciences*, *10*(12), 101464. https://doi.org/10.3390/educsci10120362

Merkert, A., Lohse-Bossenz, H., Neuber, K., & Lenske, G. (2023). Selbstbezug in videobasierten Unterrichtsreflexionen von Lehramtsstudierenden im Bachelor. *Zeitschrift für Erziehungswissenschaft*, *26*(5), 1259–1280. https://doi.org/10.1007/s11618-023-01178-4

Meschede, N., Steffensky, M., Wolters, M., & Möller, K. (2015). Professionelle Wahrnehmung der Lernunterstützung im naturwissenschaftlichen Grundschulunterricht: Theoretische Beschreibung und empirische Erfassung. *Unterrichtswissenschaft*, *43*(4), 317–335.

Messick, S. (1989). Meaning and Values in Test Validation: The Science and Ethics of Assessment. *Educational Researcher*, *18*(2), 5–11. https://doi.org/10.3102/0013189X018002005

Messick, S. (1995). Validity of Psychological Assessment: Validation of Inferences From Persons' Responses and Performances as Scientific Inquiry Into Score Meaning. *American Psychologist*, *50*(9), 741–749.

Mientus, L. (2023). *Reflexion und Reflexivität* [Dissertation]. Universität Potsdam. https://doi.org/10.25932/publishup-61000

Mientus, L., Wulff, P., Nowak, A., & Borowski, A. (2021). ReFeed: computerunterstütztes Feedback zu Reflexionstexten. In M. Kubsch, S. Sorge, J. Arnold & N. Graulich (Hrsg.), *Lehrkräftebildung neu gedacht* (S. 160–165). Waxmann.

Mientus, L., Wulff, P., Nowak, A., & Borowski, A. (2023). Fast-and-frugal means to assess reflection-related reasoning processes in teacher training—Development and evaluation of a scalable machine learning-based metric. *Zeitschrift für Erziehungswissenschaft*, *26*(3), 677–702. https://doi.org/10.1007/s11618-023-01166-8

Miller, G. E. (1990). The Assessment of Clinical Skills/Competence/Performance. *Academic Medicine*, *65*(9), 63–67.

Molloy, E., Boud, D., & Henderson, M. (2020). Developing a learning-centred framework for feedback literacy. *Assessment & Evaluation in Higher Education*, *45*(4), 527–540. https://doi.org/10.1080/02602938.2019.1667955

Moore-Russo, D. A., & Wilsey, J. N. (2014). Delving into the meaning of productive reflection: A study of future teachers' reflections on representations of teaching. *Teaching and Teacher Education*, *37*, 76–90. https://doi.org/10.1016/j.tate.2013.10.002

Moosbrugger, H., & Kelava, A. (2012). Qualitätsanforderungen an einen psychologischen Test (Testgütekriterien). In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 7–26). Springer.

Mühlhausen, U. (Hrsg.). (2006). *Unterrichten lernen mit Gespür: Szenarien für eine multimedial gestützte Analyse und Reflexion von Unterricht ; mit DVD-Beilage mit multimedialen Unterrichtsdokumenten* (2., überarb. Aufl.). Schneider Verl. Hohengehren.

Müller, A., & Ditton, H. (2014). Feedback: Begriff, Formen und Funktionen. In H. Ditton & A. Müller (Hrsg.), *Feedback und Rückmeldungen* (S. 11–28). Waxmann.

Müller-Böling, D., & Müller, M. (1986). *Akzeptanzfaktoren der Bürokommunikation* (Bd. 17). Oldenbourg.

Narciss, S. (2006). *Informatives tutorielles Feedback: Entwicklungs- und Evaluationsprinzipien auf der Basis instruktionspsychologischer Erkenntnisse* (Bd. 56). Waxmann.

Naumann, A., Musow, S., Aichele, C., Hochweber, J., & Hartig, J. (2019). Instruktionssensitivität von Tests und Items. *Zeitschrift für Erziehungswissenschaft*, *22*(1), 181–202. https://doi.org/10.1007/s11618-018-0832-0

Neuber, K., & Göbel, K. (2018). Schülerrückmeldungen zum Unterricht und Unterrichtsreflexion. https://doi.org/10.17185/duepublico/46591

Neuber, K., & Göbel, K. (2019). Reflexion im Praxissemester. Ein Forschungskonzept unter Rückgriff auf Schülerrückmeldungen zum Unterricht. In M. Degeling, N. Franken, S. Freund, S. Greiten, D. Neuhaus & J. Schellenbach-Zell (Hrsg.), *Herausforderung Kohärenz: Praxisphasen in der universitären Lehrerbildung* (S. 302–311). Verlag Julius Klinkhardt. https://doi.org/10.25656/01:17283

Neumann, K. (2013). Mit welchem Auflösungsgrad können Kompetenzen modelliert werden? In welcher Beziehung stehen Modelle zueinander, die Kompetenz in einer Domäne mit unterschiedlichem Auflösungsgrad beschreiben? *Zeitschrift für Erziehungswissenschaft*, *16*(S1), 35–39. https://doi.org/10.1007/s11618-013-0382-4

Neumann, K., Fischer, H. E., Labudde, P., & Viiri, J. (2014). Design of the Study. In H. E. Fischer, P. Labudde, K. Neumann & J. Viiri (Hrsg.), *Quality of instruction in physics* (S. 31–48). Waxmann.

Neumann, K., Kind, V., & Harms, U. (2019). Probing the amalgam: the relationship between science teachers' content, pedagogical and pedagogical content

knowledge. *International Journal of Science Education*, *41*(7), 847–861. https://doi.org/10.1080/09500693.2018.1497217

Neuweg, G. H. (2005). Emergenzbedingungen pädagogischer Könnerschaft. In H. Heid & C. Harteis (Hrsg.), *Verwertbarkeit* (S. 205–228). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-663-07736-7_10

Neuweg, G. H. (2014). Das Wissen der Wissensvermittler: Problemstellungen, Befunde und Perspektiven der Forschung zum Lehrerwissen. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (S. 583–614). Waxmann.

Neuweg, G. H. (2015a). *Das Schweigen der Könner: Gesammelte Schriften zum impliziten Wissen*. Waxmann.

Neuweg, G. H. (2015b). Kontextualisierte Kompetenzmessung. Eine Bilanz zu aktuellen Konzeptionen und forschungsmethodischen Zugängen. *Zeitschrift für Pädagogik*, *61*(3), 377–383. https://doi.org/10.25656/01:15368

Neuweg, G. H. (2020). Implizites Wissen in der Lehrerinnen- und Lehrerbildung. In C. Cramer, J. König, M. Rothland & S. Blömeke (Hrsg.), *Handbuch Lehrerinnen- und Lehrerbildung* (S. 764–769). Verlag Julius Klinkhardt. https://doi.org/10.35468/hblb2020-093

Newton, P. E. (2012). Clarifying the Consensus Definition of Validity. *Measurement: Interdisciplinary Research & Perspective*, *10*(1-2), 1–29. https://doi.org/10.1080/15366367.2012.669666

Nieminen, J. H., & Carless, D. (2023). Feedback literacy: a critical review of an emerging concept. *Higher Education*, *85*, 1381–1400. https://doi.org/10.1007/s10734-022-00895-9

Noushad, B., van Gerven, P. W. M., & de Bruin, A. B. H. (2024). Twelve tips for applying the think-aloud method to capture cognitive processes. *Medical teacher*, *46*(7), 892–897. https://doi.org/10.1080/0142159X.2023.2289847

Nowak, A. (2023). *Untersuchung der Qualität von Selbstreflexionstexten zum Physikunterricht: Entwicklung des Reflexionsmodells REIZ* (Bd. 371). Logos Verlag.

Nowak, A., Ackermann, P., & Borowski, A. (2018). Rahmenthema "Reflexion" im Praxissemester Physik. In A. Borowski, H. Prechtl & A. Ehlert (Hrsg.), *PSI-Potsdam* (S. 217–230). Universitätsverlag Potsdam.

Nowak, A., Kempin, M., Kulgemeyer, C., & Borowski, A. (2019). Reflexion von Physikunterricht. In C. Maurer (Hrsg.), *Naturwissenschaftliche Bildung als Grundlage für berufliche und gesellschaftliche Teilhabe* (S. 838–841). Universität Regensburg.

Nowak, A., Liepertz, S., & Borowski, A. (2018). Reflexionskompetenz von Praxissemesterstudierenden im Fach Physik. In C. Maurer (Hrsg.), *Qualitätsvoller Chemie- und Physikunterricht - normative und empirische Dimensionen* (S. 142–145). Universität Regensburg.

Ohle, A., Fischer, H. E., & Kauertz, A. (2011). Der Einfluss des physikalischen Fachwissens von Primarstufenlehrkräften auf Unterrichtsgestaltung und

Schülerleistung. *Zeitschrift für Didaktik der Naturwissenschaften*, *17*, 357–389.

Olbrecht, T. (2010). *Akzeptanz von E-Learning. Eine Auseinandersetzung mit dem Technologieakzeptanzmodell zur Analyse individueller und sozialer Einflussfaktoren* [Dissertation]. Friedrich-Schiller-Universität. Fakultät für Sozial- und Verhaltenswissenschaften. https://www.db-thueringen.de/servlets/MCRFileNodeServlet/dbt_derivate_00021996/Olbrecht/Dissertation.pdf

Ollesch, J., Dörfler, T., & Vogel, M. (2018). Die inhaltliche Validierung von Unterrichtsvignetten durch eine mehrstufige Expertenbefragung. In J. Rutsch, M. Rehm, M. Vogel, M. Seidenfuß & T. Dörfler (Hrsg.), *Effektive Kompetenzdiagnose in der Lehrerbildung* (S. 129–151). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-20121-0_7

Oner, D., & Adadan, E. (2011). Use of Web-Based Portfolios as Tools for Reflection in Preservice Teacher Education. *Journal of Teacher Education*, *62*(5), 477–492. https://doi.org/10.1177/0022487111416123

Oser, F., & Baeriswyl, F. J. (2001). Choreographies of Teaching: Bridging Instruction to Learning. In V. Richardson (Hrsg.), *Handbook of research on teaching* (S. 1031–1065). American Educational Research Association.

Oser, F., Heinzer, S., & Salzmann, P. (2010). Die Messung der Qualität von professionellen Kompetenzprofilen von Lehrpersonen mit Hilfe der Einschätzung von Filmvignetten. *Unterrichtswissenschaft*, *38*(1), 5–28.

Panadero, E., Jonsson, A., & Botella, J. (2017). Effects of self-assessment on self-regulated learning and self-efficacy: Four meta-analyses. *Educational Research Review*, *22*, 74–98. https://doi.org/10.1016/j.edurev.2017.08.004

Park, S., & Chen, Y.-C. (2012). Mapping out the integration of the components of pedagogical content knowledge (PCK): Examples from high school biology classrooms. *Journal of Research in Science Teaching*, *49*(7), 922–941. https://doi.org/10.1002/tea.21022

Parkin, H. J., Hepplestone, S., Holden, G., Irwin, B., & Thorpe, L. (2012). A role for technology in enhancing students' engagement with feedback. *Assessment & Evaluation in Higher Education*, *37*(8), 963–973. https://doi.org/10.1080/02602938.2011.592934

Pellegrino, J. W., Chudowsky, N., & Glaser, R. (Hrsg.). (2001). *Knowing what students know: The science and design of educational*. National Academies Press.

Peterson, P. L., & Comeaux, M. A. (1987). Teachers' schemata for classroom events: The mental scaffolding of teachers' thinking during classroom instruction. *Teaching and Teacher Education*, *3*(4), 319–331. https://doi.org/10.1016/0742-051X(87)90024-2

Plöger, W., Scholl, D., & Seifert, A. (2015). Analysekompetenz - ein zweidimensionales Konstrukt?! *Unterrichtswissenschaft*, *43*(2), 166–184. https://www.digizeitschriften.de/dms/img/?PID=PPN513613439_0043%7CLOG_0025

Poldner, E., van der Schaaf, M., Simons, P. R.-J., van Tartwijk, J., & Wijngaards, G. (2014). Assessing student teachers' reflective writing through quantitative content analysis. *European Journal of Teacher Education*, *37*(3), 348–373. https://doi.org/10.1080/02619768.2014.892479

Pollmanns, M., Griewatz, H.-P., Hüning, R., Jornitz, S., Kabel, S., Leser, C., & Mayer, B. (2022). Wie (angehende) Lehrpersonen über Unterricht nachdenken bzw. nachdenken sollen. Professionalisierungstheoretische Analysen von Reflexionskrisen. In C. Reintjes & I. Kunze (Hrsg.), *Reflexion und Reflexivität in Unterricht, Schule und Lehrer:innenbildung* (S. 58–76). Verlag Julius Klinkhardt.

Pospeschill, M. (2010). *Testtheorie, Testkonstruktion, Testevaluation*. utb GmbH. https://doi.org/10.36198/9783838534312

Postholm, M. B. (2008). Teachers developing practice: Reflection as key activity. *Teaching and Teacher Education*, *24*(7), 1717–1728. https://doi.org/10.1016/j.tate.2008.02.024

Pultorak, E. G. (1996). Followling the Developmental Process of Reflection in Novice Teachers: Three Years of Investigation. *Journal of Teacher Education*, *47*(4), 283–291. https://doi.org/10.1177/0022487196474006

R Core Team. (2022). R: A Language and Environment for Statistical Computing. https://www.R-project.org

Raaflaub, M., Wyss, C., & Hüsler, N. (2019). Kollegiale Unterrichtsreflexion im Lehramtsstudium. *Journal für LehrerInnenbildung*, *19*(3), 50–57. https://doi.org/10.25656/01:18149

Radkowitsch, A., Sommerhoff, D., Nickl, M., Codreanu, E., Ufer, S., & Seidel, T. (2023). Exploring the diagnostic process of pre-service teachers using a simulation – A latent profile approach. *Teaching and Teacher Education*, *130*, 104172. https://doi.org/10.1016/j.tate.2023.104172

Rasch, B., Friese, M., Hofmann, W., & Naumann, E. (2014a). *Quantitative Methoden 1*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-43524-3

Rasch, B., Friese, M., Hofmann, W., & Naumann, E. (2014b). *Quantitative Methoden 2*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-43548-9

Rath, V. (2017). *Diagnostische Kompetenz von angehenden Physiklehrkräften. Modellierung, Testinstrumentenentwicklung und Erhebung der Performanz bei der Diagnose von Schülervorstellungen in der Mechanik: Modellierung, Testinstrumententwicklung und Erhebung der Performanz bei der Diagnose von Schülervorstellungen in der Mechanik: Dissertation* (Bd. 220). Logos.

Reagan, E. M., Schram, T., McCurdy, K., Chang, T.-H., & Evans, C. M. (2016). Politics of policy: Assessing the implementation, impact, and evolution of the Performance Assessment for California Teachers (PACT) and edTPA. *Education Policy Analysis Archives*, *24*, 9. https://doi.org/10.14507/epaa.24.2176

Reckase, M. D. (1998). Consequential Validity From the Test Developer's Perspective. *Educational Measurement: Issues and Practice*, *17*(2), 13–16. https://doi.org/10.1111/j.1745-3992.1998.tb00827.x

Reh, S. (2004). Abschied von der Profession, von Professionalität oder vom Professionellen? Theorien und Forschungen zur Lehrerprofessionalität: Zeitschrift für Pädagogik 50 (2004) 3, S. 358-372. *Zeitschrift für Pädagogik*, *50*. https://doi.org/10.25656/01:4815

Rehm, M., & Bölsterli, K. (2014). Entwicklung von Unterrichtsvignetten. In D. Krüger, I. Parchmann & H. Schecker (Hrsg.), *Methoden in der naturwissenschaftsdidaktischen Forschung* (S. 213–225). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-37827-0_18

Reinders, H. (2006). Kausalanalysen in der Längsschnittforschung. Das Crossed-Lagged-Panel Design: Diskurs Kindheits- und Jugendforschung 1 (2006) 4, S. 569-587. *Diskurs Kindheits- und Jugendforschung*, *1*. https://doi.org/10.25656/01:1001

Reintjes, C., & Kunze, I. (2022). Editorial: Reflexion und Reflexivität in Unterricht, Schule und Lehrer:innenbildung. In C. Reintjes & I. Kunze (Hrsg.), *Reflexion und Reflexivität in Unterricht, Schule und Lehrer:innenbildung* (S. 9–16). Verlag Julius Klinkhardt. https://doi.org/10.35468/5969-01

Riel, M. (2022). *Empirische Überprüfung eines Verfahrens zur Messung von Reflexion bei Lehramtsstudierenden* [Dissertation]. Universität Passau. Philosophische Fakultät.

Riese, J. (2009). *Professionelles Wissen und professionelle Handlungskompetenz von (angehenden) Physiklehrkräften: Dissertation* (Bd. 97). Logos.

Riese, J., Gramzow, Y., & Reinhold, P. (2017). Die Messung fachdidaktischen Wissens bei Anfängern und Fortgeschrittenen im Lehramtsstudiengang Physik. *Zeitschrift für Didaktik der Naturwissenschaften*, *23*(1), 99–112. https://doi.org/10.1007/s40573-017-0059-2

Riese, J., Kulgemeyer, C., Zander, S., Borowski, A., Fischer, H. E., Gramzow, Y., Reinhold, P., Schecker, H., & Tomczyszyn, E. (2015). Modellierung und messung des Professionswissens in der Lehramtsausbildung Physik. In S. Blömeke & O. Zlatkin-Troitschanskaia (Hrsg.), *Modellierung und Messung des Professionswissens in der Lehramtsausbildung Physik* (S. 55–79). Beltz Juventa.

Riese, J., & Reinhold, P. (2008). Entwicklung und Validierung eines Instruments zur Messung professioneller Handlungskompetenz bei (angehenden) Physiklehrkräften. *Lehrerbildung auf dem Prüfstand*, *1*(2), 625–640. https://doi.org/10.25656/01:15868

Riese, J., & Reinhold, P. (2010). Empirische Erkenntnisse zur Struktur professioneller Handlungskompetenz von angehenden Physiklehrkräften. *Zeitschrift für Didaktik der Naturwissenschaften*, *16*, 167–187.

Riese, J., & Reinhold, P. (2012). Die professionelle Kompetenz angehender Physiklehrkräfte in verschiedenen Ausbildungsformen. *Zeitschrift für Erziehungswissenschaft*, *15*(1), 111–143. https://doi.org/10.1007/s11618-012-0259-y

Riese, J., Vogelsang, C., Schröder, J., Borowski, A., Kulgemeyer, C., Reinhold, P., & Schecker, H. (2022). Entwicklung von Unterrichtsplanungsfähigkeit im Fach Physik: Welchen Einfluss hat Professionswissen? *Zeitschrift für Erziehungswissenschaft*, *25*(4), 843–867. https://doi.org/10.1007/s11618-022-01112-0

Rogge, T. (2023). Reflexion – besser digital? Ergebnisse einer Studie zur onlinebasierten Videoannotation zur Reflexion eigenen Unterrichts im Praxissemester. In L. Mientus, C. Klempin & A. Nowak (Hrsg.), *PSI-Potsdam* (S. 301–307). Universitätsverlag Potsdam. https://doi.org/10.25932/publishup-63150

Rosenberg, J., Beymer, P., Anderson, D., van Lissa, C., & Schmidt, J. (2018). tidyLPA: An R Package to Easily Carry Out Latent Profile Analysis (LPA) Using Open-Source or Commercial Software. *Journal of Open Source Software*, *3*(30), 978. https://doi.org/10.21105/joss.00978

Roters, B. (2011). *Professionalisierung durch Reflexion in der Lehrerbildung: Eine empirische Studie an einer deutschen und einer US-amerikanischen Universität: Dissertation* (Bd. 12). Waxmann.

Russell, T. (2005). Can reflective practice be taught? *Reflective Practice*, *6*(2), 199–204. https://doi.org/10.1080/14623940500105833

Rutsch, J., Vogel, M., Rehm, M., & Dörfler, T. (2018). Modellierung der Testletstruktur bei vignettenbasierten Testverfahren mit geschlossenem Antwortformat. In J. Rutsch, M. Rehm, M. Vogel, M. Seidenfuß & T. Dörfler (Hrsg.), *Effektive Kompetenzdiagnose in der Lehrerbildung* (S. 27–46). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-20121-0_2

Rutsch, J., Vogel, M., Seidenfuß, M., Dörfler, T., & Rehm, M. (2018). Professionalisierungsprozesse angehender Lehrkräfte untersuchen. In J. Rutsch, M. Rehm, M. Vogel, M. Seidenfuß & T. Dörfler (Hrsg.), *Effektive Kompetenzdiagnose in der Lehrerbildung* (S. 9–25). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-20121-0_1

Sadler, P. M., Sonnert, G., Coyle, H. P., Cook-Smith, N., & Miller, J. L. (2013). The Influence of Teachers' Knowledge on Student Learning in Middle School Physical Science Classrooms. *American Educational Research Journal*, *50*(5), 1020–1049. https://doi.org/10.3102/0002831213477680

Sandmann, A. (2014). Lautes Denken – die Analyse von Denk-, Lern- und Problemlöseprozessen. In D. Krüger, I. Parchmann & H. Schecker (Hrsg.), *Methoden in der naturwissenschaftsdidaktischen Forschung* (S. 179–188). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-37827-0_15

Schäfer, S., & Seidel, T. (2015). Noticing and reasoning of teaching and learning components by pre-service teachers. *Journal for educational research online*, *7*, 34–58. https://doi.org/10.25656/01:11489

Schaper, N. (2014). Validitätsaspekte von Kompetenzmodellen und -tests für hochschulische Kompetenzdomänen. In F. Musekamp & G. Spöttl (Hrsg.), *Kompetenz im Studium und in der Arbeitswelt* (S. 21–48). Peter Lang. https://doi.org/10.25656/01:12883

Schecker, H. (2014). Überprüfung der Konsistenz von Itemgruppen mit Cronbachs Alpha: Zusatzmaterial. In D. Krüger, I. Parchmann & H. Schecker (Hrsg.), *Methoden in der naturwissenschaftsdidaktischen Forschung* (S. 1–7). Springer Berlin Heidelberg.

Schecker, H., & Duit, R. (2018). Schülervorstellungen und Physiklernen. In H. Schecker, T. Wilhelm, M. Hopf & R. Duit (Hrsg.), *Schülervorstellungen und Physikunterricht* (S. 1–21). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-57270-2_1

Scheeler, M. C., Ruhl, K. L., & McAfee, J. K. (2004). Providing Performance Feedback to Teachers: A Review. *Teacher Education and Special Education: The Journal of the Teacher Education Division of the Council for Exceptional Children*, *27*(4), 396–407. https://doi.org/10.1177/088840640402700407

Schermelleh-Engel, K., & Schweizer, K. (2012). Multitrait-Multimethod-Analysen. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 345–362). Springer. https://doi.org/10.1007/978-3-642-20072-4_14

Schermelleh-Engel, K., & Werner, C. S. (2012). Methoden der Reliabilitätsbestimmung. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 119–141). Springer. https://doi.org/10.1007/978-3-642-20072-4_6

Schild, N. (2021). *Eignung von domänenspezifischen Studieneingangsvariablen als Prädiktoren für Studienerfolg im Fach und Lehramt Physik: Dissertation* (Bd. 307). Logos. https://doi.org/10.30819/5226

Schmalz, I. M. (2019). *Akzeptanz von Großprojekten*. Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-23639-7

Schmelzing, S., Wüsten, S., Sandmann, A., & Neuhaus, B. (2010). Fachdidaktisches Wissen und Reflektieren im Querschnitt der Biologielehrerbildung. *Zeitschrift für Didaktik der Naturwissenschaften*, *16*, 189–207.

Schmiemann, P., & Lücken, M. (2014). Validität – Misst mein Test, was er soll? In D. Krüger, I. Parchmann & H. Schecker (Hrsg.), *Methoden in der naturwissenschaftsdidaktischen Forschung* (S. 107–118). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-37827-0_9

Schön, D. A. (1983). *The Reflective Practitioner: How Professionals Think in Action*. Basic Books.

Schön, D. A. (1987). *Educating the reflective practitioner: Toward a new design for teaching and learning in the professions*. Jossey-Bass.

Schröder, J., Riese, J., Vogelsang, C., Borowski, A., Buschhüter, D., Enkrott, P., Kempin, M., Kulgemeyer, C., Reinhold, P., & Schecker, H. (2020). Die Messung der Fähigkeit zur Unterrichtsplanung im Fach Physik mit

Hilfe eines standardisierten Performanztests. *Zeitschrift für Didaktik der Naturwissenschaften*, *26*(1), 103–122. https://doi.org/10.1007/s40573-020-00115-w

Schulz, W. (1966). Unterricht - Analyse und Planung. In P. Heimann, G. Otto & W. Schulz (Hrsg.), *Unterricht* (S. 13–47). Schroedel.

Schützeichel, R. (2020). Professionswissen. In C. Schnell & M. Pfadenhauer (Hrsg.), *Handbuch Professionssoziologie* (S. 1–23). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-13154-8_15-1

Seidel, T. (2014). Angebots-Nutzungs-Modelle in der Unterrichtspsychologie. Integration von Struktur- und Prozessparadigma. *Zeitschrift für Pädagogik*, *60*(6), 850–866. https://doi.org/10.25656/01:14686

Seidel, T., Blomberg, G., & Stürmer, K. (2010). "Oberserver" - Validierung eines videobasierten Instruments zur Erfassung der professionellen Wahrnehmung von Unterricht. Projekt OBSERVE. In E. Klieme, D. Leutner & M. Kenk (Hrsg.), *Kompetenzmodellierung. Zwischenbilanz des DFG-Schwerpunktprogramms und Perspektiven des Forschungsansatzes* (S. 296–306). Beltz.

Seidel, T., Kobarg, M., & Rimmele, R. (2003). Aufbereitung der Videodaten. In T. Seidel (Hrsg.), *Technischer Bericht zur Videostudie "Lehr-Lern-Prozesse im Physikunterricht" ; BIQUA* (S. 77–98). IPN.

Seidel, T., & Prenzel, M. (2007). Wie Lehrpersonen Unterricht wahrnehmen und einschätzen — Erfassung pädagogisch-psychologischer Kompetenzen mit Videosequenzen. In M. Prenzel, I. Gogolin & H.-H. Krüger (Hrsg.), *Kompetenzdiagnostik* (S. 201–216). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-90865-6_12

Seidel, T., Prenzel, M., Dalehefte, I. M., Meyer, L., Herweg, C., Lehrke, M., & Duit, R. (2005). The IPN Video Study - An Overview. In T. Seidel, M. Prenzel & M. Kobarg (Hrsg.), *How to run a video study* (S. 7–19). Waxmann.

Seidel, T., Prenzel, M., Rimmele, R., Dalehefte, I. M., Herweg, C., Kobarg, M., & Schwindt, K. (2006). Blicke auf den Physikunterricht. Ergebnisse der IPN Videostudie. *Zeitschrift für Pädagogik*, *52*(6), 799–821. https://doi.org/10.25656/01:4489

Seidel, T., Stürmer, K., Blomberg, G., Kobarg, M., & Schwindt, K. (2011). Teacher learning from analysis of videotaped classroom situations: Does it make a difference whether teachers observe their own teaching or that of others? *Teaching and Teacher Education*, *27*(2), 259–267. https://doi.org/10.1016/j.tate.2010.08.009

Seifert, A., Hilligus, A. H., & Schaper, N. (2009). Entwicklung und psychometrische Überprüfung eines Messinstruments zur Erfassung pädagogischer Kompetenzen in der universitären Lehrerbildung. *Lehrerbildung auf dem Prüfstand*, *2*(1), 82–103.

Sekretariat der Kultusministerkonferenz. (2004). *Standards für die Lehrerbildung: Bildungswissenschaften: Beschluss der Kultusministerkonferenz vom*

16.12.2004 i. d. F. vom 16.05.2009. https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_12_16-Standards-Lehrerbildung-Bildungswissenschaften.pdf

Shaw, S., & Crisp, V. (2011). Tracing the evolution of validity in educational measurement: past issues and contemporary challenges. *Research Matters: A Cambridge Assessment publication*, (11), 14–19.

Shepard, L. A. (1997). The Centrality of Test Use and Consequences for Test Validity. *Educational Measurement: Issues and Practice*, *16*(2), 5–24. https://doi.org/10.1111/j.1745-3992.1997.tb00585.x

Shepard, L. A. (2016). Evaluating test validity: reprise and progress. *Assessment in Education: Principles, Policy & Practice*, *23*(2), 268–280. https://doi.org/10.1080/0969594X.2016.1141168

Shulman, L. S. (1986). Those Who Understand: Knowledge Growth in Teaching. *Educational Researcher*, *15*(2), 4–14. http://www.jstor.org/stable/1175860

Shulman, L. S. (1987). Knowledge and Teaching: Foundations of the New Reform. *Harvard Educational Review*, *57*(1), 1–21.

Shute, V. J. (2008). Focus on Formative Feedback. *Review of Educational Research*, *78*(1), 153–189. https://doi.org/10.3102/0034654307313795

Simon, B. (2001). *Wissensmedien im Bildungssektor: Eine Akzeptanzuntersuchung an Hochschulen* [Dissertation]. Wirtschaftsuniversität Wien.

Sippel, S. (2009). Zur Relevanz von Assessment-Feedback in der Hochschullehre. *Zeitschrift für Hochschulentwicklung*, *4*(1), 1–22. https://doi.org/10.3217/zfhe-4-01/02

Sireci, S. G. (1998). The Construct of Content Validity. *Social Indicators Research*, *45*(1/3), 83–117. https://doi.org/10.1023/A:1006985528729

Skilling, K., & Stylianides, G. J. (2020). Using vignettes in educational research: a framework for vignette construction. *International Journal of Research & Method in Education*, *43*(5), 541–556. https://doi.org/10.1080/1743727X.2019.1704243

Sluijsmans, D. M. A., Brand-Gruwel, S., van Merriënboer, J. J. G., & Bastiaens, T. J. (2003). The training of peer assessment skills to promote the development of reflection skills in teacher education. *Studies in Educational Evaluation*, *29*, 23–42.

Sorge, S., Kröger, J., Petersen, S., & Neumann, K. (2019). Structure and development of pre-service physics teachers' professional knowledge. *International Journal of Science Education*, *41*(7), 862–889. https://doi.org/10.1080/09500693.2017.1346326

Sorge, S., Petersen, S., & Neumann, K. (2016). Die Bedeutung der Studierfähigkeit für den Studienerfolg im 1. Semester in Physik. *Zeitschrift für Didaktik der Naturwissenschaften*, *22*(1), 165–180. https://doi.org/10.1007/s40573-016-0048-x

Sorge, S., Stender, A., & Neumann, K. (2019). The Development of Science Teachers' Professional Competence. In A. Hume, R. Cooper & A. Borow-

ski (Hrsg.), *Repositioning Pedagogical Content Knowledge in Teachers' Knowledge for Teaching Science* (S. 151–166). Springer Singapore. https://doi.org/10.1007/978-981-13-5898-2_6

Spurk, D., Hirschi, A., Wang, M., Valero, D., & Kauffeld, S. (2020). Latent profile analysis: A review and "how to" guide of its application within vocational behavior research. *Journal of Vocational Behavior*, *120*, 1–21. https://doi.org/10.1016/j.jvb.2020.103445

Stahnke, R., & Blömeke, S. (2021). Novice and expert teachers' noticing of classroom management in whole-group and partner work activities: Evidence from teachers' gaze and identification of events. *Learning and Instruction*, *74*, 101464. https://doi.org/10.1016/j.learninstruc.2021.101464

Stecher, B., Le, V.-N., Hamilton, L., Ryan, G., Robyn, A., & Lockwood, J. R. (2006). Using Structured Classroom Vignettes to Measure Instructional Practices in Mathematics. *Educational Evaluation and Policy Analysis*, *28*(2), 101–130. https://doi.org/10.3102/01623737028002101

Steiner, P. M., & Atzmüller, C. (2006). Experimentelle vignettendesigns in faktoriellen surveys. *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie*, *58*(1), 117–146. https://doi.org/10.1007/s11575-006-0006-9

Stender, A. (2014). *Unterrichtsplanung: Vom Wissen zum Handeln: Theoretische Entwicklung und empirische Überprüfung des Transformationsmodells der Unterrichtsplanung: Dissertation* (Bd. 170). Logos.

Stender, A., Brückmann, M., & Neumann, K. (2015). Vom Professionswissen zum kompetenten Handeln im Unterricht: Die Rolle der Unterrichtsplanung: Beiträge zur Lehrerinnen- und Lehrerbildung 33 (2015) 1, S. 121-133. *Beiträge zur Lehrerinnen- und Lehrerbildung*, *33*. https://doi.org/10.25656/01:13901

Stender, J., Watson, C., Vogelsang, C., & Schaper, N. (2021). Wie hängen bildungswissenschaftliches Professionswissen, Einstellungen zu Reflexion und die Reflexionsperformanz angehender Lehrpersonen zusammen? *Herausforderung Lehrer*innenbildung - Zeitschrift zur Konzeption, Gestaltung und Diskussion*, *4*(1), 229–248. https://doi.org/10.11576/hlz-4057

Stichweh, R. (1996). Professionen in einer funktional differenzierten Gesellschaft. In A. Combe & W. Helsper (Hrsg.), *Pädagogische Professionalität* (S. 49–69). Suhrkamp.

Stobart, G. (2009). Determining validity in national curriculum assessments. *Educational Research*, *51*(2), 161–179. https://doi.org/10.1080/00131880902891305

Szogs, M., Große, A., & Korneck, F. (2020). Wie bedingen sich die Unterrichtsqualität und Reflexivität angehender Physiklehrkräfte? In S. Habig (Hrsg.), *Naturwissenschaftliche Kompetenzen in der Gesellschaft von morgen* (S. 499–502). Universität Duisburg-Essen.

Szogs, M., Große, A., & Korneck, F. (2021). Wie bedingen sich Reflexivität und Unterrichtsqualität, die angehende Physiklehrkräfte zeigen? In S. Habig

(Hrsg.), *Naturwissenschaftlicher Unterricht und Lehrerbildung im Umbruch* (S. 234–237). Universität Duisburg-Essen.

Szogs, M., Kobl, C., Volmer, M., & Korneck, F. (2019). Bedeutsamkeit von Reflexion und Reflexivität in der Professionalisierung von Lehrkräften sowie ihre Beziehung zu anderen Prozessen und Konstrukten. In C. Maurer (Hrsg.), *Naturwissenschaftliche Bildung als Grundlage für berufliche und gesellschaftliche Teilhabe* (S. 317–320). Universität Regensburg.

Taherdoost, H. (2018). A review of technology acceptance and adoption models and theories. *Procedia Manufacturing*, *22*, 960–967. https://doi.org/10.1016/j.promfg.2018.03.137

Tenorth, H.-E. (2006). Professionalität im Lehrerberuf. *Zeitschrift für Erziehungswissenschaft*, *9*(4), 580–597. https://doi.org/10.1007/s11618-006-0169-y

Tepner, O., Borowski, A., Dollny, S., Fischer, H. E., Jüttner, M., Kirschner, S., Leutner, D., Neuhaus, B. J., Sandmann, A., Sumfleth, E., Thillmann, H., & Wirth, J. (2012). Modell zur Entwicklung von Testitems zur Erfassung des Professionswissens von Lehrkräften in den Naturwissenschaften. *Zeitschrift für Didaktik der Naturwissenschaften*, *18*, 7–28.

Tepner, O., & Dollny, S. (2014). Entwicklung eines Testverfahrens zur Analyse fachdidaktischen Wissens. In D. Krüger, I. Parchmann & H. Schecker (Hrsg.), *Methoden in der naturwissenschaftsdidaktischen Forschung* (S. 311–323). Springer Berlin Heidelberg.

Tepner, O., & Sumfleth, E. (2019). Postscript. In A. Hume, R. Cooper & A. Borowski (Hrsg.), *Repositioning Pedagogical Content Knowledge in Teachers' Knowledge for Teaching Science* (S. 317–344). Springer Singapore.

Terhart, E. (2012). Wie wirkt Lehrerbildung? Forschungsprobleme und Gestaltungsfragen. *Zeitschrift für Bildungsforschung*, *2*(1), 3–21. https://doi.org/10.1007/s35834-012-0027-3

Thompson, N., & Pascal, J. (2012). Developing critically reflective practice. *Reflective Practice*, *13*(2), 311–325. https://doi.org/10.1080/14623943.2012.657795

Trapmann, S., Hell, B., Weigand, S., & Schuler, H. (2007). Die Validität von Schulnoten zur Vorhersage des Studienerfolgs - eine Metaanalyse. *Zeitschrift für Pädagogische Psychologie*, *21*(1), 11–27. https://doi.org/10.1024/1010-0652.21.1.11

Turner, M., Kitchenham, B., Brereton, P., Charters, S., & Budgen, D. (2010). Does the technology acceptance model predict actual use? A systematic literature review. *Information and Software Technology*, *52*(5), 463–479. https://doi.org/10.1016/j.infsof.2009.11.005

Ullmann, T. D. (2019). Automated Analysis of Reflection in Writing: Validating Machine Learning Approaches. *International Journal of Artificial Intelligence in Education*, *29*(2), 217–257. https://doi.org/10.1007/s40593-019-00174-2

Valli, L. (1997). Listening to other voices: A description of teacher reflection in the United States. *Peabody Journal of Education*, *72*(1), 67–88. https://doi.org/10.1207/s15327930pje7201_4

van Beveren, L., Roets, G., Buysse, A., & Rutten, K. (2018). We all reflect, but why? A systematic review of the purposes of reflection in higher education in social and behavioral sciences. *Educational Research Review*, *24*, 1–9. https://doi.org/10.1016/j.edurev.2018.01.002

van Es, E. A., & Sherin, M. G. (2008). Mathematics teachers' "learning to notice" in the context of a video club. *Teaching and Teacher Education*, *24*(2), 244–276. https://doi.org/10.1016/j.tate.2006.11.005

van Manen, M. (1977). Linking Ways of Knowing with Ways of Being Practical. *Curriculum Inquiry*, *6*(3), 205–228. https://doi.org/10.2307/1179579

Veal, W. R. (2002). Content Specific Vignettes as Tools for Research and Teaching. *Electronic Journal of Science Education*, *6*(4), 1–37.

Venkatesh, V., & Davis, F. D. (2000). A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. *Management Science*, *46*(2), 186–204. https://doi.org/10.1287/mnsc.46.2.186.11926

Vogelsang, C. (2014). *Validierung eines Instruments zur Erfassung der professionellen Handlungskompetenz von (angehenden) Physiklehrkräften: Zusammenhangsanalysen zwischen Lehrerkompetenz und Lehrerperformanz* (Bd. 174). Logos Verlag.

Vogelsang, C., Borowski, A., Buschhüter, D., Enkrott, P., Kempin, M., Kulgemeyer, C., Reinhold, P., Riese, J., Schecker, H., & Schröder, J. (2019). Entwicklung von Professionswissen und Unterrichtsperformanz im Lehramtsstudium Physik. Analysen zu valider Testwertinterpretation. *Zeitschrift für Pädagogik*, (4), 473–491. https://doi.org/10.25656/01:23990

Vogelsang, C., & Cauet, E. (2017). Wie valide sind Porfessionswissenstests? Zum Zusammenhang von erfasstem Wissen, Unterrichtshandeln und Unterrichtserfolg. In H. E. Fischer & E. Sumfleth (Hrsg.), *Professionelle Kompetenz von Lehrkräften der Chemie und Physik* (S. 77–96). Logos.

Vogelsang, C., Kulgemeyer, C., & Riese, J. (2022). Learning to Plan by Learning to Reflect?—Exploring Relations between Professional Knowledge, Reflection Skills, and Planning Skills of Preservice Physics Teachers in a One-Semester Field Experience. *Education Sciences*, *12*(7), 479. https://doi.org/10.3390/educsci12070479

Vogelsang, C., & Rehfeldt, D. (2021). Relationierung von Theorie und Praxis in der Lehrkräftebildung im Fach Physik – eine Übersicht über Forschungen und Formate. In C. Caruso, C. Harteis & A. Gröschner (Hrsg.), *Theorie und Praxis in der Lehrerbildung* (S. 333–348). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-32568-8_19

Vogelsang, C., & Reinhold, P. (2013). Zur Handlungsvalidität von Tests zum professionellen Wissen von Lehrkräften. *Zeitschrift für Didaktik der Naturwissenschaften*, *19*, 103–128.

Vogelsang, C., Riese, J., Kulgemeyer, C., & Borowski, A. (2017). Profile-P+ - Professionskompetenz und Unterrichtsperformanz im Lehramtsstudium Physik. In C. Maurer (Hrsg.), *Implementation fachdidaktischer Innovation im Spiegel von Forschung und Praxis* (S. 752–755). Universität Regensburg.

von Aufschnaiter, C., Fraij, A., & Kost, D. (2019). Reflexion und Reflexivität in der Lehrerbildung. *Herausforderung Lehrer_innenbildung - Zeitschrift zur Konzeption, Gestaltung und Diskussion*, *2*(1), 144–159. https://doi.org/10.4119/UNIBI/hlz-144

von Aufschnaiter, C., Hofmann, C., Geisler, M., & Kirschner, S. (2019). Möglichkeiten und Herausforderungen der Förderung von Reflexivität in der Lehrerbildung. *SEMINAR*, *25*(1), 49–60.

von Aufschnaiter, C., & Blomberg, G. (2010). Professionelle Kompetenz von (angehenden) Lehrkräften erfassen – Desiderata. *Zeitschrift für Didaktik der Naturwissenschaften*, *16*, 361–367.

von Aufschnaiter, C., Cappell, J., Dübbelde, G., Ennemoser, M., Mayer, J., Stiensmeier-Pelster, J., Sträßer, R., & Wolgast, A. (2015). Diagnostische Kompetenz - Theoretische Überlegungen zu einem zentralen Konstrukt der Lehrerbildung. *Zeitschrift für Pädagogik*, *61*(5), 738–758. https://doi.org/10.25656/01:12723

Voss, T., Kunina-Habenicht, O., Hoehne, V., & Kunter, M. (2015). Stichwort Pädagogisches Wissen von Lehrkräften: Empirische Zugänge und Befunde. *Zeitschrift für Erziehungswissenschaft*, *18*(2), 187–223. https://doi.org/10.1007/s11618-015-0626-6

Voss, T., & Kunter, M. (2011). Pädagogisch-psychologisches Wissen von Lehrkräften. In M. Kunter, J. Baumert & W. Blum (Hrsg.), *Professionelle Kompetenz von Lehrkräften* (S. 193–214). Waxmann.

Voss, T., Kunter, M., Seiz, J., Hoehne, V., & Baumert, J. (2014). Die Bedeutung des pädagogisch-psychologischen Wissens von angehenden Lehrkräften für die Unterrichtsqualität. *Zeitschrift für Pädagogik*, *60*(2), 184–201. https://doi.org/10.25656/01:14653

Wahl, D. (2001). Nachhaltige Wege vom Wissen zum Handeln: Beiträge zur Lehrerbildung 19 (2001) 2, S. 157-174. *Beiträge zur Lehrerbildung*, *19*. https://doi.org/10.25656/01:13453

Watermann, R., & Klieme, E. (2006). Modellierung von Kompetenzstufen mit Hilfe der latenten Klassenanalyse. *Empirische Pädagogik*, *20*(3), 312–336.

Weber, K. E., Elstrodt-Wefing, N., & Hoge, K. (2022). Expertise- und stimulusbedingte Unterschiede in schriftlichen videobasierten Reflexionen. In E. Gläser, J. Poschmann, P. Büker & S. Miller (Hrsg.), *Reflexion und Reflexivität im Kontext Grundschule. Perspektiven für Forschung, Lehrer:innenbildung und Praxis* (S. 77–89). Verlag Julius Klinkhardt : Bad Heilbrunn. https://doi.org/10.25656/01:25554

Weber, K. E., Neuber, K., & Prilop, C. N. (2023). Videobasierte Reflexion von klassenführungsspezifischen Ereignissen – Welche Rolle spielen Wissen und

Reflexionsbereitschaft von Lehramtsstudierenden? *Zeitschrift für Erziehungswissenschaft*, *26*(5), 1235–1257. https://doi.org/10.1007/s11618-023-01195-3

Weinert, F. E. (2001). Concept of competence: a conceptual clarification. In D. S. Rychen & L. Hersh Salganik (Hrsg.), *Defining and selecting key competencies* (S. 45–65). Hogrefe & Huber Publishers.

Weinert, F. E. (Hrsg.). (2002a). *Leistungsmessungen in Schulen* (2. unveränd. Aufl.). Beltz.

Weinert, F. E. (2002b). Vergleichende Leistungsmessung in Schulen - eine umstrittene Selbstverständlichkeit. In F. E. Weinert (Hrsg.), *Leistungsmessungen in Schulen* (S. 17–31). Beltz.

Weißbach, A., & Kulgemeyer, C. (2022a). Reflexion von Physikunterricht - ein Online Assessment mit Feedback. *PhyDid B - Didaktik Der Physik - Beiträge Zur DPG-Frühjahrstagung*, 203–210.

Weißbach, A., & Kulgemeyer, C. (2022b). Reflexion von Physikunterricht - ein Online-Assessment mit Feedback. In S. Habig & H. van Vorst (Hrsg.), *Unsicherheit als Element von naturwissenschaftsbezogenen Bildungsprozessen* (S. 756–759). Universität Duisburg-Essen.

Weißbach, A., & Kulgemeyer, C. (2023a). Physikunterricht reflektieren: Eine Lernumgebung für die Lehramtsausbildung. In L. Mientus, C. Klempin & A. Nowak (Hrsg.), *Reflexion in der Lehrkräftebildung* (S. 241–247). Universitätsverlag Potsdam.

Weißbach, A., & Kulgemeyer, C. (2023b). Reflection on Physics Teaching: An Online-Assessment Offering Automated Feedback. https://www.researchgate.net/publication/385172169_Reflection_On_Physics_Teaching_An_Online-Assessment_Offering_Automated_Feedback

Weißbach, A., & Kulgemeyer, C. (2023c). Reflexion von Physikunterricht: Ein Online-Assessment mit Feedback. In H. van Vorst (Hrsg.), *Lernen, Lehren und Forschen in einer digital geprägten Welt* (S. 861–864). Universität Duisburg-Essen.

Weißbach, A., & Kulgemeyer, C. (2024a). Reflection on Physics Teaching - A comparison of a performance assessment and a multiple-choice assessment. https://www.researchgate.net/publication/379483629_Reflection_on_Physics_Teaching_-A_comparison_of_a_performance_assessment_and_a_multiple-choice_assessment

Weißbach, A., & Kulgemeyer, C. (2024b). Reflexionsfähigkeit von Physiklehramtsstudierenden: Ein Online-Self-Assessment mit Feedback. *Unterrichtswissenschaft*, *52*(1), 135–161. https://doi.org/10.1007/s42010-024-00195-4

Weißbach, A., & Kulgemeyer, C. (2024c). Reflexionsfähigkeit: Validierungsstudien zu einem Test mit Feedback. In H. van Vorst (Hrsg.), *Frühe naturwissenschaftliche Bildung* (S. 266–269). Universität Duisburg-Essen.

Weißbach, A., & Kulgemeyer, C. (2024d). Wie hängen Unterrichtsreflexion und Noticing zusammen? https://www.researchgate.net/publication/384925967_Wie_hangen_Unterrichtsreflexion_und_Noticing_zusammen

Weller, B. E., Bowen, N. K., & Faubert, S. J. (2020). Latent Class Analysis: A Guide to Best Practice. *Journal of Black Psychology*, *46*(4), 287–311. https://doi.org/10.1177/0095798420930932

Weston, T. L., & Amador, J. M. (2023). Teacher noticing: A literature review of mathematics and science teacher noticing conceptualizations. *School Science and Mathematics*, *123*(7), 293–308. https://doi.org/10.1111/ssm.12629

Weyers, J., König, J., Santagata, R., Scheiner, T., & Kaiser, G. (2023). Measuring teacher noticing: A scoping review of standardized instruments. *Teaching and Teacher Education*, *122*, 103970. https://doi.org/10.1016/j.tate.2022.103970

Windt, A., & Lenske, G. (2016). Qualität der Sachunterrichtsreflexion im Vorbereitungsdienst. In C. Maurer (Hrsg.), *Authentizität und Lernen - das Fach in der Fachdidaktik* (S. 284–286). Universität Regensburg.

Wirtz, M. A. (Hrsg.). (2017). *Dorsch - Lexikon der Psychologie* (18., überarbeitete Auflage). Hogrefe.

Wirtz, M. A., & Caspar, F. (2002). *Beurteilerübereinstimmung und Beurteilerreliabilität: Methoden zur Bestimmung und Verbesserung der Zuverlässigkeit von Einschätzungen mittels Kategoriensystemen und Ratingskalen*. Hogrefe Verlag für Psychologie.

Wöhlke, C. (2020). *Entwicklung und Validierung eines Instruments zur Erfassung der professionellen Unterrichtswahrnehmung angehender Physiklehrkräfte: Dissertation*. Logos.

Woitkowski, D. (2015). *Fachliches Wissen Physik in der Hochschulausbildung: Konzeptualisierung, Messung, Niveaubildung: Dissertation* (Bd. 185). Logos.

Woitkowski, D., & Borowski, A. (2017). Fachwissen im Lehramtsstudium Physik. In H. E. Fischer & E. Sumfleth (Hrsg.), *Professionelle Kompetenz von Lehrkräften der Chemie und Physik* (S. 57–74). Logos.

Woitkowski, D., & Riese, J. (2017). Kriterienorientierte Konstruktion eines Kompetenzniveaumodells im physikalischen Fachwissen. *Zeitschrift für Didaktik der Naturwissenschaften*, *23*(1), 39–52. https://doi.org/10.1007/s40573-016-0054-z

Wolming, S., & Wikström, C. (2010). The concept of validity in theory and practice. *Assessment in Education: Principles, Policy & Practice*, *17*(2), 117–132. https://doi.org/10.1080/09695941003693856

Wubbels, T., & Korthagen, F. A. J. (1990). The Effects of a Pre-service Teacher Education Program for the Preparation of Reflective Teachers. *Journal of Education for Teaching*, *16*(1), 29–43. https://doi.org/10.1080/0260747900160102

Wulff, P., Mientus, L., Nowak, A., & Borowski, A. (2021). Stärkung praxisorientierter Hochschullehre durch computerbasierte Rückmeldung zu Reflexionstexten. *die hochschullehre*, *7*(1), 93–99. https://doi.org/10.3278/HSL2111W

Wulff, P., Mientus, L., Nowak, A., & Borowski, A. (2023a). KI-basierte Auswertung von schriftlichen Unterrichtsreflexionen im Fach Physik und automatisierte Rückmeldung. In L. Mientus, C. Klempin & A. Nowak (Hrsg.), *PSI-Potsdam* (S. 103–115). Universitätsverlag Potsdam.

Wulff, P., Mientus, L., Nowak, A., & Borowski, A. (2023b). Utilizing a Pretrained Language Model (BERT) to Classify Preservice Physics Teachers' Written Reflections. *International Journal of Artificial Intelligence in Education*, *33*, 439–466. https://doi.org/10.1007/s40593-022-00290-6

Wulff, P., Westphal, A., Mientus, L., Nowak, A., & Borowski, A. (2023). Enhancing writing analytics in science education research with machine learning and natural language processing—Formative assessment of science and non-science preservice teachers' written reflections. *Frontiers in Education*, *7*. https://doi.org/10.3389/feduc.2022.1061461

Wyss, C. (2008). Zur Reflexionsfähigkeit und -praxis der Lehrperson. *Bildungsforschung*, *5*(2), 1–15.

Wyss, C. (2013). *Unterricht und Reflexion: Eine mehrperspektivische Untersuchung der Unterrichts- und Reflexionskompetenz von Lehrkräften: Dissertation* (Bd. 44). Waxmann.

Wyss, C., & Mahler, S. (2021). Mythos Reflexion. Theoretische und praxisbezogene Erkenntnisse in der Lehrer*innenbildung. *Journal für LehrerInnenbildung*, *21*, 16–25. https://doi.org/10.25656/01:22102

Zulliger, S., Buholzer, A., & Ruelmann, M. (2022). Observed Quality of Formative Peer and Self-Assessment in Everyday Mathematics Teaching and its Effects on Student Performance. *European Journal of Educational Research*, *11*(2), 663–680. https://doi.org/10.12973/eu-jer.11.2.663

Abbildungsverzeichnis

5.1 Wirkungskette der schulischen Bildung nach Terhart (2012, S. 7) in schwarz sowie (mit grauen Ergänzungen) Modell des Einflusses der Lehrkräftebildung auf das Schüler*innenlernen nach Diez (2010, S. 442) (deutsche Übersetzung angelehnt an Terhart (2012, S. 7)) . . 12

5.2 Abgrenzung von Schulwissen und universitärem Wissen abhängig vom curricularen und mathematischen Fortschritt nach Enkrott (2021, S. 45) 21

5.3 Komponenten des PCK nach Magnusson, Krajcik und Borko (1999, S. 99) (schwarze gestrichelte und durchgezogene Pfeile beschreiben eine gegenseitige Beeinflussung) sowie die darauf aufbauende Modellierung des PCK von Park und Chen (2012, S. 925) (durchgezogene graue und schwarze Pfeile markieren die gegenseitige Beeinflussung der Facetten). 23

5.4 Modell des fachdidaktischen Wissens nach Gramzow, Riese und Reinhold (2013, S. 21) 25

5.5 Modell der professionellen Handlungskompetenz nach Riese und Reinhold (2010, S. 170) 30

5.6 Modell von Kompetenz als Kontinuum von Blömeke, Gustafsson und Shavelson (2015, S. 7) 31

5.7 Vereinfachte Professionalisierungskette von Lehrkräften nach Vogelsang et al. (2019, S. 474) 32

5.8 Modell der Transformation von Professionswissen in Handeln von Stender (2014, S. 33) 33

5.9 Refined Consensus Model of PCK von Carlson et al. (2019, S. 84) 34

6.1 Darstellung der erforderlichen Übersetzungsschritte, um Ausprägungen einer Zieldomäne Testwerte zuzuweisen nach Dickmann (2016, S. 67) 55

6.2 Darstellung der TRA (nur dunkle Blöcke) sowie leicht vereinfachte Darstellung der TPB nach Ajzen (2002 (überarbeitet 2006), S. 1) unter Auslassung des Faktors der tatsächlichen Verhaltenskontrolle 61

6.3 Darstellung des TAM (Davis, 1989) nach Turner et al. (2010, S. 472). 63

7.1 Modell zur Bewertung der Qualität von Reflexionen von Physikunterricht in drei Dimensionen nach Nowak, Kempin, Kulgemeyer und Borowski (2019, S. 839) 95
7.2 Struktur des Reflexionsperformanztests (Darstellung in Anlehnung an Kulgemeyer et al., 2021, S. 3044) 107

8.1 Feedback-Modell nach Hattie und Timperley (2007, S. 87) (Übersetzung von Hosenfeld, 2010, S. 55) 123

10.1 Struktur des Reflexionstests (Darstellung aufbauend auf Kulgemeyer et al., 2021, S. 3044) 140

11.1 Exemplarische Darstellung des Boxplots zur Rückmeldung des Gesamtergebnisses 156
11.2 Screenshot aus dem Selbstlernmaterial: Überblick zur Reflexion der zweiten eingebundenen Videovignette 159

12.1 Darstellung der erforderlichen Übersetzungsschritte, zur Zuweisung und Rückmeldung von Testwerten für die Zieldomäne der Fähigkeit zur Unterrichtsreflexion 164
12.2 Darstellung des INA für die Evaluation des Reflexionstests mit Assessment-Feedback ergänzt um die jeweils gestellten Anforderungen und Indikatoren zur Überprüfung dieser 172

13.1 Adaption des TAM/TAM2 für die Akzeptanzbefragung zur Lernumgebung 178

14.1 Darstellung der Untersuchungen entlang des INA mit den jeweiligen Anforderungen und Indikatoren sowie (fett gedruckt) den Bewertungsgrundlagen zu deren Überprüfung (Teil 1/2) 186
14.2 Darstellung der Untersuchungen entlang des INA mit den jeweiligen Anforderungen und Indikatoren sowie (fett gedruckt) den Bewertungsgrundlagen zu deren Überprüfung (Teil 2/2) 188

17.1 Histogramm der erreichten Testwerte (in %) mit Dichtefunktion der Normalverteilung ($N = 192$) 225
17.2 Scree-Plot zur explorativen Faktorenanalyse von den Aufgaben im Reflexionstest 231

19.1 Darstellung des INA für die Interpretation der Testwerte des Reflexionstests mit Assessment-Feedback, ergänzt um die Bewertung der jeweils überprüften Indikatoren auf Grundlage der durchgeführten Validierungsstudien 318

20.1 Basismodell zum Zusammenhang von Noticing und Unterrichtsreflexion 324

21.1 Latente Profile der Reflexionsfähigkeit (in aufsteigender Reihenfolge: 1 *(blau ▲)*, 2 *(grün ●)*, 3 *(rot ■)*, 4 *(lila ◆)*) 330
21.2 Entwicklung der Zuweisung der Personen zu den Fähigkeitsstufen über beide Messzeitpunkte 332
21.3 Modell zum Zusammenhang von Noticing und Unterrichtsreflexion (erste Erweiterung) 338

22.1 Histogramme zu den jeweils erreichten Testwerten (in %) in den Aufgaben zur Reflexion von in Rahmen der Aufgaben zum Noticing (a) als bedeutsam bzw. (b) nicht als bedeutsam eingeschätzten Aspekten 343
22.2 Modell zum Zusammenhang von Noticing und Unterrichtsreflexion (zweite Erweiterung) 346

23.1 CLP-Design für die Untersuchung von Entwicklungszusammenhängen zwischen den Fähigkeiten zum Noticing und zur Unterrichtsreflexion 348
23.2 Ergebnisse des PLS-SEM zu Zusammenhängen zwischen den Fähigkeiten zur Reflexion und zum Noticing im CLP-Design für (a) alle vorliegenden Prä-Post-Datensätze und (b) die Teilstichprobe der Personen in den oberen drei Fähigkeitsstufen ($^{*}p < 0,05$; $^{***}p < 0,001$) 350
23.3 Übersicht über die Analysen zum CLP-Design für die Gesamtstichprobe (bzw. die Teilstichprobe der Personen in den oberen drei Fähigkeitsstufen); bestimmt wurden jeweils die bivariaten Korrelationen zwischen den Fähigkeiten zum Noticing und zur Unterrichtsreflexion zu den beiden MZP sowie für beide Fähigkeiten zwischen den MZP und die partiellen Kreuzkorrelationen für die Kreuzpfade ($^{*}p < 0,05$; $^{**}p < 0,01$) 352
23.4 Modell zum Zusammenhang von Noticing und Unterrichtsreflexion (dritte Erweiterung) 354

24.1 Erweitertes Modell zum Zusammenhang von Noticing und Unterrichtsreflexion 358

Tabellenverzeichnis

6.1 Differenzierung des holistischen Validitätskonzepts in Argumentationsgrundlagen für den Testwertinterpretationen und deren Nutzung nach Messick (1995, S. 748). 46

7.1 Übersicht über Prozessmodelle zur (Unterrichts-)Reflexion (mit ↻ markierte Modelle sind explizit zyklisch modelliert, die Umsetzung einer Handlung löst demnach eine neue Reflexion aus) 90

7.2 Bepunktung einer Reflexion im Reflexionsperformanztest anhand der erreichten Stufe (Kempin, Kulgemeyer & Schecker, 2019, S. 358); Beispiele übersetzt aus Kulgemeyer et al. (2021, S. 2046) 109

10.1 Übersicht über die Videovignetten und reflektierten Aspekte im Reflexionstest 144

11.1 Leifragen zur Selbstreflexion von Unterricht nach Nowak, Liepertz und Borowski (2018, S. 223) 158

14.1 Übersicht über die Zusammensetzung der Stichprobe 184

16.1 Übersicht über die Kategorien zur Kodierung der Think-Aloud-Interviews 207

16.2 Übersicht über Anteile der verschiedenen Kategorien an allen Kodierungen (unter Auslassung der Kategorie *Vorlesen*) je Interview (als „≈ 0“ sind diejenigen Anteile der Kategorien *eigene Erfahrung* und *Professionalisierung* eingetragen, die nur zu sehr geringen Anteilen in den jeweiligen Think-Aloud-Protokollen vorkommen; „0“ beschreibt Kategorien die im jeweiligen Interview nicht angesprochen werden) sowie die Beurteilerübereinstimmung für die Kategorien (κ_{Cohen}); die Kategorie *Konstruktrelevante Überlegungen* fasst alle Kategorien bis auf *Sonstiges* zusammen 222

17.1 Untersuchung der Zusammenhänge zwischen den Testwerten und Variablen der Demografie mit Bezug zur schulischen Ausbildung . 241

18.1 Übersicht über die Kategorien zur Kodierung der Interviews zur Pilotierung des Assessment-Feedbacks. 255

18.2 Übersicht über die Kategorien zur Kodierung der Interviews zum überarbeiteten Assessment-Feedback inklusive Selbstlernmaterial (dargestellt ist das vollständige Kategoriensystem für alle Interviewstudien zum Assessment-Feedback, wobei die Kategorien, die im Rahmen der Pilotierung auftraten und hier nicht angesprochen wurden, grau gesetzt sind) 276

19.1 Übersicht über Bewertung der Indikatoren und Limitationen zum Teil 1 des INA (fett gedruckt sind Limitationen, die als die Testwertinterpretation einschränkend zu bewerten sind) 304

19.2 Übersicht über Bewertung der Indikatoren und Limitationen zum Teil 2 des INA 306

19.3 Übersicht über Bewertung der Indikatoren und Limitationen zum Teil 3 des INA (fett gedruckt sind Limitationen, die als die Testwertinterpretation einschränkend zu bewerten sind) 308

19.4 Übersicht über Bewertung der Indikatoren und Limitationen zum Teil 4 des INA (fett gedruckt sind Limitationen, die als die Testwertinterpretation einschränkend zu bewerten sind) 311

21.1 Übersicht über die Modellgüteparameter und ihre Interpretation für die verschiedenen Modelle der LPA 329

21.2 Übersicht über vorliegende signifikante Unterschiede mit mittlerer („X“) bzw. großer Effektstärke („**X**“) in den Aufgabenscores benachbarter Fähigkeitsstufen mit Zuordnung der Aufgaben zu den Wissensbasen (f - fachliches, fd - fachdidaktisches und p - pädagogisches Wissen). 333

21.3 Übersicht über die Zusammensetzung der Fähigkeitsstufen nach Ausbildungsphasen in Prozent- und Absolutwerten 335

21.4 Durchschnittliche Motivation und Bearbeitungsdauer des Reflexionstests (mit Standardabweicheungen) in den einzelnen Fähigkeitsstufen 337

G.1 Übersicht über die Kategorien zur Kodierung der Think-Aloud-Interviews. 472

G.2 Schema zur Standardisierung des Vorgehens bei der Kodierung (mit „AO“ wird „Antwortoption(en)“ abgekürzt) 475

H.1 Übersicht über Anteile der Kategorien Bewertungen (mit und ohne Begründung) sowie Alternativen (mit und ohne Begründung) aufgeschlüsselt nach Aufgaben zu Bewertungen (Bew.) und Alternativen (Alt.) 478

I.1 Kennwerte zur Beurteilung der Reliabilität (α_{Cronbach}) und Trennschärfe (TS) der verwendeten Skalen für die Gesamtstichprobe sowie die darin enthaltenen Teilstichproben der Bachelor- (BA) und Masterstudierenden (MA) 479

I.2 Ergebnisse des Kolmogorow-Smirnow-Tests für den Reflexionstest und den Reflexionsperformanztest (Gesamtscore sowie Teilscores zu Reflexionen mit Bezug zum fachlichen, fachdidaktischen und pädagogischen Wissen) für die Gesamtstichprobe und die betrachteten Teilstichproben der Physik-Lehramtsstudierenden im Bachelor- (BA-Phy) bzw. Masterstudium (MA-Phy) sowie der Studierenden des Grundschullehramts im Fach Sachunterricht mit Physik als Wahlfach (SU) 480

I.3 Ergebnisse der Tests auf Varianzhomogenität (Levene-Tests) und einfaktoriellen Varianzanalysen (ANOVAs) zur Prüfung von Gruppenunterschieden für Testwerte zum Reflexionstest und Reflexionsperformanztest 481

I.4 Post-hoc-Analyse zur Prüfung von Gruppenunterschieden für Testwerte zum fachlichen Wissen im Reflexionsperformanztest für die betrachteten Teilstichproben der Physik-Lehramtsstudierenden im Bachelor- (BA-Phy) bzw. Masterstudium (MA-Phy) sowie der Studierenden des Grundschullehramts im Fach Sachunterricht mit Physik als Wahlfach (SU) 481

I.5 Übersicht über die Kennwerte zum Reflexionstest (*Refl.*) und Test zum fachdidaktischen Wissen (*FDW*) 490

I.6 Ergebnisse der Kolmogorow-Smirnow-Tests (*KS-Test*) zu den Facetten im Test zum fachdidaktischen Wissen Instruktionsstrategien (*Instr.*), Schülervorstellungen (*SV*), Experimente und Wissenschaftsverständnis (*Exp. & Wiss.*) und Fachdidaktische Konzepte (*FD-Konzepte*) für alle vorliegenden Datensätze ($N = 41$) 490

K.1 Übersicht über die Einschätzungen der Studierenden zur Akzeptanz der Lernumgebung (kursiv gesetzte Items sind invertiert formuliert); Teil 2 wurde von den Studierenden der Interviewkohorte C (nach Erhalt der Rückmeldung vor Durchführung der Interviews) beantwortet und Teil 3 eingebettet in der *LimeSurvey*-Umgebung vorrangig von den sechs Studierenden der Interviewkohorte B nach Bearbeitung des Selbstlernmaterials (im Rahmen der Entwicklung und Befragung noch als „Fördermaterial“ bezeichnet) 514

L.1 Ergebnisse der paarweisen Wilcoxon-Tests mit Bonferroni-Korrektur für den Gesamtscore im Reflexionstest sowie für die Teilscores zum Bewerten und Vorschlagen von Alternativen (*Eff.-st.: Effektstärke*) 515

L.2 Ergebnisse der paarweisen Wilcoxon-Tests mit Bonferroni-Korrektur für die Aufgabenscores zu Vignette 1 516
L.3 Ergebnisse der paarweisen Wilcoxon-Tests mit Bonferroni-Korrektur für die Aufgabenscores zu Vignette 2 516
L.4 Ergebnisse der paarweisen Wilcoxon-Tests mit Bonferroni-Korrektur für die Aufgabenscores zu Vignette 3 (*Eff.-st.: Effektstärke*) 517
L.5 Ergebnisse der paarweisen Wilcoxon-Tests mit Bonferroni-Korrektur für die Aufgabenscores zu Vignette 4 (*Eff.-st.: Effektstärke*) 518
L.6 Ergebnisse der paarweisen Wilcoxon-Tests mit Bonferroni-Korrektur für die Aufgabenscores zu Vignette 5 (*Eff.-st.: Effektstärke*) 519
L.7 Ergebnisse der paarweisen Wilcoxon-Tests mit Bonferroni-Korrektur für die Aufgabenscores zu Vignette 6 (*Eff.-st.: Effektstärke*) 520
L.8 Ergebnisse der paarweisen Wilcoxon-Tests mit Bonferroni-Korrektur für die Aufgabenscores zu Vignette 7 (*Eff.-st.: Effektstärke*) 521
L.9 Ergebnisse der paarweisen Wilcoxon-Tests mit Bonferroni-Korrektur für die Aufgabenscores zur globalen Reflexion (*Eff.-st.: Effektstärke*) 522
L.10 Ergebnisse der paarweisen Wilcoxon-Tests mit Bonferroni-Korrektur für die Zusammensetzung der Fähigkeitsstufen nach Ausbildungsphase und Umfang absolvierter Pflichtpraktika 523
L.11 Ergebnisse der paarweisen Wilcoxon-Tests mit Bonferroni-Korrektur für die Zusammensetzung der Fähigkeitsstufen nach Testmotivation und Bearbeitungsdauer 524

M.1 Ergebnisse der paarweisen t-Tests mit Bonferroni-Korrektur zu Hypothese 2 526

Abkürzungsverzeichnis

Terminologie

Alt	Aufgaben zu Alternativen im Reflexionstest
AO	Antwortoption
BA	Bachelorstudierende
Bew	Aufgaben zu Bewertungen im Reflexionstest
CLP	Cross-Lagged-Panel
cPCK	collective Pedagogical Content Knowledge
ePCK	enacted Pedagogical Content Knowledge
FDW	fachdidaktisches Wissen
INA	Interpretations-Nutzungs-Argument
MA	Masterstudierende
MC	Multiple-Choice
MZP	Messzeitpunkt
PCK	Pedagogical Content Knowledge
pPCK	personal Pedagogical Content Knowledge
PLS	Partial Least Square
Phy	Physik
Ref	Referandar*innen
SEM	Strukturgleichungsmodell
SU	Sachunterrichtsstudierende (Grundschullehramt)

Forschungsprojekte

COACTIV	Cognitive Activation in the Classroom
COACTIV -R	Cognitive Activation in the Classroom - Ergänzungsstudie mit Referendar*innen
MT21	Mathematics Teaching in the 21st Century
PLUS	Professionswissen von Lehrkräften, naturwissenschaftlicher Unterricht und Zielerreichung im Übergang von der Primar- zur Sekundarstufe
ProfiLe-P	Professionskompetenz im Lehramtsstudium Physik
ProfiLe-P+	zweite Projektphase zu *ProfiLe-P*
ProfiLe-P-Transfer	Transferprojekt zu *ProfiLe-P*
ProwiN	Professionswissen von Lehrkräften in den Naturwissenschaften
QuiP	Quality of Instruction in Physics
TEDS-M	The Teacher Education Study in Mathematics
TIMSS	Trends in International Mathematics and Science Study

Symbole

$\alpha_{\mathbf{Cronbach}}$	Reliabilitätskoeffizient nach Cronbach
ANOVA	einfaktorielle Varianzanalyse
$d_{\mathbf{Cohen}}$	Effektstärke nach Cohen
$d_{\mathbf{Cohen},z}$	angepasste Effektstärke nach Cohen für gepaarte Stichproben
$\kappa_{\mathbf{Cohen}}$	Maß der Beurteilerübereinstimmung nach Cohen
D	Statistik des Kolmogorow-Smirnow-Tests
F	Statistik des Levene-Tests bzw. zu ANOVAs
H	Statistik des Kruskal-Wallis-Tests
KI	Konfidenzintervall
M	Mittelwert
N	Stichprobengröße

n. s.	nicht signifikant
p	Signifikanzwert
SD	Standardabweichung
$r_{\textbf{Pearson}}$	Pearson-Korrelationskoeffizient
$\rho_{\textbf{Spearman}}$	Spearman-Rangkorrelationskoeffizient
t	Statistik des t-Tests
W	Statistik des Wilcoxon-Tests

Teil VI

Anhang

Der Anhang ist Bestandteil der Open-Access-Publikation dieser Arbeit. Diese ist unter dem folgenden Link abrufbar: https://doi.org/10.30819/5977

A Aufgaben im Reflexionstest

Der vollständige Reflexionstest kann unter https://umfragen.physikdidaktik.uni-bremen.de/index.php/516586?lang=de abgerufen werden. Nachfolgend sind alle MC-Aufgaben zum Noticing und zur Unterrichtsreflexion aufgeführt.

Die **fett** gedruckten Antwortoptionen stellen jeweils die richtigen Antwortoptionen dar.

Hinweise zur Benennung der Aufgaben:

- Die Aufgaben zum Noticing sind mit *No* benannt und aufsteigend (entsprechend der Videovignetten bzw. *G* für die globale Reflexion der Stunde) nummeriert.
- Die Benennung der Reflexionsaufgaben setzt sich wie folgt zusammen: *R*[*Nummer der Vignette* bzw. *G*][*Bew* (Bewertungen) bzw. *Alt* (Alternativen)][Nummer des Reflektierten Aspekts je Vignette]

A.1 Zu Videovignette 1

Inhalt: Besprechung von Quizfragen zur Wiederholung der Newtonschen Axiome

No1 - Noticing-Aufgabe

Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung (Lob und/oder Kritik) erhalten, um etwas für sich bzw. seinen Unterricht zu lernen?

☐ **Formulierung des Merksatzes zum 1. Axiom**
☐ **Formulierung des Merksatzes zum 2. Axiom**
☐ **Formulierung des Merksatzes zum 3. Axiom**
☐ **Ergebnissicherung des Quiz**
☐ **Gesprächsführung**

R1Bew1 - Formulierung des Merksatzes zum 3. Axiom (Bewertung)

Welches Feedback möchten Sie Robert zu seiner Formulierung des Merksatzes zum 3. Axiom am ehesten geben?

☐ Deine Formulierung des dritten Axioms ist gut, da alles Wichtige enthalten und es gut nachvollziehbar ist.
☐ **Deine Formulierung des 3. Axioms ist fachlich falsch.**
☐ **Die Formulierung des 3. Axioms ist für die Schüler*innen nicht angemessen, da sie zu komplex ist.**
☐ **Die Formulierung des 3. Axioms ist nicht gut, da die Vermischung von Beispiel und Merksatz die Schüler*innen verwirren kann.**
☐ Es ist nicht gut, dass nur die Kurzformel „Kraft=Gegenkraft“, aber nicht „actio=reactio“ im Merksatz vorkommt. Diese Abkürzung ist nämlich etwas, was die Schüler*innen vielleicht kennen.
☐ Es ist gut, dass du im Merksatz deutlich machst, dass die Kräfte an verschiedenen Körpern angreifen, da du so einer Verwechslung mit dem Kräftegleichgewicht vorbeugst.
☐ **Es ist nicht gut, dass die Kurzformel „Kraft=Gegenkraft“ im Merksatz vorkommt, da Schüler*innen diese oft mit dem Kräftegleichgewicht assoziieren.**

R1Alt1 - Formulierung des Merksatzes zum 3. Axiom (Alternative)

Welche der unten stehenden Hinweise zu seiner Formulierung des Merksatzes zum 3. Axiom möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde am ehesten geben?

☐ Deine Formulierung des 3. Axioms ist gut so, du kannst sie wieder so verwenden.
☐ **Es wäre besser, zuerst das Axiom (ohne Beispiel) zu notieren und anschließend ein Beispiel, um den Merksatz zu vereinfachen.**
☐ Es wäre besser, das Beispiel im Merksatz ausführlicher darzustellen, damit die Schüler*innen die Anwendung vom 3. Axiom verstehen können.
☐ **Es wäre sinnvoll, das Axiom um eine Skizze zu ergänzen, damit es anschaulicher wird.**
☐ **Es wäre besser, im Merksatz zu verdeutlichen, dass Kraft und Gegenkraft gleichzeitig wirken. Deine Formulierung („anschließend“ übt Körper B eine Gegenkraft auf Körper A aus) ist fachlich falsch.**
☐ **Du solltest im Merksatz über den Kräften Vektorpfeile notieren, damit die Wirkrichtung der Kräfte deutlich wird.**

☐ Du solltest im Merksatz verdeutlichen, dass Kraft und Gegenkraft in genau entgegengesetzte Richtungen wirken, damit die Schüler*innen die Bedeutung des Vorzeichens im Beispiel verstehen.

A.2 Zu Videovignette 2

Inhalt: Diskussion eines Comics zum 3. Newtonschen Axiom

Noticing-Aufgabe

Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung (Lob und/oder Kritik) erhalten, um etwas für sich bzw. seinen Unterricht zu lernen?

☐ **Gesprächsführung**
☐ **Einsatz des Comics**
☐ **Schülerfehlvorstellungen**

R2Bew1 - Schülerfehlvorstellungen (Beobachtung)

Welche der folgenden Schülerfehlvorstellungen sind während des Unterrichtsausschnitts sichtbar geworden?

☐ Es sind keine Schülerfehlvorstellungen sichtbar geworden.
☐ Die Schüler*innen denken, dass Kräfte nur dann auftreten, wenn sich zwei Körper berühren.
☐ **Die Schüler*innen schreiben dem Apfel eine Kraft als Eigenschaft zu.**
☐ **Die Schüler*innen denken, dass nur aktive (d. h. sich bewegende) Körper eine Kraft ausüben können, passive leisten nur Widerstand.**
☐ Die Schüler*innen denken, dass einer Kraft immer eine Beschleunigung zugrunde liegen muss und schließen daher eine Kraftwirkung bei ruhenden Gegenständen aus.
☐ **Die Schüler*innen verwechseln Kraft und Gegenkraft mit den wirkenden Kräften, die einen Körper im Kräftegleichgewicht halten.**
☐ Die Schüler*innen sehen nicht, dass der Apfel über den Baum mit der Erde verbunden ist und der Baum durch das „Festhalten“ an der Erde eine Gegenkraft zum am Baum „ziehenden“ Apfel ausübt.

R2Alt1 - Schülerfehlvorstellungen (Umgang)

Welche der unten stehenden Hinweise zum Umgang mit Schülerfehlvorstellungen möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde am ehesten geben?

☐ Dein Umgang mit den Schülerfehlvorstellungen ist angemessen, das kannst du wieder so machen.
☐ Du solltest den Prozess der Kraftwirkung thematisieren, um die Schülerfehlvorstellungen aus dem Weg zu räumen.
☐ **Du solltest den Unterschied zwischen Kraft und Energie thematisieren, um zu verdeutlichen, warum die Schülerfehlvorstellungen falsch sind.**
☐ **Du solltest erklären/verdeutlichen, dass Kräfte auch dann wirken (können), wenn man nichts aktiv tut. Dazu kannst du das Rollbrettexperiment verwenden: Zwei Schüler*innen stehen auf je einem Rollbrett und halten je ein Ende einer Schnur. Egal ob beide Schüler*innen ziehen oder nur eine*r, treffen sich die Rollbretter in der Mitte.**
☐ **Du solltest mit den Schüler*innen die verwendeten Begriffe und den Unterschied zwischen dem Kräftegleichgewicht und dem 3. Newtonschen Axiom klären, damit ihnen klar wird, dass warum ihre Vorstellungen falsch sind.**
☐ Du solltest die Schüler*innen das Wechselwirkungsprinzip ausführlich mündlich oder schriftlich formulieren lassen, damit du siehst, was die Schüler*innen verstanden haben bzw. wo Fehlvorstellungen bestehen.

A.3 Zu Videovignette 3

Inhalt: Einführung des Impulsbegriffs am Beispiel der Titanic

Noticing-Aufgabe

Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung (Lob und/oder Kritik) erhalten, um etwas für sich bzw. seinen Unterricht zu lernen?

☐ **Einsatz des Kontexts „Titanic“**
☐ Gesprächsführung
☐ **Fachliche Korrektheit**
☐ **Fachsprache**
☐ **Phasenübergang**

R3Bew1 - Fachsprache (Bewertung)

Welches Feedback möchten Sie Robert zur Nutzung der Fachsprache am ehesten geben?

- ☐ **Es ist nicht gut, dass die Schüler*innen den Begriff Schwung nutzen, da die Bedeutung des Begriffs noch nicht geklärt und ist so unklar ist, was genau die Schüler*innen beschreiben.**
- ☐ **Es ist gut, dass die Schüler*innen zunächst ihre Alltagssprache verwenden können um zu erklären, warum es zur Kollision gekommen ist. So sinkt die Hemmschwelle für die Schüler*innen, ihre Ideen zu äußern.**
- ☐ **Es ist gut, zunächst mit dem alltagssprachlichen Begriff „Schwung“ einzusteigen, da der Begriff „Schwung“ ein tragfähiges Konzept für den Impulsbegriff aus dem Alltag der Schüler*innen ist.**
- ☐ Die Formulierung „stärker bremsen“ ist nicht gut gewählt, da auch Körper mit großem Impuls, wie z. B. die Titanic, stärker bremsen könnten, um eine Kollision zu vermeiden.
- ☐ **Es ist nicht gut, dass die Schüler*innen mit dem Begriff „Schwung“ arbeiten. So werden sie sich diesen Begriff merken und zukünftig nicht mit dem Impulsbegriff arbeiten.**
- ☐ Es ist nicht gut, dass du die Schüler*innen in der Benutzung des Begriffs „Schwung“ bestärkst. So förderst du ihre Fehlvorstellungen (z. B. dass beide Begriffe gleichgesetzt werden können und dass Körper mehr Kraft haben, wenn sie mehr Schwung besitzen).
- ☐ Es ist nicht gut, dass du so stockend redest und viele Füllwörter nutzt. Das lenkt die Schüler*innen vom Inhalt ab.

R3Alt1 - Fachsprache (Alternative)

Welche der unten stehenden Hinweise zu der Nutzung von Fachsprache möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde am ehesten geben?

- ☐ **Es wäre besser, die Begriffe Masse und Impuls (ggf. mit der zugehörigen Formel) einzuführen, damit die Schüler*innen die korrekte Fachsprache lernen.**
- ☐ **Es wäre besser, wenn du die Schüler*innen dazu aufforderst, mehr Fachbegriffe zu nutzen (z. B. dem Massebegriff), damit sie sich so dem Impulsbegriff annähern können.**
- ☐ **Es wäre besser, den Impulsbegriff (für den Begriff „Schwung“) schon früher einzuführen, um zu verdeutlichen, um welches physikalische Konzept es geht.**

☐ Du solltest gegen die Schülervorstellung vorgehen, dass Körper „Schwung haben“ (vergleichbar mit der Schülervorstellung „Kraft haben“), damit diese sich nicht verfestigt
☐ **Die Verwendung der Begrifflichkeiten war angemessen, das kannst du wieder so durchführen.**
☐ Du solltest dir in der Vorbereitung genauer überlegen, was du den Schüler*innen erzählen möchtest, damit du flüssiger und motivierender vortragen kannst.
☐ Es wäre besser, die Schüler*innen zur Verwendung der Fachsprache aufzufordern, um zu vermeiden, dass sie die Konzepte Schwung und Impuls weiterhin verwechseln.

R3Bew2 - Phasenübergang (Bewertung)

Welches Feedback möchten Sie Robert zur Gestaltung des Phasenübergangs am ehesten geben?

☐ **Der Phasenübergang ist dir nicht gut gelungen, da er zu abrupt war.**
☐ Der Phasenübergang ist dir nicht gut gelungen, da er zu kurz war
☐ Der Phasenübergang ist die nicht gut gelungen, da die Inhalte zum Wechselwirkungsprinzip nicht ausreichend verstanden und gesichert wurden.
☐ Es ist unproblematisch, dass der Phasenübergang abrupt war, da du mit dem Impulsbegriff wirklich ein komplett neues Thema beginnst.
☐ **Es ist nicht gut, dass du so einen abrupten Übergang zwischen den Themen gemacht hast, da der rote Faden der Stunde so für die Schüler*innen nicht erkennbar ist.**
☐ **Es ist nicht gut, dass du sagst, das Thema ist „komplett neu“, da den Schüler*innen so das bisher Gelernte nicht mehr relevant erscheint.**
☐ **Es ist nicht gut, dass du sagst, das Thema ist „komplett neu“, da es einen fachlichen Bezug zwischen den Newtonschen Axiomen und dem Impulsbegriff gibt.**

R3Alt2 - Phasenübergang (Alternative)

Welche der unten stehenden Hinweise zur Gestaltung des Phasenübergangs möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde am ehesten geben?

☐ Den Phasenübergang hast du gut gestaltet, das kannst du wieder so machen.
☐ **Es wäre besser, den Zusammenhang zwischen den Newtonschen Axiomen und dem Impulsbegriff zu verdeutlichen, da die Schüler*innen den Zusammenhang nicht unbedingt selbst erkennen können.**

☐ Es wäre besser, den Einstieg in den Impulsbegriff auf die nächste Stunde zu verschieben, um den Schüler*innen zunächst noch mehr Möglichkeiten zu geben, das Wechselwirkungsprinzip zu verstehen und anzuwenden.
☐ Es wäre besser, vor dem Einstieg in das neue Thema auch die Tafel zu wischen, damit die Schüler*innen sich voll auf das neue Thema konzentrieren können und nicht noch über das alte Thema nachdenken.
☐ **Es wäre besser, einen glatteren Übergang zu gestalten, z. B. durch eine kurze Verknüpfung, wie „Wir finden jetzt etwas Neues mit dem gelernten Wissen heraus“, damit der rote Faden der Stunde für die Schüler*innen nachvollziehbar ist.**
☐ **Es wäre besser, den Übergang durch das Zeigen eines Experiments zu gestalten, um den Bruch zwischen den Themen für die Schüler*innen weniger hart zu gestalten.**

A.4 Zu Videovignette 4

Inhalt: Durchführung eines Freihandexperiments (Werfen von zwei Tennisbällen mit unterschiedlicher Masse bzw. unterschiedlicher Geschwindigkeit) zur Veranschaulichung der Zusammenhänge zwischen Masse bzw. Geschwindigkeit und Impuls

Noticing-Aufgabe

Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung (Lob und/oder Kritik) erhalten, um etwas für sich bzw. seinen Unterricht zu lernen?

☐ **Kontext/Kontextwechsel**
☐ **Fachsprache**
☐ Gesprächsführung
☐ **Wahl des Experiments**
☐ **Sicherung**

R4Bew1 - Wahl des Experiments (Bewertung)

Welches Feedback möchten Sie Robert zur Wahl des Experiments am ehesten geben?

☐ **Das Experiment ist gut gewählt, da die beiden Schüler, die das Experiment durchführen, den Impuls am eigenen Körper erfahren können.**

☐ **Das Experiment ist gut gewählt, da es den Einfluss von Masse und Geschwindigkeit auf den Impuls gut darstellt.**
☐ **Es ist nicht gut, dass nur über die Masse und die Geschwindigkeit gesprochen wurde, da unklar bleibt, ob der Impuls noch von anderen Größen, wie z. B. der Form abhängen kann.**
☐ **Das Experiment ist nicht gut gewählt, da es den Zusammenhang bzw. die Abgrenzung des Impulses zur kinetischen Energie nicht veranschaulicht.**
☐ Es ist gut, dass du die beiden Schüler in das Experiment einbezogen hast. So motivierst und aktivierst du sie.
☐ Es ist gut, dass du vorgibst, wer nach vorne kommen soll. So vermeidest zu lange Organisations-/Wartezeiten nach der Frage „Wer will denn...?“.
☐ Es ist gut, dass du nur zwei Schüler das Experiment durchführen lässt und nicht alle Schüler*innen, da du so Zeit sparst und das Experiment so einfach ist, dass es auch von den sitzenden Schüler*innen gut nachvollzogen werden kann.

R4Alt1 - Wahl des Experiments (Alternative)

Welche der unten stehenden Hinweise zur Wahl des Experiments möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde am ehesten geben?

☐ **Das Experiment ist gut gewählt, du kannst es erneut einsetzen.**
☐ Es wäre besser, statt des Experiments ein Video zu zeigen, in welchem gezeigt wird, welchen Einfluss verschiedene Massen und Geschwindigkeiten auf den Impuls haben, da das die Schüler*innen eher anspricht, als ein Experiment.
☐ **Es wäre besser, ein anderes Experiment durchzuführen, das für alle Schüler*innen anschaulich ist (z. B. eine Person, die auf einem Skateboard steht und einen Ball wegwirft, oder das Werfen beider Bälle in den Sand zur Betrachtung der „Krater“).**
☐ Es wäre besser, ein Stoßexperiment durchzuführen, da du so gut auf die Impulserhaltung überleiten kannst.
☐ **Es wäre besser, ein Experiment zu wählen, zu dem eine Messreihe aufgenommen werden kann, damit die Formel aus diesen objektiven Messungen abgeleitet werden kann.**
☐ Es wäre besser, wenn du das Experiment von allen Schüler*innen durchführen ließest, da es den Schüler*innen gut tut, sich zwischendurch zu bewegen.
☐ Du solltest darauf achten, dass die Schüler vorne so stehen, dass alle sitzenden Schüler*innen erkennen können, wie leicht/schwer es ist, den Ball zu fangen.

R4Bew2 - Sicherung (Bewertung)

Welches Feedback möchten Sie Robert zur Sicherung am ehesten geben?

- ☐ **Es ist nicht gut, dass du den Schüler*innen einen fertigen Merksatz vorgibst, da das reine Lesen und Abschreiben nicht verständnisfördernd ist.**
- ☐ Es ist gut, dass die Schüler*innen den Merksatz ins Heft übertragen, damit alle Schüler*innen den Merksatz als Sicherung notiert haben.
- ☐ **Es ist nicht gut, dass du die Ergebnisse des Experiments nicht sicherst, so ist unklar, was die Schüler*innen aus dem Experiment mitnehmen sollen.**
- ☐ Der Merksatz ist gut gelungen, da die „Je-umso"-Sätze die Zusammenhänge von Masse, Geschwindigkeit und Impuls nachvollziehbar darstellen.
- ☐ **Es ist nicht gut, dass du den Impuls als skalare Größe einführst. So grenzt du den Impuls nicht ausreichend von der kinetischen Energie ab.**
- ☐ Es ist gut, dass du den Begriff der Energie nicht erwähnst. So vermeidest du Verwirrung bei den Schüler*innen, die die verschiedenen Fachbegriffe auf Grundlage ihrer Erfahrung nicht trennen können.
- ☐ **In deiner Sicherung fehlt eine Abgrenzung des Impulsbegriffs von der kinetischen Energie und der Kraft, da die Begriffe den Schüler*innen zwar bekannt sind, aber Unterschiede für sie nicht ersichtlich sind.**

R4Alt2 - Sicherung (Alternative)

Welche der unten stehenden Hinweise zur Sicherung möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde am ehesten geben?

- ☐ Die Ergebnisse des Experiments hast du gut gesichert, das kannst du wieder so machen.
- ☐ Es wäre besser, die Beobachtungen des Experiments schriftlich an der Tafel zu sammeln (z. B. in einer Tabelle), damit du im Verlauf der Stunde darauf zurückgreifen kannst.
- ☐ Es wäre besser, die Schüler*innen in die Sicherung mit einzubeziehen (z. B. einen Merksatz formulieren zu lassen), damit sie stärker aktiviert und involviert werden.
- ☐ **Es wäre besser, zusätzlich zum Merksatz noch ein Beispiel zu zeigen, um auch die Formel anschaulich darzustellen.**
- ☐ **Es wäre besser, in der Sicherung auch auf das Experiment einzugehen (z. B. indem du auf der Folie zum Merksatz auch auf das Experiment eingehst oder Skizzen ergänzt), um die Inhalte stärker zu vernetzen.**

☐ **Es wäre besser, den Impuls als vektorielle Größe einzuführen, um ihn von der kinetischen Energie abzugrenzen.**
☐ **Es wäre besser, auch auf die kinetische Energie der Bälle einzugehen, um an das Vorwissen der Schüler*innen anzuknüpfen.**

A.5 Zu Videovignette 5

Inhalt: Durchführung eines Demonstrationsexperiments zur Impulserhaltung (Zusammenstoß zweier Wagen auf der Luftkissenbahn)

Noticing-Aufgabe

Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung (Lob und/oder Kritik) erhalten, um etwas für sich bzw. seinen Unterricht zu lernen?

☐ **Umgang mit den Schüler*innen**
☐ **Notation der Messwerte**
☐ **Durchführung/Auswertung des Experiments**
☐ **Messunsicherheiten**
☐ Wahl des Experiments/Experimenteller Aufbau
☐ **Aktivierung der Schüler*innen**
☐ **Kontext/Kontextwechsel**

R5Bew1 - Umgang mit den Schüer*innen (Bewertung)

Welches Feedback möchten Sie Robert zu seinem Umgang mit den Schüler*innen am ehesten geben?

☐ **Es ist nicht gut, dass du sagst, dass Mädchen eine schöne Schrift haben und Jungen besser Werte ablesen können, da du die Mädchen so auf eine schöne Handschrift reduzierst.**
☐ **Es ist nicht gut, dass du die Mädchen nur für Fleißarbeiten im Unterricht einbeziehst, da du so Stereotype förderst.**
☐ Es ist gut, dass du die Mädchen für ihr Handeln im Unterricht lobst, um ihr Selbstkonzept im Physikunterricht zu stärken.
☐ **Es ist nicht gut, dass du das Lob für die Handschrift von Mädchen mit einer inhaltlichen Kritik (vergessene Einheiten) kombinierst, da die Ursache hier nicht das Geschlecht ist.**
☐ Es ist nicht gut, dass du die Schülerin schon vor dem Notieren des 2. Messwertes erneut auf die Einheit hinweist, da du so das Bild festigst, Mädchen wären in Physik nicht kompetent.

☐ Es ist nicht gut, dass du der Schülerin an der Tafel direkt sagst, welche Einheit fehlt, da du ihr so unterstellst, sie wüsste das nicht.
☐ Es ist in Ordnung dass du sagst, dass Paula aufgrund ihrer schönen Handschrift die Werte notieren soll, da es wichtig ist, dass alle die Werte an der Tafel gut lesen können.

R5Alt1 - Umgang mit den Schüer*innen (Alternative)

Welche der unten stehenden Hinweise zum Umgang mit den Schüler*innen möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde am ehesten geben?

☐ Dein Umgang mit den Mädchen und Jungen ist in Ordnung so, da Paula wirklich eine gut lesbare Handschrift hat und du daher keine Vorurteile bedienst.
☐ Du solltest eine Geschichte um das Experiment herum konstruieren (und z. B. in Form eines Videos zeigen), um die Mädchen zu motivieren.
☐ **Es wäre besser, die Schüler*innen zu fragen, wer freiwillig schreiben bzw. die Werte ablesen möchte, um den Schüler*innen nicht nach Geschlecht eine Aufgabe zuzuweisen.**
☐ Es wäre besser, die Aufgaben (aufschreiben, Werte ablesen) durch zu tauschen, damit alle Schüler*innen eine Aufgabe wahrnehmen.
☐ **Du solltest darauf achten, die Mädchen nicht für ihre Handschrift, sondern explizit für ihre physikalische Kompetenz zu loben, da du sonst das Bild vermittelst, Physik wäre eine „Jungsfach".**
☐ **Du solltest auf geschlechtsspezifische Zuschreibungen verzichten, da du so ein unangemessenes Verhältnis zwischen den Geschlechtern im Physikunterricht erzeugst.**
☐ **Du solltest grundlegend an deiner Einstellung gegenüber Mädchen und Jungen arbeiten, um dein stereotypisches Verhalten zu ändern.**

R5Bew2 - Messunsicherheiten (Bewertung)

Welches Feedback möchten Sie Robert zu seinem Umgang mit Messunsicherheiten am ehesten geben?

☐ **Es ist nicht gut, dass du sagst, du könntest „exakt" oder „genau" messen, da du so die Schülervorstellung beförderst, man könnte Größen ohne Messunsicherheiten bestimmen.**
☐ **Du gehst nicht auf die Messunsicherheiten ein. Das ist nicht gut, da diese eine wichtige Rolle in der Physik spielen.**
☐ Es ist nicht gut, dass du nur einen der beiden Wagen wiegst, da die Schüler*innen so nicht sicher sein können, dass die Wagen wirklich gleich schwer sind.
☐ **Es ist nicht gut, dass du sagst, die Wagen wären gleich schwer, da das – auch wenn sie baugleich sind – nicht stimmen muss.**

☐ **Es ist nicht gut, dass du das Gewicht mit 0,1 kg angibst, da du hier nicht berücksichtigst, welches die signifikanten Nachkommastellen sind.**
☐ Es ist gut, dass du darauf verzichtest, auch auf Messunsicherheiten einzugehen, um die Komplexität des Experiments überschaubar zu halten.
☐ Es ist gut, dass du das Gewicht des Wagens mit 0,1 kg angibst, um Fehler in der Berechnung des Impulses aufgrund vieler Nachkommastellen zu vermeiden.

R5Alt2 - Messunsicherheiten (Alternative)

Welche der unten stehenden Hinweise zu seinem Umgang mit Messunsicherheiten möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde am ehesten geben?

☐ Es ist gut, wie du mit Messunsicherheiten umgegangen bist. Das kannst du erneut so machen.
☐ **Es wäre besser, alle Nachkommastellen anzugeben, die die Laborwaage dir anzeigt, da du den Fokus auf genaue Messungen legst und die Laborwaage genauer als 0,1 kg wiegen kann.**
☐ **Es wäre besser, anstelle der Aussage, du könntest „exakt“ messen, auch Messunsicherheiten einzubeziehen, da Messungen nie wirklich exakt sind.**
☐ Es wäre besser, beide Wagen zu wiegen, damit du den Schüler*innen ein korrektes Bild von genauem wissenschaftlichen Arbeiten vermittelst.
☐ **Es wäre besser, beide Wagen zu wiegen, da die Wagen – auch wenn sie baugleich sind – nicht gleich schwer sein müssen.**
☐ **Es wäre besser, beide Wagen zu wiegen, da du im Vergleich der Massen auf Messunsicherheiten bzw. Ungenauigkeiten eingehen kannst.**
☐ Es wäre besser, beide Wagen zu wiegen, damit die Schüler*innen sich selbst überzeugen können, dass die Wagen gleich schwer sind.

R5Bew3 - Aktivierung der Schüler*innen (Bewertung)

Welches Feedback möchten Sie Robert zur Aktivierung der Schüler*innen am ehesten geben?

☐ **Es ist nicht gut, dass du die Schüler*innen nur oberflächlich involvierst, da die Schüler*innen trotz dieser Aktivierung nicht mitdenken müssen.**
☐ Es ist gut, dass du Schüler*innen in das Experiment einbeziehst, um sie im Unterricht zu aktiveren.
☐ Es ist gut, dass du das Experiment selbst durchführst, da du in der Durchführung im Vergleich zu den Schüler*innen dabei sehr sicher bist.

☐ **Es ist nicht gut, dass die Schüler*innen dir ohne Arbeitsauftrag zusehen, da die Schüler*innen so keinen Anreiz haben, aufmerksam zu sein.**
☐ Es ist nicht gut, dass du die Masse des Wagens und insbesondere die Geschwindigkeit selbst bestimmst, da du den Schüler*innen vorher schon erklärt hast, wie sie das selbst machen könnten.
☐ Es ist nicht gut, dass du das Experiment so lehrerzentriert durchführst, da dies für die Schüler*innen langweilig ist.
☐ **Es ist nicht gut, dass Paula von der Tafel aus und die Schüler*innen aus den hinteren Reihen den experimentellen Aufbau nicht gut erkennen können, da es für sie schwer ist, dir zu folgen.**

R5Alt3 - Aktivierung der Schüler*innen (Alternative)

Welche der unten stehenden Hinweise zur Aktivierung der Schüler*innen möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde am ehesten geben?

☐ **Es wäre besser, die Schüler*innen zu fragen, wer freiwillig schreiben bzw. die Werte ablesen möchte, um den Schüler*innen nicht nach Geschlecht eine Aufgabe zuzuweisen.**
☐ Es wäre besser, die Schüler*innen zu fragen, ob sie den Ablauf verstanden oder noch Schwierigkeiten haben, um den Schüler*innen die Möglichkeit für Rückfragen zu geben.
☐ Du hast die Schüler*innen ausreichend aktiviert. Das kannst du wieder so machen.
☐ Es wäre besser, Hypothesen der Schüler*innen abzufragen, um sie zur Beobachtung des Experiments zu motivieren.
☐ **Es wäre besser, den Schüler*innen Fragen zum Vorgehen während des Experiments oder zur Berechnung der Geschwindigkeit zu stellen, um sie zum Mitdenken anzuregen.**
☐ Es wäre besser, die Schüler*innen stärker in die Durchführung des Experiments einzubeziehen (z. B. um Wagen anzustoßen, Werte abzulesen oder die Geschwindigkeit zu bestimmen), um die Schüler*innen aktiv einzubinden.
☐ **Es wäre besser, die Schüler*innen zum Experiment nach vorn kommen zu lassen, damit sie alles gut erkennen können.**

A.6 Zu Videovignette 6

Inhalt: Experimentelle Überprüfung der Impulserhaltung mit den Messwerten des Demonstrationsexperiments

Noticing-Aufgabe

Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung (Lob und/oder Kritik) erhalten, um etwas für sich bzw. seinen Unterricht zu lernen?

☐ **Umgang mit den Schüler*innen**
☐ **Umgang mit dem Einwand von Marie/Messunsicherheiten**
☐ **Einheit des Impulses**

R6Bew1 - Umgang mit den Schüler*innen (Bewertung)

Welches Feedback möchten Sie Robert zu Umgang mit den Schüler*innen am ehesten geben?

☐ Es ist nicht gut, dass du „Mädels“ sagst, da das ein sexistischer Ausdruck ist.
☐ **Es ist nicht gut, dass du die Mädchen und Jungen so unterschiedlich behandelst, da das nicht wertschätzend gegenüber den Mädchen ist.**
☐ **Es ist nicht gut, dass du die Mädchen und Jungen in ihrem Können so unterschiedlich einschätzt, da du so die Vorstellung förderst, Physik wäre ein „Jungsfach“.**
☐ **Du thematisierst die Einheit des Impulses nicht. Das ist nicht gut, da dies inkonsistent zur früheren Szene ist, in der du bei Paula auf die Notation der Einheit der Geschwindigkeit bestanden hast.**
☐ Es ist gut, dass du die Mädchen dazu aufforderst, ihre Tabelle zu korrigieren, damit du sicher sein kannst, dass sie die gleichen Werte, wie die Jungen notiert haben.

R6Alt1 - Umgang mit den Schüler*innen (Alternative)

Welche der unten stehenden Hinweise zu seinem Umgang mit den Schüler*innen möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde am ehesten geben?

☐ Dein Umgang mit den Schüler*innen ist in Ordnung, du kannst ihn so beibehalten.

☐ **Du solltest die Mädchen stärker in den Unterricht einbeziehen (z. B. berechnete Werte vorlesen lassen), um ihr physikalisches Selbstkonzept zu stärken.**
☐ **Du solltest auf geschlechtsspezifische Zuschreibungen verzichten, um die Schülerinnen nicht zu diskriminieren.**
☐ Du solltest von jedem*r Schüler*in je nur einen berechneten Wert erfragen, da du so die berechneten Werte von mehr Schüler*innen kontrollieren kannst.
☐ Du solltest verstärkt schwächere Schüler*innen in den Unterricht einbinden, da wichtig ist, dass sie die Inhalte verstehen.
☐ **Du solltest die Schüler*innen unabhängig vom Geschlecht dazu auffordern, falsche Ergebnisse in der Tabelle zu korrigieren, damit am Ende alle Schüler*innen die richtigen Ergebnisse in der Tabelle notiert haben.**
☐ **Du solltest hinterfragen, woher die geschlechtsspezifischen Zuschreibungen stammen, die du vertrittst, um sie zu überwinden.**

R6Bew1 - Umgang mit dem Einwand von Marie/Messunsicherheiten (Bewertung)

Welches Feedback möchten Sie Robert zu seinem Umgang mit Maries Einwand bzw. mit Messunsicherheiten am ehesten geben?

☐ Die Reaktion auf Maries Einwand ist dir gut gelungen, da du so möglichst effizient zum zentralen Ergebnis der Impulserhaltung gelangst.
☐ Die Reaktion auf Maries Einwand ist dir gut gelungen, da Rundungsfehler eine zentrale Fehlerquelle für die Unterschiede der Ergebnisse sind.
☐ Die Reaktion auf Maries Einwand ist dir nicht gut gelungen, da du den Fokus von der experimentellen Bestätigung der Impulserhaltung hin zur Diskussion von leichten Ungenauigkeiten verschiebst.
☐ **Es ist nicht gut, dass du nur so knapp auf den Einwand von Marie eingehst, da die Begründung nicht ausreicht, um den Schüler*innen zu verdeutlichen, dass die Unterschiede aufgrund von Messunsicherheiten zustande kommen.**
☐ Es ist nicht gut, dass du nur so knapp auf den Einwand von Marie eingehst, da so unklar bleibt, ob bzw. wo die Schüler*innen Rundungsfehler gemacht haben.
☐ **Es ist nicht gut, dass du die Ungleichheit der Messwerte mit „Rundungsfehlern“ begründest, da diese Begründung falsch ist.**
☐ **Es ist nicht gut, dass du die Ungleichheit der Messwerte mit „Rundungsfehlern“ begründest, da du so den Schüler*innen die Schuld dafür gibst, dass die Ergebnisse nicht exakt gleich sind.**

R6Alt1 - Umgang mit dem Einwand von Marie/Messunsicherheiten (Alternative)

Welche der unten stehenden Hinweise zu seinem Umgang mit Maries Einwand bzw. mit Messunsicherheiten möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde am ehesten geben?

☐ Deine Reaktion auf den Einwand von Marie ist angemessen, das kannst du erneut so machen.
☐ Du solltest den Schüler*innen angeben, auf wie viele Nachkommastellen sie runden sollen (1 oder 2 Stellen), damit du diese Situation umgehen kannst.
☐ Es wäre besser, den Schüler*innen am Anfang mitzuteilen, dass Messunsicherheiten und Rundungsfehler nicht berücksichtigt werden, um diese Situation zu vermeiden.
☐ Es wäre besser, ausführlicher zu begründen, wo die Rundungsfehler entstehen, damit die Schüler*innen deiner Erklärung besser folgen können.
☐ **Du solltest bereits bei der Vorstellung des Experiments über die vorhandene Reibung und Energieumwandlung beim Aufprall sprechen, damit den Schüler*innen in dieser Phase klar ist, dass die Ergebnisse aufgrund dieser Einflüsse nicht exakt gleich sein können.**
☐ **Es wäre besser, zu thematisieren, dass bei der Durchführung des Versuches Messunsicherheiten berücksichtigt werden müssen, damit die Schüler*innen nachvollziehen können, wieso es zu Abweichungen kommt.**
☐ **Du solltest das Thema Messunsicherheiten für dich selbst noch einmal aufbereiten, damit du von solchen Fragen nicht vom Thema abgebracht wirst.**

R6Bew2 - Einheit des Impulses (Bewertung)

Welches Feedback möchten Sie Robert zu seinem Umgang mit der Einheit des Impulses am ehesten geben?

☐ **Es ist nicht gut, dass du nicht über die Einheit des Impulses sprichst, da die Einheit ein wichtiger Bestandteil der Angabe des Impulses ist.**
☐ **Es ist nicht gut, dass du nicht über die Einheit des Impulses sprichst, da dies inkonsistent zur früheren Szene ist, in der du auf die Notation der Einheit der Geschwindigkeit bestanden hast.**
☐ Es ist gut, dass du die Einheit des Impulses nicht so stark in den Fokus gerückt hast, um die Schüler*innen nicht durch die neue Einheit zu verwirren.
☐ Es ist gut, dass du den Fokus auf die berechneten Werte und nicht auf die Einheit gelegt hast, da die gleichen Zahlen ausreichen, um die Impulserhaltung auf Grundlage des Experiments zu bestätigen.

☐ **Es ist nicht gut, dass du die Schüler*innen nicht dazu aufforderst, die Einheiten anzugeben, da so unklar ist, ob die Impulse wirklich gleich groß sind.**

R6Alt2 - Einheit des Impulses (Alternative)

Welche der unten stehenden Hinweise zu seinem Umgang mit der Einheit des Impulses möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde am ehesten geben?

☐ Es ist in Ordnung, dass du die Einheit des Impulses zunächst vernachlässigst, das kannst du erneut so machen.
☐ **Du solltest die Einheit des Impulses angeben, da die Einheit ein wichtiger Bestandteil der Angabe des Impulses ist.**
☐ **Du solltest darauf achten, dass die Schüler beim Vorlesen der Werte auch die Einheit angeben, um in deinem eigenen Verhalten konsistent zu bleiben.**
☐ Du solltest die Einheit des Impulses oben über die Spalte der Tabelle notieren, um die Schreibarbeit in jeder Zeile zu sparen.
☐ **Du solltest die Schüler*innen dazu auffordern, die Einheit des Impulses zu nennen, damit die Schüler*innen sich an die Einheit des Impulses gewöhnen.**

A.7 Zu Videovignette 7

Inhalt: Sicherung der Erkenntnisse aus dem Demonstrationsexperiment zur Impulserhaltung mit einem vorgefertigten Merksatz

Noticing-Aufgabe

Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung (Lob und/oder Kritik) erhalten, um etwas für sich bzw. seinen Unterricht zu lernen?

☐ **Formulierung des Merksatzes zur Impulserhaltung**
☐ **Interaktion mit Schüler*innen**

R7Bew1 - Formulierung des Merksatzes zur Impulserhaltung (Bewertung)

Welches Feedback möchten Sie Robert zu seiner Formulierung des Merksatzes zur Impulserhaltung am ehesten geben?

☐ Die Formulierung des Merksatzes ist dir gut gelungen, da alle wichtigen Informationen korrekt enthalten sind.
☐ Es ist gut, dass du im Merksatz auf das durchgeführte Experiment verweist, da die Schüler*innen den Merksatz direkt mit der Realität verbinden können.
☐ **Die Formulierung des Merksatzes ist dir nicht gut gelungen, da sie für die Schüler*innen zu anspruchsvoll ist.**
☐ **Es ist nicht gut, dass du deinen vorgefertigten Merksatz nutzt und die Merksätze der Schüler*innen nicht einbeziehst, da das den Schüler*innen gegenüber nicht wertschätzend ist.**

R7Alt1 - Formulierung des Merksatzes zur Impulserhaltung (Alternative)

Formulierung des Merksatzes zur Impulserhaltung – Alternative: Welche der unten stehenden Hinweise zu seiner Formulierung des Merksatzes zur Impulserhaltung möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde am ehesten geben?

☐ Die Formulierung des Merksatzes ist angemessen, du kannst sie erneut verwenden.
☐ **Es wäre besser, Beispiel und Merksatz zu trennen, um die Komplexität des Merksatzes zu reduzieren.**
☐ Es wäre besser, zusätzlich zum langen Merksatz eine kurze Version zu formulieren, da sich die Schüler*innen eine kurze Version besser merken können.
☐ **Es wäre besser, den Merksatz nur in Bezug auf einen Stoßprozess zu formulieren, damit der Merksatz näher am besprochenen Experiment bleibt.**
☐ Es wäre besser, den Merksatz um die Einheit des Impulses zu ergänzen, da die Einheit stets zur Angabe einer physikalischen Größe gehört.
☐ **Du solltest den Merksatz schülernäher gestalten (z. B. indem du eine Skizze ergänzt oder die Formulierung sprachlich vereinfachst), um die Schüler*innen bei Verständnis zu unterstützen.**
☐ **Es wäre besser, eine Formulierung der Schüler*innen als Merksatz zu verwenden, da diese Formulierungen für sie besser verständlich sind.**

R7Bew2 - Interaktion mit den Schüler*innen (Bewertung)

Welches Feedback möchten Sie Robert zu seiner Interaktion mit den Schüler*innen am ehesten geben?

☐ Es ist gut, dass die Schüler*innen ihre Merksätze vorlesen, da du so erkennen kannst, ob sie dir Impulserhaltung umfassend verstanden haben.
☐ **Es ist gut, dass die Schüler*innen einen Merksatz formulieren sollen, um ein tiefergehendes Verständnis zu fördern.**
☐ **Es ist nicht gut, dass du den Schüler*innen keine Rückmeldung auf die Merksätze gibst, da die Schüler*innen so nicht einordnen können, ob ihre Merksätze angemessen formuliert sind.**
☐ **Es ist nicht gut, dass du deinen vorgefertigten Merksatz nutzt und die Merksätze der Schüler*innen nicht einbeziehst, da das den Schüler*innen gegenüber nicht wertschätzend ist.**
☐ Es ist nicht gut, dass du deine Formulierung als „besser“ bezeichnest, da du die Formulierungen der Schüler*innen so abwertest
☐ Es ist nicht gut, dass nur zwei Schüler*innen ihre Merksätze vorgestellt haben, da du so nicht weißt, welche Merksätze die restlichen Schüler*innen formuliert haben.
☐ **Es ist nicht gut, dass nur zwei Jungen ihre Merksätze vorgestellt haben, da die Schüler*innen so lernen, dass du nur Jungen drannimmst.**

R7Alt2 - Interaktion mit den Schüler*innen (Alternative)

Welche der unten stehenden Hinweise zu seiner Interaktion mit den Schüler*innen möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde am ehesten geben?

☐ Du solltest den Merksatz ausführlicher erklären bzw. mit den Schüler*innen besprechen, da einige Aspekte des Merksatzes nicht selbsterklärend sind.
☐ **Du solltest gemeinsam mit den Schüler*innen einen Merksatz erstellen, um sie stärker in den Unterricht einzubeziehen.**
☐ Es wäre besser, insgesamt mehr Schüler*innen ihre Merksätze vorstellen zu lassen, da du aus den Antworten Rückschlüsse auf die Vorstellungen der Schüler*innen ziehen kannst.
☐ **Du solltest die Schüler*innen sich melden lassen, wenn sie ihren Merksatz vorstellen wollen, da sich möglicherweise auch ein Mädchen gemeldet hätte.**
☐ Der Umgang mit den Schüler*innen war angemessen, so kannst du dich wieder verhalten.
☐ Du solltest deinen Umgang mit den Schüler*innen überdenken. Dass du sie und ihre Arbeit nicht lobst, kann dem Verhältnis der Schüler*innen zum Physikunterricht nachhaltig schaden.
☐ **Du solltest darauf aufpassen, keine Rollenklischees in den Unterricht einzubringen, um die Schüler*innen nicht in diese Muster zu drängen.**

A.8 Zur globalen Reflexion

Globale Reflexion der Doppelstunde; keine Videovignette

Noticing-Aufgabe

Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung (Lob und/oder Kritik) erhalten, um etwas für sich bzw. seinen Unterricht zu lernen?

☐ **Ziele**
☐ **Ablauf der Unterrichtsstunde**
☐ **Auswahl und Umfang der Stundeninhalte**
☐ Medieneinsatz
☐ Auftreten der Lehrkraft
☐ **Umgang mit Schüler*innen/Interaktion**
☐ **Fachliche Korrektheit**

RGBew1 - Interaktion mit den Schüler*innen (Bewertung)

Welches Feedback möchten Sie Robert zu seiner Interaktion mit den Schüler*innen am ehesten geben?

☐ Es ist nicht gut, dass du die Beiträge von Schüler*innen kaum berücksichtigst, da dies demotivierend für die Schüler*innen ist.
☐ **Es ist nicht gut, dass du so lehrerzentriert unterrichtest, da die Schüler*innen so nicht aktiv ein Verständnis für die Inhalte entwickeln.**
☐ **Es ist nicht gut, dass du die Jungen und Mädchen unterschiedlich behandelst, da das demotivierend ist.**
☐ Es ist gut, dass du Sicherungsphasen in deinen Unterricht eingebaut hast, damit du sichergehst, dass die Schüler*innen die zentralen Inhalte aus der Stunde mitnehmen.
☐ **Es ist gut, dass du verschiedene Sozialformen im Unterricht einplanst (Einzelarbeit, Partnerarbeit usw.), da du den Unterricht so abwechslungsreicher gestaltest.**
☐ Es ist nicht gut, dass du im Unterricht nur von wenige Schüler*innen Beiträge anhörst, da die restlichen Schüler*innen so nicht aktiviert werden.
☐ **Deine Verständnisversicherung mit einer Frage, wie „Habt ihr noch Fragen?“ ist nicht zielführend, da Schüler*innen auf diese Frage eher nicht antworten.**
☐ **Du thematisierst die auftretenden Schülervorstellungen nicht. Das ist nicht gut, da sich diese so verfestigen.**

RGAlt1 - Interaktion mit den Schüler*innen (Alternative)

Welche der unten stehenden Hinweise zu seiner Interaktion mit den Schüler*innen möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde am ehesten geben?

☐ Dein Umgang mit den Schüler*innen ist angemessen, du kannst ihn so beibehalten.
☐ **Du solltest die Schüler*innen grundsätzlich stärker in den Unterricht einbeziehen (z. B. in Diskussionen, bei der Durchführung von Experimenten oder der Formulierung von Merksätzen), um die Impulserhaltung gemeinsam mit den Schüler*innen zu erarbeiten.**
☐ Du solltest den Schüler*innen ausführlicher Feedback zu ihren Beiträgen geben, damit die Schüler*innen einordnen können, ob ihre Aussagen korrekt sind.
☐ **Du solltest dir die Schülervorstellungen zum Thema vor dem Unterricht ansehen, damit du im Unterricht angemessen auf diese eingehen kannst. um das Verständnis der Schüler*innen zu fördern.**
☐ **Du solltest dir grundsätzlich Gedanken darüber machen, wie (z. B. mit welchen Aufgabenstellungen) du das Verständnis der Schüler*innen abprüfen kannst, um sicherzugehen, dass die Schüler*innen die Ziele der Stunde erreichen.**
☐ Du solltest deinen Umgang mit den Schüler*innen grundsätzlich überdenken. Dass du sie nicht lobst und wertschätzt, kann dem Verhältnis der Schüler*innen zum Physikunterricht nachhaltig schaden.
☐ **Es ist wichtig, dass du darauf aufpasst, keine Rollenklischees in den Unterricht einzubringen, um die Schüler*innen nicht in diese Muster zu drängen.**

RGBew2 - Fachliche Korrektheit (Bewertung)

Welches Feedback möchten Sie Robert zur fachlichen Korrektheit der Stunde am ehesten geben?

☐ **Deinen Unterricht hast du insgesamt gut durchgeführt, da er überwiegend fachlich korrekt war.**
☐ **Du hast fachliche Fehler in der Unterrichtsstunde gemacht. Das ist nicht gut, da du den Schüler*innen so falsche Inhalte beibringst.**
☐ Es ist nicht gut, dass du die Abweichungen beim experimentellen Nachweis der Impulserhaltung als Rundungsfehler darstellst. Das ist fachlich falsch.
☐ Es ist nicht gut, dass du die Einheit des Impulses nicht besprichst, da die Einheit ein wichtiger Bestandteil der Angabe eines Impulses ist.
☐ **Du thematisierst die auftretenden Schülervorstellungen nicht. Das ist nicht gut, da sich diese so verfestigen.**

☐ Es ist gut, dass du so großen Wert auf die Einheit der Geschwindigkeit in der Tabelle legst, da die Einheit zentraler Bestandteil einer Größenangabe ist.
☐ Es ist nicht gut, dass du bei der Durchführung des Demonstrationsexperimentes sagst, du würdest „exakt“ messen, da du den Schüler*innen das Bild vermittelst, du könntest Messunsicherheiten vermeiden.

RGAlt2 - Fachliche Korrektheit (Alternative)

Welche der unten stehenden Hinweise zur fachlichen Korrektheit der Stunde möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde am ehesten geben?

☐ Du solltest auch bei der Besprechung der Impulse auf die Einheiten achten, um einheitlich vorzugehen.
☐ **Es wäre besser, die Schülervorstellungen, die im Unterricht auftreten (z. B. dass ein Apfel Kraft „hat“), aufzugreifen, um den Schüler*innen die fachlich korrekte Sichtweise näher zu bringen.**
☐ **Du solltest dir die Schülervorstellungen zum Thema vor dem Unterricht angucken, damit du mit ihnen vertraut bist und Ideen entwickeln kannst, wie du ihnen begegnen kannst.**
☐ Es wäre besser, auch auf Messunsicherheiten beim Fahrbahn-Experiment einzugehen, damit die Schüler*innen die Impulserhaltung trotz der leicht unterschiedlichen Werte akzeptieren.
☐ **Du solltest darauf achten, dass deine Merksätze fehlerfrei formuliert sind, damit die Schüler*innen keine Fehler aus deinen Unterlagen übernehmen.**
☐ Du solltest stärker auf deine Wortwahl achten, um nicht aus Versehen falsche Vorstellungen zu befördern.

B *R*-Skript zur Erstellung der Rückmeldungen

Die Rückmeldungen werden in Form von PDF-Dateien bereitgestellt. Die folgenden beiden Programme zeigen die vorbereitende Auswertung der Ergebnisse und Erstellung der Diagramme (Kapitel B.1,) die anschließend gemeinsam mit den Textelementen in die PDF-Dateien eingebunden werden (Kapitel B.2). Die Zuordnung der Ergebnisse und Diagramme zu den Dateien erfolgt über den im Testinstrument erhobenen persönlichen Code.[206]

B.1 Auswertung und Erstellung der Diagramme

```
1  library(tidyverse)
2  library(ggbrace)
3  library(showtext)
4  font_add("Arial", regular="arial.ttf", bold = "arialbd.ttf",
       italic = "ariali.ttf", bolditalic = "arialbi.ttf")
5  showtext_auto()
6  options(OutDec= ",")
7
8  results <- read.csv("./results-survey516586.csv", header=
       TRUE, sep="\t", encoding="UTF-8")
9  results[is.na(results)]<-0
10 results <- mutate(results,
11                   ScoreRefl1 = (Refl02bIIIBew.SQ002.+
                        Refl02bIIIBew.SQ003.+Refl02bIIIBew.SQ004
                        .+Refl02bIIIBew.SQ007.)/4+(Refl02bIIIAlt
                        .SQ002.+Refl02bIIIAlt.SQ004.+
                        Refl02bIIIAlt.SQ005.+Refl02bIIIAlt.SQ006
                        .)/4,
```

[206] Die Rückmeldung für Dozierende wird grundsätzlich analog erstellt. Änderungen liegen bzgl. der erstellten Diagramme und Textelemente vor; die Zuordnung der Kohorte erfolgt über einen „Kohortencode“, der den Teilnehmenden vor der Bearbeitung des Reflexionstests über die Dozierenden mitgeteilt und von ihnen im Reflexionstest angegeben wird.

```
12                    ScoreRefl2 = (Refl03bIVBeob.SQ003.+
                         Refl03bIVBeob.SQ004.+Refl03bIVBeob.SQ006
                         .)/3+(Refl03IVUmgang.SQ003.+
                         Refl03IVUmgang.SQ004.+Refl03IVUmgang.
                         SQ005.)/3,
13                    ScoreRefl3 = (Refl06bIVBew.SQ001.+
                         Refl06bIVBew.SQ002.+Refl06bIVBew.SQ003.+
                         Refl06bIVBew.SQ005.)/4+(Refl06bIVAlt.
                         SQ001.+Refl06bIVAlt.SQ002.+Refl06bIVAlt.
                         SQ003.+Refl06bIVAlt.SQ005.)/4+(
                         Refl06bVBew.SQ001.+Refl06bVBew.SQ005.+
                         Refl06bVBew.SQ006.+Refl06bVBew.SQ007.)/
                         4+(Refl06bVAlt.SQ002.+Refl06bVAlt.SQ005
                         .+Refl06bVAlt.SQ006.)/3,
14                    ScoreRefl4 = (Refl07bIVBew.SQ001.+
                         Refl07bIVBew.SQ002.+Refl07bIVBew.SQ003.+
                         Refl07bIVBew.SQ004.)/4+(Refl07bIVAlt.
                         SQ001.+Refl07bIVAlt.SQ003.+Refl07bIVAlt.
                         SQ005.)/3+(Refl07bVBew.SQ001.+
                         Refl07bVBew.SQ003.+Refl07bVBew.SQ005.+
                         Refl07bVBew.SQ007.)/4+(Refl07bVAlt.SQ004
                         .+Refl07bVAlt.SQ005.+Refl07bVAlt.SQ006.+
                         Refl07bVAlt.SQ007.)/4,
15                    SocreRefl5 = (Refl09bIBew.SQ001.+
                         Refl09bIBew.SQ002.+Refl09bIBew.SQ004.)/
                         3+(Refl09bIAlt.SQ003.+Refl09bIAlt.SQ005
                         .+Refl09bIAlt.SQ006.+Refl09bIAlt.SQ007.)
                         /4+(Refl09bIVBew.SQ001.+Refl09bIVBew.
                         SQ002.+Refl09bIVBew.SQ004.+Refl09bIVBew.
                         SQ005.)/4+(Refl09bIVAlt.SQ002.+
                         Refl09bIVAlt.SQ003.+Refl09bIVAlt.SQ005.+
                         Refl09bIVAlt.SQ006.)/4+(Refl09bVIBew.
                         SQ001.+Refl09bVIBew.SQ004.+Refl09bVIBew.
                         SQ007.)/3+(Refl09bVIAlt.SQ001.+
                         Refl09bVIAlt.SQ005.+Refl09bVIAlt.SQ007.)
                         /3,
16                    ScoreRefl6 = (Refl12bIBew.SQ002.+
                         Refl12bIBew.SQ003.+Refl12bIBew.SQ004.)/
                         3+(Refl12bIAlt.SQ002.+Refl12bIAlt.SQ003
                         .+Refl12bIAlt.SQ006.+Refl12bIAlt.SQ007.)
                         /4+(Refl12bIIBew.SQ004.+Refl12bIIBew.
                         SQ006.+Refl12bIIBew.SQ007.)/3+(
                         Refl12bIIAlt.SQ005.+Refl12bIIAlt.SQ006.+
                         Refl12bIIAlt.SQ007.)/3+(Refl12bIIIBew.
                         SQ001.+Refl12bIIIBew.SQ002.+
                         Refl12bIIIBew.SQ005.)/3+(Refl12bIIIAlt.
```

```
                    SQ002.+Refl12bIIIAlt.SQ003.+
                    Refl12bIIIAlt.SQ005.)/3,
17                  ScoreRefl7 = (Refl13bIBew.SQ003.+
                    Refl13bIBew.SQ004.)/2+(Refl13bIAlt.SQ002
                    .+Refl13bIAlt.SQ004.+Refl13bIAlt.SQ006.+
                    Refl13bIAlt.SQ007.)/4+(Refl13bIIBew.
                    SQ002.+Refl13bIIBew.SQ003.+Refl13bIIBew.
                    SQ004.+Refl13bIIBew.SQ007.)/4+(
                    Refl13bIIAlt.SQ002.+Refl13bIIAlt.SQ004.+
                    Refl13bIIAlt.SQ007.)/3,
18                  ScoreReflG = (GesbVIBew.SQ002.+GesbVIBew.
                    SQ003.+GesbVIBew.SQ005.+GesbVIBew.SQ007
                    .+GesbVIBew.SQ008.)/5+(GesVIAlt.SQ002.+
                    GesVIAlt.SQ004.+GesVIAlt.SQ005.+GesVIAlt
                    .SQ007.)/4+(GesVIIBew.SQ001.+GesVIIBew.
                    SQ002.+GesVIIBew.SQ005.)/3+(GesVIIAlt.
                    SQ002.+GesVIIAlt.SQ003.+GesVIIAlt.SQ005
                    .)/3,
19                  Score = ScoreRefl1+ScoreRefl2+ScoreRefl3+
                    ScoreRefl4+SocreRefl5+ScoreRefl6+
                    ScoreRefl7+ScoreReflG,
20                  ScoreProz = Score/32*100,
21                  ScoreBew = (Refl02bIIIBew.SQ002.+
                    Refl02bIIIBew.SQ003.+Refl02bIIIBew.SQ004
                    .)/4+(Refl06bIVBew.SQ001.+Refl06bIVBew.
                    SQ002.+Refl06bIVBew.SQ003.+Refl06bIVBew.
                    SQ005.)/4+(Refl06bVBew.SQ001.+
                    Refl06bVBew.SQ005.+Refl06bVBew.SQ006.+
                    Refl06bVBew.SQ007.)/4+(Refl07bIVBew.
                    SQ001.+Refl07bIVBew.SQ002.+Refl07bIVBew.
                    SQ003.+Refl07bIVBew.SQ004.)/4+(
                    Refl07bVBew.SQ001.+Refl07bVBew.SQ003.+
                    Refl07bVBew.SQ005.+Refl07bVBew.SQ007.)/
                    4+(Refl09bIBew.SQ001.+Refl09bIBew.SQ002
                    .+Refl09bIBew.SQ004.)/3+(Refl09bIVBew.
                    SQ001.+Refl09bIVBew.SQ002.+Refl09bIVBew.
                    SQ004.+Refl09bIVBew.SQ005.)/4+(
                    Refl09bVIBew.SQ001.+Refl09bVIBew.SQ004.+
                    Refl09bVIBew.SQ007.)/3+(Refl12bIBew.
                    SQ002.+Refl12bIBew.SQ003.+Refl12bIBew.
                    SQ004.)/3+(Refl12bIIBew.SQ004.+
                    Refl12bIIBew.SQ006.+Refl12bIIBew.SQ007.)
                    /3+(Refl12bIIIBew.SQ001.+Refl12bIIIBew.
                    SQ002.+Refl12bIIIBew.SQ005.)/3+(
                    Refl13bIBew.SQ003.+Refl13bIBew.SQ004.)/
                    2+(Refl13bIIBew.SQ002.+Refl13bIIBew.
                    SQ003.+Refl13bIIBew.SQ004.+Refl13bIIBew.
```

```
                        SQ007.)/4+(GesbVIBew.SQ002.+GesbVIBew.
                        SQ003.+GesbVIBew.SQ005.+GesbVIBew.SQ007
                        .+GesbVIBew.SQ008.)/5+(GesVIIBew.SQ001.+
                        GesVIIBew.SQ002.+GesVIIBew.SQ005.)/3,
22                    ScoreBewProz = ScoreBew/15*100,
23                    ScoreAlt = (Refl02bIIIAlt.SQ002.+
                        Refl02bIIIAlt.SQ004.+Refl02bIIIAlt.SQ005
                        .+Refl02bIIIAlt.SQ006.)/4+(Refl06bIVAlt.
                        SQ001.+Refl06bIVAlt.SQ002.+Refl06bIVAlt.
                        SQ003.+Refl06bIVAlt.SQ005.)/4+(
                        Refl06bVAlt.SQ002.+Refl06bVAlt.SQ005.+
                        Refl06bVAlt.SQ006.)/3+(Refl07bIVAlt.
                        SQ001.+Refl07bIVAlt.SQ003.+Refl07bIVAlt.
                        SQ005.)/3+(Refl07bVAlt.SQ004.+
                        Refl07bVAlt.SQ005.+Refl07bVAlt.SQ006.+
                        Refl07bVAlt.SQ007.)/4+(Refl09bIAlt.SQ003
                        .+Refl09bIAlt.SQ005.+Refl09bIAlt.SQ006.+
                        Refl09bIAlt.SQ007.)/4+(Refl09bIVAlt.
                        SQ002.+Refl09bIVAlt.SQ003.+Refl09bIVAlt.
                        SQ005.+Refl09bIVAlt.SQ006.)/4+(
                        Refl09bVIAlt.SQ001.+Refl09bVIAlt.SQ005.+
                        Refl09bVIAlt.SQ007.)/3+(Refl12bIAlt.
                        SQ002.+Refl12bIAlt.SQ003.+Refl12bIAlt.
                        SQ006.+Refl12bIAlt.SQ007.)/4+(
                        Refl12bIIAlt.SQ005.+Refl12bIIAlt.SQ006.+
                        Refl12bIIAlt.SQ007.)/3+(Refl12bIIIAlt.
                        SQ002.+Refl12bIIIAlt.SQ003.+
                        Refl12bIIIAlt.SQ005.)/3+(Refl13bIAlt.
                        SQ002.+Refl13bIAlt.SQ004.+Refl13bIAlt.
                        SQ006.+Refl13bIAlt.SQ007.)/4+(
                        Refl13bIIAlt.SQ002.+Refl13bIIAlt.SQ004.+
                        Refl13bIIAlt.SQ007.)/3+(GesVIAlt.SQ002.+
                        GesVIAlt.SQ004.+GesVIAlt.SQ005.+GesVIAlt
                        .SQ007.)/4+(GesVIIAlt.SQ002.+GesVIIAlt.
                        SQ003.+GesVIIAlt.SQ005.)/3,
24                    ScoreAltProz = ScoreAlt/15*100,
25                    ScoreSV = (Refl03bIVBeob.SQ003.+
                        Refl03bIVBeob.SQ004.+Refl03bIVBeob.SQ006
                        .)/3+(Refl03IVUmgang.SQ003.+
                        Refl03IVUmgang.SQ004.+Refl03IVUmgang.
                        SQ005.)/3,
26                    ScoreSVProz = ScoreSV/2*100,
27                    ScoreNoticing = 1
28 )
29 results = results[-21,]
30
31 PC1 <- toupper(results$PersCode.PC1.)
```

```
32 PC1[69] <- "X"
33 PC1[78] <- "X"
34 PC2 <- toupper(results$PersCode.PC2.)
35 PC2[69] <- "X"
36 PC2[78] <- "X"
37 PC3 <- toupper(results$PersCode.PC3.)
38 PC3[69] <- "X"
39 PC3[78] <- "X"
40 PC4 <- toupper(results$PersCode.PC4.)
41 PC4[69] <- "1"
42 PC4[78] <- "2"
43 FS <- as.numeric(gsub("A", "", results$FSPhy))
44 FS_mean <- round(mean(FS),1)
45 N <- length(PC1)
46 Q <- quantile(results$ScoreProz, probs = c(0, 0.25, 0.5,
     0.75, 1))
47 code <- numeric(0)
48 einordnung <- numeric(0)
49 foerderhinweis <- numeric(0)
50 foerderhinweis_SV <- numeric(0)
51 for (i in 1:N){
52   code <- c(code,paste(PC1[i],PC2[i],PC3[i],PC4[i], sep=""))
53   einordnung <- c(einordnung, if (results$ScoreProz[i]<Q[2])
       {("der␣unteren␣25")} else if (Q[2]<=results$ScoreProz[i
     ] & results$ScoreProz[i]<Q[3]) {("der␣unteren␣25-50")}
     else if (Q[3]<=results$ScoreProz[i] & results$ScoreProz[
     i]<Q[4]) {("der␣oberen␣50-25")} else {("der␣oberen␣25")
     })
54   foerderhinweis <- c(foerderhinweis, if ((results$
     ScoreBewProz[i]-results$ScoreAltProz[i]) < 0) {print("
     Bewertungen")} else if ((results$ScoreBewProz[i]-results
     $ScoreAltProz[i]) > 0) {print("Alternativen")} else {
     print("Bewertungen␣und␣Alternativen")})
55   foerderhinweis_SV <- c(foerderhinweis_SV, if (results$
     ScoreSVProz[i] < 50) {paste("Außerdem␣empfehlen␣wir,␣
     dass␣Sie␣sich␣mit␣", "typischen␣Fehlvorstellungen␣von␣
     Schüler*innen", "␣und␣dem␣Umgang␣damit␣beschäfitgen.␣
     Hinweise␣dazu␣finden␣Sie␣u.␣a.␣im␣Buch:␣Schecker,␣H.,␣
     Wilhelm,␣T.,␣Hopf,␣M.␣&␣Duit,␣R.␣(Hrsg.)␣(2018).␣", "Sch
     ülervorstellungen␣und␣Physikunterricht", ".␣Berlin,␣
     Heidelberg:␣Springer␣Spektrum.", sep="")} else {print("
     ")})
56 }
57
58 FS.i <- numeric(0)
59 FS_mean.i <- numeric(0)
60 N.i <- numeric(0)
```

```
N.Ref.i <- numeric(0)

for (i in 1:N){
  z.sub <- subset(results[i,])
  z.sub2 <- subset(results[-i,])

  FS.i <- na.omit(as.numeric(gsub("A", "", z.sub2$FSPhy)))
  FS_mean.i <- c(FS_mean.i, round(mean(FS.i),1))
  N.Ref.i <- c(N.Ref.i, sum(z.sub2$FSPhy == ""))
  N.i <- c(N.i, length(FS.i))

  ggplot()+
    geom_boxplot(data=z.sub2, mapping = aes(x="Gesamt",y=
       ScoreProz), coef=100)+
    stat_boxplot(data=z.sub2, mapping = aes(x="Gesamt",y=
       ScoreProz), geom = "errorbar", width = 0.5, coef=100)+
    geom_point(data=z.sub, mapping=aes(x="Gesamt",y=
       ScoreProz, shape="Ihr Ergebnis"), color="red", size=3)
       +
    geom_hline(yintercept=z.sub$ScoreProz, linetype='dotted'
       , col = 'red')+
    scale_x_discrete(limits=c("Gesamt",""))+
    ylim(0,100)+
    xlab("")+
    ylab("Ergebnis in %")+
    theme(legend.position="right")+
    theme(legend.title=element_blank())+
    theme(text = element_text(family = "Arial"))+
    annotate("text", x=0.5, y=Q[3], label="Median", size=3)+
    geom_brace(aes(c(1.4,1.5), y=c(Q[1],Q[2]), label="untere
       25% der Ergebnisse\nin der Vergleichsgruppe"),
       inherit.data=FALSE, labelsize=3, rotate=90)+
    geom_brace(aes(c(1.5,1.6), y=c(Q[2],Q[4]), label="
       mittlere 50% der Ergebnisse\nin der Vergleichsgruppe")
       , inherit.data=FALSE, labelsize=3, rotate=90)+
    geom_brace(aes(c(1.4,1.5), y=c(Q[4],Q[5]), label="obere
       25% der Ergebnisse\nin der Vergleichsgruppe"), inherit
       .data=FALSE, labelsize=3, rotate=90)
  ggsave(file=paste("./Plots/Ges_",PC1[i],PC2[i],PC3[i],PC4[
     i],".pdf", sep=""), width=6, height=4)
  graphics.off()
  ggplot()+
    geom_boxplot(data=z.sub2, mapping = aes(x="1",y=
       ScoreBewProz), coef=100)+
    stat_boxplot(data=z.sub2, mapping = aes(x="1",y=
       ScoreBewProz), geom = "errorbar", width = 0.5, coef
       =100)+
```

```
93      geom_point(data=z.sub, mapping=aes(x="1",y=ScoreBewProz,
            shape="Ihre Teilergebnisse"), color="red", size=2)+
94      geom_boxplot(data=z.sub2, mapping = aes(x="2",y=
           ScoreAltProz), coef=100)+
95      stat_boxplot(data=z.sub2, mapping = aes(x="2",y=
           ScoreAltProz), geom = "errorbar", width = 0.5, coef
           =100)+
96      geom_point(data=z.sub, mapping=aes(x="2",y=ScoreAltProz,
            shape="Ihre Teilergebnisse"), color="red", size=2)+
97      geom_boxplot(data=z.sub2, mapping = aes(x="3",y=
           ScoreSVProz), coef=100)+
98      stat_boxplot(data=z.sub2, mapping = aes(x="3",y=
           ScoreSVProz), geom = "errorbar", width = 0.5, coef
           =100)+
99      geom_point(data=z.sub, mapping=aes(x="3",y=ScoreSVProz,
           shape="Ihre Teilergebnisse"), color="red", size=2)+
100     ylim(-1,101)+
101     xlab("")+
102     scale_x_discrete(labels=c("1"="Bewertungen","2"="
           Alternativen","3"="SV"))+
103     ylab("Ergebnis in %")+
104     theme(legend.position="right")+
105     theme(legend.title=element_blank())+
106     theme(text = element_text(family = "Arial"))
107   ggsave(file=paste("./Plots/Teile_",PC1[i],PC2[i],PC3[i],
         PC4[i],".pdf", sep=""), width=5, height=3)
108   graphics.off()
109 }
110
111 for (i in 191:N){
112   rmarkdown::render("Auswertung_TN_Foerderhinweis.Rmd",
113                     output_file = paste("./TN/Ergebnis_",PC1
                          [i],PC2[i],PC3[i],PC4[i],".pdf", sep="
                          "),
114                     params=list(Person = code[i],
                          Gesamtscore = results$ScoreProz[i],
                          Fachsemester = FS_mean.i[i], Anzahl =
                          N.i[i]))
115 }
```

B.2 Erstellung der Rückmeldungen

```
---
output:
  pdf_document:
    latex_engine: xelatex
header-includes:
  - \usepackage[ngerman]{babel}
  - \usepackage{fontspec}
  - \setmainfont{Arial}
  - \definecolor{PPblau}{RGB}{0,61,127}
  - \geometry{a4paper, left=2.5cm, right=2.5cm, top=1cm,
     bottom=2.5cm, headheight=1.5cm, includehead}
  - \usepackage{pdfpages}
  - \usepackage[labelfont=it]{caption}
  - \addto\captionsngerman{\renewcommand{\figurename}{Abb.}}
  - \usepackage{fancyhdr}
  - \pagestyle{fancy}
  - \fancyhf{}
  - \fancyhead[L]{\textcolor{PPblau}{Rückmeldung der
     Ergebnisse zum Reflexionstest}}
  - \fancyhead[C]{}
  - \fancyhead[R]{\parbox[t][5cm][t]{5cm}{\includegraphics[
     width=5cm]{transfer-bunt-mit.png}}}
  - \renewcommand{\headrulewidth}{0pt}
  - ''r␣paste0("\\fancyfoot[L]{","\\textcolor{PPblau}{Ihr␣
     Testcode:␣",params$Person,"}}")''
  - \fancyfoot[R]{\textcolor{PPblau}{\thepage}}
  - \renewcommand{\footrulewidth}{0pt}
  - \setlength{\parskip}{11pt}
  - \newcommand{\Titel}[1]{\begin{center}\fontsize{22pt}{15
     pt}\selectfont\textcolor{PPblau}{\textbf{#1}}\end{center
     }}
  - \newcommand{\Ueberschrift}[1]{\vspace{6pt}\par\fontsize
     {13pt}{0pt}\selectfont\textcolor{PPblau}{\textbf{#1}}\
     vspace{-6pt}\fontsize{11pt}{15pt}\selectfont}
params:
        Person: ""
        Gesamtscore: ""
        Fachsemester: ""
        Anzahl: ""
---
\thispagestyle{empty}
\includepdf{Titelseite_Reflexion_Teilnehmende.pdf}

```

```
\newpage
\setcounter{page}{1}
\fontsize{11pt}{15pt}\selectfont
```{r setup, include=FALSE}
knitr::opts_chunk$set(echo = FALSE, fig.align="center")
```
\textbf{Sehr geehrte:r Teilnehmer:in,}

Sie haben kürzlich am sogenannten Reflexionstest teilgenommen, in dem eine zentrale Fähigkeit für das fundierte, theoriegeleitete Handeln von Lehrer:innen - die Reflexionsfähigkeit - erhoben wird. Diese Rückmeldung soll Ihnen dabei helfen, sich in Bezug auf diese Fähigkeit realistisch einschätzen zu können.

\fbox{
\begin{minipage}{\linewidth}
Insgesamt haben Sie \textbf{`r round(params$Gesamtscore,1)`\% der maximal möglichen Punktzahl} erreicht. Damit liegt ihr Ergebnis im Bereich `r einordnung[i]`\% der Testergebnisse der Teilnehmenden aus der Vergleichsgruppe (bestehend aus `r N.i[i]` Studierenden im durchschnittlich `r FS_mean.i[i]`. Fachsemester und `r N.Ref.i[i]` Referendar*innen).

\end{minipage}
}

Im Folgenden finden Sie eine detailliertere Darstellung Ihrer Ergebnisse sowie erläuternde Informationen.

\Ueberschrift{Reflektieren - Was ist das genau \& Wie kommt das im Test vor?}

Ziel einer Reflexion ist die Verbesserung der Unterrichtsqualität und bzw. oder die Weiterentwicklung persönlicher Kompetenzen.\footnote{Das dem Test und dieser Rückmeldung zugrundeliegende Reflexionsmodell ist beschrieben in: Nowak, A., Kempin, M., Kulgemeyer, C. \& Borowski, A. (2019). Reflexion von Physikunterricht. In: Christian Maurer (Hrsg.). Naturwissenschaftliche Bildung als Grundlage für berufliche und gesellschaftliche Teilhabe. Regensburg, S. 838-841.} Im Rahmen einer Reflexion können unterschiedliche Aussagen getätigt werden:

* Beobachtete Aspekte können **beschrieben** werden,
```

```
60 * diese können **bewertet** werden,
61 * es können kurz- und langfristige **Alternativen**
      vorgeschlagen werden: Kurzfristige Alternativen beinhalten
       alternative Vorgehensweisen (oder eine Bestätigung der
      gewählten Vorgehensweise) zu den einzelnen Aspekten, lä
      ngerfristige Alternativen entsprechen Konsequenzen für die
       eigene Professionalisierung bzw. die Professionalisierung
       des Gegenübers sowie für den Folgeunterricht.
62
63 Von diesen Aspekten wurden im Testinstrument das **Bewerten*
      * und **Vorschlagen von Alternativen** fokussiert.
64
65 Die Aspekte, über die reflektiert werden kann, sind vielfä
      ltig. Im Test wurden Ihnen zu jedem Unterrichtsausschnitt
      Aspekte vorgegeben, zu denen Sie in je einer Aufgabe (1)
      das Vorgehen der Lehrkraft **bewerten** und (2) **
      Alternativen** zum Vorgehen vorschlagen sollten.
66
67 Auch die Diagnose von und der Umgang mit **Schü
      lervorstellungen** sind zentrale Aspekte der
      Physikdidaktik, die im Testinstrument in Bezug auf einen
      Unterrichtsausschnitt explizit adressiert werden. Zu
      diesem Ausschnitt waren Sie aufgefordert, **beobachtete
      Schülervorstellungen zu benennen** und **Strategien zum
      Umgang** mit diesen vorzuschlagen - hier gilt also
      ausnahmsweise nicht die Struktur von je einer Frage zum
      Bewerten und Vorschlagen von Alternativen.
68
69 \newpage
70 \Ueberschrift{Rückmeldung zur Ihrer Testbearbeitung:}
71
72 Im Rahmen der Testbearbeitung konnten Sie Rückmeldung zu
      verschiedenen Aspekten der beobachteten
      Unterrichtsausschnitte gegeben. Das folgende Diagramm (s.
      Abb. 1) zeigt Ihnen Ihr **Gesamtergebnis** (roter Punkt)
      im Verhältnis zu den Gesamtergebnissen aller bisherigen
      Testbearbeitungen von `r N.i[i]` Studierenden im
      durchschnittlich `r FS_mean.i[i]`. Fachsemester und `r N.
      Ref.i[i]` Referendar*innen (dargestellt durch den Boxplot)
      . Die Angabe der Ergebnisse erfolgt in Prozent und zeigt
      Ihnen, wie viele der maximal erreichbaren Punkte Sie (und
      alle bisherigen Testbearbeitungen) erreicht haben.
73
74 Zum einfacheren Verständnis ist dieser erste Boxplot zu
      Ihrem Gesamtergebnis beschriftet. Weitere Hinweise zur
      Interpretation der Diagramme finden Sie untenstehend.
75
```

```
\begin{figure}
\centering
\includegraphics[width=\linewidth]{./Plots/Ges_`r params$
   Person`.pdf}
\vspace{-0.5cm}
\caption{Ihr Gesamtergebnis}
\end{figure}

Im nachfolgenden Diagramm (s. Abb. 2) ist ihr Testergebnis (
   rote Punkte) aufgeschlüsselt in die **Teilergebnisse** zu
   den Aspekten \glqq Bewertungen\grqq{}, \glqq Alternativen\
   grqq{} und \glqq Schülervorstellungen\grqq{}, ebenfalls im
    Verhältnis zu allen bisherigen Testbearbeitungen von `r N
   .i[i]` Studierenden im durchschnittlich `r FS_mean.i[i]`.
   Fachsemester und `r N.Ref.i[i]` Referendar*innen (
   dargestellt durch die Boxplots) und ebenfalls als
   prozentualer Anteil der maximal erreichbaren Punkte,
   abgebildet. Eine Erläuterung der Aufschlüsselung der
   Teilergebnisse finden Sie untenstehend.

\begin{figure}
\centering
\includegraphics[width=\linewidth]{./Plots/Teile_`r params$
   Person`.pdf}
\vspace{-0.5cm}
\caption{Ihre Teilergebnisse}
\end{figure}

**Bewertungen:** Unter dem Aspekt \glqq Bewertungen\grqq{}
   sind alle Items zusammengefasst, in dessen Rahmen
   verschiedene Aspekte der Unterrichtsausschnitte bewertet
   werden können (jeweils die erste Frage zu einem Aspekt).

**Alternativen/Konsequenzen:** Unter dem Aspekt \glqq
   Alternativen\grqq{} sind alle Items zusammengefasst, in
   dessen Rahmen Alternativen zu verschiedenen Aspekten der
   Unterrichtsausschnitte benannt werden können (jeweils die
   zweite Frage zu einem Aspekt).

**Schülervorstellungen (SV):** Unter den Aspekt \glqq Schü
   lervorstellungen\grqq{} fallen die beiden Items, in denen
   Sie beobachtete Schülervorstellungen benennen und
   Strategien zum Umgang mit diesen vorschlagen sollten.
\textit{Hinweis: Die Anzahl der Items zu Schü
   lervorstellungen ist also deutlich geringer, als die zu
   Bewertungen und Alternativen - diesem Aspekt kommt also
   deutlich weniger Gewicht in der Gesamtwertung zu.}
```

```
98
99  \Ueberschrift{Wie kann ich meine Reflexionsfähigkeit schulen
       ?}
100
101 Im Rahmen Ihres Studiums erlangen Sie Wissen (z. B. zu Schü
       lervorstellungen oder zu Experimenten), welches Sie für
       eine fundierte Reflexion benötigen. Dieses Wissen dient
       Ihnen als (theoretische) Grundlage für Ihre Bewertungen
       oder alternativen Handlungsvorschläge zu gesehenem oder
       selbst durchgeführtem Unterricht.
102 Reflektieren ist praktisches Handeln und muss daher auch
       durch die Anwendung trainiert werden. Fördermaterial mit
       Informationen und Hinweisen, um das Reflektieren praktisch
        zu üben, finden Sie hier: \textcolor{PPblau}{\url{https:/
       /umfragen.physikdidaktik.uni-bremen.de/index.php/149738?
       lang=de}}.
103
104 Auf Grundlage Ihrer Ergebnisse im Diagnose-Instrument
       empfehlen wir Ihnen, bei der Bearbeitung des Fö
       rdermaterials einen \textbf{Fokus auf die Entwicklung von
       `r foerderhinweis[i]`} zu legen.
105
106 `r foerderhinweis_SV[i]`
107
108 \Ueberschrift{Hinweise zur Interpretation der Diagramme:}
109
110 * Ihr Testergebnis wird jeweils durch den roten Punkt
       markiert.
111 * Der ebenfalls eingezeichnete Boxplot stellt jeweils eine Ü
       bersicht über alle bisherigen Testergebnisse von
       Studierenden unterschiedlicher Fachsemester (überwiegend
       Studierende in der Mitte des Bachelorstudiums sowie
       Studierende im Praxissemester) dar.
112 * Alle bisherigen Testbearbeitungen liegen im Bereich
       zwischen der oberen und unteren Grenze der \glqq Antennen\
       grqq{} im Diagramm.
113 * Im Rahmen der Box liegen die mittleren 50% der bisherigen
       Testergebnisse.
114 * Die horizontale Linie markiert den Median der bisherigen
       Testergebnisse. Der Median teilt die Testergebnisse in der
        Mitte. Es liegt also genau eine Hälfte der bisherigen
       Testergebnisse über bzw. unter dem Median.
115
116 \
117 \
118 \fbox{
119 \begin{minipage}{\linewidth}
```

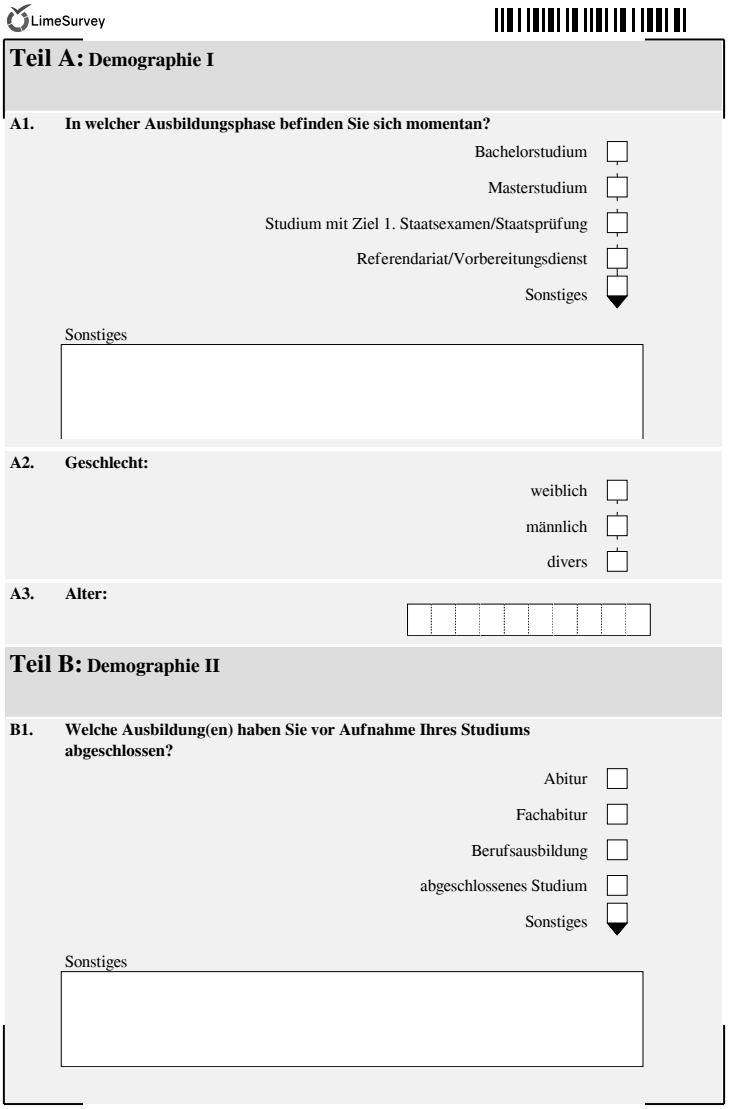
```
120 Wir freuen uns über Ihr \textbf{Feedback zum Testinstrument
      und insbesondere auch zu dieser Rückmeldung}! Folgen Sie
      dafür einfach diesem Link (\url{https://umfragen.
      physikdidaktik.uni-bremen.de/index.php/851195?lang=de})
      oder wenden Sie sich direkt an  \url{anna.weissbach@uni-
      bremen.de}.
121 \end{minipage}
122 }
```

C Zugang zum Selbstlernmaterial

Das Selbstlernmaterial ergänzt als Teil des Assessment-Feedbacks die Rückmeldung der Testwerte um die Möglichkeit zur weiterführenden Auseinandersetzung mit Unterrichtsreflexion (im Sinne des Feed Up nach Hattie und Timperley (2007)). Es umfasst Informationen zur Konzeptualisierung von Unterrichtsreflexion, Hinweise zur Durchführung einer Unterrichtsreflexion sowie die Möglichkeit, drei weitere Videovignetten angeleitet zu reflektieren und die eigene Reflexion mit einer Beispiellösung abzugleichen. Dazu sind die drei Vignetten analog zur Darstellung im Reflexionstest mit Hilfestellungen eingebunden, die Beispiellösung kann schrittweise eingesehen werden (einen ersten Eindruck bietet Abbildung 11.2 auf Seite 159).

Das Selbstlernmaterial ist vollständig in der Software *LimeSurvey* umgesetzt und unter folgendem Link abzurufen: https://umfragen.physikdidaktik.uni-bremen.de/index.php/149738?lang=de.

D Fragebogen zur Demografie



LimeSurvey

Teil A: Demographie I

A1. In welcher Ausbildungsphase befinden Sie sich momentan?

- ☐ Bachelorstudium
- ☐ Masterstudium
- ☐ Studium mit Ziel 1. Staatsexamen/Staatsprüfung
- ☐ Referendariat/Vorbereitungsdienst
- ☐ Sonstiges

Sonstiges

A2. Geschlecht:

- ☐ weiblich
- ☐ männlich
- ☐ divers

A3. Alter:

Teil B: Demographie II

B1. Welche Ausbildung(en) haben Sie vor Aufnahme Ihres Studiums abgeschlossen?

- ☐ Abitur
- ☐ Fachabitur
- ☐ Berufsausbildung
- ☐ abgeschlossenes Studium
- ☐ Sonstiges

Sonstiges

LimeSurvey

B2. **In welchem Beruf haben Sie Ihre Ausbildung abgeschlossen?**

B3. **In welchem Fach haben Sie Ihr Studium abgeschlossen?**

B4. **Haben Sie Ihre Hochschulreife in Deutschland erlangt?**

Ja ☐

Nein ☐

B5. **Meine Abiturnote betrug *(z.B. 2,4)*:**

Alternativ: Welche Abschlussnote haben Sie in Ihrer letzten Ausbildung erreicht?

B6. **Welches Abitur haben Sie abgelegt?**

Abitur nach Klasse 12 (G8) ☐

Abitur nach Klasse 13 (G9) ☐

Sonstiges ☐

Sonstiges

B7. **Meine Abschlussnote für die Hochschulzugangsberechtigung betrug:**

LimeSurvey

B8. **Bitte geben Sie Ihre Schulnoten als Punkte im Bereich von 15 bis 0 an, wobei 15 Punkte der Note 1+ und 0 Punkte der Note 6 entsprechen.**

15 Punkte: 1+, 14 Punkte: 1, 13 Punkte: 1-,

12 Punkte: 2+, 11 Punkte: 2, 10 Punkte: 2-,

9 Punkte: 3+, 8 Punkte: 3, 7 Punkte: 3-,

6 Punkte: 4+, 5 Punkte: 4, 4 Punkte: 4-,

3 Punkte: 5+, 2 Punkte: 5, 1 Punkt: 5-,

0 Punkte: 6

Meine letzte Physiknote war:

Meine letzte Mathematiknote war:

Meine letzte Deutschnote war:

B9. **Physik war ein Prüfungsfach in meiner Abschlussprüfung (Abitur oder alternative Prüfung zur Hochschulzugangsberechtigung).**

Ja ☐

Nein ☐

B10. **Mathematik war ein Prüfungsfach in meiner Abschlussprüfung (Abitur oder alternative Prüfung zur Hochschulzugangsberechtigung).**

Ja ☐

Nein ☐

LimeSurvey

B11. **Bitte geben Sie die Anzahl der im Physikstudium (Bachelor und Master zusammengezählt) verbrachten Semester inklusive dem aktuellen Semester an. Andere Fächer zählen nicht!**

- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5
- ☐ 6
- ☐ 7
- ☐ 8
- ☐ 9
- ☐ 10
- ☐ 11
- ☐ 12
- ☐ 13
- ☐ 14
- ☐ 15
- ☐ 16
- ☐ 17
- ☐ 18
- ☐ 19
- ☐ 20+

B12. **Welchen Abschluss streben Sie momentan an?**

- ☐ Physik-Lehramt (Bachelor/Master/Staatsexamen)
- ☐ Physik-Mono (Bachelor/Master/Diplom)
- ☐ Sonstiges

Sonstiges

LimeSurvey

B13. **Welches Fach bzw. welche Fächer studieren Sie *neben* Physik?**

- ☐ Keine
- ☐ Sonstiges

Sonstiges

B14. **Anzahl der Fachsemester im zweiten Fach (Bachelor und Master zusammengezählt) inklusive dem aktuellen Semester:**

- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5
- ☐ 6
- ☐ 7
- ☐ 8
- ☐ 9
- ☐ 10
- ☐ 11
- ☐ 12
- ☐ 13
- ☐ 14
- ☐ 15
- ☐ 16
- ☐ 17
- ☐ 18
- ☐ 19
- ☐ 20+

B15. **Haben Sie im Rahmen Ihres Studiums bereits Schulpraktika absolviert?**

Ja ☐
Nein ☐

B16. **Anzahl der Wochen:**

B17. **Haben Sie neben den Pflichtpraktika noch weitere freiwillige Schulpraktika absolviert?**

Ja ☐
Nein ☐

B18. **Anzahl der Wochen:**

B19. **Waren Sie als Aushilfslehrkraft/Vertretungslehrkraft tätig?**

Ja ☐
Nein ☐

B20. **In welchem Umfang waren Sie als Aushilfslehrkraft/Vertretungslehrkraft tätig?**

Anzahl der Unterrichtsstunden pro Woche:

Anzahl der Schulhalbjahre:

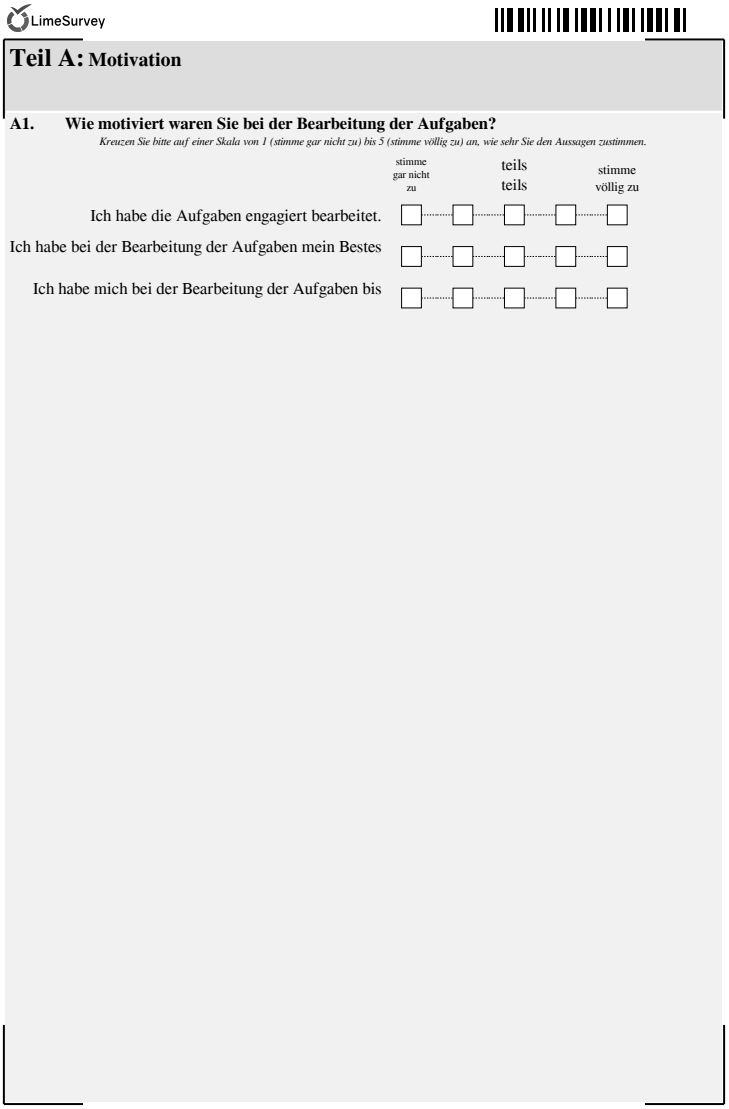
B21. **Haben Sie außerschulische Lehrerfahrungen (als Nachhilfelehrer, Trainingsleiter im Sportverein etc.)?**

Ja ☐
Nein ☐

B22. **Haben Sie Lehr-Erfahrung an der Universität (als Tutor, Übungsleiter, Betreuer im Praktikum etc.)?**

Ja ☐
Nein ☐

E Items zur Motivation während der Testbearbeitung



LimeSurvey

Teil A: Motivation

A1. Wie motiviert waren Sie bei der Bearbeitung der Aufgaben?

Kreuzen Sie bitte auf einer Skala von 1 (stimme gar nicht zu) bis 5 (stimme völlig zu) an, wie sehr Sie den Aussagen zustimmen.

	stimme gar nicht zu		teils teils		stimme völlig zu
Ich habe die Aufgaben engagiert bearbeitet.	☐	☐	☐	☐	☐
Ich habe bei der Bearbeitung der Aufgaben mein Bestes	☐	☐	☐	☐	☐
Ich habe mich bei der Bearbeitung der Aufgaben bis	☐	☐	☐	☐	☐

F Vollständiger Akzeptanzfragebogen

Nachfolgend sind die zur Erfassung der Akzeptanz verwendeten Items zugeordnet zu den Faktoren des TAM dargestellt. Zusätzlich ist markiert, zu welchem Zeitpunkt die Beantwortung der Items intendiert ist: (1) Am Ende der Testbearbeitung, (2) nach Erhalt der Rückmeldung, vor der Bearbeitung des Selbstlernmaterials sowie (3) am Ende der Bearbeitung des Selbstlernmaterials (hier im Rahmen der Entwicklung als „Fördermaterial“ bezeichnet). Eingeschätzt werden die Items auf einer vierstufigen Likert-Skala mit der zusätzlichen Option „keine Antwort“.

Relevanz für den Beruf

(1) Die Unterrichtsausschnitte zeigen authentische Situationen.
(1) Die Unterrichtsausschnitte sind gut geeignet, um über Unterricht zu reflektieren.
(1) Ich finde es wichtig, Unterricht zu reflektieren.
(1) Die Fragen zu den Unterrichtsausschnitten haben einen angemessenen Schwierigkeitsgrad.
(1) Die Fragen zu den Unterrichtsausschnitten waren zu umfangreich.
(1) Die Fragen zu den Unterrichtsausschnitten waren zu einfach.
(2) Die Inhalte der Rückmeldung sind für mein Studium relevant.
(3) Die Inhalte der des Materials sind für mein Studium relevant.

Qualität der Ergebnisse

(2) Die Rückmeldung gibt mir neue Informationen in Bezug auf meine Reflexionsfähigkeit.
(2) Die Rückmeldung hilft mir dabei, meine Reflexionsfähigkeit besser einschätzen zu können.
(1) Die Rückmeldung zu meiner Testbearbeitung interessiert mich.
(2) Es hat sich gelohnt, den Test gemacht und die Rückmeldung erhalten zu haben.
(2) Der Aufwand für Testbearbeitung und Arbeit mit der Rückmeldung ist im Vergleich zum Nutzen zu hoch

Subjektive Norm

(2) Mein*e Dozent*in würde es gutheißen, falls ich das Fördermaterial nutze.
(2) Mein*e Dozent*in würde es gutheißen, falls ich den Test zukünftig erneut durchführe, um meinen Fortschritt zu überprüfen
(2) Viele meiner Kommiliton*innen sehen einen Mehrwert in der Nutzung der Lernumgebung (Test, Rückmeldung und Fördermaterial).

Wahrgenommener Nutzen

(2) Ich finde die Rückmeldung nützlich für mein Studium .
(3) Die Nutzung des Materials unterstützt mich bei der Verbesserung meiner Reflexionsfähigkeit.
(3) Ich finde das Material nützlich für mein Studium.

Wahrgenommene Bedienbarkeit

(1) Bei der Testdurchführung traten technische Schwierigkeiten auf.
(2) Beim Herunterladen der Rückmeldung traten Schwierigkeiten auf.
(3) Bei der Arbeit mit dem Material traten technische Schwierigkeiten auf.
(1) Bei der Bearbeitung der Fragen habe ich immer eine Antwortoption gefunden, die zu meinen Überlegungen gepasst hat.
(1) Bei der Bearbeitung der Fragen fehlte mir die Möglichkeit, selbst weitere Antwortoptionen zu ergänzen.

Wahrung der Anonymität

(1) Ich befürchte, dass meine Testergebnisse mir im Nachhinein zugeordnet werden können.

Nutzungsabsicht

(2) Ich möchte den Test in Zukunft erneut durchführen und eine Rückmeldung erhalten, um meine Entwicklung zu überprüfen.
(2) Ich werde den Test (mit Rückmeldung) meinen Kommiliton*innen weiterempfehlen.
(3) Ich werde das Material meinen Kommiliton*innen weiterempfehlen.
(2) Ich beabsichtige, mich mit dem Fördermaterial zu beschäftigen.

G Manual zur Auswertung der Think-Aloud-Interviews

G.1 Transkriptionsleitfaden

Die nachfolgend dargestellten Hinweise zur Transkription sind in Anlehnung an Seidel, Kobarg und Rimmele (2003, S. 85ff.) entstanden und für das Ziel der Arbeit angepasst.[207]

Transkriptionsbeginn und -ende: Die Transkription beginnt mit dem Start der inhaltlichen Bearbeitung des Reflexionstests (ausgelassen wurden jeweils die Begrüßung, die Einverständniserklärung zur Aufnahme der Interviews und organisatorische Absprachen) Die Transkription endet mit dem Abschluss der Testbearbeitung.

Sprechercodes: Die Sprechercodes der Teilnehmenden lauten entsprechend der Reihenfolge der Interviews „LD1“ bis „LD7“; die Interviewleitung erhält als Sprechercode „I“.

Zeitmarken: werden jeweils ans Ende einer Sprechsequenz gesetzt.

Überschneidungen von Sprechsequenzen: Überschneidungen von Sprechsequenzen werden durch // gekennzeichnet. Der Beginn einer Überschneidung wird mit // gekennzeichnet; die Sprechsequenz wird in einer neuen Zeile mit // eingeleitet. Beispiel aus Interview mit LD7 (Z. 19f.):

I:	Ne, mitschreiben musst du // gar nicht, ne. Alles klar. #00:07:18-1#
LD7:	Alles klar. #00:07:16-5#

Pausen: Pausen von bis zu drei Sekunden werden durch (.), (..) bzw. (...) gekennzeichnet. Für längere Pausen durch das Setzen eines neuen Sprechercodes.

Satzzeichen: Sätze und andere Phrasen werden mit einem Punkt beendet sofern sie keine Frage darstellen (Fragen werden dabei auch dann unterstellt, wenn die Stimme gehoben wird). Zur besseren Lesbarkeit werden weitere Satzzeichen verwendet („,“, „:“ oder (einfache) Anführungszeichen).

[207] Beispielsweise werden die Aufzeichnungen der Bildschirminteraktion nur da berücksichtigt, wo sie für die Nachvollziehbarkeit des Transkripts bedeutsam sind.

Abkürzungen: Abkürzungen werden ausgeschrieben (z. B. „okay“ statt O.K.), sofern sie nicht von den interviewten Personen abgekürzt werden.

Weitere Hinweise:

-	Stocken/Stottern des Sprechenden (Zur besseren Lesbarkeit der Transkripte wird lediglich der Begriff „Schüler*innen“ mit * und nicht als stockend notiert sofern er als Teil von Antwortoptionen vorgelesen wird.)
(unv.)	Wort ist nicht verstanden worden
(Wort?)	Wort kann nur vermutet werden
(Anmerkung)	Anmerkungen der transkribierenden Person, die für das Verständnis bzw. die Einordnung des Transkripts als wichtig erachtet werden (z. B. um die Zuordnung einer Äußerung zur Antwortoptionen im Reflexionstest herzustellen).
„Backchannels“	weisen auf Zustimmung oder Ablehnung hin, zusätzlich zur folgenden Schreibweise wird der Sinn in eckigen Klammern angefügt Zustimmung: Mhm [ja] Ablehnung: Hm-m [nein]
„Discount markers“ und „Hesitation indicators“	dienen der Einleitung oder Verstärkung von Aussagen (z. B. „also“, „ähm“) oder zeigen Zögern an; sie werden in den Text integriert. Beispiel aus Interview mit LD6 (Z. 33): „Ähm (.) drittes Axiom mhm fand ich auch okay. “

G.2 Segmentierung der Think-Aloud-Protokolle

Die Transkripte (Think-Aloud-Protokolle) wurden zunächst eventbasiert segmentiert. Der Vorteil der eventbasierten Segmentierung gegenüber der Orientierung an Zeitintervallen liegt in der Übereinstimmung von Segmenten und Sinneinheiten (Brückmann & Duit, 2014, S. 192). In der Segmentierung der vorliegenden Interviews waren folgende Events ausschlaggebend für den Abschluss bzw. Beginn eines neuen Segments:

- Sprecher*innenwechsel (ausgenommen sind lediglich kurze Einwürfe, wie „Mhm“, die den Redefluss der anderen Person nicht unterbrechen und somit keinen inhaltlichen Einfluss nehmen; eine Teilung der Aussage in zwei Segmente würde deren Nachvollziehbarkeit mindern und wäre insofern nicht zielführend; vgl. Kuckartz, 2018, S. 43),
- das Vorlesen eines neuen Elements aus dem Testinstrument (z. B. eine Antwortoption)
- die Bezugnahme auf eine neue Antwortoption im Testinstrument, auch wenn hier ggf. inhaltliche Verknüpfungen bestehen (z. B. bezieht sich die Aus-

sage „Die Auswertung des Experiments, dazu kann man jetzt irgendwie nichts sagen“ (LD7, Z. 135) auf die Antwortoption „Auswertung des Experiments“, die direkt anschließende Aussage „und Messunsicherheiten hat er zum Beispiel auch nicht thematisiert“ (LD7, Z. 136) auf die Antwortoption „Messunsicherheiten“) oder

- die Äußerung einer inhaltlich neuen und eigenständigen Überlegung (z. B. wenn eine eigene Erfahrung beschrieben wird, nicht aber eine Begründung für die Auswahl einer Antwortoption)

Falls inhaltliche Bezüge zu Antwortoptionen im Testinstrument unklar sind, wurde zusätzlich die Bildschirmaufnahme zu Rate gezogen. Die Bildschirmaufnahme bildet so zusammen mit dem Interviewtranskript die Kontexteinheit, die zum Verständnis der Segmente herangezogen werden kann (vgl. Kuckartz, 2018, S. 44). Das vollständige Manual zur Auswertung (zur Segmentierung sowie der nachfolgend beschriebenen Kodierung) der Think-Aloud-Protokolle befindet sich im Anhang in Kapitel G. Zur Herstellung von Bezügen zu den Transkripten werden die Segmente nummeriert.

G.3 Kodierung der Think-Aloud-Protokolle

Kodiert werden alle in einem Segment auftretenden Aussagen. Dazu wird zunächst jeder auftretende Aussagetyp identifiziert. Sollten Aussagetypen mehrfach vorkommen (auch Begründungen), wird geprüft, ob sich die entsprechenden Aussagen z.B. bei der Zustimmung/Ablehnung von Antwortoptionen auf den gleichen Inhalt beziehen (eine Kodierung) oder verschiedene Inhalte adressieren, z. B. verschiedene Handlungsoptionen vorschlagen oder Bewertungen mit unterschiedlichen Begründungen äußern (eine Kodierung je Inhalt). Beispiel:

> `Und es ist nicht gut, dass nur zwei Schüler ihre Merksätze vorgestellt haben, da du so nicht weißt, welche Merksätze die restlichen Schülerinnen formuliert haben. [Vorlesen]` Ja, [Zustimmung zur Antwortoption] *hier hätte er auf jeden Fall noch mehr Beispiele einbeziehen können [Alternative]* und dann vielleicht eine Mischung aus den Merksätzen dann (.) nochmal zu einem neuen Merksatz dann zusammenfassen können. [Alternative]“ (LD3, Z. 293)

In diesem Segment wird zunächst eine Antwortoption vorgelesen, dann wird ihr zugestimmt und anschließend werden zwei voneinander unabhängig umsetzbare alternative Handlungsoptionen vorgeschlagen, kodiert wird also: Vorlesen, Alternative (2x), Zustimmung zur Antwortoption.

Weitere Hinweise zur Kodierung:

- Begründungen liefern Informationen über Gründe für eine getätigte Aussage; Signalwörter für Begründungen sind u. a. „weil“, „denn“, „insofern“. Begründungen werden nur als solche kodiert, wenn sie sich inhaltlich auf den Unterricht (d. h. auf eine Bewertung, Alternative oder die Zustimmung bzw. Ablehnung einer Antwortoption) beziehen bzw. auf die eigene Professionalität beziehen (für Beispiele s. u.in der Tabelle).
- Die Korrektheit bzw. Sinnhaftigkeit der einzelnen Aussagen wird nicht bewertet.
- Einschränkungen werden nicht explizit kodiert, z. B. in Bezug auf die Gewichtung der einzelnen Kritikpunkte „[...] Es könnte be- benutzt werden, ja, aber es ist nicht das, was ich am größten kritisieren würde, aber dennoch etwas kritisierwürdiges.“ (LD1, Z. 59) oder fachliche Überlegungen „[...] die Formulierung sprachlich vereinfachen (.) auf (jeden?) kann gut sein, solange es (nachher?) noch fachlich korrekt ist und trotzdem die wichtigsten Fachbegriffe genannt werden.“ (LD3, Z. 285)

G.3.1 Übersicht über die Kategorien

Die Kategorien für die Kodierung der Think-Aloud-Protokolle sind in Tabelle G.1 beschrieben und mit Beispielen illustriert.

Kategorie	Erläuterung	Beispiele
Vorlesen	Text wird ggf. inkl. Füllwörter wie „äh“ oder Verhaspeln vorgelesen.	*Vorlesen mit Verhaspeln:* „Messunsicherheiten, Umgang mit der- mit dem Einwand von Marie. [...]“ (LD6, Z. 271)
Beschreibung	Es wird Unterricht (Lehrer- oder Schüler*innen handeln) oder der Unterrichtsablauf beschrieben.	*Beschreibung des Unterrichtsablaufs:* „Okay. (..) Grober Verlauf also: Quiz, dieser kleine Comic, (...) und zwei kleine- zwei Experimente? Ein Experiment. Ja.“ (LD2, Z. 8)

Bewertung von Unterricht (mit/ohne Begründung)	Es wird Lehrerhandeln bzw. die Unterrichtsgestaltung (z. B. die Fülle an Unterrichtsinhalten) als gut/nicht gut/... bewertet.	*Bewertung:* „[...] ich finde, Beispiele sind da ganz gut.“ (LD5, Z. 266) *Bewertung mit Begründung:* „Das Experiment fand ich irgendwie auch nicht so (.) super gewählt, weil es dann irgendwie was ganz anderes war, als mit der Titanic, das hat dann auch sehr voneinander abgewichen.“ (LD7, Z. 97)
Alternativen zum Vorgehen (mit/ohne Begründung)	Es werden alternative Vorgehensweisen zum Lehrerhandeln bzw. zu Unterrichtsinhalten oder dem Unterrichtsablauf formuliert.	*Alternative mit 2 Begründungen:* „[...] jedoch hätte ich es, glaube ich, besser gefunden, wenn er von Anfang auch sagt, gut, je schneller der Ball ist, das bedeutet, er hat eine höhere Geschwindigkeit, schwerer bedeutet, dass die Masse größer ist, und nicht so sehr noch in der Alltagssprache geblieben wäre. (.)“ (LD3, Z. 101) *Alternative (in Kombination mit der Überlegung, wieso die Lehrkraft wie beobachtet vorgeht)*: „Wieso stellt er das jetzt alleine vor und nicht mit dem Schüler?“ (LD5, Z. 10)

Zustimmung bzw. Ablehnung einer Antwortoption (mit/ohne Begründung)	Vorgegebenen Antwortoptionen wird zugestimmt bzw. diese werden abgelehnt. Das kann sowohl durch kurze Bemerkungen, wie „Ja“ oder „stimmt nicht“, als auch durch elaborierte Aussagen geschehen.	*Zustimmung mit Begründung:* „[...] es könnten Rundungsfehler sein bei den Schülern, es könnten aber auch eben Mess- Messunsicherheiten sein, die sollte er erwähnen, deswegen würde ich das ankreuzen.“ (LD5, Z. 226) *Ablehnung einer Antwortoption (trotz Signalwort „Ja“ für Zustimmung):* „[...] Ja, das kann nicht sein, ansonsten wären sie sich ja alle einig. (..)“ (LD1, Z. 81)
Eigene Erfahrung	Es werden eigene Erfahrungen aus der Perspektive als Lernende*r oder Lehrende*r geschildert bzw. es wird auf andere Erfahrungen Bezug genommen (z. B. geführte Gespräche).	„[...] Wenn ich mir vorstelle, ich hätte jetzt so ein Experiment in der Uni gemacht, dann hätte ich ja auch beide Wagen messen müssen.“ (LD7, Z. 166)

Weitere Überlegungen zum Unterricht bzw. den Inhalten	Es werden sonstige Überlegungen (ggf. offene Fragen) zum beobachteten Unterricht geäußert, die sich auf folgende Aspekte beziehen können: • Vorwissen der Schüler*innen und Schülergemäßheit, • Einordnungen in das Stundenraster bzw. den Unterrichtsablauf, • Vermutungen über zukünftigen Unterricht oder Intentionen der handelnden Person(en), • „allgemeingültige“, d. h. von der konkret beobachteten Situation abstrahierte fachliche (physikalische oder didaktische) Äußerungen oder Überlegungen • Erinnerungen an die beobachtete Situation • sowie sonstige (offene) Fragen, die sich den Teilnehmenden in Bezug auf den beobachteten Unterricht stellen	*Schülergemäßheit:* „[...] Ähm die Formulierung fand ich an einigen Stellen, [...] aber weiß ich nicht, wie sehr Schülerinnen und Schüler das verstehen würden.“ (LD2, Z. 29) *Einordnung:* „[...] Und dann ist das diese Phase, wo noch ein Experiment durchgeführt wird mit der Luftkissen- äh Fahrbahn. [...]“ (LD7, Z. 128) *Vermutung:* „[...] ein Beispiel kommt wahrscheinlich auch noch später.“ (LD5, Z. 133) *Fachliche Überlegung:* „[...] ja, da wären wir (dann?) wieder bei dem Trägheits- äh gesetz. (unv.) [...]“ (LD1, Z. 19) *Erinnerung:* „Ähm (.) wie hat er nochmal das erste Axiom formuliert? (7s) Ne, ich komme nicht mehr drauf“ (LD1, Z. 36) *Ergebnisoffene Überlegung:* „[...] ich weiß nicht, ob ein anderes Experiment das nicht besser machen könnte.“ (LD2, Z. 137)

Eigene Professionalisierung	Es werden Einschätzungen zur eigenen Professionalisierung vorgenommen (Einordnung des eigenen Wissens, Verhaltens o. ä. in Bezug auf den Lehrberuf).	„[...] Kenne ich mich leider aus dem Kopf nicht gut genug aus, [...]“ (LD2, Z. 358) „[...] Manchmal denke ich, (.) ist Physik das richtige Fach, bin mir selbst nicht richtig sicher.“ (LD6, Z. 144) „[...] da müsste ich nochmal nachgucken, werde ich aber höchstwahrscheinlich nicht machen, weil ich heute noch genügend zu tun habe.“ (LD1, Z. 86)

Sonstiges	Aussagen, die keiner der bisherigen Kategorien zugeordnet werden können; dazu gehören u. a. • Aussagen oder Fragen zur Interviewsituation oder der Methode des Lauten Denkens, • Aussagen, die sich auf die Testdurchführung (nicht auf den beobachteten Unterricht) beziehen, z. B. Verweise auf frühere oder spätere Stellen der Testdurchführung bzw. des Testinstruments oder den empfundenen Schwierigkeitsgrad, • Assoziationen, die bei der Testdurchführung auftreten sowie • Aussagen zum Testinstrument, die potentielle Überarbeitungsbedarfe kennzeichnen (z. B. zu Rechtschreibfehlern, zur Authentizität, zu inhaltlichen Wiederholungen in Antwortoptionen, zum Wunsch nach zusätzlichen Antwortoptionen, zur Länge oder zu (Verständnis-) Schwierigkeiten); Kennzeichnung mit „Sonstiges*“.	*Methode:* „Die Aufgaben muss ich jetzt aber nicht laut vorlesen, die also das zählt dann nicht zum lauten Denken, ne?“ (LD7, Z. 4) *Testinstrument:* „[...] Also dieser Test Tex- Limesurvey oder wie auch immer das nennt ist definitiv zu lang für meinen Geschmack.“(LD1, Z. 310) *Methode:* „Die Aufgaben muss ich jetzt aber nicht laut vorlesen, die also das zählt dann nicht zum lauten Denken, ne?“ (LD7, Z. 4) *Testdurchführung:* „ganz interessant und den Rest habe ich nicht gelesen, also machen wir einfach mal weiter.“ (LD1, Z. 19) oder „Ähm ja wie ich gerade schon gesagt habe, [...]“ (LD6, Z. 31) *Assoziation:* „[...] Wie in dem berühmt berüchtigten Beispiel von Sir Isaac Newton, was nebenbei gesagt historisch falsch ist.“ (LD1 Z. 73) *Sonstiges* (Wiederholung als möglicher Indikator für Verständnisschwierigkeit):* „[...] Mein Hirn hat das nicht verarbeitet, nochmal. [...]“ (LD1, Z. 365) *Länge des Testinstruments:* „[...] Also dieser Test Tex- Limesurvey oder wie auch immer das nennt ist definitiv zu lang für meinen Geschmack. [...]“ (LD1, Z. 310) *Ergänzung des Testinstruments:* „[...] da hätte ich gern noch eine Option, dass [...]“ (LD5, Z. 123)

	Hinweis: Im Interview LD1 wird der Prompt „Robert hat schon einige Rückmeldungen erhalten [...]“ jeweils nach Auswahl der Aspekte, zu denen Feedback gegeben werden sollte, vermutlich nie gelesen (zumindest nie vorgelesen und teilweise falsch wiedergegeben, z. B. „das habe ich nicht ausgewählt, ich habe keine Ahnung, warum ich hier bin.“; LD1, Z. 451), sodass die Bedeutung für den Testverlauf nicht verstanden wurde. Äußerungen zu diesem Prompt werden daher auch hier als „Sonstiges*“ kodiert.	*Authentizität des Testinstruments:* „Was ich mich jetzt frage ist, ob der Unterricht wirklich so stattgefunden hat oder ob er einfach nur so (.) exemplarisch ausgewählt wurde, [...]“ (LD6, Z. 268) „Äh (..) (ah?) ich weiß nicht“ (LD7, Z. 271)

Tabelle G.1: Übersicht über die Kategorien zur Kodierung der Think-Aloud-Interviews.

G.3.2 Kodierschema

Zur Strukturierung der Vorgehensweise bei der Kodierung der Think-Aloud-Protokolle wird obige Kategorienbeschreibung um ein Kodierschema ergänzt, dass sicherstellt, dass jeweils alle auftretenden Kategorien berücksichtigt werden (s. Tabelle G.2).

	Für jedes Segment einzeln zu prüfen: *Hinweis: Einzelne Aussagetypen können auch mehrfach vorkommen; werden unterschiedliche Aspekte bewertet/beschrieben/..., so werden mehrere Kodierungen vorgenommen*	**JA**	**NEIN**	**Kodierung**	**weiter bei...**
1	Wird Text vorgelesen?	X		Vorlesen	2
			X		2
2	Wird Zustimmung/Ablehnung zu einer AO geäußert?	X			3
			X		10
3	Wird die Zustimmung/Ablehnung begründet?	X			4
			X	Zustimmung bzw. Ablehnung AO	10
Hinweis: Enthält die Begründung mehrere der folgenden Elemente in den Zeilen 4-9 (d. h. mehrere Elemente aus Bewertung, Alternative, Erfahrungen und Professionalisierung), wird die Zustimmung bzw. Ablehnung der Antwortoption nur einmal kodiert. *Beispiel:* Enthält die Begründung für eine Zustimmung eine Bewertung und Alternative wird kodiert: Zustimmung AO, Bewertung, Alternative; NICHT: Zustimmung AO, Bewertung, Zustimmung AO, Alternative					
4	Enthält die Begründung eine Bewertung?	X			8
			X		5
5	Enthält die Begründung eine Alternative?	X			9
			X		6
6	Enthält die Begründung eigene Erfahrungen?	X		Zustimmung bzw. Ablehnung AO, eigene Erfahrung	7

			X		7
7	Enthält die Begründung Aussagen zur (eigenen) Professionalisierung?	X		Zustimmung bzw. Ablehnung AO, Professionalisierung	10
			X	Zustimmung bzw. Ablehnung AO + Begründung	10
8	Wird die enthaltene Bewertung begründet?	X		Zustimmung bzw. Ablehnung AO, Bewertung + Begründung	5
			X	Zustimmung bzw. Ablehnung AO, Bewertung	5
9	Wird die enthaltene Alternative begründet?	X		Zustimmung bzw. Ablehnung AO, Alternative + Begründung	6
			X	Zustimmung bzw. Ablehnung AO, Alternative	6
10	Wird eine Situation/ein Verhalten bewertet?	X			11
			X		12
11	Wird die Bewertung begründet?	X		Bewertung + Begründung	12
			X	Bewertung	12
12	Werden Alternativen zu einer Situation/einem Verhalten vorgeschlagen?	X			13
			X		14
13	Wird die Alternative begründet?	X		Alternative + Begründung	14
			X	Alternative	14

14	Wird eine Situation/ein Verhalten beschrieben?	X		Beschreibung	15
			X		15
15	Werden eigene Erfahrungen berichtet?	X		eigene Erfahrung	16
			X		16
16	Werden Aussagen über die eigene Professionalisierung getätigt?	X		Professionalisierung	17
			X		17
17	Falls obige Kategorien nicht zutreffen: Ist ein inhaltlicher Bezug zum Unterricht bzw. der Professionalisierung der Lehrkraft gegeben?	X		Überlegung zum Unterricht	
			X	Sonstiges	

Tabelle G.2: Schema zur Standardisierung des Vorgehens bei der Kodierung (mit „AO“ wird „Antwortoption(en)“ abgekürzt)

H Ergänzende Auswertung der Think-Aloud-Protokolle zum Teil 2 des INA: Kodierungen von Bewertungen und Alternativen zu den entsprechenden Aufgaben

Tabelle H.1 stellt eine Übersicht über die Anteile der Kodierungen zu den Kategorien Bewertungen und Alternativen an allen Kodierungen (unter Auslassung der Kategorie Vorlesen) dar. Es wird deutlich, dass die Studierenden während der Bearbeitung von Aufgaben zu Bewertungen und zu Alternativen mehr Alternativen als Bewertungen äußern.

Kodierung	LD1		LD2		LD3	
	Aufgaben zu		Aufgaben zu		Aufgaben zu	
	Bew.	Alt.	Bew.	Alt.	Bew.	Alt.
Bewertungen mit/ohne Begründung	2 %	8 %	6 %	15 %	2 %	3 %
Alternativen mit/ohne Begründung	1 %	5 %	1 %	2 %	1 %	15 %

Kodierung	LD4		LD5		LD6	
	Aufgaben zu		Aufgaben zu		Aufgaben zu	
	Bew.	Alt.	Bew.	Alt.	Bew.	Alt.
Bewertungen mit/ohne Begründung	4 %	10 %	12 %	2 %	3 %	16 %
Alternativen mit/ohne Begründung	0 %	8 %	3 %	20 %	3 %	15 %

Kodierung	LD7			
	Aufgaben zu			
	Bew.	Alt.		
Bewertungen mit/ohne Begründung	5 %	8 %		
Alternativen mit/ohne Begründung	2 %	9 %		

Tabelle H.1: Übersicht über Anteile der Kategorien Bewertungen (mit und ohne Begründung) sowie Alternativen (mit und ohne Begründung) aufgeschlüsselt nach Aufgaben zu Bewertungen (Bew.) und Alternativen (Alt.)

I Vollständige statistische Untersuchungen zum Teil 3 des INA

I.1 Indikator 3.2: Übersicht über Kennwerte für die Reliabilität (α_{Cronbach}) und Trennschärfen

Stichprobe	α_{Cronbach}	durchschnittliche TS	minimale TS	maximale TS
Gesamtscore				
Gesamt ($N = 192$)	$0,95$	$0,58$	$0,27$ (R3Bew1)	$0,75$ (RGAlt1)
BA ($N = 90$)	$0,96$	$0,61$	$0,33$ (R4Alt1)	$0,79$ (R7Bew2)
MA ($N = 86$)	$0,92$	$0,48$	$0,22$ (R3Bew1)	$0,66$ (R5Bew3)
Score für Bewertungen				
Gesamt ($N = 192$)	$0,90$	$0,55$	$0,23$ (R3Bew1)	$0,71$ (R7Bew2)
BA ($N = 90$)	$0,90$	$0,58$	$0,30$ (R3Bew1)	$0,78$ (R7Bew2)
MA ($N = 86$)	$0,83$	$0,45$	$0,18$ (R3Bew1)	$0,64$ (R5Bew3)
Score für Alternativen				
Gesamt ($N = 192$)	$0,90$	$0,57$	$0,32$ (R3Alt1)	$0,75$ (RGAlt1)
BA ($N = 90$)	$0,92$	$0,62$	$0,35$ (R4Alt1)	$0,80$ (RGAlt1)
MA ($N = 86$)	$0,83$	$0,46$	$0,24$ (R3Alt1)	$0,60$ (R6Alt3)

Tabelle I.1: Kennwerte zur Beurteilung der Reliabilität (α_{Cronbach}) und Trennschärfe (TS) der verwendeten Skalen für die Gesamtstichprobe sowie die darin enthaltenen Teilstichproben der Bachelor- (BA) und Masterstudierenden (MA)

I.2 Indikator 3.4

I.2.1 Ergebnisse des Kolmogorow-Smirnow-Tests

Stichprobe	Reflexionstest	Performanztest (Gesamtscore)
Gesamt ($N = 39$)	$D(39) = 0,07; p = 0,98$	$D(39) = 0,08; p = 0,96$
BA-Phy ($N = 10$)	$D(10) = 0,22; p = 0,66$	$D(10) = 0,11; p \approx 1,0$
MA-Phy ($N = 12$)	$D(12) = 0,21; p = 0,60$	$D(12) = 0,16; p = 0,84$
SU ($N = 17$)	$D(17) = 0,16; p = 0,71$	$D(17) = 0,19; p = 0,56$
Stichprobe	Performanztest (Testwert fachliches Wissen)	Performanztest (Testwert fachdidaktisches Wissen)
Gesamt ($N = 39$)	$D(39) = 0,10; p = 0,80$	$D(39) = 0,12; p = 0,64$
BA-Phy ($N = 10$)	$D(10) = 0,16; p = 0,96$	$D(10) = 0,19; p = 0,86$
MA-Phy ($N = 12$)	$D(12) = 0,14; p = 0,97$	$D(12) = 0,23; p = 0,57$
SU ($N = 17$)	$D(17) = 0,27; p = 0,16$	$D(17) = 0,14; p = 0,91$
Stichprobe	Performanztest (Testwert pädagogisches Wissen)	
Gesamt ($N = 39$)	$D(39) = 0,13; p = 0,48$	
BA-Phy ($N = 10$)	$D(10) = 0,21; p = 0,77$	
MA-Phy ($N = 12$)	$D(12) = 0,21; p = 0,69$	
SU ($N = 17$)	$D(17) = 0,15; p = 0,86$	

Tabelle I.2: Ergebnisse des Kolmogorow-Smirnow-Tests für den Reflexionstest und den Reflexionsperformanztest (Gesamtscore sowie Teilscores zu Reflexionen mit Bezug zum fachlichen, fachdidaktischen und pädagogischen Wissen) für die Gesamtstichprobe und die betrachteten Teilstichproben der Physik-Lehramtsstudierenden im Bachelor- (BA-Phy) bzw. Masterstudium (MA-Phy) sowie der Studierenden des Grundschullehramts im Fach Sachunterricht mit Physik als Wahlfach (SU)

I.2.2 Levene-Tests und ANOVAs

Skala	Levene-Test	ANOVA
Reflexionstest	$F(2;36) = 1,89;\ p = 0,17$	$F(2;36) = 2,38;\ p = 0,11$
Performanztest (Gesamt)	$F(2;36) = 0,15;\ p = 0,87$	$F(2;36) = 2,78;\ p = 0,08$
Performanztest (Testwert fachliches Wissen)	$F(2;36) = 2,27;\ p = 0,12$	$F(2;36) = 12,63;\ p < 0,01$
Performanztest (Testwert fachdidaktische Wissen)	$F(2;36) = 0,39;\ p = 0,68$	$F(2;36) = 2,86;\ p = 0,07$
Performanztest (Testwert pädagogisches Wissen)	$F(2;36) = 0,40;\ p = 0,68$	$F(2;36) = 1,08;\ p = 0,35$

Tabelle I.3: Ergebnisse der Tests auf Varianzhomogenität (Levene-Tests) und einfaktoriellen Varianzanalysen (ANOVAs) zur Prüfung von Gruppenunterschieden für Testwerte zum Reflexionstest und Reflexionsperformanztest

Gruppen	p	Effektstärke
MA-SU	$< 0,001$	$2,09$
MA-BA	$1,0$	-
SU-BA	$0,004$	$-1,23$

Tabelle I.4: Post-hoc-Analyse zur Prüfung von Gruppenunterschieden für Testwerte zum fachlichen Wissen im Reflexionsperformanztest für die betrachteten Teilstichproben der Physik-Lehramtsstudierenden im Bachelor- (BA-Phy) bzw. Masterstudium (MA-Phy) sowie der Studierenden des Grundschullehramts im Fach Sachunterricht mit Physik als Wahlfach (SU)

I.3 Indikator 3.5

I.3.1 Indikator 3.5: Quasi-Längsschnitt

Für alle Tests wird $p < 0,05$ als Signifikanzniveau angelegt. In Fällen, in denen dieser Wert bei der Prüfung der Normalverteilung der betrachteten Daten mit dem Kolmogorow-Smirnow-Test nur knapp überschritten wird (bis $p \leq 0,10$), werden zusätzlich nicht parametrische Alternativen berechnet und hier berichtet.

Gruppenunterschiede: Ausbildungsphase (Studierende im Bachelor- und Masterstudium und Referendariat)

- $N = 187$
- Gruppe Bachelorstudierende: Prüfung der Normalverteilung der Testwerte mit Kolmogorow-Smirnow-Test: $D(90) = 0,13; p = 0,10 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung
- Gruppe Masterstudierende: Prüfung der Normalverteilung der Testwerte mit Kolmogorow-Smirnow-Test: $D(86) = 0,09; p = 0,56 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung
- Gruppe Referendar*innen: Prüfung der Normalverteilung der Testwerte mit Kolmogorow-Smirnow-Test: $D(11) = 0,16; p = 0,90 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung
- Prüfung der Varianzhomogenität mit dem Levene-Test: $F(2; 184) = 5,58; p = 0,004 < 0,05 \Rightarrow$ *Keine* Annahme der Varianzhomogenität
- Welch-ANOVA: $F(2; 35,48) = 10,52; p = 0,0003 < 0,05$
- Post-hoc-Analyse zur Prüfung von Gruppenunterschieden für die Testwerte im Reflexionstest:

Gruppen	p	Gruppen	p	Gruppen	p
BA, MA	$< 0,001$	BA, Ref	$0,14$	MA, Ref	$1,0$

Zusammenhang zum Fachsemester

- $N = 178$
- Prüfung der Normalverteilung der Testwerte mit Kolmogorow-Smirnow-Test: $D(178) = 0,10; p = 0,052 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung
- Prüfung der Normalverteilung der Fachsemester mit Kolmogorow-Smirnow-Test: $D(178) = 0,13; p = 0,01 < 0,05 \Rightarrow$ *Keine* Annahme der Normalverteilung
- $\rho_{\text{Spearman}} = 0,18; p = 0,049 < 0,05 \Rightarrow$ Korrelation vorhanden (kleine Effektstärke)

Zusätzliche Bestimmung des Zusammenhangs unter Ausschluss der Studierenden des Grundschullehramts:

- $N = 159$
- Prüfung der Normalverteilung der Testwerte mit Kolmogorow-Smirnow-Test: $D(159) = 0,10; p = 0,10 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung
- Prüfung der Normalverteilung der Fachsemester mit Kolmogorow-Smirnow-Test: $D(178) = 0,13; p = 0,0498 < 0,05 \Rightarrow$ *Keine* Annahme der Normalverteilung
- $\rho_{\text{Spearman}} = 0,21; p = 0,007 < 0,05 \Rightarrow$ Korrelation vorhanden (kleine Effektstärke)

Gruppenunterschiede: absolvierte Pflichtpraktika

Prüfung von Unterschieden in den Testwerten zwischen beiden Gruppen (Pflichtpraktika absolviert bzw. nicht absolviert):

- Gruppe *Absolvierte Pflichtpraktika ($N = 131$)*: Prüfung der Normalverteilung der Testwerte mit Kolmogorow-Smirnow-Test: $D(131) = 0,12; p = 0,06 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung
- Gruppe *Keine absolvierten Pflichtpraktika ($N = 47$)*: Prüfung der Normalverteilung der Testwerte mit Kolmogorow-Smirnow-Test: $D(47) = 0,12; p = 0,53 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung
- Prüfung der Varianzhomogenität mit dem Levene-Test: $F(1; 176) = 0,13; p = 0,72 \geq 0,05 \Rightarrow$ Annahme der Varianzhomogenität
- t-Test: $t(176) = -0,270;\ p = 0,008 < 0,01 \Rightarrow$ Unterschied zwischen den Gruppen
- $d_{\text{Cohen}} = 0,46$

- Zusätzliche Bestimmung des Wilcoxon-Tests (nicht-parametrisch): $W = 2148,5$; $p = 0,002 < 0,05 \Rightarrow$ Unterschied zwischen den Gruppen

Zusätzliche Untersuchung des Zusammenhangs zwischen dem Umfang der absolvierten Pflichtpraktika (in Wochen) und den Testwerten für die Gruppe der Personen mit absolvierten Pflichtpraktika:

- $N = 131$
- Prüfung der Normalverteilung der Wochen im Pflichtpraktikum mit Kolmogorow-Smirnow-Test: $D(178) = 0,13; p = 0,0498 < 0,05 \Rightarrow$ *Keine* Annahme der Normalverteilung
- $\rho_{\text{Spearman}} = 0,17$; $p = 0,056 \geq 0,05 \Rightarrow$ *Keine* Korrelation vorhanden

Gruppenunterschiede: freiwillig absolvierte Praktika

Prüfung von Unterschieden in den Testwerten zwischen beiden Gruppen (freiwillige Praktika absolviert bzw. nicht absolviert)

- Gruppe *Absolvierte freiwillige Praktika ($N = 20$)*: Prüfung der Normalverteilung der Testwerte mit Kolmogorow-Smirnow-Test: $D(20) = 0,17$; $p = 0,53 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung
- Gruppe *Keine absolvierten freiwilligen Praktika ($N = 158$)*: Prüfung der Normalverteilung der Testwerte mit Kolmogorow-Smirnow-Test: $D(158) = 0,10$; $p = 0,09 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung
- Prüfung der Varianzhomogenität mit dem Levene-Test: $F(1; 176) = 0,24$; $p = 0,63 \geq 0,05 \Rightarrow$ Annahme der Varianzhomogenität
- t-Test: $t(176) = 1,14$; $p = 0,26 \geq 0,05 \Rightarrow$ *Kein* Unterschied zwischen den Gruppen
- Zusätzliche Bestimmung des Wilcoxon-Tests (nicht-parametrisch): $W = 1794$; $p = 0,33 \geq 0,05 \Rightarrow$ *Kein* Unterschied zwischen den Gruppen

Gruppenunterschiede: vorliegende Unterrichtserfahrung als Vertretungslehrkraft

Prüfung von Unterschieden in den Testwerten zwischen beiden Gruppen (Erfahrung als Vertretungslehrkraft liegt bzw. liegt nicht vor)

- Gruppe *Vorliegende Unterrichtserfahrung ($N = 51$)*: Prüfung der Normalverteilung der Testwerte mit Kolmogorow-Smirnow-Test: $D(51) = 0,13; p = 0,34 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung

- Gruppe *Keine vorliegende Unterrichtserfahrung (N = 127)*: Prüfung der Normalverteilung der Testwerte mit Kolmogorow-Smirnow-Test: $D(127) = 0,12; p = 0,066 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung
- Prüfung der Varianzhomogenität mit dem Levene-Test: $F(1; 176) = 4,56; p = 0,03 < 0,05 \Rightarrow$ *Keine* Annahme der Varianzhomogenität
- Welch-t-Test: $t(75,49) = -0,21$; $p = 0,84 \geq 0,05 \Rightarrow$ *Kein* Unterschied zwischen den Gruppen
- Zusätzliche Bestimmung des Wilcoxon-Tests (nicht-parametrisch): $W = 2980$; $p = 0,41 \geq 0,05 \Rightarrow$ *Kein* Unterschied zwischen den Gruppen

Gruppenunterschiede: vorliegende außerschulische Lehrerfahrung

Prüfung von Unterschieden in den Testwerten zwischen beiden Gruppen (außerschulische Lehrerfahrung liegt bzw. liegt nicht vor)

- Gruppe *Vorliegende außerschulische Erfahrung (N = 153)*: Prüfung der Normalverteilung der Testwerte mit Kolmogorow-Smirnow-Test: $D(153) = 0,10; p = 0,09 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung
- Gruppe *Keine vorliegende außerschulische Erfahrung (N = 25)*: Prüfung der Normalverteilung der Testwerte mit Kolmogorow-Smirnow-Test: $D(25) = 0,10; p = 0,95 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung
- Prüfung der Varianzhomogenität mit dem Levene-Test: $F(1; 176) = 3,93; p = 0,049 < 0,05 \Rightarrow$ *Keine* Annahme der Varianzhomogenität
- Welch-t-Test: $t(29,13) = -2,78$; $p = 0,009 < 0,01 \Rightarrow$ Unterschied zwischen den Gruppen
- $d_{\text{Cohen}} = 0,65$
- Zusätzliche Bestimmung des Wilcoxon-Tests (nicht-parametrisch): $W = 1217$; $p = 0,003 < 0,05 \Rightarrow$ Unterschied zwischen den Gruppen

Gruppenunterschiede: vorliegende universitäre Lehrerfahrung

Prüfung von Unterschieden in den Testwerten zwischen beiden Gruppen (universitäre Lehrerfahrung liegt bzw. liegt nicht vor)

- Gruppe *Vorliegende universitäre Erfahrung (N = 43)*: Prüfung der Normalverteilung der Testwerte mit Kolmogorow-Smirnow-Test: $D(43) = 0,13; p = 0,39 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung

- Gruppe *Keine vorliegende universitäre Erfahrung ($N = 135$)*: Prüfung der Normalverteilung der Testwerte mit Kolmogorow-Smirnow-Test: $D(135) = 0,10; p = 0,13 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung
- Prüfung der Varianzhomogenität mit dem Levene-Test: $F(1; 176) = 3,64; p = 0,058 \geq 0,05 \Rightarrow$ Annahme der Varianzhomogenität
- t-Test: $t(176) = -2,98$; $p = 0,003 < 0,05 \Rightarrow$ Unterschied zwischen den Gruppen
- $d_{\text{Cohen}} = 0,57$

Gruppenunterschiede: Lehramtstyp (Primar- oder Sekundarstufe)

Prüfung von Unterschieden in den Testwerten zwischen beiden Gruppen (Lehramt für die Primar- oder Sekundarstufe)

- Gruppe *Primarstufe ($N = 19$)*: Prüfung der Normalverteilung der Testwerte mit Kolmogorow-Smirnow-Test: $D(19) = 0,17; p = 0,60 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung
- Gruppe *Sekundarstufe ($N = 159$)*: Prüfung der Normalverteilung der Testwerte mit Kolmogorow-Smirnow-Test: $D(159) = 0,10; p = 0,10 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung
- Prüfung der Varianzhomogenität mit dem Levene-Test: $F(1; 176) = 0,0001; p = 0,99 \geq 0,05 \Rightarrow$ Annahme der Varianzhomogenität
- t-Test: $t(176) = 0,04$; $p = 0,97 \geq 0,05 \Rightarrow$ *Kein* Unterschied zwischen den Gruppen

Gruppenunterschiede: Zweitfach

Prüfung von Unterschieden in den Testwerten zwischen den drei Gruppen mit einem Zweitfach aus den Naturwissenschaften, dem Zweitfach Mathematik oder sonstigem Zweitfach

- Gruppe *Naturwissenschaftliches Zweitfach ($N = 25$)*: Prüfung der Normalverteilung der Testwerte mit Kolmogorow-Smirnow-Test: $D(25) = 0,13; p = 0,74 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung
- Gruppe *Zweitfach Mathematik ($N = 111$)*: Prüfung der Normalverteilung der Testwerte mit Kolmogorow-Smirnow-Test: $D(111) = 0,10; p = 0,27 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung
- Gruppe *Sonstiges Zweitfach ($N = 40$)*: Prüfung der Normalverteilung der Testwerte mit Kolmogorow-Smirnow-Test: $D(40) = 0,16; p = 0,23 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung

- Prüfung der Varianzhomogenität mit dem Levene-Test: $F(2; 173) = 1,39; p = 0,25 \geq 0,05 \Rightarrow$ Annahme der Varianzhomogenität
- ANOVA: $F(2, 173) = 1,10;\ p = 0,33 \geq 0,05 \Rightarrow$ *Kein* Unterschied zwischen den Gruppen

Zusammenhang zum Alter

- $N = 192$
- Prüfung der Normalverteilung der Testwerte mit Kolmogorow-Smirnow-Test: $D(192) = 0,10; p = 0,053 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung
- Prüfung der Normalverteilung des Alters mit Kolmogorow-Smirnow-Test: $D(192) = 0,20; p < 0,001 < 0,05 \Rightarrow$ *Keine* Annahme der Normalverteilung
- $\rho_{\text{Spearman}} = 0,04; p = 0,56 \geq 0,05 \Rightarrow$ *Keine* Korrelation vorhanden

Gruppenunterschiede: Geschlechtsidentität

- Gruppe *Weiblich ($N = 72$)*: Prüfung der Normalverteilung der Testwerte mit Kolmogorow-Smirnow-Test: $D(72) = 0,12; p = 0,26 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung
- Gruppe *Männlich ($N = 120$)*: Prüfung der Normalverteilung der Testwerte mit Kolmogorow-Smirnow-Test: $D(120) = 0,12; p = 0,066 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung
- Prüfung der Varianzhomogenität mit dem Levene-Test: $F(1; 191) = 0,21; p = 0,22 \geq 0,05 \Rightarrow$ Annahme der Varianzhomogenität
- t-Test: $t(191) = 2,02;\ p = 0,045 < 0,05 \Rightarrow$ Unterschied zwischen den Gruppen
- $d_{\text{Cohen}} = 0,30$
- Zusätzliche Bestimmung des Wilcoxon-Tests (nicht-parametrisch): $W = 5049$; $p = 0,05 \geq 0,05 \Rightarrow$ *Knapp kein* Unterschied zwischen den Gruppen

Zusammenhang zur Abiturnote

- $N = 175$
- Prüfung der Normalverteilung der Testwerte mit Kolmogorow-Smirnow-Test: $D(175) = 0,10; p = 0,057 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung
- Prüfung der Normalverteilung der Abiturnote Kolmogorow-Smirnow-Test: $D(175) = 0,09; p = 0,13 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung

- $r_{\text{Pearson}} = -0,22; p = 0,004 < 0,05 \Rightarrow$ Korrelation vorhanden (kleine Effektstärke)
- $\rho_{\text{Spearman}} = -0,17; p = 0,02 < 0,05 \Rightarrow$ Korrelation vorhanden (kleine Effektstärke)

Gruppenunterschiede: Abitur nach G8/G9

- Gruppe *G8 ($N = 107$)*: Prüfung der Normalverteilung der Testwerte mit Kolmogorow-Smirnow-Test: $D(120) = 0,13; p = 0,28 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung
- Gruppe *G9 ($N = 58$)*: Prüfung der Normalverteilung der Testwerte mit Kolmogorow-Smirnow-Test: $D(58) = 0,12; p = 0,09 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung
- Prüfung der Varianzhomogenität mit dem Levene-Test: $F(1; 163) = 0,31; p = 0,58 \geq 0,05 \Rightarrow$ Annahme der Varianzhomogenität
- t-Test: $t(163) = 1,36$; $p = 0,17 \geq 0,05 \Rightarrow$ *Kein* Unterschied zwischen den Gruppen
- Zusätzliche Bestimmung des Wilcoxon-Tests (nicht-parametrisch): $W = 3508$; $p = 0,17 \geq 0,05 \Rightarrow$ *Kein* Unterschied zwischen den Gruppen

Zusammenhang zur letzten Physik-, Mathematik- und Deutschnote

- $N = 178$
- Prüfung der Normalverteilung der Testwerte mit Kolmogorow-Smirnow-Test: $D(178) = 0,10; p = 0,052 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung

Physiknote:

- Prüfung der Normalverteilung der letzten Physiknote Kolmogorow-Smirnow-Test: $D(178) = 0,14; p = 0,001 < 0,05 \Rightarrow$ *Keine* Annahme der Normalverteilung
- $\rho_{\text{Spearman}} = 0,02; p = 0,76 \geq 0,05 \Rightarrow$ *Keine* Korrelation vorhanden

Mathematiknote:

- Prüfung der Normalverteilung der letzten Mathematiknote Kolmogorow-Smirnow-Test: $D(178) = 0,14; p = 0,002 < 0,05 \Rightarrow$ *Keine* Annahme der Normalverteilung
- $\rho_{\text{Spearman}} = 0,09; p = 0,23 \geq 0,05 \Rightarrow$ *Keine* Korrelation vorhanden (kleine Effektstärke)

Deutschnote:

- Prüfung der Normalverteilung der letzten Deutschnote Kolmogorow-Smirnow-Test: $D(178) = 0,09; p = 0,10 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung
- $r_{\text{Pearson}} = 0,15; p = 0,045 < 0,05 \Rightarrow$ Korrelation vorhanden (kleine Effektstärke)
- $\rho_{\text{Spearman}} = 0,14; p = 0,07 \geq 0,05 \Rightarrow$ *Keine* Korrelation vorhanden

Gruppenunterschiede: Physik als Prüfungsfach im Abitur

- Gruppe *Physik als Prüfungsfach ($N = 96$)*: Prüfung der Normalverteilung der Testwerte mit Kolmogorow-Smirnow-Test: $D(96) = 0,07; p = 0,68 \geq 0,05$ $\Rightarrow$ Annahme der Normalverteilung
- Gruppe *Physik nicht als Prüfungsfach ($N = 82$)*: Prüfung der Normalverteilung der Testwerte mit Kolmogorow-Smirnow-Test: $D(82) = 0,12; p = 0,19 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung
- Prüfung der Varianzhomogenität mit dem Levene-Test: $F(1; 176) = 4,24; p = 0,04 < 0,05 \Rightarrow$ *Keine* Annahme der Varianzhomogenität
- Welch-t-Test: $t(149,1) = -1,51;\ p = 0,13 \geq 0,05 \Rightarrow$ *Kein* Unterschied zwischen den Gruppen

I.3.2 Indikator 3.5: Längsschnitt

Gesamtstichprobe ($N = 41$)

- Prüfung der Normalverteilung der Testwertdifferenzen zwischen beiden MZP mit Kolmogorow-Smirnow-Test: $D(41) = 0,18; p = 0,14 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung
- Gepaarter t-Test: $t(40) = -3,70; p < 0,01 \Rightarrow$ Unterschied zwischen den MZP
- $d_{\text{Cohen},z} = 0,58$

Bachelorstudierende ($N = 16$):

- Prüfung der Normalverteilung der Testwertdifferenzen zwischen beiden MZP mit Kolmogorow-Smirnow-Test: $D(16) = 0,16; p = 0,77 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung
- Gepaarter t-Test: $t(15) = -2,88; p < 0,05 \Rightarrow$ Unterschied zwischen den MZP
- $d_{\text{Cohen},z} = 0,72$

Masterstudierende ($N = 25$):

- Prüfung der Normalverteilung der Testwertdifferenzen zwischen beiden MZP mit Kolmogorow-Smirnow-Test: $D(25) = 0,20; p = 0,24 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung
- Gepaarter t-Test: $t(24) = -2,38; p < 0,05 \Rightarrow$ Unterschied zwischen den MZP
- $d_{\text{Cohen},z} = 0,48$

I.4 Indikator 3.6

Test	MZP	M	SD	KS-Test (p-Wert)	α_{Cronbach}
Refl.	Prä + Post ($N = 41$)	73 %	11 %	$D(41) = 0,07; p = 0,98$	0,91
	Prä ($N = 18$)	70 %	11 %	$D(18) = 0,11; p = 0,98$	0,90
	Post ($N = 13$)	75 %	11 %	$D(13) = 0,14; p = 0,78$	0,92
FDW	Prä + Post ($N = 41$)	53 %	11 %	$D(41) = 0,12; p = 0,63$	0,66
	Prä ($N = 18$)	51 %	12 %	$D(18) = 0,12; p = 0,95$	0,71
	Post ($N = 13$)	55 %	10 %	$D(13) = 0,14; p = 0,78$	0,59

Tabelle I.5: Übersicht über die Kennwerte zum Reflexionstest (*Refl.*) und Test zum fachdidaktischen Wissen (*FDW*)

- Prä + Post: $r_{\text{Pearson}} = -0,16; p = 0,32$
- Prä: $r_{\text{Pearson}} = -0,31; p = 0,21$
- Post: $r_{\text{Pearson}} = -0,11; p = 0,63$

Facette	KS-Test	Korrelation zu Testwerten des Reflexionstests
Instr.	$D(41) = 0,10; p = 0,78$	$r_{\text{Pearson}} = -0,08; p = 0,64$
SV	$D(41) = 0,10; p = 0,82$	$r_{\text{Pearson}} = -0,12; p = 0,46$
Exp. & Wiss.	$D(41) = 0,24; p = 0,02$	$\rho_{\text{Spearman}} = 0,03; p = 0,87$
FD-Konzepte	$D(41) = 0,13; p = 0,48$	$r_{\text{Pearson}} = -0,17; p = 0,28$

Tabelle I.6: Ergebnisse der Kolmogorow-Smirnow-Tests (*KS-Test*) zu den Facetten im Test zum fachdidaktischen Wissen Instruktionsstrategien (*Instr.*), Schülervorstellungen (*SV*), Experimente und Wissenschaftsverständnis (*Exp. & Wiss.*) und Fachdidaktische Konzepte (*FD-Konzepte*) für alle vorliegenden Datensätze ($N = 41$)

J Interviewleitfäden und Beispielrückmeldungen zur Evaluation des Assessment-Feedbacks

Nachfolgend werden die Interviewleitfäden zur Evaluation des Assessment-Feedbacks sowie (anonymisierte) jeweils aktuelle Versionen der Beispielrückmeldungen eingebunden.

Im Rahmen der Interviews wird das Assessment-Feedback als „Rückmeldung“ und das Selbstlernmaterial als „(online-)Material“ bezeichnet.

J.1 Interviewleitfaden zur Pilotierung

J.1.1 Einleitung/Einverständnis

Vielen Dank, dass Sie sich Zeit für dieses Interview nehmen.

Ich möchte zunächst um Ihr Einverständnis für die Verwendung Ihrer Daten bitten. Dazu lese ich Ihnen noch einmal die Einverständniserklärung vor und bitte Sie danach um Ihr mündliches Einverständnis.

Ich bin damit einverstanden, dass dieses Interview aufgezeichnet wird. Dieses Interview wird für eine Verschriftlichung, Verwendung oder Weitergabe anonymisiert. Diese Einverständniserklärung zur Nutzung der erhobenen Daten kann jederzeit widerrufen werden.

Sind Sie damit einverstanden?

Sie haben vor einigen Wochen den Reflexionstest bearbeitet und im Anschluss eine Rückmeldung zu Ihrer Testbearbeitung erhalten. Ich würde gleich gerne etwas darüber erfahren, wie Sie die Rückmeldung zum Reflexionstest wahrnehmen bzw. wahrgenommen haben. Dabei geht es explizit **nicht** darum, Sie bzw. Ihre Aussagen zu bewerten, sondern darum, das Format und die Gestaltung der Rückmeldung zu verbessern.

Welche Fragen haben Sie noch, bevor wir mit dem Interview beginnen?

J.1.2 Fragen 1: Eigene Rückmeldung

Zum Einstieg geht es um Ihre eigene Rückmeldung

1. Können Sie sich an Ihre Rückmeldung zur Bearbeitung des Reflexionstests erinnern?
2. Welche Schlüsse haben Sie daraus gezogen?

J.1.3 Fragen 2: Beispielrückmeldung

Ich zeige Ihnen jetzt ein Beispiel einer Rückmeldung zum Reflexionstest. Nehmen Sie sich einen Moment Zeit, um die Rückmeldung anzuschauen/zu lesen. Wenn Sie möchten, können Sie ihre Gedanken, offene Fragen usw. laut aussprechen.

Im Anschluss habe ich einige Fragen an Sie.

1. Haben Sie zuvor mit Boxplots/diese Art von Diagramm gearbeitet? (In welchem Rahmen?)
2. Welche Schlüsse ziehen Sie aus dieser Rückmeldung zum Reflexionstest?
3. Welche Informationen sind besonders hilfreich oder interessant für Sie?
4. Sind Informationen überflüssig? Welche?
5. Welche Fragen bleiben für Sie offen?
6. Welche weiteren Informationen hätten Sie gerne?
    a) Informationen über Ziel des Tests/Projekts
    b) (Explizite) Information über „Zielzustand“ Reflexionsfähigkeit
    c) Feinere Aufschlüsselung der Bewertung
    d) Andere Diagramme
    e) Beispielaufgabe
    f) Musterlösungen mit Begründung
    g) Literaturhinweise / Hinweise auf physikdidaktische Grundlagentexte
    h) Material/Hinweise, um Reflexionsfähigkeit zu schulen

J.1.4 Fragen 3: Bereitstellung der Rückmeldung

Sie haben Ihre Rückmeldung über einen Download erhalten.

1. Gab es Schwierigkeiten beim Erhalt der Rückmeldung? Welche?
2. In welchem Form wünschen Sie sich die Rückmeldung?

J.1.5 Abschluss

Haben Sie sonstige Fragen/Anmerkungen?

J.2 Beispielrückmeldung zur Pilotierung

UNIVERSITÄT PADERBORN
Die Universität der Informationsgesellschaft

Rückmeldung zum Reflexionstest

Ihr Testcode: [redacted]

Sehr geehrte:r Teilnehmer:in,

Sie haben kürzlich am Reflexionstest teilgenommen. Im Rahmen dieses Tests haben Sie Rückmeldungen zu Ausschnitten einer gehaltenen Unterrichtsstunde gegeben. Hiermit erhalten Sie eine individuelle Rückmeldung zu Ihrer Bearbeitung des Test. Ihre Ergebnisse finden Sie neben einigen erläuternden Informationen in den untenstehenden Diagrammen.

Vielen Dank für Ihre Teilnahme!

Reflexionskompetenz - Was ist das?

Ziel einer Reflexion ist die Verbesserung der Unterrichtsqualität und bzw. oder die Weiterentwicklung persönlicher Kompetenzen. Im Rahmen einer Reflexion können unterschiedliche Aussagen getätigt werden:

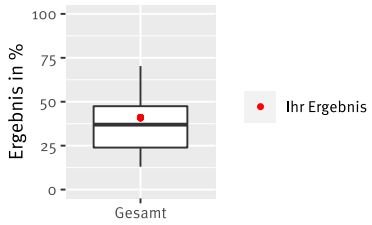
- Beobachtete Aspekte können **beschrieben** werden,
- diese können **bewertet** werden,
- es können **alternative Vorgehensweisen** (oder eine Bestätigung der gewählten Vorgehensweise) zu diesen Aspekten genannt werden und
- es können **Konsequenzen** für die eigene Professionalisierung bzw. die Professionalisierung des Gegenübers sowie für den Folgeunterricht formuliert werden.

Von diesen vier Aspekten wurden im Testinstrument das **Bewerten** und **Vorschlagen von Alternativen** bzw. **Formulieren von Konsequenzen** fokussiert.

Rückmeldung zur Ihrer Testbearbeitung:

Im Rahmen der Testbearbeitung konnten Sie Rückmeldung zu verschiedenen Aspekten der beobachteten Unterrichtsausschnitte gegeben. Das folgende Diagramm zeigt Ihnen Ihr **Gesamtergebnis** (roter Punkt) im Verhältnis zu den Gesamtergebnissen aller bisherigen Testbearbeitungen von 21 Studierenden im durchschnittlich 5. Fachsemester (dargestellt durch den Boxplot). *Hinweis: Der Vergleichsdatensatz befindet sich aktuell noch im Aufbau und ist daher noch klein.*

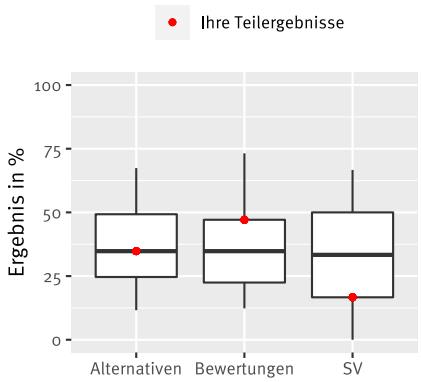
Einige Hinweise zur Interpretation des Diagramms finden Sie untenstehend.



1

UNIVERSITÄT PADERBORN
Die Universität der Informationsgesellschaft

Im nachfolgenden Diagramm ist ihr Testergebnis (rote Punkte) aufgeschlüsselt in die **Teilergebnisse** zu den Aspekten „Bewertungen“, „Alternativen“ und „Schülervorstellungen“, ebenfalls im Verhältnis zu allen bisherigen Testbearbeitungen von 21 Studierenden im durchschnittlich 5. Fachsemester (dargestellt durch die Boxplots) abgebildet.



Bewertungen: Unter dem Aspekt „Bewertungen“ sind alle Items zusammengefasst, in dessen Rahmen verschiedene Aspekte der Unterrichtsausschnitte bewertet werden können.

Alternativen/Konsequenzen: Unter dem Aspekt „Alternativen“ sind alle Items zusammengefasst, in dessen Rahmen Alternativen zu verschiedenen Aspekten der Unterrichtsausschnitte benannt werden bzw. Konsequenzen gezogen werden können.

Schülervorstellungen (SV): Die Diagnose von und der Umgang mit Schülervorstellungen ist ein zentraler Aspekt in der Physikdidaktik. Im Rahmen der beobachteten Unterrichtsausschnitte konnten verschiedene Schülervorstellungen beobachtet werden. Zu einem Unterrichtsausschnitt gab es die Möglichkeit, beobachtete Schülervorstellungen zu benennen und Strategien zum Umgang mit diesen vorzuschlagen.

Hinweise zur Interpretation der Diagramme:

- Ihr Testergebnis wird jeweils durch den roten Punkt markiert.
- Der ebenfalls eingezeichnete Boxplot stellt jeweils eine Übersicht über alle bisherigen Testergebnisse von Studierenden unterschiedlicher Fachsemester (überwiegend Studierende in der Mitte des Bachelorstudiums sowie Studierende im Praxissemester) dar.
- Alle bisherigen Testbearbeitungen liegen im Bereich zwischen der oberen und unteren Grenze der „Antennen“ im Diagramm.
- Im Rahmen der Box liegen die mittleren 50% der bisherigen Testergebnisse.
- Die horizontale Linie markiert den Median der bisherigen Testergebnisse. Der Median teilt die Testergebnisse in der Mitte. Es liegt also genau eine Hälfte der bisherigen Testergebnisse über bzw. unter dem Median.
- Etwaige schwarze Punkte zeigen „Ausreißer“, also Ergebnisse, die im Vergleich zu den mitt-

UNIVERSITÄT PADERBORN
Die Universität der Informationsgesellschaft

leren 50% der bisherigen Testergebnisse besonders hoch oder niedrig sind.

Über Feedback zum Testinstrument und insbesondere auch zu dieser Rückmeldung freuen wir uns. Richten Sie dieses gern an anna.weissbach@uni-paderborn.de.

J.3 Interviewleitfaden zur Erhebung mit Kohorte B

J.3.1 Einleitung/Einverständnis

Vielen Dank, dass Sie sich Zeit für dieses Interview nehmen.

Ich möchte zunächst um Ihr Einverständnis für die Verwendung Ihrer Daten bitten. Dazu lese ich Ihnen noch einmal die Einverständniserklärung vor und bitte Sie danach um Ihr mündliches Einverständnis.

Ich bin damit einverstanden, dass dieses Interview aufgezeichnet wird. Dieses Interview wird für eine Verschriftlichung, Verwendung oder Weitergabe anonymisiert. Diese Einverständniserklärung zur Nutzung der erhobenen Daten kann jederzeit widerrufen werden.

Sind Sie damit einverstanden?

Sie haben vor einigen Wochen den Reflexionstest bearbeitet und im Anschluss eine Rückmeldung zu Ihrer Testbearbeitung erhalten. Ich würde gleich gerne etwas darüber erfahren, wie Sie die Rückmeldung zum Reflexionstest wahrnehmen bzw. wahrgenommen haben. Dabei geht es explizit **nicht** darum, Sie bzw. Ihre Aussagen zu bewerten, sondern darum, das Format und die Gestaltung der Rückmeldung zu verbessern. *Ihre Aussagen in diesem Interview nehmen also auch **keinen** Einfluss auf Ihre Bewertung in universitären Veranstaltungen.*

Welche Fragen haben Sie noch, bevor wir mit dem Interview beginnen?

J.3.2 Fragen 1: Eigene Rückmeldung

Zum Einstieg geht es um Ihre eigene Rückmeldung

1. Können Sie sich an Ihre Rückmeldung zur Bearbeitung des Reflexionstests erinnern?
2. Welche Schlüsse haben Sie daraus gezogen?

J.3.3 Fragen 2: Beispielrückmeldung

Ich zeige Ihnen jetzt ein Beispiel einer Rückmeldung zum Reflexionstest. Nehmen Sie sich einen Moment Zeit, um die Rückmeldung anzuschauen/zu lesen. Wenn Sie möchten, können Sie ihre Gedanken, offene Fragen usw. laut aussprechen.

Im Anschluss habe ich einige Fragen an Sie.

1. Haben Sie zuvor mit Boxplots/diese Art von Diagramm gearbeitet? (In welchem Rahmen?)

2. Haben Sie zuvor mit der Effektstärke gearbeitet? (In welchem Rahmen?)
3. Welche Schlüsse ziehen Sie aus dieser Rückmeldung zum Reflexionstest? Falls nicht selbständig erwähnt: Welche Schlüsse ziehen Sie aus der Information über die Effektstärke?
4. Welche Informationen in der Rückmeldung sind besonders hilfreich oder interessant für Sie?
5. Welche Inhalte sind aus Ihrer Perspektive überflüssig?
6. Welche Fragen bleiben für Sie offen?
7. Welche weiteren Informationen hätten Sie gerne?

J.3.4 Fragen 3: Selbstlernmaterial

In der letzten Seminarsitzung (der Veranstaltung Ziele und Konzeptionen) haben Sie sich mit Reflexion beschäftigt und dabei auch mit einem online-Material gearbeitet. Die folgenden Fragen beziehen sich auf dieses Material:

1. Wie haben Sie mit dem Material gearbeitet? [Vorgehen, Fokussierung, ...]
2. Was nehmen Sie für sich aus der Arbeit mit dem Material mit? [vorrangig inhaltlich]
3. Bietet das Material aus Ihrer Perspektive einen Mehrwert und warum?
4. Welche Inhalte waren besonders hilfreich oder interessant für Sie?
5. Welche Inhalte sind aus Ihrer Perspektive überflüssig?
6. Welche Fragen bleiben für Sie offen?

J.3.5 Abschluss

Haben Sie sonstige Fragen/Anmerkungen zum Diagnose-Instrument, zur Rückmeldung oder dem online-Material?

J.4 Beispielrückmeldung zur Erhebung mit Kohorte B



Rückmeldung der Ergebnisse zum Reflexionstest

ProfileP Transfer
Professionswissen in der Lehramtsausbildung Physik

Sehr geehrte:r Teilnehmer:in,

Sie haben kürzlich am sogenannten Reflexionstest teilgenommen, in dem eine zentrale Fähigkeit für das fundierte, theoriegeleitete Handeln von Lehrer:innen – die Reflexionsfähigkeit – erhoben wird. Diese Rückmeldung soll Ihnen dabei helfen, sich in Bezug auf diese Fähigkeit realistisch einschätzen zu können.

Insgesamt haben Sie **68,8% der maximal möglichen Punktzahl** erreicht. Damit liegt ihr Ergebnis im Bereich der oberen 50-25% der Testergebnisse der Teilnehmenden aus der Vergleichsgruppe (bestehend aus 80 Studierenden im durchschnittlich 4,9. Fachsemester).

Im Folgenden finden Sie eine detailliertere Darstellung Ihrer Ergebnisse sowie erläuternde Informationen.

Reflektieren - Was ist das genau & Wie kommt das im Test vor?

Ziel einer Reflexion ist die Verbesserung der Unterrichtsqualität und bzw. oder die Weiterentwicklung persönlicher Kompetenzen.[1] Im Rahmen einer Reflexion können unterschiedliche Aussagen getätigt werden:

- Beobachtete Aspekte können **beschrieben** werden,
- diese können **bewertet** werden,
- es können kurz- und langfristige **Alternativen** vorgeschlagen werden: Kurzfristige Alternativen beinhalten alternative Vorgehensweisen (oder eine Bestätigung der gewählten Vorgehensweise) zu den einzelnen Aspekten, längerfristige Alternativen entsprechen Konsequenzen für die eigene Professionalisierung bzw. die Professionalisierung des Gegenübers sowie für den Folgeunterricht.

Von diesen Aspekten wurden im Testinstrument das **Bewerten** und **Vorschlagen von Alternativen** fokussiert.

Die Aspekte, über die reflektiert werden kann, sind vielfältig. Im Test wurden Ihnen zu jedem Unterrichtsausschnitt Aspekte vorgegeben, zu denen Sie in je einer Aufgabe (1) das Vorgehen der Lehrkraft **bewerten** und (2) **Alternativen** zum Vorgehen vorschlagen sollten.

Auch die Diagnose von und der Umgang mit **Schülervorstellungen** sind zentrale Aspekte der Physikdidaktik, die im Testinstrument in Bezug auf einen Unterrichtsausschnitt explizit adressiert werden. Zu diesem Ausschnitt waren Sie aufgefordert, **beobachtete Schülervorstellungen zu benennen** und **Strategien zum Umgang** mit diesen vorzuschlagen - hier gilt also ausnahmsweise nicht die Struktur von je einer Frage zum Bewerten und Vorschlagen von Alternativen.

[1] Das dem Test und dieser Rückmeldung zugrundeliegende Reflexionsmodell ist beschrieben in: Nowak, A., Kempin, M., Kulgemeyer, C. & Borowski, A. (2019). Reflexion von Physikunterricht. In: Christian Maurer (Hrsg.). Naturwissenschaftliche Bildung als Grundlage für berufliche und gesellschaftliche Teilhabe. Regensburg, S. 838-841.

[redacted]

Rückmeldung zur Ihrer Testbearbeitung:

Im Rahmen der Testbearbeitung konnten Sie Rückmeldung zu verschiedenen Aspekten der beobachteten Unterrichtsausschnitte gegeben. Das folgende Diagramm (s. Abb. 1) zeigt Ihnen Ihr **Gesamtergebnis** (roter Punkt) im Verhältnis zu den Gesamtergebnissen aller bisherigen Testbearbeitungen von 80 Studierenden im durchschnittlich 4,9. Fachsemester (dargestellt durch den Boxplot). Die Angabe der Ergebnisse erfolgt in Prozent und zeigt Ihnen, wie viele der maximal erreichbaren Punkte Sie (und alle bisherigen Testbearbeitungen) erreicht haben. Zum Vergleich Ihres Ergebnisses und des Mittelwertes der Vergleichsgruppe (blauer Punkt), können Sie zusätzlich für das Gesamtergebnis die Effektstärke (Cohens d) auf der rechten, vertikalen Achse ablesen (Der Wert 0,0 entspricht dem Mittelwert der Vergleichsgruppe).[2]

Zum einfacheren Verständnis ist dieser erste Boxplot zu Ihrem Gesamtergebnis beschriftet. Weitere Hinweise zur Interpretation der Diagramme finden Sie untenstehend.

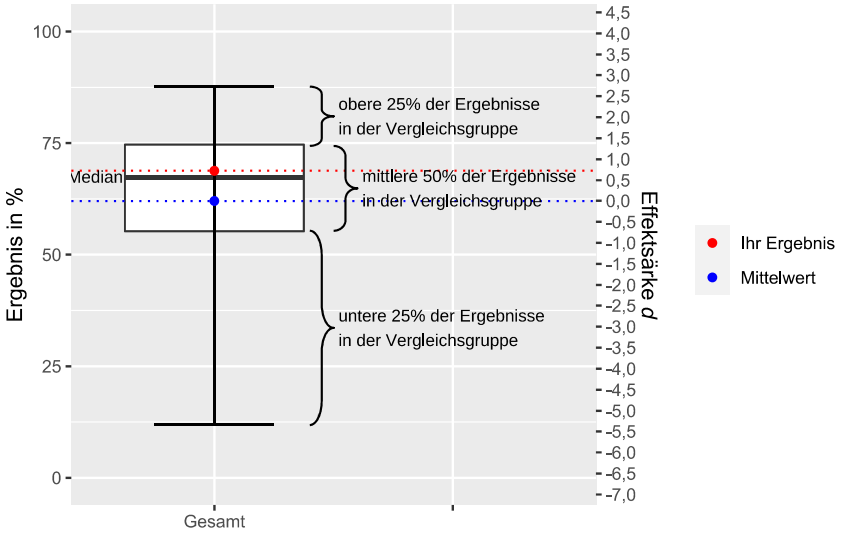


Abb. 1: Ihr Gesamtergebnis

Im nachfolgenden Diagramm (s. Abb. 2) ist ihr Testergebnis (rote Punkte) aufgeschlüsselt in die **Teilergebnisse** zu den Aspekten „Bewertungen", „Alternativen" und „Schülervorstellungen", ebenfalls im Verhältnis zu allen bisherigen Testbearbeitungen von 80 Studierenden im durchschnittlich 4,9. Fachsemester (dargestellt durch die Boxplots) und ebenfalls als prozentualer Anteil der maximal erreichbaren Punkte, abgebildet. Eine Erläuterung der Aufschlüsselung der Teilergebnisse finden Sie untenstehend.

Bewertungen: Unter dem Aspekt „Bewertungen" sind alle Items zusammengefasst, in dessen

[2] Nach Cohen (1988, S. 40) entsprechen Effektstärken zwischen $d = 0,2$ und $d = 0,5$ kleinen, von $d = 0,5$ bis $d = 0,8$ mittleren und ab $d = 0,8$ großen Effekten.
Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Science* (2. Auflage). New Jersey: Lawrence Erlbaum Associates, Hillsdale.

[redacted]

Rückmeldung der Ergebnisse zum Reflexionstest

ProfileP Transfer
Professionswissen in der Lehramtsausbildung Physik

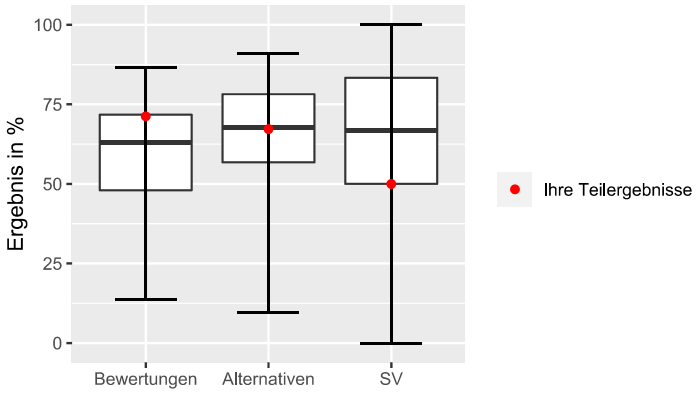


Abb. 2: Ihre Teilergebnisse

Rahmen verschiedene Aspekte der Unterrichtsausschnitte bewertet werden können (jeweils die erste Frage zu einem Aspekt).

Alternativen/Konsequenzen: Unter dem Aspekt „Alternativen" sind alle Items zusammengefasst, in dessen Rahmen Alternativen zu verschiedenen Aspekten der Unterrichtsausschnitte benannt werden können (jeweils die zweite Frage zu einem Aspekt).

Schülervorstellungen (SV): Unter den Aspekt „Schülervorstellungen" fallen die beiden Items, in denen Sie beobachtete Schülervorstellungen benennen und Strategien zum Umgang mit diesen vorschlagen sollten. *Hinweis: Die Anzahl der Items zu Schülervorstellungen ist also deutlich geringer, als die zu Bewertungen und Alternativen – diesem Aspekt kommt also deutlich weniger Gewicht in der Gesamtwertung zu.*

Wie kann ich meine Reflexionsfähigkeit schulen?

Im Rahmen Ihres Studiums erlangen Sie Wissen (z. B. zu Schülervorstellungen oder zu Experimenten), welches Sie für eine fundierte Reflexion benötigen. Dieses Wissen dient Ihnen als (theoretische) Grundlage für Ihre Bewertungen oder alternativen Handlungsvorschläge zu gesehenem oder selbst durchgeführtem Unterricht. Reflektieren ist praktisches Handeln und muss daher auch durch die Anwendung trainiert werden. Fördermaterial mit Informationen und Hinweisen, um das Reflektieren praktisch zu üben, finden Sie hier: https://go.upb.de/Foerdermaterial.

Hinweise zur Interpretation der Diagramme:

- Ihr Testergebnis wird jeweils durch den roten Punkt markiert.
- Der ebenfalls eingezeichnete Boxplot stellt jeweils eine Übersicht über alle bisherigen Testergebnisse von Studierenden unterschiedlicher Fachsemester (überwiegend Studierende in der Mitte des Bachelorstudiums sowie Studierende im Praxissemester) dar.
- Alle bisherigen Testbearbeitungen liegen im Bereich zwischen der oberen und unteren

[redacted]

3

Grenze der „Antennen“ im Diagramm.

- Im Rahmen der Box liegen die mittleren 50% der bisherigen Testergebnisse.
- Die horizontale Linie markiert den Median der bisherigen Testergebnisse. Der Median teilt die Testergebnisse in der Mitte. Es liegt also genau eine Hälfte der bisherigen Testergebnisse über bzw. unter dem Median.

Wir freuen uns über Ihr **Feedback zum Testinstrument und insbesondere auch zu dieser Rückmeldung**! Folgen Sie dafür einfach diesem Link (go.upb.de/RueckmeldungReflexion) oder wenden Sie sich direkt an anna.weissbach@uni-paderborn.de.

[redacted]

J.5 Interviewleitfaden zur Erhebung mit Kohorte C

J.5.1 Einleitung/Einverständnis

Vielen Dank, dass Sie sich Zeit für dieses Interview nehmen.

Ich möchte zunächst um Ihr Einverständnis für die Verwendung Ihrer Daten bitten. Dazu lese ich Ihnen noch einmal die Einverständniserklärung vor und bitte Sie danach um Ihr mündliches Einverständnis.

Ich bin damit einverstanden, dass dieses Interview aufgezeichnet wird. Dieses Interview wird für eine Verschriftlichung, Verwendung oder Weitergabe anonymisiert. Diese Einverständniserklärung zur Nutzung der erhobenen Daten kann jederzeit widerrufen werden.

Sind Sie damit einverstanden?

Sie haben vor einigen Wochen den Reflexionstest bearbeitet und im Anschluss eine Rückmeldung zu Ihrer Testbearbeitung erhalten. Ich würde gleich gerne etwas darüber erfahren, wie Sie die Rückmeldung zum Reflexionstest wahrnehmen bzw. wahrgenommen haben. Dabei geht es explizit **nicht** darum, Sie bzw. Ihre Aussagen zu bewerten, sondern darum, das Format und die Gestaltung der Rückmeldung zu verbessern. *Ihre Aussagen in diesem Interview nehmen also auch* ***keinen*** *Einfluss auf Ihre Bewertung in universitären Veranstaltungen.*

Welche Fragen haben Sie noch, bevor wir mit dem Interview beginnen?

J.5.2 Vorab: Mehrfache Testbearbeitung

Wurde der Test schon einmal bearbeitet?

Wurde der Test mehrfach bearbeitet, werden die kursiv gesetzten Fragen ergänzt.

J.5.3 Fragen 1: Eigene Rückmeldung

Zum Einstieg geht es um Ihre eigene Rückmeldung:

1. Können Sie sich an Ihre Rückmeldung zur Bearbeitung des Reflexionstests erinnern?
2. Welche Schlüsse haben Sie daraus gezogen?
3. *Inwiefern haben Sie bei der Beschäftigung mit der Rückmeldung die letzte Bearbeitung des Instruments bzw. Ihre damalige Rückmeldung berücksichtigt? (Ggf. umformuliert: Hat sich Ihre Perspektive auf die Rückmeldung oder Ihre persönlichen Ergebnisse durch die wiederholte Bearbeitung der Umgebung verändert?)*

J.5.4 Fragen 2: Beispielrückmeldung

Ich zeige Ihnen jetzt ein Beispiel einer Rückmeldung zum Reflexionstest. Nehmen Sie sich einen Moment Zeit, um die Rückmeldung anzuschauen/zu lesen. Wenn Sie möchten, können Sie ihre Gedanken, offene Fragen usw. laut aussprechen.

Im Anschluss habe ich einige Fragen an Sie.

1. Haben Sie zuvor mit Boxplots/dieser Art von Diagramm gearbeitet? (In welchem Rahmen?)
2. Welche Schlüsse ziehen Sie aus dieser Rückmeldung zum Reflexionstest?
3. Welche Informationen in der Rückmeldung sind besonders hilfreich oder interessant für Sie?
4. Welche Inhalte sind aus Ihrer Perspektive überflüssig?
5. Welche Fragen bleiben für Sie offen?
6. Welche weiteren Informationen hätten Sie gerne?

J.5.5 Fragen 3: Selbstlernmaterial

In der Rückmeldung, die Sie erhalten haben, ist ein online-Material verlinkt. Haben Sie dieses Material bearbeitet (ganz oder zumindest teilweise)? Bzw. planen Sie, mit dem Material zu arbeiten?

Falls nein:

1. Warum nicht?
2. *Hat die Tatsache, dass Sie das Material schon kennen, Ihre Entscheidung beeinflusst, nicht mit dem Material zu arbeiten?/arbeiten zu wollen?*
3. Welche Rahmenbedingungen würden Sie benötigen, damit Sie sich mit dem Material beschäftigen?

Falls ja: Die folgenden Fragen beziehen sich auf dieses Material.

1. *Hat die Tatsache, dass Sie das Material schon kennen, Ihre Entscheidung beeinflusst, mit dem Material zu arbeiten?/arbeiten zu wollen?*
   Die nachfolgenden Fragen werden lediglich gestellt, wenn das Material bearbeitet wurde.
2. Wie haben Sie mit dem Material gearbeitet? [Vorgehen, Fokussierung, . . .]
3. Was nehmen Sie für sich aus der Arbeit mit dem Material mit? [vorrangig inhaltlich]
4. Bietet das Material aus Ihrer Perspektive einen Mehrwert und warum?

5. Welche Inhalte waren besonders hilfreich oder interessant für Sie?
6. Welche Inhalte sind aus Ihrer Perspektive überflüssig?
7. Welche Fragen bleiben für Sie offen?
8. *Inwiefern hat sich Ihre Arbeit mit dem Material dadurch verändert, dass Sie es jetzt ein zweites Mal bearbeitet haben?*

J.5.6 Fragen 4: Zur wiederholten Testbearbeitung

Kommen wir zurück zum Diagnose-Instrument. Sie haben das Instrument jetzt (mindestens) das zweite Mal bearbeitet.

1. *Woran haben Sie sich vor und bei der Bearbeitung des Instruments noch erinnert?*
2. *Inwiefern hat es die Bearbeitung des Instruments beeinflusst, dass Sie das Instrument vorher schon bearbeitet hatten?*

J.5.7 Abschluss

Haben Sie sonstige Fragen/Anmerkungen zum Diagnose-Instrument, zur Rückmeldung oder dem online-Material?

J.6 Beispielrückmeldung zur Erhebung mit Kohorte C



Rückmeldung der Ergebnisse zum Reflexionstest

ProfileP Transfer
Professionswissen in der Lehramtsausbildung Physik

Sehr geehrte:r Teilnehmer:in,

Sie haben kürzlich am sogenannten Reflexionstest teilgenommen, in dem eine zentrale Fähigkeit für das fundierte, theoriegeleitete Handeln von Lehrer:innen – die Reflexionsfähigkeit – erhoben wird. Diese Rückmeldung soll Ihnen dabei helfen, sich in Bezug auf diese Fähigkeit realistisch einschätzen zu können.

Insgesamt haben Sie **57,9% der maximal möglichen Punktzahl** erreicht. Damit liegt ihr Ergebnis im Bereich der unteren 25% der Testergebnisse der Teilnehmenden aus der Vergleichsgruppe (bestehend aus 100 Studierenden im durchschnittlich 5,9. Fachsemester und 3 Referendar*innen).

Im Folgenden finden Sie eine detailliertere Darstellung Ihrer Ergebnisse sowie erläuternde Informationen.

Reflektieren - Was ist das genau & Wie kommt das im Test vor?

Ziel einer Reflexion ist die Verbesserung der Unterrichtsqualität und bzw. oder die Weiterentwicklung persönlicher Kompetenzen.[1] Im Rahmen einer Reflexion können unterschiedliche Aussagen getätigt werden:

- Beobachtete Aspekte können **beschrieben** werden,
- diese können **bewertet** werden,
- es können kurz- und langfristige **Alternativen** vorgeschlagen werden: Kurzfristige Alternativen beinhalten alternative Vorgehensweisen (oder eine Bestätigung der gewählten Vorgehensweise) zu den einzelnen Aspekten, längerfristige Alternativen entsprechen Konsequenzen für die eigene Professionalisierung bzw. die Professionalisierung des Gegenübers sowie für den Folgeunterricht.

Von diesen Aspekten wurden im Testinstrument das **Bewerten** und **Vorschlagen von Alternativen** fokussiert.

Die Aspekte, über die reflektiert werden kann, sind vielfältig. Im Test wurden Ihnen zu jedem Unterrichtsausschnitt Aspekte vorgegeben, zu denen Sie in je einer Aufgabe (1) das Vorgehen der Lehrkraft **bewerten** und (2) **Alternativen** zum Vorgehen vorschlagen sollten.

Auch die Diagnose von und der Umgang mit **Schülervorstellungen** sind zentrale Aspekte der Physikdidaktik, die im Testinstrument in Bezug auf einen Unterrichtsausschnitt explizit adressiert werden. Zu diesem Ausschnitt waren Sie aufgefordert, **beobachtete Schülervorstellungen zu benennen** und **Strategien zum Umgang** mit diesen vorzuschlagen - hier gilt also ausnahmsweise nicht die Struktur von je einer Frage zum Bewerten und Vorschlagen von Alternativen.

[1] Das dem Test und dieser Rückmeldung zugrundeliegende Reflexionsmodell ist beschrieben in: Nowak, A., Kempin, M., Kulgemeyer, C. & Borowski, A. (2019). Reflexion von Physikunterricht. In: Christian Maurer (Hrsg.). Naturwissenschaftliche Bildung als Grundlage für berufliche und gesellschaftliche Teilhabe. Regensburg, S. 838-841.

[redacted]

Rückmeldung zur Ihrer Testbearbeitung:

Im Rahmen der Testbearbeitung konnten Sie Rückmeldung zu verschiedenen Aspekten der beobachteten Unterrichtsausschnitte gegeben. Das folgende Diagramm (s. Abb. 1) zeigt Ihnen Ihr **Gesamtergebnis** (roter Punkt) im Verhältnis zu den Gesamtergebnissen aller bisherigen Testbearbeitungen von 100 Studierenden im durchschnittlich 5,9 Fachsemester und 3 Referendar*innen (dargestellt durch den Boxplot). Die Angabe der Ergebnisse erfolgt in Prozent und zeigt Ihnen, wie viele der maximal erreichbaren Punkte Sie (und alle bisherigen Testbearbeitungen) erreicht haben.

Zum einfacheren Verständnis ist dieser erste Boxplot zu Ihrem Gesamtergebnis beschriftet. Weitere Hinweise zur Interpretation der Diagramme finden Sie untenstehend.

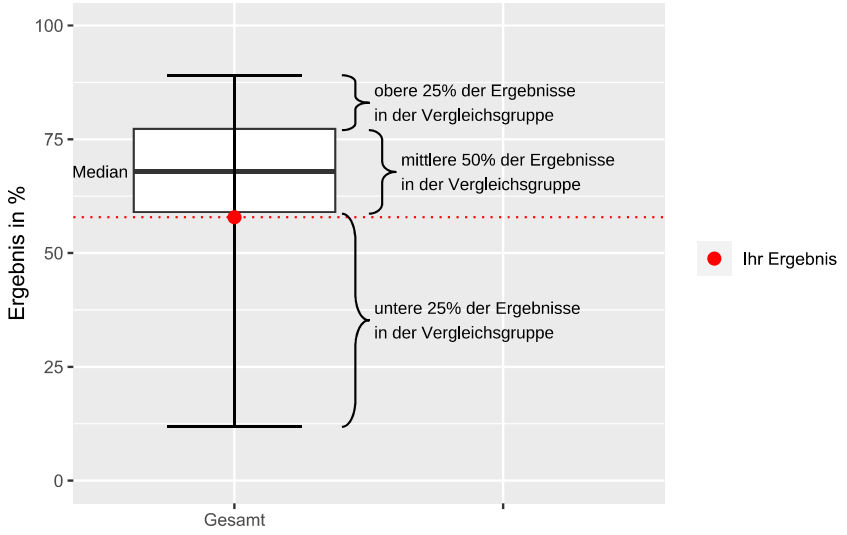


Abb. 1: Ihr Gesamtergebnis

Im nachfolgenden Diagramm (s. Abb. 2) ist ihr Testergebnis (rote Punkte) aufgeschlüsselt in die **Teilergebnisse** zu den Aspekten „Bewertungen“, „Alternativen“ und „Schülervorstellungen“, ebenfalls im Verhältnis zu allen bisherigen Testbearbeitungen von 100 Studierenden im durchschnittlich 5,9. Fachsemester und 3 Referendar*innen (dargestellt durch die Boxplots) und ebenfalls als prozentualer Anteil der maximal erreichbaren Punkte, abgebildet. Eine Erläuterung der Aufschlüsselung der Teilergebnisse finden Sie untenstehend.

Bewertungen: Unter dem Aspekt „Bewertungen“ sind alle Items zusammengefasst, in dessen Rahmen verschiedene Aspekte der Unterrichtsausschnitte bewertet werden können (jeweils die erste Frage zu einem Aspekt).

Alternativen/Konsequenzen: Unter dem Aspekt „Alternativen“ sind alle Items zusammengefasst, in dessen Rahmen Alternativen zu verschiedenen Aspekten der Unterrichtsausschnitte

[redacted]

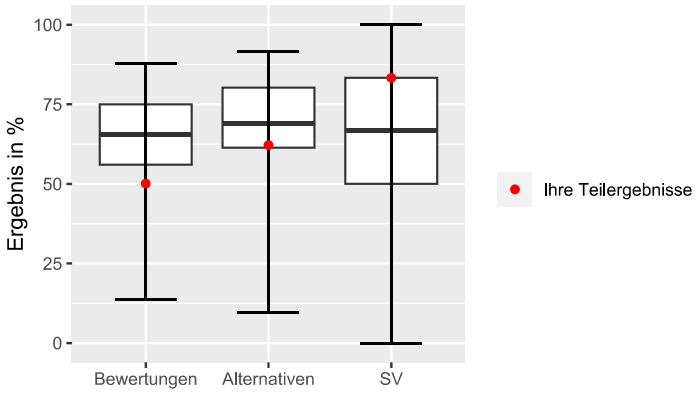


Abb. 2: Ihre Teilergebnisse

benannt werden können (jeweils die zweite Frage zu einem Aspekt).

Schülervorstellungen (SV): Unter den Aspekt „Schülervorstellungen" fallen die beiden Items, in denen Sie beobachtete Schülervorstellungen benennen und Strategien zum Umgang mit diesen vorschlagen sollten. *Hinweis: Die Anzahl der Items zu Schülervorstellungen ist also deutlich geringer, als die zu Bewertungen und Alternativen – diesem Aspekt kommt also deutlich weniger Gewicht in der Gesamtwertung zu.*

Wie kann ich meine Reflexionsfähigkeit schulen?

Im Rahmen Ihres Studiums erlangen Sie Wissen (z. B. zu Schülervorstellungen oder zu Experimenten), welches Sie für eine fundierte Reflexion benötigen. Dieses Wissen dient Ihnen als (theoretische) Grundlage für Ihre Bewertungen oder alternativen Handlungsvorschläge zu gesehenem oder selbst durchgeführtem Unterricht. Reflektieren ist praktisches Handeln und muss daher auch durch die Anwendung trainiert werden. Fördermaterial mit Informationen und Hinweisen, um das Reflektieren praktisch zu üben, finden Sie hier: https://go.upb.de/Foerdermaterial.

Auf Grundlage Ihrer Ergebnisse im Diagnose-Instrument empfehlen wir Ihnen, bei der Bearbeitung des Fördermaterials einen **Fokus auf die Entwicklung von Bewertungen** zu legen.

Hinweise zur Interpretation der Diagramme:

- Ihr Testergebnis wird jeweils durch den roten Punkt markiert.
- Der ebenfalls eingezeichnete Boxplot stellt jeweils eine Übersicht über alle bisherigen Testergebnisse von Studierenden unterschiedlicher Fachsemester (überwiegend Studierende in der Mitte des Bachelorstudiums sowie Studierende im Praxissemester) dar.
- Alle bisherigen Testbearbeitungen liegen im Bereich zwischen der oberen und unteren Grenze der „Antennen" im Diagramm.
- Im Rahmen der Box liegen die mittleren 50% der bisherigen Testergebnisse.

[redacted]

Rückmeldung der Ergebnisse zum Reflexionstest

ProfileP Transfer
Professionswissen in der Lehramtsausbildung Physik

- Die horizontale Linie markiert den Median der bisherigen Testergebnisse. Der Median teilt die Testergebnisse in der Mitte. Es liegt also genau eine Hälfte der bisherigen Testergebnisse über bzw. unter dem Median.

Wir freuen uns über Ihr **Feedback zum Testinstrument und insbesondere auch zu dieser Rückmeldung**! Folgen Sie dafür einfach diesem Link (go.upb.de/RueckmeldungReflexion) oder wenden Sie sich direkt an anna.weissbach@uni-bremen.de.

[redacted]

4

K Ergebnisse der Akzeptanzbefragung zum Assessment-Feedback und zum Selbstlernmaterial

Aufgrund der geringen Anzahl an Einschätzungen werden die absoluten Häufigkeiten der Einschätzungen entlang der Likert-Skala (*1 - stimme gar nicht zu, 2 - stimme eher nicht zu, 3 - stimme eher zu, 4 stimme voll und ganz zu, 0 - keine Antwort*) berichtet.

		Item	stimme gar nicht zu	stimme eher nicht zu	stimme eher zu	stimme voll und ganz zu	keine Antwort
Teil des Akzeptanzfragebogens	2 ($N = 5$)	Ich möchte den Test in Zukunft erneut durchführen und eine Rückmeldung erhalten, um meine Entwicklung zu überprüfen.	0	2	2	1	0
		Ich werde den Test (mit Rückmeldung) meinen Kommiliton*innen weiterempfehlen.	0	2	3	0	0
		Ich beabsichtige, mich mit dem Fördermaterial zu beschäftigen.	0	0	2	3	0
		Beim Herunterladen der Rückmeldung traten Schwierigkeiten auf.	*5*	*0*	*0*	*0*	*0*
		Ich finde die Rückmeldung nützlich für mein Studium/Referendariat.	0	0	4	1	0
		Die Inhalte der Rückmeldung sind für mein Studium/Referendariat relevant.	0	1	3	1	0
		Die Rückmeldung gibt mir neue Informationen in Bezug auf meine Reflexionsfähigkeit.	0	2	1	2	0
		Die Rückmeldung hilft mir dabei, meine Reflexionsfähigkeit besser einschätzen zu können.	0	1	2	2	0
		Es hat sich gelohnt, den Test gemacht und die Rückmeldung erhalten zu haben.	0	2	3	0	0
		Der Aufwand für die Testbearbeitung und Arbeit mit der Rückmeldung ist im Vergleich zum Nutzen zu hoch.	*0*	*4*	*0*	*1*	*0*
		Mein*e Dozent*in würde es gutheißen, falls ich das Fördermaterial nutze.	0	0	2	3	0
		Mein*e Dozent*in würde es gutheißen, falls ich den Test zukünftig erneut durchführe, um meinen Fortschritt zu überprüfen.	0	0	2	3	0
		Viele meiner Kommiliton*innen sehen einen Mehrwert in der Nutzung der Lernumgebung (Test, Rückmeldung und Fördermaterial).	0	1	2	2	0
	3 ($N = 7$)	Ich werde das Lernmaterial meinen Kommiliton*innen weiterempfehlen.	0	0	5	2	1
		Bei der Arbeit mit dem Material traten technische Schwierigkeiten auf.	*8*	*0*	*0*	*0*	*0*
		Die Nutzung des Materials unterstützt mich bei der Verbesserung meiner Reflexionsfähigkeit.	0	0	6	2	0
		Die Inhalte des Materials sind für mein Studium relevant.	0	0	4	4	0

Tabelle K.1: Übersicht über die Einschätzungen der Studierenden zur Akzeptanz der Lernumgebung (kursiv gesetzte Items sind invertiert formuliert); Teil 2 wurde von den Studierenden der Interviewkohorte C (nach Erhalt der Rückmeldung vor Durchführung der Interviews) beantwortet und Teil 3 eingebettet in der *LimeSurvey*-Umgebung vorrangig von den sechs Studierenden der Interviewkohorte B nach Bearbeitung des Selbstlernmaterials (im Rahmen der Entwicklung und Befragung noch als „Fördermaterial“ bezeichnet)

L Vollständige statistische Untersuchungen zur Profilanalyse (Forschungsfrage 2)

L.1 Gruppenunterschiede in Testwerten (gesamt und Teilfähigkeiten Bewerten und Vorschlagen von Alternativen)

- Kruskal-Wallis-Test zum Gesamtscore Reflexionsfähigkeit: $H(3) = 167,47$; $p < 0,001$
- Kruskal-Wallis-Test zum Teilscore Bewerten: $H(3) = 155,35$; $p < 0,001$
- Kruskal-Wallis-Test zum Teilscore Vorschlagen von Alternativen: $H(3) = 162,42$; $p < 0,001$

Fähigkeitsstufen	Gesamtscore		Bewertungen		Alternativen	
	p	Eff.-st.	p	Eff.-st.	p	Eff-.st.
1,2	$< 0,001$	$0,75$	$< 0,001$	$0,72$	$< 0,001$	$0,75$
1,3	$< 0,001$	$0,68$	$< 0,001$	$0,68$	$< 0,001$	$0,68$
1,4	$< 0,001$	$0,73$	$< 0,001$	$0,73$	$< 0,001$	$0,73$
2,3	$< 0,001$	$0,81$	$< 0,001$	$0,70$	$< 0,001$	$0,79$
2,4	$< 0,001$	$0,86$	$< 0,001$	$0,86$	$< 0,001$	$0,86$
3,4	$< 0,001$	$0,85$	$< 0,001$	$0,81$	$< 0,001$	$0,81$

Tabelle L.1: Ergebnisse der paarweisen Wilcoxon-Tests mit Bonferroni-Korrektur für den Gesamtscore im Reflexionstest sowie für die Teilscores zum Bewerten und Vorschlagen von Alternativen (*Eff.-st.: Effektstärke*)

L.2 Gruppenunterschiede in Aufgabenscores

L.2.1 Aufgaben zu Vignette 1

- Kruskal-Wallis-Test zu R1Bew1: $H(3) = 37,60$; $p < 0,001$
- Kruskal-Wallis-Test zu R1Alt1: $H(3) = 40,34$; $p < 0,001$

Fähigkeitsstufen	R1Bew1		R1Alt1	
	p	Effektstärke	p	Effektstärke
1,2	$0,08$	-	$0,26$	-
1,3	$0,004$	$0,34$	$0,01$	$0,33$
1,4	$< 0,001$	$0,54$	$< 0,001$	$0,58$
2,3	$1,0$	-	$1,0$	-
2,4	$< 0,001$	$0,42$	$< 0,001$	$0,44$
3,4	$< 0,01$	$0,36$	$< 0,001$	$0,42$

Tabelle L.2: Ergebnisse der paarweisen Wilcoxon-Tests mit Bonferroni-Korrektur für die Aufgabenscores zu Vignette 1

L.2.2 Aufgaben zu Vignette 2

- Kruskal-Wallis-Test zu R2Bew1: $H(3) = 35,78$; $p < 0,001$
- Kruskal-Wallis-Test zu R2Alt1: $H(3) = 49,39$; $p < 0,001$

Fähigkeitsstufen	R2Bew1		R2Alt1	
	p	Effektstärke	p	Effektstärke
1,2	$0,14$	-	$0,02$	$0,37$
1,3	$0,004$	$0,36$	$< 0,001$	$0,50$
1,4	$< 0,001$	$0,53$	$< 0,001$	$0,65$
2,3	$0,17$	-	$0,24$	-
2,4	$< 0,001$	$0,46$	$< 0,001$	$0,46$
3,4	$0,007$	$0,29$	$< 0,001$	$0,36$

Tabelle L.3: Ergebnisse der paarweisen Wilcoxon-Tests mit Bonferroni-Korrektur für die Aufgabenscores zu Vignette 2

L.2.3 Aufgaben zu Vignette 3

- Kruskal-Wallis-Test zu R3Bew1: $H(3) = 14,22$; $p = 0,002 < 0,05$
- Kruskal-Wallis-Test zu R3Alt1: $H(3) = 22,26$; $p < 0,001$
- Kruskal-Wallis-Test zu R3Bew2: $H(3) = 70,69$; $p < 0,001$
- Kruskal-Wallis-Test zu R3Alt2: $H(3) = 40,4$; $p < 0,001$

Fähigkeits-stufen	R3Bew1		R3Alt1		R3Bew2		R3Alt2	
	p	Eff.-st.	p	Eff.-st.	p	Eff.-st.	p	Eff.-st.
1,2	0,12	0,30	1,0	-	0,002	0,45	0,04	0,34
1,3	0,02	0,31	0,57	-	$< 0,001$	0,47	$< 0,001$	0,41
1,4	0,002	0,42	0,005	0,39	$< 0,001$	0,77	$< 0,001$	0,63
2,3	1,0	-	0,72	-	0,81	-	0,16	-
2,4	0,35	-	$< 0,001$	0,39	$< 0,001$	0,62	$< 0,001$	0,49
3,4	1,0	0,11	0,01	0,27	$< 0,001$	0,53	0,04	0,24

Tabelle L.4: Ergebnisse der paarweisen Wilcoxon-Tests mit Bonferroni-Korrektur für die Aufgabenscores zu Vignette 3 (*Eff.-st.: Effektstärke*)

L.2.4 Aufgaben zu Vignette 4

- Kruskal-Wallis-Test zu R4Bew1: $H(3) = 21,12$; $p < 0,001$
- Kruskal-Wallis-Test zu R4Alt1: $H(3) = 23,95$; $p < 0,001$
- Kruskal-Wallis-Test zu R4Bew2: $H(3) = 97,31$; $p < 0,001$
- Kruskal-Wallis-Test zu R4Alt2: $H(3) = 83,91$; $p < 0,001$

Fähigkeitsstufen	R4Bew1		R4Alt1	
	p	Eff.-st.	p	Eff.-st.
1,2	0,02	0,37	1,0	-
1,3	0,04	0,29	0,19	-
1,4	$< 0,001$	0,51	0,005	0,39
2,3	1,0	-	0,25	-
2,4	0,03	0,28	$< 0,001$	0,41
3,4	0,12	-	0,05	0,23
Fähigkeitsstufen	R4Bew2		R4Alt2	
	p	Eff.-st.	p	Eff.-st.
1,2	0,01	0,39	0,13	-
1,3	$< 0,001$	0,56	0,01	0,34
1,4	$< 0,001$	0,76	$< 0,001$	0,69
2,3	0,05	0,25	1,0	-
2,4	$< 0,001$	0,74	$< 0,001$	0,71
3,4	$< 0,001$	0,66	$< 0,001$	0,67

Tabelle L.5: Ergebnisse der paarweisen Wilcoxon-Tests mit Bonferroni-Korrektur für die Aufgabenscores zu Vignette 4 (*Eff.-st.: Effektstärke*)

L.2.5 Aufgaben zu Vignette 5

- Kruskal-Wallis-Test zu R5Bew1: $H(3) = 60,54$; $p < 0,001$
- Kruskal-Wallis-Test zu R5Alt1: $H(3) = 65,06$; $p < 0,001$
- Kruskal-Wallis-Test zu R5Bew2: $H(3) = 87,62$; $p < 0,001$
- Kruskal-Wallis-Test zu R5Alt2: $H(3) = 70,91$; $p < 0,001$
- Kruskal-Wallis-Test zu R5Bew3: $H(3) = 81,41$; $p < 0,001$
- Kruskal-Wallis-Test zu R5Alt3: $H(3) = 84,63$; $p < 0,001$

Fähigkeitsstufen	R5Bew1		R5Alt1		R5Bew2	
	p	Eff.-st.	p	Eff.-st.	p	Eff.-st.
1,2	0,05	-	< 0,001	0,69	0,01	0,40
1,3	< 0,001	0,53	< 0,001	0,65	< 0,001	0,65
1,4	< 0,001	0,75	< 0,001	0,83	< 0,001	0,73
2,3	< 0,001	0,36	1,0	-	< 0,001	0,53
2,4	< 0,001	0,61	< 0,001	0,43	< 0,001	0,70
3,4	< 0,001	0,28	< 0,001	0,36	0,003	0,31
Fähigkeitsstufen	R5Alt2		R5Bew3		R5Alt3	
	p	Eff.-st.	p	Eff.-st.	p	Eff.-st.
1,2	0,21	-	0,05	0,34	0,01	0,42
1,3	< 0,001	0,57	< 0,001	0,63	< 0,001	0,65
1,4	< 0,001	0,69	< 0,001	0,80	< 0,001	0,83
2,3	< 0,001	0,43	< 0,001	0,50	< 0,001	0,47
2,4	< 0,001	0,64	< 0,001	0,67	< 0,001	0,69
3,4	0,002	0,32	0,03	0,25	0,01	0,29

Tabelle L.6: Ergebnisse der paarweisen Wilcoxon-Tests mit Bonferroni-Korrektur für die Aufgabenscores zu Vignette 5 (*Eff.-st.: Effektstärke*)

L.2.6 Aufgaben zu Vignette 6

- Kruskal-Wallis-Test zu R6Bew1: $H(3) = 67,68$; $p < 0,001$
- Kruskal-Wallis-Test zu R6Alt1: $H(3) = 71,36$; $p < 0,001$
- Kruskal-Wallis-Test zu R6Bew2: $H(3) = 44,49$; $p < 0,001$
- Kruskal-Wallis-Test zu R6Alt2: $H(3) = 68,50$; $p < 0,001$
- Kruskal-Wallis-Test zu R6Bew3: $H(3) = 76,07$; $p < 0,001$
- Kruskal-Wallis-Test zu R6Alt3: $H(3) = 85,15$; $p < 0,001$

Fähigkeitsstufen	R6Bew1		R6Alt1		R6Bew2	
	p	Eff.-st.	p	Eff.-st.	p	Eff.-st.
1,2	$< 0,001$	0,62	$< 0,001$	0,76	$< 0,001$	0,48
1,3	$< 0,001$	0,69	$< 0,001$	0,72	$< 0,001$	0,56
1,4	$< 0,001$	0,85	$< 0,001$	0,85	$< 0,001$	0,70
2,3	0,52	0,16	0,34	0,18	0,20	0,20
2,4	0,005	0,33	$< 0,001$	0,42	0,001	0,37
3,4	0,27	0,18	0,02	0,26	0,28	0,18
Fähigkeitsstufen	R6Alt2		R6Bew3		R6Alt3	
	p	Eff.-st.	p	Eff.-st.	p	Eff.-st.
1,2	0,001	0,47	$< 0,001$	0,53	$< 0,001$	0,56
1,3	$< 0,001$	0,60	$< 0,001$	0,61	$< 0,001$	0,70
1,4	$< 0,001$	0,76	$< 0,001$	0,82	$< 0,001$	0,94
2,3	$< 0,001$	0,37	0,01	0,28	$< 0,001$	0,38
2,4	$< 0,001$	0,60	$< 0,001$	0,64	$< 0,001$	0,63
3,4	0,02	0,26	$< 0,001$	0,40	0,01	0,29

Tabelle L.7: Ergebnisse der paarweisen Wilcoxon-Tests mit Bonferroni-Korrektur für die Aufgabenscores zu Vignette 6 (*Eff.-st.: Effektstärke*)

L.2.7 Aufgaben zu Vignette 7

- Kruskal-Wallis-Test zu R7Bew1: $H(3) = 50,04$; $p < 0,001$
- Kruskal-Wallis-Test zu R7Alt1: $H(3) = 67,43$; $p < 0,001$
- Kruskal-Wallis-Test zu R7Bew2: $H(3) = 91,74$; $p < 0,001$
- Kruskal-Wallis-Test zu R7Alt2: $H(3) = 81,87$; $p < 0,001$

Fähigkeitsstufen	R7Bew1		R7Alt1	
	p	Eff.-st.	p	Eff.-st.
1,2	$< 0,001$	0,50	$< 0,001$	0,52
1,3	$< 0,001$	0,51	$< 0,001$	0,52
1,4	$< 0,001$	0,69	$< 0,001$	0,72
2,3	1,0	-	0,28	-
2,4	$< 0,001$	0,40	$< 0,001$	0,60
3,4	$< 0,001$	0,41	$< 0,001$	0,42
Fähigkeitsstufen	R7Bew2		R7Alt2	
	p	Eff.-st.	p	Eff.-st.
1,2	$< 0,001$	0,66	$< 0,001$	0,50
1,3	$< 0,001$	0,64	$< 0,001$	0,64
1,4	$< 0,001$	0,86	$< 0,001$	0,92
2,3	0,003	0,32	$< 0,001$	0,38
2,4	$< 0,001$	0,75	$< 0,001$	0,67
3,4	$< 0,001$	0,46	$< 0,001$	0,34

Tabelle L.8: Ergebnisse der paarweisen Wilcoxon-Tests mit Bonferroni-Korrektur für die Aufgabenscores zu Vignette 7 (*Eff.-st.: Effektstärke*)

L.2.8 Aufgaben zur globalen Reflexion

- Kruskal-Wallis-Test zu RGBew1: $H(3) = 76,66$; $p < 0,001$
- Kruskal-Wallis-Test zu RGAlt1: $H(3) = 107,52$; $p < 0,001$
- Kruskal-Wallis-Test zu RGBew2: $H(3) = 51,10$; $p < 0,001$
- Kruskal-Wallis-Test zu RGAlt2: $H(3) = 65,80$; $p < 0,001$

Fähigkeitsstufen	RGBew1		RGAlt1	
	p	Eff.-st.	p	Eff.-st.
1,2	$< 0,001$	$0,60$	$< 0,001$	$0,65$
1,3	$< 0,001$	$0,65$	$< 0,001$	$0,75$
1,4	$< 0,001$	$0,77$	$< 0,001$	$0,91$
2,3	$0,01$	$0,29$	$< 0,001$	$0,56$
2,4	$< 0,001$	$0,61$	$< 0,001$	$0,75$
3,4	$< 0,001$	$0,38$	$0,02$	$0,27$
Fähigkeitsstufen	RGBew2		RGAlt2	
	p	Eff.-st.	p	Eff.-st.
1,2	$0,61$	$0,21$	$0,12$	$0,30$
1,3	$< 0,001$	$0,41$	$< 0,001$	$0,51$
1,4	$< 0,001$	$0,63$	$< 0,001$	$0,80$
2,3	$0,001$	$0,35$	$< 0,001$	$0,38$
2,4	$< 0,001$	$0,61$	$< 0,001$	$0,65$
3,4	$0,01$	$0,28$	$0,002$	$0,32$

Tabelle L.9: Ergebnisse der paarweisen Wilcoxon-Tests mit Bonferroni-Korrektur für die Aufgabenscores zur globalen Reflexion (*Eff.-st.: Effektstärke*)

L.3 Gruppenunterschiede in demografischen Variablen

- Kruskal-Wallis-Test zur Ausbildungsphase: $H(3) = 18,10$; $p =< 0,001$
- Kruskal-Wallis-Test zum Fachsemester: $H(3) = 3,66$; $p = 0.30 \geq 0,05$
- Kruskal-Wallis-Test zum Umfang absolvierter Pflichtpraktika: $H(3) = 13,06$; $p = 0,005 < 0,05$

Fähigkeitsstufen	Ausbildungsphase		Praktika	
	p	Effektstärke	p	Effektstärke
1,2	$0,03$	$0,36$	$1,0$	-
1,3	$< 0,001$	$0,41$	$0,26$	-
1,4	$< 0,001$	$0,49$	$0,03$	$0,33$
2,3	$1,0$	-	$0,32$	-
2,4	$0,77$	-	$0,02$	$0,29$
3,4	$1,0$	-	$1,0$	-

Tabelle L.10: Ergebnisse der paarweisen Wilcoxon-Tests mit Bonferroni-Korrektur für die Zusammensetzung der Fähigkeitsstufen nach Ausbildungsphase und Umfang absolvierter Pflichtpraktika

L.4 Gruppenunterschiede in Testmotivation und Bearbeitungsdauer

- Kruskal-Wallis-Test zur Testmotivation: $H(3) = 15,74$; $p = 0,001 < 0,05$
- Kruskal-Wallis-Test zur Bearbeitungsdauer: $H(3) = 23,65$; $p < 0,01$

Fähigkeitsstufen	Testmotivation		Bearbeitungsdauer	
	p	Effektstärke	p	Effektstärke
1,2	$1,0$	-	$0,003$	$0,43$
1,3	$1,0$	-	$< 0,001$	$0,47$
1,4	$0,02$	$0,35$	$0,001$	$0,44$
2,3	$1,0$	-	$0,11$	-
2,4	$0,01$	$0,32$	$1,0$	-
3,4	$0,02$	$0,27$	$0,74$	-

Tabelle L.11: Ergebnisse der paarweisen Wilcoxon-Tests mit Bonferroni-Korrektur für die Zusammensetzung der Fähigkeitsstufen nach Testmotivation und Bearbeitungsdauer

M Vollständige statistische Untersuchungen zum Zusammenhang zwischen Noticing und Unterrichtsreflexion

M.1 Zu Hypothese 1

- $N = 189$
- Prüfung der Normalverteilung der Testwerte zur Reflexion mit Kolmogorow-Smirnow-Test: $D(189) = 0,10; p = 0,056 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung
- Prüfung der Normalverteilung der Testwerte zum Noticing mit Kolmogorow-Smirnow-Test: $D(189) = 0,52; p = 0,68 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung
- $r_{\text{Pearson}} = 0,61; p < 0,001 \Rightarrow$ Korrelation vorhanden (große Effektstärke)

M.2 Zu Hypothese 2

- Prüfung der Normalverteilung der Testwerte zum Noticing (Profil 1) mit Kolmogorow-Smirnow-Test: $D(17) = 0,20; p = 0,50 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung
- Prüfung der Normalverteilung der Testwerte zum Noticing (Profil 2) mit Kolmogorow-Smirnow-Test: $D(45) = 0,10; p = 0,74 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung
- Prüfung der Normalverteilung der Testwerte zum Noticing (Profil 3) mit Kolmogorow-Smirnow-Test: $D(71) = 0,09; p = 0,68 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung

- Prüfung der Normalverteilung der Testwerte zum Noticing (Profil 4) mit Kolmogorow-Smirnow-Test: $D(56) = 0,08; p = 0,88 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung
- Prüfung der Varianzhomogenität mit dem Levene-Test: $F(3; 185) = 0,85; p = 0,47 \geq 0,05 \Rightarrow$ Annahme der Varianzhomogenität
- ANOVA: $F(3, 185) = 30,04; p < 0,001$
- Post-hoc-Analyse zur Prüfung von Gruppenunterschieden für die Testwerte zum Noticing zwischen den Profilen:

Fähigkeitsstufen	p	Effektstärke
1,2	$0,01$	$0,77$
1,3	$< 0,001$	$-1,11$
1,4	$< 0,001$	$-2,00$
2,3	$0,15$	–
2,4	$< 0,001$	$-1,52$
3,4	$< 0,001$	$-1,0$

Tabelle M.1: Ergebnisse der paarweisen t-Tests mit Bonferroni-Korrektur zu Hypothese 2

M.3 Zu Hypothese 3

Vorab: Betrachtung der beiden Testwerte Reflexionsscore$_{\textit{nicht}\text{ angewählt im Noticing}}$ und Reflexionsscore$_{\text{angewählt im Noticing}}$

- Prüfung der Normalverteilung der Testwerte des Reflexionsscore$_{\text{angewählt im Noticing}}$ mit Kolmogorow-Smirnow-Test: $D(177) = 0,059; p = 0,57 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung
- Prüfung der Normalverteilung der Testwerte des Reflexionsscore$_{\textit{nicht}\text{ angewählt im Noticing}}$ mit Kolmogorow-Smirnow-Test: $D(177) = 0,15; p = 0,001 < 0,05 \Rightarrow$ *Keine* Annahme der Normalverteilung

Überprüfen von Unterschieden zwischen den beiden Testwerten über alle Profile:

- Prüfung der Normalverteilung der Testwertdifferenzen zwischen beiden Testwerten mit Kolmogorow-Smirnow-Test: $D(177) = 0,07; p = 0,44 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung

- Gepaarter t-Test: $t(176) = 12,03; p < 0,01 \Rightarrow$ Unterschied zwischen den Testwerten
- $d_{\text{Cohen},z} = 0,90$

Überprüfen von Unterschieden zwischen beiden Testwerten für Profil 1:

- Prüfung der Normalverteilung der Testwertdifferenzen zwischen beiden Testwerten mit Kolmogorow-Smirnow-Test: $D(17) = 0,12; p = 0,94 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung
- Gepaarter t-Test: $t(16) = 0,07; p = 0,94 \geq 0,05 \Rightarrow$ *Kein* Unterschied zwischen den Testwerten

Überprüfen von Unterschieden zwischen beiden Testwerten für Profil 2:

- Prüfung der Normalverteilung der Testwertdifferenzen zwischen beiden Testwerten mit Kolmogorow-Smirnow-Test: $D(45) = 0,10; p = 0,77 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung
- Gepaarter t-Test: $t(44) = 7,32; p < 0,01 \Rightarrow$ Unterschied zwischen den Testwerten
- Erreichte Testwerte
  - für als relevant einschätzte Aspekte: $M = 61\,\%$ $(SD = 8\,\%)$
  - für nicht als relevant einschätzte Aspekte: $M = 47\,\%$ $(SD = 10\,\%)$
- $d_{\text{Cohen},z} = 1,09$

Überprüfen von Unterschieden zwischen beiden Testwerten für Profil 3:

- Prüfung der Normalverteilung der Testwertdifferenzen zwischen beiden Testwerten mit Kolmogorow-Smirnow-Test: $D(70) = 0,15; p = 0,07 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung
- Gepaarter t-Test: $t(69) = 7,06; p < 0,01 \Rightarrow$ Unterschied zwischen den Testwerten
- Erreichte Testwerte
  - für als relevant einschätzte Aspekte: $M = 72\,\%$ $(SD = 5\,\%)$
  - für nicht als relevant einschätzte Aspekte: $M = 59\,\%$ $(SD = 13\,\%)$
- $d_{\text{Cohen},z} = 0,84$

Überprüfen von Unterschieden zwischen beiden Testwerten für Profil 4:

- Prüfung der Normalverteilung der Testwertdifferenzen zwischen beiden Testwerten mit Kolmogorow-Smirnow-Test: $D(45) = 0,07; p = 0,98 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung

- Gepaarter t-Test: $t(44) = 8,92; p < 0,01 \Rightarrow$ Unterschied zwischen den Testwerten
- Erreichte Testwerte
  - für als relevant einschätzte Aspekte: $M = 85\,\%$ ($SD = 4\,\%$)
  - für nicht als relevant einschätzte Aspekte: $M = 61\,\%$ ($SD = 18\,\%$)
- $d_{\text{Cohen},z} = 1,33$

N Vorausgehende statistische Untersuchungen zum Entwicklungszusammenhang zwischen den Fähigkeiten zu Noticing und Unterrichtsreflexion

N.1 Gesamtstichprobe ($N = 41$)

Prüfung der Normalverteilungen:

- Testwerte zur Reflexion zum 1. MZP mit Kolmogorow-Smirnow-Test: $D(41) = 0,11; p = 0,62 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung
- Testwerte zur Reflexion zum 2. MZP mit Kolmogorow-Smirnow-Test: $D(41) = 0,11; p = 0,72 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung
- Testwerte zum Noticing zum 1. MZP mit Kolmogorow-Smirnow-Test: $D(41) = 0,11; p = 0,72 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung
- Testwerte zum Noticing zum 1. MZP mit Kolmogorow-Smirnow-Test: $D(41) = 0,08; p = 0,94 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung

N.2 Teilstichprobe der oberen drei Profile ($N = 37$)

Prüfung der Normalverteilungen:

- Testwerte zur Reflexion zum 1. MZP mit Kolmogorow-Smirnow-Test: $D(37) = 0,07; p = 0,98 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung
- Testwerte zur Reflexion zum 2. MZP mit Kolmogorow-Smirnow-Test: $D(37) = 0,14; p = 0,44 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung
- Testwerte zum Noticing zum 1. MZP mit Kolmogorow-Smirnow-Test: $D(37) = 0,12; p = 0,71 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung
- Testwerte zum Noticing zum 1. MZP mit Kolmogorow-Smirnow-Test: $D(37) = 0,08; p = 0,97 \geq 0,05 \Rightarrow$ Annahme der Normalverteilung

O Segmentierte und kodierte Transkripte der Think-Aloud-Interviews zur Prüfung der Anforderung an Teil 2 des INA (Untersuchung kognitiver Prozesse)

Zur optischen Unterstützung sind in den Transkripten die Zeilen grau hinterlegt, die Beiträge der Interviewleitung enthalten sowie kennzeichnen, dass Videos aus dem Reflexionstest abgespielt wurden. Zusätzlich sind Kodierungen der Kategorie *Sonstiges* mit einem * versehen, wenn die jeweiligen Segmente auf Unklarheiten bei der Bearbeitung des Testinstruments hindeuten.

Interview 5 wurde vollständig durch eine zweite Person kodiert, auch diese Kodierungen sind dem Transkript beigefügt. Gleiches gilt für die ersten Segmente von Interview 4. Zur Auswertung wurde die erste Kodierung verwendet.

O.1 Interview 1: LD1

Segment	Kodierung	Transkript
0	-	[Einverständniserklärung, Demographie (Anm: Dass in Zahlenfeldern die Zahlen rechts stehen ist verwirrend)] #00:07:33-2#
1	-	I: Okay. #00:07:34-7#
2	Sonstiges	B1: Oh, okay. #00:07:36-3#

3	-	I: Das ist ähm genau, das ist der Punkt, ab dem wir gleich quasi laut Denken, also mit wir meine ich dich, (unv.) schon wieder. Also es ist jetzt so, dass während der Testdurchführung, die quasi inhaltlich mit mit dieser Situationsbeschreibung beginnt, werden gleich immer wieder Videos kommen, die angeschaut werden und zu denen im Nachhinein Fragen gestellt werden. Und ich würde dich jetzt bitten, dass du da alle deine Gedanken aussprichst laut, während du die Fragen beziehungsweise Aufgaben bearbeitest und auch die Texte und Aufgabenstellungen laut vorliest. Und wie eben auch, es geht nicht darum, irgendwie was zu erklären, das nachvollziehbar zu machen, sondern diese Vorstellung von ich bin alleine im Raum, äh genau ich spreche laut aus, worüber ich nachdenke und weiterhin gilt, genau, wenn du das länger nicht machst, dann melde ich mich kurz mit einer Aufforderung. // (.) Ansonsten bin ich jetzt quasi erstmal ruhig. #00:08:40-9#
4	Sonstiges (2x)	B1: (unv.) #00:08:38-9#
		B1: Soll ich hier gleich anfangen mit der Situationsbeschreibung? #00:08:44-5#
5	-	I: Genau, also das ist der inhaltliche Start vom Test und jetzt bist du quasi alleine, machst den Test und ich bin nicht mehr da. #00:08:51-0#
6	Vorlesen, Sonstiges (6x), eigene Erfahrung	B1: (Okay?) (.) Okay, Situationsbeschreibung. Stellen Sie sich folgendes vor. Schön formell. Sie- schon wieder so formell, absolvieren zurzeit ein Praktikum an der Schule. Ihr Mitpraktikant Robert- Roobert, (Werbung?) immer mal wieder, hat eine Doppelstunde seines Physikunterrichts aufgezeichnet und bittet Sie, ihm dazu eine Rückmeldung zu geben. Er zeigt Ihnen Ausschnitte seines Unterrichts, das Ganze kenne ich doch, ja, dieser eine Kurs, da hatten wir das genauso gemacht, und stellt Ihnen dazu Fragen. Um diese Fragen beantwort- Um diese Fragen zu beantworten, ah lern mal Lesen du Honk, können Sie Robert unter anderem darum bitten, Ihnen noch einmal das Unterrichtsvideo zu zeigen. Aber beachten Sie bitte die Zeit, denn Robert hat nur 70 min Zeit. Deswegen sollen die Videos maximal einmal wiederholt werden. Schauen Sie sich also nur in dringenden Fällen die Videos ein weiteres Mal an. Robert ist mit seinem Unterricht nicht zufrieden, kenne ich, und bittet Sie um ein Feedback. Weiter. #00:09:55-5#
7	Sonstiges	B1: Ach, ist sogar derselbe Typ, wie in dem Kurs. #00:10:01-6#
8	-	[Video Lerngruppe] // [Weiterhin Video Lerngruppe] // [Weiterhin Video Lerngruppe] #00:10:26-4#

9	Sonstiges, Beschreibung, Überlegung zum Unterricht	B1: (unv.). #00:10:14-3# B1: Ja, das nennt man Heterogenität. #00:10:24-9# B1: Mhm (.) ja wenn sie ruhig sind, dann werden sie nicht angesprochen, das ist eine ganz einfache Taktik. #00:10:32-1#
10		[Video Vorwissen] // [Weiterhin Video Vorwissen] // [Weiterhin Video Vorwissen] #00:13:38-6#
11	Alternative + Begründung	B1: Ja, da könnte man den Kontext f- äh des Vakuums am besten anwenden anwenden äh weil innerhalb des Weltraumes, dort ist so gut wie gar keine Luft und somit kaum bis gar keine Körper, die sich (ihm?) entgegenstellen können, sei es nun Partikel, sei es (.) irgendwas ähm da ist es also am einfachsten sich vorzustellen, diese gleichmäßige Bewegung, dieses Trägheits- gesetzt definitiv. #00:12:37-4#
12	Bewertung, Alternative + Begründung	B1: Der erste Satz scheint mir ein bisschen komplex formuliert, versuch mal ein bisschen runterzustufen, was die Fachwörter angeht. Fachwörter sitzen sehr selten und einfache Sprache kommt viel besser an den Schüler ran, (.) (so?) aus eigener Erfahrung #00:13:08-5#
13	Alternative (2x), Sonstiges	B1: Auch den könnte man ein wenig vereinfachen, zum Beispiel die grundlegenste Vereinfachung wäre actio gleich reactio oder ein bisschen breiter gefächert wäre immer, wenn ein Körper eine Kraft ausübt, wird eine gleichstarke, aber entgegengesetzte äh in entgegengesetzte Richung äh Kraft auf den äh Ausüber ausgeübt. Selbst das könnte man noch (feilen?), ja, da sprechen wir nachher nochmal drüber #00:14:06-4#
14	-	[Video Unterrichtsziele] #00:14:55-0#
15	Bewertung + Begründung, Überlegung zum Unterricht (2x), Sonstiges	B1: Mhm ja, ja, gar nicht mal so eine doofe Idee, das Fahrbahexperiment, das macht schon eine Menge Spaß, das dauert nur eben aufzubauen und man muss halt eben sich ein oder zwei Leute extra raussuchen, es ist allerdings auch ganz wichtig, dass der Physiklehrer nicht einfach alles vormacht, weil das ist kräftig scheiße, (.) dann haben die schüler absolut gar kein Interesse mehr dadran, es fühlt sich wieder so an, als Physik jetzt nichts- nicht etwas, was außerhalb des Unterrichts existieren würde, sondern nur auf diesen Bereich konsequent fokussiert ist. (.) Ja. #00:15:29-6#
16	Sonstiges, Vorlesen	B1: Okay, okay, okay, hier lädt einiges (B1 scrollt durch das Stundenraster). (.) Lehreverhalten und Schülerverhalten,
17	Sonstiges	L steht- (unv.) nein, ich sollte erst das Video gucken, bevor ich hier zu vorgreife #00:15:48-3#
18	-	[Video Stundenraster] // [Weiterhin Video Stundenraster] #00:16:11-7#
		B1: Okay, das erklärt das Bild. #00:15:57-7#

19	Sonstiges (4x), Vorlesen, Überlegung zum Unterricht (3x)	B1: Mhm wenn das man gut geht. Okay L erklärt Ablauf (der?) Quizrunde, äh jo. SUS beantworten Quizfragen, absolut. Partnerarbeit. Partnerarbeit? Quizze werden doch normalerweise in Einzelarbeit gemacht, aber okay. Beamer, ja der gute alte Overheadprojektor, Klassiker in der deutschen Schule. (.) Äh L moderiert Besprechung der Quizfragen SUS beantwortet Quizfragen im Plenumsgespräch, siehst du, das habe ich schon eher erwartet, (.) zeigt den Comic (unv./Geräusche). L führt in die Thematik Impuls mit Fragestellung ein, warum konnte die Titanic nicht bremsen, ja, da wären wir (dann?) wieder bei dem Trägheits- äh gesetz. (unv.) ganz interessant und den Rest habe ich nicht gelesen, also machen wir einfach mal weiter. #00:17:04-2#
20	Vorlesen (2x), Sonstiges (3x), Überlegung zum Unterricht	B1: Äh (..) Reflexion null eins. Ich habe die Stunde mit der Vorstellung des Ablaufs begonnen. Als nächstes habe ich ein Quiz zur Wiederholung der Newtonschen Axiome angeleitet, welches die Schüler Schülerinnen in Partnerarbeit durchgeführt haben. Die Durchführung des Quiz habe ich selbst- selbst habe ich zwar nicht gefilmt, dafür aber die Besprechung des Quiz mit den Schüler Schülerinnen, die ich dir jetzt zeige. Das ist groß gedruckt, also ist es wichtig, merk dir das. (.) Ah ja das Plenumsgespräch (zwischen?) Fragen. Da würde ich jetzt erwarten, dass einige Schüler Probleme mit den Antworten haben oder generell (hinterfragen?) äh bereits haben (.) ein paar Ideen, äh wo ich letztens auch schon mit einem Professor drüber gesprochen habe, diese Idee, dass statische Objekte keine Kraft ausüben können, das könnte man zum Beispiel dort erwarten. (.)
21	Sonstiges	B1: (...) Es lädt, es lädt, es lädt, jetzt. #00:18:12-5#
22	-	[Video Einleitung Reflexion 01] // [Weiterhin Video Einleitung Reflexion 01] #00:18:32-1#
23	Sonstiges	B1: Oh ja. #00:18:26-6#
24	eigene Erfahrung	B1: Mhm (.) (ach?) ja, dieses Vo- Verhalten kenne ich zu gut aus meiner eigenen Schulzeit, immer schön Merksätze auswendig lernen. Was bedeuten sie? Pff, mir doch egal. #00:18:43-5#
25	Sonstiges	B1: (4s) Ja, jetzt doch.
26	Vorlesen, Alternative	Unter welchem Namen ist das zweite Newtonsche Axiom auch bekannt und was besagt dieses Axiom. (Uh?) zwei Fragen in einem, hätte ich lieber getrennt. #00:18:58-3#
27	-	[Video Szene Reflexion 01] // [Weiterhin Video Szene Reflexion 01] // [Weiterhin Video Szene Reflexion 01] // [Weiterhin Video Szene Reflexion 01] // [Weiterhin Video Szene Reflexion 01] // [Weiterhin Video Szene Reflexion 01] // [Weiterhin Video Szene Reflexion 01] #00:21:49-6#
28	Sonstiges	B1: Das Newtonsche Axiom. #00:19:08-4#
29	Beschreibung	B1: Ach der vorn spricht. #00:19:15-9#

30	Sonstiges	B1: Mhm. #00:19:22-2#
31	Sonstiges (3x)	B1: Ich glaube, das ist jetzt ziemlich doof, mein Grafiktreiber ist abgeschmiert, ich müsste die Seite eben kurz neu laden. (.) Dreiunddreißig Prozent, Reflexion null eins. (...) Sind Sie noch da, Frau Weißbach? (.) Gut. (..) So, jetzt. #00:19:52-2#
32	Beschreibung	B1: Mhm[Ja] okay, er vorne spricht. #00:20:05-5#
33	Vorlesen	B1: das bedeutet, dass ohne die (unv.) (B1 liest vmtl. den Merksatz zum ersten Axiom aus dem Video vor). #00:20:15-3#
34	Beschreibung, Alternative, Sonstiges	B1: Pff alles an diesem Typ sieht gelangweilt aus, sein Gesicht sieht gelangweilt aus, seiner Körperhaltung fehlt jegliche Energie. (.) Wenn man Enthusiasmus bei den Schülern entwickeln will, dann muss man selber mit ein bisschen mehr Energie, ein bisschen mehr Elan, ein bisschen mehr AAH (gerufen) (.) an die Sache rangehen und er überhaupt nichts. Da ist keine Energie in dieser Körperhaltung. #00:21:46-5# B1: Aber wer weiß, vielleicht hat es auch funktioniert und ich äh überinterpretiere mal wieder, wäre nicht das erst Mal. #00:21:55-4#
35	Vorlesen, Sonstiges	B1: Okay, zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung, Lob (und/oder?) Kritik, erhalten- warte mal kurz, diesen gesamten Test kenne ich doch, den habe ich doch schonmal gemacht. (.) Ja, ja, das war in diesem einen [Dozent]-Kurs, wie hieß der nochmal? [Name des Kurses]? (..) Ja, ich glaube, das war der Kurs, in dem Kurs habe ich genau diese Aufgabe schon einmal gemacht, wo es auch mit dem Roboter äh Robert, nicht Roboter, dem Robert dann drum herum (gehen?), kategorisieren, (.) Kritiken geben (.) dadadadada okay.
36	Überlegung zum Unterricht	Ähm (.) wie hat er nochmal das erste Axiom formuliert? (7s) Ne, ich komme nicht mehr drauf,
37	Zustimmung AO	Formulierung Nummer zwei und drei allerdings und Gesprächsführung. (..)
38	Vorlesen	Hier können Sie zu bestimmten Fragen Antworten von Robert bekommen.
39	Vorlesen	(unv.) Informationen (unv.) noch einmal bekommen?
40	Vorlesen	Kannst du mir noch einmal sagen, welches Vorwissen die Schülerinnen und Schüler haben?
41	Vorlesen	Kannst du mir die Unterrichtsziele noch einmal be nennen?
42	Vorlesen	Kannst du mir noch einmal das Stundenraster zeigen?
43	Vorlesen	Kannst du mir den Unterrichtsausschnitt noch einmal zeigen?

44	Vorlesen (4x), Sonstiges (2x), Bewertung (3x), Bewertung + Begründung, Alternative + Begründung	Kannst du mir die PowerPoint-Folien nochmal zeigen? Ja, ich glaube, das letzte wäre ganz nett. (B1 öffnet Hilfe 'Kannst du mir die PowerPoint-Folien nochmal zeigen?') (.) Newtonsche Axiom, warum lädt hier nichts? (..) Ist mein Internet heute wirklich so beschissen? (.) ne, jetzt lädt er was. (unv.) (B1 liest murmelnd das erste Axiom vor) (.) Ja, okay, (unv.) ich da andere Maßstäbe, aber für deinen Physikkurs reicht es definitiv. Zweites Newtonsches Axiom, Beschleunigung a, die ein Körper erfährt, ist der Quotient- nope, das war dieses Ding, wo er mit Fachwörtern ein bisschen zu (.) sehr rumgeschlagen hat und auch die äh Form dieser Formel ist ein bisschen quer, weil die meisten Schüler werden mit ähm gleich F durch m weniger anfangen können, als jetzt mit der gant klassischen (.) äh Fma-Formel, also F gleich m mal a. (..) Wenn ein Kö- äh drittes newtosches Axiom, Wechselwirkungsprinzip, wenn ein Körper A eine Kraft F A zu B, zum Beispiel F A zu B gleich zehn Newton auf (B?) Körper B ausübt, dann übt Körper B anschließend auch immer eine Gegenkraft F B zu A, zum Beispiel minus zehn Newton, die Kräfte greifen an verschiedenen Körpern an und sind genau gleich groß, aber einander entgegengerichtet, Kraft gleich Gre- Gegenkraft. Das kann man definitiv kürzen, Merksätze soll man ja merken- sich merken können und dementsprechend sehr klei- klein und kurz sein. (.) Okay, also eins war ganz gut,
45	Bewertung, Sonstiges	Ergebnissi- sicherung war auch ganz gut, ja, dabei bleibt es. #00:24:59-6#
46	Sonstiges (2x)	B1: Ja und auch diese Meldung erinnere ich mich sehr gut dran, da war es nämlich andauernd so beim Test beim [Dozent], hat mir damals eine Menge geholfen, weil das weniger Arbeit für mich bedeutet hat. Okay. #00:25:17-9#
47	Vorlesen	B1: Welches Feedback möchten Sie Robert zu seiner Formulierung des Merksatzes zum dritten Axiom geben? (..)
48	Vorlesen, Ablehnung AO	Äh ist gut, nope,
49	Vorlesen, Ablehnung AO	fachlich falsch, das auch nicht.
50	Vorlesen, Zustimmung AO	Formulierung des dritten Axioms ist für die Schülerinnen Schüler nicht angemessen, da sie zu komplex ist. Ah ja, genau das.
51	Vorlesen, Zustimmung AO	Nicht gut, da die Vermischung von Beispiel und Merksatz- Ja, absolut.

52	Vorlesen, Zustimmung AO + Begründung	Es ist nicht gut, dass nur die Kurzformel Kraft gleich Gegenkraft, aber nicht actio gleich reactio im Merksatz vorkommt. Diese Abbildung Abkürzung ist nämlich etwas, was die Schüler Schülerinnen kennen vielleicht kennen. Äh ja das Argument kann man verstehen, actio gleich reactio ist die- (wirklich?) das bekanntere und weit verbreitere Formulierung des dritten Newtoschen Axioms, das kann man auch kritisieren. (.)
53	Vorlesen, Zustimmung AO + Begründung	Es ist gut, dass du im Me- Merksatz deutlich machst, dass die Kräfte an verschiedenen Körpern angreifen, da du so einer Verwechslung mit dem mit dem Kräftegleichgewicht vorbeugst. Ja, das ist auch ganz gut, auf den äh auf die ganze Geschichte, sind wir nämlich letztens in [Name der Veranstaltung] auch reingefallen, das war eine riesen Dis- äh kussion innerhalb von Semester fünf Physikstudenten also ist ganz gut, dass es mit dabei ist.
54	Vorlesen, Ablehnung AO + Begründung	Es ist nicht gut, dass äh (unv.) äh ne, ne also so clever schätze ich die Leute ein, dass sie zwischen den ein- beiden Sachen differenzieren können. (.)
55	Sonstiges	Weiter. #00:26:41-9#
56	Vorlesen, Sonstiges	B1: Welche der unten stehenden Hinweise zu seiner Formulierung des Merksatzes zum dritten Axiom möchten Sie Robert für eine erneute Durchführung des Unterrichtsstunde geben? (.) (Hä?) ja doch, macht Sinn.
57	Vorlesen	Okay wären Sie- wählen Sie die Antwortoptionen aus, die Ihrer Einschätzung am nächsten kommen.
58	Vorlesen, Ablehnung AO	Deine Formulierung des dritten Axioms ist gut so, du kannst es wieder so verwenden. Nope, absolut nein. (Ach?) okay vielleicht absolut nicht, das wäre vielleicht ein bisschen fies formuliert, aber es gibt definitiv kritisierbare (würdiges).
59	Vorlesen, Zustimmung AO	Es wäre besser, zuerst das Axiom ohne Beispiel zu notieren und anschließend ein Beispiel, um den Merksatz zu vereinfachen. Es könnte be- benutzt werden, ja, aber es ist nicht das, was ich am größten kritisieren würde, aber dennoch etwas kritisierwürdiges.
60	Vorlesen, Ablehnung AO	Es wäre besser, das Beispiel im Merksatz ausführlicher Darzustellen, damit die Schüler Schülerinnen die Anwendung des- vom dritten Axiom verstehen können. Als- äh nein, würde ich eher von abraten.
61	Vorlesen, Zustimmung AO + Begründung	Es wäre sinnvoll, das Axiom um eine Skizze zu er- ergänzen, damit es anschaulicher wird. Ja, das ist tatsächlich eine gute Idee, weil ein bildlicher Impuls ist ja immer besser als nur ein verbaler und oder ein schriftlicher Impuls. Erneut, ich spreche aus eigener Erfahrung.

62	Vorlesen, eigene Erfahrung, Professionalisierung, Sonstiges	Äh es wäre besser, im Merksatz zu verdeutlichen, dass Kraft und Gegenkraft gleichzeitig wirken. Deine Formulierung anschließend übt Körper B eine äh Gegenkraft auf Körper A aus ist fachlich falsch. Uh das war auch wieder so eine Diskussion letztens in einem Kurs, ob es da eine Zeitverzögerung zwischen Aktion und Reaktion gibt oder es direkt nacheinander ist. Ähm ich weiß jetzt aus dem Stegreif gar nicht, ob ob das jetzt fachlich falsch ist, oder fachlich korrekt ist, ob da jetzt eine Zeitverzögerung ist, so hatte der Herr [Dozent] ja auch das äh Pferdeproblem erklärt, also mhm schwer zu sagen, ich ignoriere es erstmal und tue so, als wäre es nicht da, um mir selber den Kopf zu retten.
63	Vorlesen, Ablehnung AO + Begründung	Du solltest im Merksatz äh über den Kräften Vektorpfeile notieren, damit die Wechsel- Wirkrichtung der Kräfte deutlich wird. Ne, ich glaube entgegengesetzte Kraft ist bereits ausreichend was die Wirkrichung der Kräfte angeht.
64	Vorlesen, Zustimmung AO + Begründung, Überlegung zum Unterricht, Sonstiges	Du solltest im Merksatz verdeutlichen, dass Kraft und Gegenkraft in genau entgegengesetzte Richtungen wirken, damit die Schüler Schülerinnen die Bedeutung des Vorzeichens im Beispiel verstehen. Ja, (.) wobei ich mir jetzt gerade gar nicht sicher bin, ob das vielleicht schon auf der Fo- Folie war, aber da diese Antwortmöglichkeit hier ist, muss es nicht da gewesen sein, weil ansonsten wäre diese Antwortmöglichkeit nicht hier, weswegen ich einfach mal so tue, als wäre es ganz korrekt. Weiter. #00:29:09-8#
65	Sonstiges (3x)	B1: (unv.) wieder bei (dem?) Plan, interessant. Jetzt geht es um das Comic, aber das Video weigert sich, zu laden. (.) Doch nicht, da bist du ja wieder. #00:29:26-8#
66	-	[Video Einleitung Reflexion 02] #00:29:50-8#
67	eigene Erfahrung, Sonstiges	B1: Ah ganz klassisch Lehrer, so wenn die Stunde vorbei ist im Lehrerzimmer so ein bisschen Smacktalken über die Schüler. Äh das kenne ich noch. Okay das Comic, Physik, Comics, das sehe ich immer so ein bisschen schwierig, weil das im Endeffekt nicht Comics sind, sondern einfach nur Erklärung mit Bild, deswegen bin ich da immer ein bisschen vorsichtig, was das äh was die Bezeichnung Comic angeht, allerdings ist (unv.) mein Maßstab von Comics auch X-Men, Marvel, DC, also vielleicht nicht der beste Maßstab.
68	Sonstiges	Video. #00:30:21-3#
69	-	[Video Szene Reflexion 02] // [Weiterhin Video Szene Reflexion 02] // [Weiterhin Video Szene Reflexion 02] // [Weiterhin Video Szene Reflexion 02] // [Weiterhin Video Szene Reflexion 02] #00:32:00-6#
70	Beschreibung	B1: (Ja?) siehst du, wie bereits gesagt, (.) (eine?) Erklärung mit Bild, mehr nicht. #00:30:34-4#

71	Sonstiges	B1: Welches Kind spricht so? (.) Antwort niemand. #00:31:07-5#
72	Beschreibung, eigene Erfahrung	B1: Kollektive Stille, sehr bekannt im Schulraum. #00:31:16-1#
73	Beschreibung, Überlegung zum Unterricht, Sonstiges	B1: Mhm wieder dieses klassische Missverständnis von Kräfte können nur (von?) bewegten Körpern ausgehen. (4s) Mhm. #00:31:42-9# B1: Okay, ganz interessant, (sehr?) interessant, scheint ja dieses oft bereits erwähnte äh Missverständnis (unv.) zu sein, dass Kraft nur ausgeübt werden kann, wenn es eine Bewegung gibt. (.) Äh ein paar scheinen die richtige Idee zu haben äh und zwar, dass die Gegenkraft vom Baum ausgeht. (.) Genauer (unv.) durch den wie heißt das auf Deutsch nochmal? Das englische Wort ist stem, (.) the stem of the apple, der (dieser?) kleine Holzstümmel da ganz oben auf dem Apfel, dass darüber die Verbindung zum Baum ist und somit hält dementsprechend die Kraft äh vom Baum auf den Apfel ausgewirkt wird, bis dann der Apfel äh den Reifeprozess beendet hat und sich dieser Stümmel trennt und das (des Baumes?) ab wann dann der Baum keine Kraft mehr auf den Apfel an sich ausübt und der Apfel somit zu Boden fällt. Wie in dem berühmt berüchtigten Beispiel von Sir Isaac Newton, was nebenbei gesagt historisch falsch ist. #00:33:01-0#
74	Vorlesen	B1: Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung abababababa.
75	Vorlesen	Gesprächsführung,
76	Vorlesen	Einsatz des Comics,
77	Vorlesen, Zustimmung AO	Schülerfehlvorstellungen. Ja
78	Sonstiges (3x), Zustimmung AO	und Einsatz des Comics, das ist mehr so ein generelles Problem, da würde ich jetzt ihn nicht an den Karren fahren, (weil?) das wäre ein bisschen zu unfreundlich. Andererseits andererseits ist Robert auch nur eine Simulation und es lohnt sich nicht wirklich, einer Simulation Empathie zu zeigen, anderer- andererseits (.) wer weiß, was dann für eine KI da hier dahinter sitzt und dann wirklich keine Empathie kriegt und bevor ich jetzt mit dieser Gedankenspirale komplett außer Kontrolle gehe, wie ich es oft genug tue, sage ich einfach Schülervorstellungen und bela- belasse es dabei. #00:33:40-7#
79	Sonstiges	B1: Natürlich. Natürlich hat er da bereits genügend Kritäh Kritik für gekriegt, das bedeutet dann weniger Arbeit für mich. #00:33:49-7#
80	Vorlesen	B1: Welche der Fehl- folgenden Schülervo- fehlvorstellungen sind während des Unterrichtsschnitts- ausschnitts sichtbar geworden? (.)
81	Vorlesen, Zustimmung AO + Begründung	Es sind keine Schülerfo- äh fehlvorstellungen sichtbar geworden. Ja, das kann nicht sein, ansonsten wären sie sich ja alle einig. (..)

82	Vorlesen	Schüler Schülerinnen denken, dass Kräfte nur dann auftreten, wenn sich zwei Körper berühren.
83	Vorlesen	Schüler Schülerinnen schreiben dem Apfel eine Kraft als Eigenschaft zu.
84	Vorlesen, Zustimmung AO	Schüler Schülerinnen denken, dass nur aktive, sich bewegende Körper eine Kraft ausüben können, passive leisten nur Widerstand. Mhm[Ja].
85	Vorlesen, Sonstiges*	Schüler Schülerinnen denken, dass einer Kraft immer die Beschleunigung zu- immer eine Beschleunigung zugrunde liegen muss und schließen daher eine Kraftwirkung bei ruhenden Gegenständen aus. Ist das nicht das selbe, wie das davor? Aber okay.
86	Vorlesen, Professionalisierung	Die Schüler Schülerinnen verwechseln Kraft und Gegenkraft mit den wirkenden Kräften, die einen Körper au- im Kräftegleichgewicht halten. Uh, da weiß ich jetzt gerade selber nicht, ob das generell betrachtet korrekt ist, da müsste ich nochmal nachgucken, werde ich aber höchstwahrscheinlich nicht machen, weil ich heute noch genügend zu tun habe. (.)
87	Vorlesen (3x), Sonstiges* (6x)	Die Schüler Schülerinnen sehen nicht, dass der Apfel über den Baum mit der Erde verbunden ist und der Baum durch das Festhalten an der Erde eine Gegenkraft zum am ziehenden am Baum ziehenden Apfel ausübt. (..) Nochmal. Die Schüler Schülerinnen sehen nicht, dass der Apfel über den Baum mit der Erde verbunden ist. Okay, wir haben eine Apfel, Baum, Verbindung zu der Erde (B1 gestikuliert), gut. Und der Baum durch das Festhalten, ja, an der Erde ein- nein falsch, hier unten festhalten. An der Erde eine Gegenkraft zum am Baum ziehenden Apfel ausübt. (..) Was? Ich verstehe den Satz immernoch nicht. (.) Der Baum durch das Festhalten an der Erde eine Gegenkraft zum am Baum ziehenden Apfel ausübt. Ist das so (unv.) so wie äh eine Hebelwirkung gemeint, oder (..) ich verstehe den Satz nicht. (...)
88	Sonstiges	Äh gucken wir nochmal die anderen an,
89	Vorlesen, Zustimmung AO	verwechseln Kraft und Gegenkraft mit den wirkenden Kräften, die einen Körper im Gleich- äh Kräftegleichgewicht halten. (.) Äh technisch gesehen auch. (..)
90	Vorlesen, Ablehnung AO	Äh (unv., liest vmtl. murmeld AO 'Es sind keine Schülerfehlvorstellungen sichbar geworden') nein.
91	Vorlesen, Ablehnung AO + Begründung	Schüler Schülerinnen denken, dass Kräfte nur dann auftreten, wenn sich zwei Körper (unv.) äh berühren, also quasi auf (.) ähm der Vorstellung hängen geblieben sind, dass die Erde und der Apfel keine Kräfte aufeinander ausüben würden, aber diese Vorstellung haben sie doch, also kann das nicht sein. (.)

92	Vorlesen, Ablehnung AO + Begründung, Sonstiges (2x)	Äh schreiben dem Apfel eine Kraft als Eigenschaft zu. (..) Nein, diesen Eindruck hatte ich nicht, die haben nicht davon gesprochen, dass der Apfel Kraft hat oder irgendwie sowas, also (.) vielleicht (unv.) bin ich ja auch einfach nur bescheuert und habe es nicht mitgekriegt, erneut, wäre nicht das erste Mal, dass das passiert ist. Belassen wir es einfach bei diesen drei. #00:36:48-9#
93	Vorlesen	B1: Welche der unten stehenden Hinweise zum Umgang mit Schülerfehlvorstellungen möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde gebengeben?
94	Vorlesen	Wählen Sie die Antwortoptionen aus, die Ihrer Einschätzung am nächsten kommen. (.)
95	Vorlesen, Überlegung zum Unterricht, Beschreibung, Bewertung, Sonstiges	Dein Umgang mit den Schülervorstellungen Fehlvorstellungen ist angemessen, das kannst du wieder so machen. (.) Hat er in dem Video überhaupt auf die reagiert? Er hat, glaube ich, nur gesagt, dass (..) dass sie äh dass da verschiedene Vorstellungen sind, aber er hat jetzt keine Arbeit- äh oder zumindest keine dokumentierte Arbeit geleistet, um diese Schülervorstellung irgendwie entgegenzuwirken, also meines Wissens nach hat er momentan absolut gar nicht gemacht und (.) das ist nie eine gute Idee. (.) Okay, vielleicht manchmal schon, aber dann gehen wir schon in die Philosohpie rein und das hier ist ja immerhin äh geht es jetzt um das Limesurvey diesen und nicht um meine philosophische Einstellung, was (das?) Tun gegen das nichts Tun (unv.) angeht.
96	Vorlesen, Überlegung zum Unterricht, Alternative + Begründung	Du solltest den Prozess der Kraftwirkung thematisieren, um die Schülerfehlvorstellungen aus dem Weg zu räumen. (..) Hatte er das nicht schon gesagt, dass äh dieser diese ganze Kraftwirkung bereits durchgemacht wurde? Eine Wiederholungsstunde (könnte?) er vielleicht einwenden, das ist so gut wie immer eine gute Idee. (.)
97	Vorlesen, Bewertung + Begründung, Zustimmung AO + Begründung	Du solltest den Unterschied zwischen Kraft und Energie thematisieren, um ver- zu verdeutlichen, warum die Schülervorstellungen falsch sind. (.) Ein bisschen harsch formuliert, weil einfach den zu Schülerinnen zu sagen, (unv.) falsch, das ist äh dumme Idee in (unv.) äh in meiner Erfahrung, aber Unterschied zwischen Kraft und Energie the- zu thematisieren, das ist etwas, wo selbst äh hochstudierte Studenten noch drüber stolpern, also ist das (.) äh nie etwas, was man auf die leichte Schulter nehmen sollte.

98	Vorlesen, Zustimmung AO + Begründung (2x)	Du solltest erklären Schrägstrich ver- äh deutlichen, dass Kräfte auch dann wirken können, wenn man nichts aktiv tut. Dazu kannst du da- das Rollbrettexperiment verwenden: Zwei Schüler Schülerinnen stehen auf je einem Rollbrett und halten je eine ein Ende einer Schnur. Egal ob beide Schüler Schülerinnen ziehen oder nur eine einer, treff- treffen sich die Rollbretter in der Mitte. Das ist eine gute Idee, generell ist das ein ziemlich cooles Experiment, das auch noch eine Menge Spaß macht, kann ich immer nur empfehlen. (.)
99	Vorlesen, Zustimmung AO + Begründung	Du solltest mit den Schüler Schülerinnen die verwendeten Begriffe und den Unterschied zwischen dem Kräftegleichgewicht und dem dritten Newtonschen Axiom klären, damit ihnen klar wird, warum ihre Vorstellungen äh falsch sind. Wie bereits in dem [Dozent]-Kurs von letzte Woche, nein sogar vorletzte Woche mittlerweise, durchgenommen, der Unterschied zwischen Kräftegleichgewicht und drittes Newtonsches Axiom ist ein generelles Problem für Physikstudenten, auch das sollte man noch einmal durchgehen.
100	Vorlesen, (Ablehnung AO) + Begründung (2x)	Du solltest die Schüler Schülerinnen das Wechselwirkungsprinzip ausführlich mündlich oder schriftlich formulieren lassen, damit du siehst, was die Schüler Schülerinnen verstanden haben beziehungsweise wo Fehlvorstellungen bestehen. Meiner Meinung nach konnte man das bereits sehr gut aus dem Quiz rauslesen, wo jetzt die Fehlvorstellungen sind und wo nicht, also das würde für die Schüler nur (.) mehr Arbeit bedeuten und dementsprechend äh die- sen Kampf Lehrer gegen Schüler noch ein bisschen verstärken, dann würde äh das Bild verfestigt, der Lehrer, das ist einfach nur jemand, der mir Aufgaben gibt, der mir Probleme gibt, die ich ohne ihn nicht hätte und äh das ist generell etwas, wo man gegen arbeiten sollte oder zumindest ich gegen arbeiten will.
101	Sonstiges	Weiter. #00:40:12-4#
102	Sonstiges	B1: (unv.) wie lang wird dieses Ding? (.)
103	Vorlesen	Bei der Besprechung des Comics hatte ich das Gefühl, dass die Schüler Schülerinnen einige Schülervorstellungen haben. Deswegen habe ich die Schüler Schülerinnen gebeten, Stellung zu einer Seite des Comics zu beziehen, also Argumente entweder für die Seite des Mädchens oder des Jungens zu sammeln. Diese Argumente haben wir dann besprochen und dann hat Anna hat fachlich korrekt erklärt, warum die Position des Mädchens die richtige ist. Anschließend hatten die Schüler Schülerinnen die Möglichkeit, offen gebliebene Fragen zum Comic zu stellen. Da die Schüler Schülerinnen hier keine Fragen hatten, habe ich den Unterricht fortgeführt und zum Thema Impuls übergeleitet.

104	eigene Erfahrung, Bewertung + Begründung (2x), Überlegung zum Unterricht	Ja, diese einfache Frage hat noch jemand Fragen hat in meiner Erfahrung noch nie geklappt, weil die Leute immer dann entweder so wenig wissen, dass sie noch nicht mal wissen, wo sie fragen sollen (.) oder (einfach?) schlichtweg zu schüchtern sind und sich selber nicht blamieren wollen, also (.) mhm (.) dumme Frage, ich weiß jetzt auch ni- keine bessere Möglichkeit, aber (.) äh in meiner Erfahrung hat die noch nie funktioniert. (.)
105	Sonstiges	Weiter. #00:41:17-3#
106	Sonstiges	B1: Ja, du wieder. #00:41:21-2#
107	-	[Unterbrechung aufgrund technischer Schwierigkeiten, B1 fällt für ca. 08:40 aus dem Meeting, kehrt um #00:50:03-9# ins Meeting zurück, neu öffnen des Tests bis zur aktuellen Video Einleitung Reflexion 03] #00:54:25-2#
108	-	[Video Einleitung Reflexion 03] #00:54:34-0#
109	Sonstiges	B1: Mhm lässt sich immer was, selbst (bei?) Perfektionismus lässt sich was verbessern. #00:54:39-4#
110	Sonstiges	B1: Es lädt, es lä- jetzt. #00:54:43-5#
111	-	[Video Szene Reflexion 03] // [Weiterhin Video Szene Reflexion 03] // [Weiterhin Video Szene Reflexion 03] // [Weiterhin Video Szene Reflexion 03] // [Weiterhin Video Szene Reflexion 03] // [Weiterhin Video Szene Reflexion 03] #00:56:30-0#
112	Sonstiges	B1: (unv.) #00:55:15-1#
113	Sonstiges	B1: Ja gut, was soll das heißen? #00:55:49-4#
114	Sonstiges	B1: Mehr bremsen, (was?) denn für eine Formulierung. #00:55:59-4#
115	Sonstiges	B1: Mhm mhm. #00:56:09-0#
116	Sonstiges	B1: Fußball, natürlich, natürlich ist es Fußball. #00:56:24-3#
117	Sonstiges	B1: Ich glaube, die erste Hiobsbotschaft ist gekommen, mein Grafiktreiber ist mal wieder abgestürzt. (.) Mal schauen, wie lange wir dieses Mal durchhalten. #00:56:38-6#
118	Sonstiges, Vorlesen, Sonstiges*	B1: Okay, Kritik, lass mal schnell machen, bevor die ganze Kiste hier mal wider absäuft. Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung, Lob oder- und/oder Kritik, erhalten, um etwas für sich beziehungsweise seinen Unterricht zu lernen? Oh kann hier auch Lob äh geben, das habe ich gar nicht gewusst.
119	Vorlesen, Bewertung	Naja (.) Einsatz des Beispiels Titanic, ja von mir aus, jedes andere Schiff hätte auch ge- geklappt, aber die Titanic ist schlichtweg das berühmteste, also ist es ganz gut.
120	Vorlesen, Alternative	Gesprächsführung selbe Kritik wie se- äh wie sonst auch, pack mal ein bisschen Elan in die Stimme min Jung.
121	Vorlesen	Fachliche Korrektheit,
122	Vorlesen	Fachsprache,
123	Vorlesen	Übergang der Unterrichtsphasen (.)

124	Zustimmung AO	ja, da gibt es wirklich was für- zur Gesprächführung zu kritisieren. #00:57:25-8#
125	Sonstiges	B1: Natürlich hat er bereits genügend äh Feedback gegeben. #00:57:30-8#
126	Vorlesen	B1: Welches Feeba- Feedback möchten Sie Robert zur Nutzung der Fachsprache geben?
127	Vorlesen, Sonstiges*, Zustimmung AO, Alternative	Es ist nicht gut, dass die Schüler Schülerinnen den Begriff Schwung nutzen, da die Bedeutug des Begriffs noch n- da ist ein Schreibfehler, das nicht ist falsch geschrieben nicht schreibe ich selber oft genug falsch, also (unv.) Fehler und so unklar ist, was genau die Schüler- innen beschreiben. Äh es ist tatsächlich schwierig, etwas zu benutzen, was noch nicht definiert wurde, was man vielleicht machen könnte, ist die Definition von Schwung da also so ein quasi Addendum oder (.) Exkurs, Exkurs war das Wort, da (in?) einzufügen.
128	Vorlesen, Zustimmung AO + Begründung	Es ist gut, dass die Schü- Schüler Schülerinnen zunächst ihre Alltagssprache verwenden können um zu erklären, warum es zur Kollision gekommen ist. So sinkt die Hemmschwelle für die Schüler Schülerinnen, ihre Ideen auszu äh zu äußern. Definitiv, d- das linguistische Level ist etwas, das viele Leute, besonders generell bei Wissenschaften abschreckt, äh irgendwie teilzunehmen und äh jegliche Möglichkeit, um um diese Hemmschwelle vorbei zu kommen, ist (eine?) sehr gute Möglichkeit.
129	Vorlesen, Sonstiges*, Zustimmung AO	Es ist gut, zunächst mit dem allba- alltagssprachlichen Begriff Schwung einzusteigen, da der Begriff Schwung ein tragfähiges Konzept für den Impulsbegriff aus dem Alltag der Schüler Schülerinnen ist. Das klingt so s- äh das klingt ziemlich ähnlich zu dem vorherigen, aber von mir aus.
130	Vorlesen, Zustimmung AO	Die Formulierung stärker bremsen ist nicht gut gewählt, da auch Körper mit großem Impuls, wie zum Beispiel die Titanic, s- stärker bremsen können könnten, um eine Kollision zu vermeiden. Ja, stärker bremsen, hatte ich ja (vorher?) schon gesagt, was für eine Formulierung. (..)
131	Vorlesen, Sonstiges, (Ablehnung AO) + Begründung	Es ist nicht gut, dass die Schüler Schülerinnen mit dem Begriff Schwung arbeiten. So (.) werden äh so werden sie mit diesem Begriff me- s- so werden sie sich, behalte mal deine Zunge unter Kontrolle, so werden sie sich diesen Begriff merken und zu- zukünftig nicht mit dem Impulsbegriff arbeiten. Ah ich schätze die Schüler doch schon clever genug ein, um den Unterschied zwischen A und B zu wissen. (.)

132	Vorlesen, Sonstiges (keine Hürde), Ablehnung AO, Alternative	Es ist nicht gut, dass du die Schüler Schülerinnen mit dem mit der Benutzung des Begriffs Schwung- die gesamte Seite scheint sich nur um dieses eine Wort zu drehen, ganz interessant. So förderst du ihre Fehlvorstellungen, zum Beispiel dass beide Begriffe gleichgesetzt werden können und dass Körper mehr Kraft haben, wenn sie mehr Schwung besitzen. Oh ne, dem kann man relativ einfach entgegenwirken, indem man einfach später äh ei- äh die Definition von Schwung mit hinten dran hängt, also das würde ich jetzt nicht kritikwü- würdig führen.
133	Vorlesen, Zustimmung AO + Begründung, Alternative	Es ist nicht gut, dass du so stockend redest und viele Füllwörter nutzt. So lenkst du die Schüler Schülerinnen vom Inhalt ab. Ja, das ist so ein generelles sprachliches Problem, das kommt aber mit der Übung, (.) also da würde ich mir- es ist kritikwürdig, ja, aber denoch sollte man sich da nicht (nachher?) zu viele Gedanken machen, das kommt schlichtweg mit der Übung. #01:00:23-2#
134	Vorlesen	B1: Welche der unten stehenden Hinweise zur Nutzung von- zu der Nutzung von Fachsprache möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben?
135	Vorlesen, Ablehnung AO, Alternative (2x) + Begründung (2x), Überlegung zum Unterricht	Es wäre Begriffe- besser, die Begriffe Masse und Impuls, gegebenenfalls mit der zugehörigen Formel, einzuführen, damit die Schüler Schülerinnen die korrekte Fachsprache lernen. Ich würde tatsächlich davon abraten und erst bei der alltäglichen Sprache bleiben, weil du- man hat gerade eben mit einem alltäglichen Beispiel angefangen. Alltäglich, so alltäglich, wie die Titanic nun mal ist, (.) äh und dann würde ich den Schwung von diese- von diesem Beispiel nehmen und dann in die physikalischen Sachen überleiten anstatt direkt bumm harte Grenze, hier ist phyhier ist Alltag, hier ist Physik, so einen- so einen sanften Übergang zu geben, dass- um besser zu zeigen, dass äh halt eben es gibt einen Übergang zwischen Alltag und Physik und es fließt ein bisschen ineinander über, anstatt einfach zu sagen bumm, hier ist die Grenze, da das ist Physik, das ist Alltag (..) (unv.) entgegensetzen. (.)

136	Vorlesen, Sonstiges, Sonstiges*, Ablehnung AO, Alternative + Begründung	Wäre besser, wenn du die Schüler Schülerinnen dazu aufforderst, mehr Fachbegriffe zu nutzen, zum Beispiel dem Massebegriff, damit sie sich so dem Impulsbegriff annähern können. (Äh?) selbe Kritik wie vorhin auch, ist wieder so, das sind wieder so zwei Punkte, die sehr ähnlich zueinander sind, (.) ne aber (.) es ist schwierig das zu sagen, weil klar muss man das alles äh kritisieren und sagen, benutz doch lieber die Fachwörter, weil Physik als Wissenschaft ist ein sehr fokussiertes und sehr korrektes Fach, aber wie bereits gesagt, bisschen später vielleicht. Gib den Schülern erstmal einen Moment Zeit, (.) um sich da ran zu tasten, wenn du jemanden in einen dunklen Raum packst, dann gibst du- dann willst du auch erstmal, dass die einen Moment haben, um sich zu orientieren und ihnen nicht gleich sofort einen äh Rubics Cube, wie auch immer die Dinger auf Deutsch heißen, in die Hand drücken und zu erwarten, dass sie den gleich lösen, also ja, aber später.
137	Vorlesen, Sonstiges*	Es wäre besser, den Impulsbegriff für den Begriff Schwung schon früher einzuführen, um zu verdeutlichen, um welches physikalische Konzept es geht. Erneut, warum sind diese Aussagen alle so ähnlich zueinander. (.)
138	Vorlesen, Zustimmung AO, Alternative	Du solltest gegen die Schülervorstellung vorgehen, dass Körper Schwung haben, vergleichbar mit der Schülervorstellung Kraft haben, damit diese sich nicht verfestigt. Da stimme ich zu. Äh Schwung haben als Eigenschaft, das ist tatsächlich etwas, das sollte man im Keim ersticken, so schnell, wie es geht. Äh gleiches gilt für Kraft haben. (.)
139	Vorlesen, Zustimmung AO, Bewertung	Die Verwendung der Begrifflichkeiten war angemessen, das kannst du so weiter durchführen. Ja, ja, ja, ja, also das war tatsächlich einer der besseren, wenn man mich fragen würde.
140	Vorlesen, Zustimmung AO + Begründung	Du solltest dir in der Vorbereitung genauer überlegen, was du den Schüler Schülerinnen erzählen möchtest, damit du flüssiger und motivierender vortragen kannst. Ja, Vorbereitung ist generell eine gute Idee, sollte man immer machen.
141	Vorlesen, Sonstiges*	Es wäre besser, die Schüler Schülerinnen zur Verwendung der Fachsprache aufzufordern, um zu vermeiden, dass die Konzepte Schwung und Impuls weiter- (.) das ist (.) vier Antworten gleiche Aussage, was soll das? (...)
142	Sonstiges	Weiter. #01:03:31-8#
143	Vorlesen, Sonstiges	B1: Reflexion null drei, wir haben erst die Hälfte vorbei? (.) Ich sitze hier seit über einer Stunde und ich habe heute noch was zu tun, meine Güte. (.)
144	Vorlesen	Mhm welches Feedback möchten Sie Robert zur Gestaltung des Phasenübergangs geben?
145	Vorlesen	Wähle Sie die Antwortoptionen aus, die Ihrer Einschätzung am nächsten kommen.

146	Vorlesen, Überlegung zum Unterricht	Der Phasenübergang ist dir nicht gut gelungen. (.) Da war ein Phasenübergang? (.) Okay.
147	Vorlesen	Nicht gut gelungen, da er zu abrupt war.
148	Vorlesen	Nicht gut gelungen, da er zu kurz war.
149	Vorlesen	Nicht gut gelungen, da die Inhalte zum Wechselwirkungsprinzip nicht ausreichend verstanden und gesichert wurden. (..)
150	Vorlesen, Zustimmung AO + Begründung	Es ist unproblema- problematisch, dass der Phasenübergang abrupt war, hm (bejahend)- da du mit dem Impulsbegriff wirklich ein komplett neues Thema beginnst. (.) Ja und ja. Es ist ein komplett anderes Thema, aber dennoch ein sehr rapider Wechsel. Schüler sind immerhin auch nur Menschen und du kannst nicht erwarten, dass sie wie ein Roboter einfach jetzt lade anderes Programm, lade anderes Programm, lade anderes Programm, ne, das kann man nicht erwarten.
151	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du so einen abrupten Übergang zwischen den Themen gemacht hast, da der rote Faden der Stunde so für die Schüler Schülerinnen nicht erkennbar ist. Ja. (.)
152	Vorlesen, Sonstiges*, Bewertung + Begründung	Es ist nicht gut, dass du sagst, das Thema ist komplett neu, da (es?) den Schüler Schülerinnen so das bisher Gelernte nicht (.) so das bisher- fehlt da ein Wort? Da den Schüler Schülerinnen so das bisher Gelernte nicht mehr relevant erscheint. Nein, ist doch ein richtiger Satz, ich bin einfach nur blöd. Äh (unv.) es ist tatsächlich generell eine doofe Idee, etwas komplett neu zu (sagen?), weil auf irgendeine Art und Weise ist es immer verbunden mit allem anderen, also (.) ja. (.)
153	Vorlesen, Zustimmung AO	Äh es ist nicht gut, dass du komplett neu sagst, da es einen fachlichen Bezug zwischen den Newtonschen Axiomen und dem Impulsbegriff gibt. Auch das. (...)
154	Sonstiges	Tschuldigung. #01:05:27-0#
155	Vorlesen	B1: Welche der unten stehenden Hinweise zur Gestaltung des Phasenübergangs möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben?
156	Vorlesen	Phasenübergang ist gut gestaltet, das kannst du wieder so machen.
157	Vorlesen	Es wäre besser, den Zusammenhang zwischen den Newtonschen Axiomen äh und dem (unv.) Impulsbegriff zu verdeutlichen, da die Schüler Schülerinnen (.) Zusammenhang nicht unbedingt selbst erkennen können. (.)

158	Vorlesen, Überlegung zum Unterricht, Zustimmung AO	Es wäre besser, den Einstieg in den Impulsbegriff für die nächste Stunde zu verschieben, um den Schüler Schülerinnen zunächst noch mehr Möglichkeiten zu geben, das Wechselwirkungsprinzip zu verstehen und anzuwenden. Okay, hier müssen wir kurz eine Annahme machen, ist es eine einzelne fünfundvierzig Minuten Physikstunde und keine eins Komma fünf Stunden Doppelphysikstunde? Dann stimme ich dem hier zu.
159	Vorlesen, Verweis auf frühere Stelle im Test, Alternative + Begründung, Sonstiges	Es wäre besser, vor dem Einstieg in das neue Thema auch die Tafel zu wischen, damit die Schüler Schülerinnen sich voll auf das neue Thema konzentrieren können und nicht noch über das alte Thema nachdenken. Hier sind wir in einem direkten Konflikt zu ein paar vorherigen Ideen und zwar, dass es alles sehr wohl miteinander verbunden ist und man könnte tatsächlich ähm die Notizen aus der vorherigen Stunde oder des vorherigen Themas dazu benutzen, um eine- um diese Verbindung zwischen den beiden Themen zu verfestigen oder generell darzustellen, weil in Physik, wenn du es richtig anstellt, ist alles mit allem verbunden. (.) Also- (.)
160	Vorlesen, Zustimmung AO	Es wäre besser, einen glatteren Übergang zu gestaltenlove it, love it jetzt schon, zum Beispiel durch eine kurze Verknüpfung,wie 'Wir finden jetzt etwas neues mit dem gelernten Wissen heraus', damit der rote Faden der Stunde für die Schüler Schülerinnen nachvollziehbar ist. Wunderbar, perfekt. (.)
161	Vorlesen, Sonstiges*	Es wäre besser, den Übergang durch das Zeigen eines Experiments zu gestalten, um den Bruch zwischen den Themen (.) für die Schüler Schülerinnen weniger hart zu gestalten. Den Satz muss ich nochmal lesen, weil ich habe jetzt schon den Anfang wieder vergessen.
162	Vorlesen, Zustimmung AO	Es wäre besser, den Übergang durch das Zeigen eines Experiments zu gestalten, um den Bruch zwischen den Themen für die Schüler Schülerinnen weniger hart zu gestalten. (..) Ja, (ich?) (ja?) das kann man auch so nennen (.) und verwenden. #01:07:24-9#
163	Sonstiges	B1: (Na?) Video wieder. (...) Hello again. #01:07:33-3#
164	-	[Einleitung Reflexion 04] #01:07:43-4#
165	Überlegung zum Unterricht	B1: Okay, kinetische Energie ist bereits bekannt, merke. #01:07:46-2#
166	Sonstiges	B1: Es lädt. (4s)
167	Sonstiges	Tschuldigung, ich weiß, es kommt- gehört nicht zum Test, aber ich komme nicht drum herum, immer wieder rüber auf ihr Bild zu sehen und ich sehe da so ein besorgtes Gesicht, als ob ich irgendwas komplett falsches gesagt hätte. (.) Sie sitzen da so. #01:08:08-6#
168	-	I: Das ist mein normales Gesicht. #01:08:11-4#
169	Sonstiges	B1: Okay. #01:08:13-0#

170	-	I: Ist wirklich alles in Ordnung, also es gibt auch quasi keine falschen Aussagen jetzt gerade. (..) // (unv.) #01:08:22-5#
171	Sonstiges	B1: (unv.) wichtigste Frage, die je gestellt wurde. Schon gut, die Pandemie hat aus mir einen Zyniker gemacht. #01:08:27-2#
172	-	I: Alles gut. #01:08:28-2#
173	Sonstiges	B1: Ja, okay. Weiter im Text äh oder im Video besser gesagt. #01:08:32-8#
174	-	[Video Szene Reflexion 04] // [Weiterhin Video Szene Reflexion 04] // [Weiterhin Video Szene Reflexion 04] // [Weiterhin Video Szene Reflexion 04] // [Weiterhin Video Szene Reflexion 04] #01:09:35-1# // [Weiterhin Video Szene Reflexion 04] // [Weiterhin Video Szene Reflexion 04] // [Weiterhin Video Szene Reflexion 04] // [Weiterhin Video Szene Reflexion 04] #01:10:27-3#
175	Sonstiges	B1: Mhm. #01:08:39-6#
176	Sonstiges	B1: Es gibt so viele schmutzige Witze, die ich jetzt reißen könnte. #01:08:45-5#
177	Sonstiges	B1: Die haben ja richtige (unv.) wie bei der (Studios?) oder was auch immer ist da kurz vorm (unten?) vor dem linken? #01:09:06-9#
178	Sonstiges	B1: Ich mag seine Stimme. #01:09:21-5#
179	Sonstiges	B1: Mhm. #01:09:53-0#
180	Bewertung	B1: Gute Überleitung. #01:10:00-4#
181	Professionalisierung	B1: Mhm (.) habe ich bis heute nicht verstanden. #01:10:06-7#
182	Sonstiges	B1: Mhm. #01:10:22-5#
183	eigene Erfahrung, Sonstiges	B1: Äh abschreiben, da werden Erinnerungen wach und jede einzelne davon hasse ich. (..) Weiter. #01:10:33-1#
184	Vorlesen, Sonstiges	B1: Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung, Lob und/oder Kritik, erhalten, um etwas für sich beziehungsweise seinen Unterricht zu lernen? (..) Äh als Zeitbegrenzung werde ich definitiv nur einen davon wählen. (.)
185	Vorlesen	Äh eingesetzte Beispiele,
186	Vorlesen	Fachsprache,
187	Vorlesen	Gesprächsführung,
188	Vorlesen	Wahl des Experiments,
189	Vorlesen	Ex- Ergebnissicherung.
190	Überlegung zum Unterricht	Das Beispiel, ja was für ein Beispiel hat er denn genommen? Was war das nochmal. (..)
191	Vorlesen, Zustimmung AO	Okay, Fachsprache, Gesprächsführung, Wahl des Experimentes, Ergebnissicherung, (..) ich glaube, ich gehe wieder auf Gesprächsführung. #01:11:13-2#

192	Sonstiges*	B1: Jedes einzelne Mal. (...) Wofür ist diese eine diese eine Seite überhaupt gut? Robert hat schon genügend äh Rückmeldung über (unv.) wofür existiert das? Ist das nur eine Ablenkung? Damit die Immersion nicht kaputt geht? Was soll das? #01:11:33-4#
193	Vorlesen	B1: Okay, welches Feedback möchten Sie Robert zur Wahl des Experiments geben?
194	Vorlesen, Sonstiges* (2x)	Das Experiment ist gut gewählt, da die beiden Schülerhabe ich mich verklickt und ausversehen, egal machen einfach weiter.
195	Vorlesen, Zustimmung AO + Begründung	Das Experiment ist gut gewählt, da die beiden Schüler die das Experiment durchführen, den Impuls am eigenen Körper erfahren können. Ja, das kann man schon sagen, äh eigene Körpererfahrung ist immer gut.
196	Vorlesen, Zustimmung AO	Das Experiment ist gut gewählt, da es den (.) den Einfluss von Masse und Geschwindigkeit auf den Impuls gut darstellt. Auch das kann ich gut finden.
197	Vorlesen, Alternative, Ablehnung AO + Begründung	Es ist nicht gut, da nur über die Masse und die Geschwindigkeit gesprochen wurde, da unklar bleibt, ob der Impuls noch von anderen Größen, wie zum Beispiel der Form abhängig (.) abhängen kann. Ja dafür könnte man jetzt einfach das Experiment erweitern, aber das macht das bereits existierende Experiment jetzt nicht kaputt oder falsch oder andersweitig doof. Also nein.
198	Vorlesen, Zustimmung AO, Überlegung zum Unterricht, Sonstiges (2x)	Das Experiment ist nicht gut gewählt, da es den Zusammenhang beziehungsweise die Abgrenzung des Impulses zur kinetischen Energie nicht veranschaulicht. (..) Ja, das könnte man schon kritisieren, aber mir fällt jetzt auch kein besseres Experiment ein, wo man den Unterschied zwischen Impuls und kinetischer Energie darstellen könnte, (.) Ah schwierig, schwierig, schwierig, ich will jetzt nicht der Typ sein, der Kritik gibt, aber es nicht se- keine bessere Lösung hat. (..) Allerdings ist es auch definitiv ein gerechtfertigtes Kritikpunkt äh fuck it, lassen wir einfach mal das Kreuz hier. (.)
199	Vorlesen, Alternative	Es ist gut, dass du die beiden Schüler in das Experiment einbezogen hast. So motivierst und aktivierst du sie. Ja, noch besser wäre es vielleicht noch ein paar mehr Schüler reinzube- äh einzubeziehen, so drei oder vier oder fünf oder sechs oder wie auch immer viele da sind.
200	Vorlesen, Ablehnung AO + Begründung	Es ist gut, dass du vorgibst, wer nach vorne kommen soll. So vermeidest zu lange Organisations-, Wartezeiten nach der Frage 'Wer will denn'. (Oh?) diese Frage wird in meiner Erfahrung tatsächlich ziemlich schnell beantwortet, also das ist jetzt eher weniger kritisch. (.)

201	Vorlesen, Zustimmung AO + Begründung, Ablehnung AO + Begründung, Sonstiges	Es ist gut, dass du nur zwei Schüler das Experiment durchführen lässt und nicht alle Schüler Schülerinnen, da du so Zeit sparst und das Experiment so einfach ist, dass es auch von den sitzenden Schüler Schülerinnen gut nachvollzogen werden kann. (.) Ah jein, jein, jein, jein ähm (.) ja, organisatorisch spart es da eine Menge Zeit an ein, aber da ist halt meine Frage, um welchen Preis, weil ähm es ist für die sitzenden Schüler tatsächlich nicht möglich, diesen diese Impulserfahrung dann zu machen, weil sie schlichtweg diese Impulserfahrung nicht haben, sondern das können nur diese beiden wie (heißen?)- ich nenne sie jetzt einfach mal privilegierten Schüler (.) machen, die du da äh die man dann zu vorne geholt hat, also ein klares jein auf diese äh Kritikmöglichkeit, da ich allerdings nur ja oder nein auswählen kann, bleibe ich erstmal bei nein. #01:14:16-9#
202	Sonstiges	B1: Siebenundfünfzig Prozent erst, ich habe noch was zu tun heute. #01:14:24-9#
203	Vorlesen	B1: Welche von den unten stehenden Hinweise des zur Wahl des Experiments möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben?
204	Vorlesen	Experiment ist gut gewählt, du kannst es erneut einsetzen.
205	Vorlesen, Ablehnung AO	Es wäre besser, statt des Experiments ein Video zu zeigen- Nope, da breche ich sofort ab, das ist eine doofe Idee.
206	Vorlesen, Zustimmung AO, Sonstiges	Es wäre besser, ein anderes Experiment ähm- durchzuführen, das für alle Schüler Schülerinnen anschaulich ist, zum Beispiel eine Person, die auf einem Skateboard steht und einen Ball wegwirft, oder das Werfen beider Bälle in den Sand zur Betrachtung der Krater. (..) Mhm das ist eine interessante Idee, das könnte tatsächlich klappen, ja. (..) Oh ich kann ja beides machen.
207	Vorlesen	Es wäre besser, ein Stoßexperiment durchzuführen, da- da du so gut auf die Impulserhaltung überleiten kannst.
208	Vorlesen, Ablehnung AO + Begründung	Es wäre besser, ein Experiment zu wöhlen- wählen, zu dem eine Messreihe aufgenommen werden kann, damit die Formel aus diesen objektiven Messungen abgeleitet werden kann. Ich würde vorsichtig sein, ein neues, sei es nun komplett oder relativ neues Thema mit bereits solcher fachlichen Korrektheit zu beginnen, also nein.
209	Vorlesen, Zustimmung AO	Es wäre besser, wenn du das Experiment von allen Schüler Schülerinnen durchführen ließest- Jo, absolut. (.)

210	Vorlesen, (Ablehnung AO) + Begründung (2x)	Du solltest darauf achten, dass (du?) die Schüler vorne so stehen, dass alle sitzenden Schüler Schülerinnen erkennen sollten, wie leicht oder schwer es ist, den Ball zu fangen. (Ha?) das ist schlichtweg unmöglich, weil wenn du da einen Sportler und einen (..) nicht Sportler stehen hast, dann wird eine andere- anders reagieren als der andere, also das ist schlichtweg mög- unmöglich, das k-korrekt darzustellen, sodass es alle problem(los?) verstehen. #01:16:02-9#
211	Sonstiges	B1: (Oh je?).
212	Vorlesen	Welches Feedback möchten Sie Robert zur Ergebnissicherung geben?
213	Vorlesen, Zustimmung AO, Ablehnung AO + Begründung	Es ist nicht gut, dass du den Schüler Schülerinnen einen fertigen Merksatz vorgibst, da das reine Lesen und Abschreiben nicht verständnisfördernd ist. (.) Äh jein wieder, ja ich stimme dem zu, nein, weil (.) ähm (unv.) man müsste sonst die einzelnen Sätze korrigieren und das würde wieder Zeit fressen und so viel hat man ja bekanntlich als Lehrer nicht, also wieder klares Jein. (.)
214	Vorlesen, Ablehnung AO, Zustimmung AO	Äh es ist gut, dass die Schüler Schülerinnen den Merksatz ins Heft übertragen, damit alle Schüler Schülerinnen den Merksatz als Sicherung notiert haben. (..) Meine Erfahrung sagt mir nein, aber mein Wissen sagt mir ja. Ich stimme dem zu, aber ich bin nicht fröhlich darüber. (.)
215	Vorlesen, Überlegung zum Unterricht, Bewertung	Es ist nicht gut, dass du die Ergebnisse des Experiments nicht sicherst, so ist unklar, was die Schüler Schüler aus dem Experiment mitnehmen sollen. Was, hört die Unterrichtseinheit einfach danach auf? Einfach so p ist gleich m mal v, so das war es. (.) Was ist denn das für (eine?) bescheuerte Idee. (...)
216	Vorlesen, Zustimmung AO + Begründung	Der Merksatz ist gut gelungen, da die Je-umso-Sätze die Zusammenhänge von Masse, Geschwindigkeit und Impuls nachvollziehbar darstellen. Ja, je- Je-umso oder Je-desto-Sätze schei- äh sind generell sehr einfach zu merken, also ja. (.)
217	Vorlesen, Zustimmung AO, Ablehnung AO + Begründung, Alternative	Es ist nicht gut, dass du den Im- Impuls als skalare Größe einführst. So grenzt du den Impuls nicht ausreichend von der kinetischen Energie ab. (.) Äh jein wieder, generell betrachtet ja, stimmt diese Aussage schon, aber nein, weil (.) d- das beste, was man machen kann, um Schüler abzuschrecken, ist sie mit ü- Informationen zu überhäufen und gleich alles sofort zu geben. Also wieder klarer Fall von Jein, zwar richtig, aber später lieber. (.)

218	Vorlesen, Zustimmung AO, Ablehnung AO + Begründung, Alternative	Es ist gut, dass du den Begriff der Energie nicht erwähnst. So vermeidest du Verwirrung bei den Schüler Schülerinnen, die die verschiedenen Fachbegriffe auf Grundlage ihrer Erfahrung nicht trennen können. (..) Oh wieder klarer Fall von Jein, ja, diese Aussage stimmt, aber nein, irgendwann muss ja dieses Thema angesprochen werden, weil irgendwann wird ein Schüler mal darüber stolpern und eine Antwort verlangen, also muss es geklärt werden. Erneut nicht sofort, aber es muss geklärt werden. (..)
219	Vorlesen, Sonstiges*	In deiner Sicherung fehlt eine Abgrenzung des Impulsbegriffs von der kinetischen Energie und der Kraft, se- siehe Kritik vom Punkt dadrüber, das ist wieder so ein klarer Fall von zwei Aussagen gleiche Aussage. (..)
220	Sonstiges	Weiter. #01:18:39-8#
221	Vorlesen, Sonstiges	B1: We- äh welche der unten stehenden Hinweise zur Ergebnissicherung (unv.) oh da ist ja noch mehr geladen.
222	Vorlesen	Die Ergebnisse des Experiments hast du gut gesichert, da du so wei- das kannst du wieder so machen. (.)
223	Vorlesen, Zustimmung AO + Begründung	Es wäre besser, die Beobachtungen des Experiments schriftlich an der Tafel zu sammeln, (.) damit du im Verlauf der Stunde darauf zugreifen zurückgreifen (.) kannst. Ja, ja, solche Sachen zu dokumentieren ist immer eine gute Idee. (..)
224	Vorlesen, Zustimmung AO, Alternative + Begründung	Es wäre besser, die Schüler Schülerinnen in die Sicherung mit einzubeziehen, zum Beispiel einen Merksatz formulieren zu lassen, damit sie stärker aktiviert und involviert werden. Ja, das wäre dann auch (eine?) super Lösung für das Problem von vorhin, wo der Merksatz vorge- äh nicht vorgegeben sein sollte, aber dennoch korrekt sein sollte, das wäre ein guter Kompromiss zwischen den beiden Sachen, kann ich nur empfehlen. (.)
225	Vorlesen, Ablehnung AO + Begründung	Es wäre besser, zus- zusätzlich noch- zum Merksatz noch ein Beispiel zu zeigen, um auch die Formel anschaulich darzustellen. Ich meine, das Experiment kann gesehen werden als bereits ein Beispiel, (.) da scheiden sich jetzt die Geister, ich sage einfach mal, Experiment reicht. (.)
226	Vorlesen, Zustimmung AO + Begründung	Es wäre besser, in der Sicherung auch auf das Experiment einzugehen, zum Beispiel in- indem du auf der Folie zum Merksatz auch auf das Experiment eingehst oder Skizzen ergänzt, um die Inhalte stärker zu vernetzen. Vernetzen ist immer eine gute Idee, da bin ich mit dabei. (.)
227	Vorlesen, Überlegung zum Unterricht, Professionalisierung, Zustimmung AO, Ablehnung AO	Es wäre besser, den Impuls als vektorielle Größe einzuführen, um ihn von der kinetischen Energie abzugrenzen. (...) Frage, kennen die Schüler bereits den Begriff Vektor und wissen, was (es?) ist, weil ich erinnere mich beim besten Willen nicht mehr daran, wann ich den Begriff Vektor kennengelernt habe. Angenommen, die wissen bereits, was ein Vektor ist, dann ja. Wenn nein, dann nein. (..)

228	Vorlesen, Verweis auf frühere Stelle im Test, Zustimmung AO	Es wäre besser, auch auf die kinetische Energie der Bälle einzugehen, um an das Vorwissen der Schüler Schülerinnen anzuknüpfen. (.) Wieder was ich vorhin gesagt habe. Jein. Dieses Mal sage ich einfach ja. #01:20:36-4#
229	Vorlesen, Sonstiges	B1: Nachdem wir das Experiment mit den Tennisbällen abgeschlossen und die Schüler Schülerinnen den Merksatz zum Impuls abgeschrieben haben, habe ich zu einem Demonstrationsexperiment übergeleitet, an dem ich die Impulserhaltung veranschaulichen möchte. Mit dem Demonstrationsexperiment möchte ich einen Auffahrunf-Auffahrunfall simulieren. (.) Den Schüler Schülerinnen habe ich dann erklärt, wie der Autounfall im Experiment nachgestellt wird und wie die Messwertaufnahme funktioniert. Den experimentellen Aufbau und eine Messwerttabelle habe ich (.) vor der Stunde vorbereitet, hier habe ich dir zwei Fotos vom Aufbau von der Ta- und von der Tabelle mitgebracht. (.) Jetzt beginnt das große Wartespiel bis das Bild lädt. (5s)
230	Sonstiges	Mal kurz Rechenaufgabe, wir sind bei sechzig Prozent und etwa (...) eineinhalb Stunden Arbeitszeit. Schätzungsweise eine Stunde vierze- fünfzehn noch oben drauf, dann haben wir es (.) viel zu spät. (.) Meine Güte. (...)
231	Beschreibung (3x), Überlegung zum Unterricht, Bewertung, Alternative + Begründung	Mhm (Gleitbahn?) mit einem Kompressor, (.) äh mit Lichtschranken zur Messung, das zeigt schonmal, dass es ziemlich genau ist. (.) Ich weiß aber nicht, was auch nur (irgend?) ach das da könnte das Display sein (.) äh viele von den Kabeln sehen ziemlich nutzlos aus, ja, die liegen einfach nur rum als Dekoration (.) ja, sieht ganz gut aus, sieht gut aus. (..) Vielleicht wäre es noch besser, äh wenn man (noch?) auf diesen äh Schiebedingern, fällt der Name gerade nicht ein, noch irgendwie so ein Gehäuse droben drüber bauen könnte, dass es tatsächlich aussieht, wie ein Auto oder irgendwas, was die äh Schüler aus dem Alltag kennen, äh was sich bewegt und was (.) buff machen kann. (...) Tabelle (...) ziemlich standard, ziemlich einfach, ja, warum nicht. #01:22:50-5#
232	Sonstiges	B1: Hello Again. #01:22:55-1#
233	-	[Video Einleitung Reflexion 05] #01:23:00-7#
234	Überlegung zum Unterricht	B1: Das ist eins der größten Probleme im Unterricht generell, egal, welches Fach. #01:23:06-5#
235	-	[Video Szene Reflexion 05] // [Weiterhin Video Szene Reflexion 05] // [Weiterhin Video Szene Reflexion 05] // [Weiterhin Video Szene Reflexion 05] // [Weiterhin Video Szene Reflexion 05] // [Weiterhin Video Szene Reflexion 05] // [Weiterhin Video Szene Reflexion 05] #01:25:36-5#
236	Bewertung	B1: Uh, mit so einer Aussage wäre ich vorsichtig. #01:23:50-1#

237	Überlegung zum Unterricht	B1: Ach ja, eins der größten Mysterien der Physik, warum die S- Ein SI-Einheit für Gewicht Kilogramm ist. #01:24:14-8#
238	Sonstiges	B1: Mhm. #01:24:26-3#
239	Sonstiges	B1: (Schnarcht) (nö, nö?) ist was passiert? #01:24:41-4#
240	Beschreibung	B1: Sie steht da auch einfach so, wie bestellt und nicht abgeholt. (4s)
241	Sonstiges	Mhm (...) äh Grundkurs für Rhetoriker und (im Unterrichten?) (unv.) äh wenn du mit einer Masse redest, immer schön mit dem Rücken zur Masse gucken. #01:25:05-2#
242	Alternative	B1: Junge, pass mal ein bisschen auf, was aus deinem Mund kommt. #01:25:33-9#
243	Beschreibung	B1: Okay, bisschen abrupter Wechsel. #01:25:43-2#
244	Vorlesen+	B1: Äh welche Kritiken.
245	Vorlesen	Umgang mit den Schüler Schülerinnen, (..)
246	Vorlesen	Durchführung, (..)
247	Vorlesen	Messunsicherheiten,
248	Vorlesen	Wahl des Experimentes,
249	Vorlesen	Aktivierung der Schüler Schülerinnen,
250	Vorlesen	eingesetzte Beispiele. //
251	Sonstiges	(Oh?) ich war eigen- #01:26:04-2#
252	-	I: Ich hab- #01:26:05-8#
253	Sonstiges	B1: Ja #01:26:05-0#
254	-	I: Ich hake einmal ganz kurz ein, du bist du oben durchgegangen, kannst du vielleicht noch einmal ganz kurz sagen, was du überlegt hast? #01:26:11-7#
255	Sonstiges	B1: Zu was jetzt? #01:26:13-9#
256	-	I: Du hattest die ähm die Antwortoptionen hier einfach nicht mehr äh also nicht mehr laut ausgesprochen, weder vorgelesen, noch, was du gedacht hast. #01:26:21-9#
257	Sonstiges	B1: Oh, ich habe das Muster mittlerweile erkannt. Nach dem Video das erst was kommt, das (ist?) die Kriterien, dann äh kommt dieser eine Disclaimer, da hat er schon genügend von gekriegt und dann äh kommt da dieses was genau willst du denn kritisieren, was willst du für die Hinweise äh geben, also dieses Muster setzt sich durch das gesamte Ding, ich hab mittlerweile den Punkt (erkannt?), wo ich das Muster erkannt habe und nicht mehr alles vorlesen muss. (6s) Oder war das nicht, was du meinst? #01:26:52-2#
258	-	I: Ne, ist okay, also das war quasi der Reminder an das laute Denken, ist alles gut. #01:26:57-6#
259	Sonstiges	B1: Ah, ja, äh ich habe da nicht mehr dran gedacht. Mein Hirn ist einfach so Video, okay, erste Folie Kategorie (4s) ja. Sonst noch was? #01:27:09-8#
260	-	I: Ne, war (unv.) ups, sorry, war nur als Reminder, dass äh dass du das weiter machst, sonst alles gut. #01:27:16-0#

261	Vorlesen	Okay. Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung, Lob und/oder Kritik, erhalten, um etwas für sich beziehungsweise sei- seinen Unterricht zu lernen?
262	Vorlesen	Der Umgang mit den Schüler Schülerinnen,
263	Vorlesen	Notation der Messwerte,
264	Vorlesen	Durchführung und Auswertung des Experiments durch die Lehrkraft,
265	Vorlesen	Messunsicherheiten,
266	Vorlesen	Wahl des Experimentes,
267	Vorlesen	experimenteller Aufbau,
268	Vorlesen	Aktivierung der Schüler Schülerinnen,
269	Vorlesen	Eingesetzte Beispiele. (.)
270	Zustimmung AO (2x)	Ich nehme jetzt mal diese beiden,
271	Sonstiges	eigentlich würde ich gerne no- über noch mehr reden, aber mir geht die Zeit (jetzt?) so langsam flöten. #01:27:44-4#
272	Sonstiges	B1: Der Disclaimer, dass er da schon von genügend von gekriegt hat. #01:27:48-8#
273	Vorlesen	B1: Und welches Feedback möchten Sie Robert zu seinem Umgang mit den Schüler Schülerinnen geben?
274	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du sagst, dass Mädchen eine schöne Schrift haben und Jungen besser Werte ablesen können, da du die Mädchen auf so auf eine schöne Handschrift reduzierst. Jo.
275	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du die Mädchen nur Fleißarbeiten im Unterricht einbeziehst, da du so Stereotype förderst. Angenommen, er macht das sonst im Unterricht auch so, ja. (.)
276	VorlesenÜberlegung zum Unterricht, Ablehnung AO + Begründung	Es ist gut, dass du die Mädchen äh für ihr Handeln im Unterricht lobst, um ihr Selbstkonzept im Physikunterricht zu stärken. (Uh?) ich äh ich weiß nicht, ob an die Tafel schreiben als Handeln im Unterricht äh zählt, ich sage jetzt einfach mal nein, weil das kann jeder machen, selbst jemand, der noch nie in dem Unterricht war, äh ja.

277	Vorlesen (2x), Sonstiges (2x), Ablehnung AO, Sonstiges*	Es ist nicht gut, dass du das Lob (unv.) für die Handschrift von Mädchen mit einer inhaltlichen Kritik, vergessene Einheiten, kombinierst, da die Ursache hier nicht über das Geschlecht- (.) Was? Es ist nicht gut, dass du das Lob für die Handschrift von Mädchen mit einer inhaltlichen Kritik kombinierst, da du die Ursache hier da die Ursache hier nicht über das Geschlecht ist. (5s) Was? (.) Okay, angenommen (...) ich interpretiere das jetzt mal so, der Satz will mir sagen, dass es doof ist, dass er erst (.) etwas lobt, dass für das Geschlecht bezogen ist und dann für etwas kritisiert, was nicht auf das Geschlecht bezogen ist und somit die Schülerin verwirrt, ob jetzt ihre Handlungen auf ihr Geschlecht bezogen sind oder nicht. Angenommen, es funktioniert so, (.) sage ich nein. Was ist das denn bitte schön für eine Aussage. (.)
278	Vorlesen, Überlegung zum Unterricht, Ablehnung AO + Begründung	Es ist nicht gut, dass du die Schülerin schon vor dem Notieren des zweiten Messwertes erneut auf die Ein- heit hinweist, da du so das Bild festigst, Mädchen wären in Physik nicht kompe- Vor dem Notieren des zweiten Messwertes? Sie hatte doch nur eins notiert. (4s) Also da sage ich nein, schlichtweg, weil es nicht passiert ist. (.)
279	Vorlesen, Ablehnung AO + Begründung	Es ist nicht gut, dass du der Schülerin an der Tafel direkt sagst, welche Einheit fehlt, da du ihr so unterstellst, sie wüsste das nicht. (..) Das war eine schlichtweg einfach nur eine Kritik, dass die falsche Einheit benutzt wurde. Erneut, ich verstehe nicht, wa- was es hier zu kritisieren gibt. (.)
280	Vorlesen, Verweis auf frühere Stelle im Test	Es ist nicht gut, dass du der Schülerin an der Tafel direkt sagst, welche Einheit fehlt, das- ah das hatten wir gerade.
281	Vorlesen, Ablehnung AO + Begründung	Äh es ist in Ordnung dass du sagst, dass Paula aufgrund ihrer schönen Handschrift ihre die Werte notieren soll, da es wichtig ist, dass alle die Werte an der Tafel gut lesen können. (...) Also wenn das ein Kriterium ist, dann darf ich kein Lehrer (sein?), weil ich habe eine Sauklaue wie sonst was. (...)
282	Sonstiges	Weiter. #01:30:41-8#
283	Sonstiges	B1: Jetzt kommen wir zu den Hinweisen.
284	Vorlesen	Welche der unten stehenden Hinweise zum Umgang mit den Schüler Schülerinnen möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben? (.)
285	Vorlesen, Ablehnung AO	Dein Umgang mit den Mädchen und Jungen ist in Ordnung so, da (äh) Paula wirklich eine gut lesbare Hand- Handschrift hat. (Äh?) nö. (.)

286	Vorlesen, Zustimmung AO + Begründung	Du solltest eine Geschichte um das Experiment herum konstruieren, zum Beispiel in Form eines Videos zeigen, um die Mädchen zu motivieren. (..) Mhm (.) ich meine, generell ein bisschen mehr Kontext zu geben ist eine gute Idee, aber (..) ich würde das eher machen, um generell was zu geben und nicht einfach nur, um die Mädchen zu (..) bezirzen, sagen wir es so. (.)
287	Vorlesen, Zustimmung AO	Es wäre besser, die Schüler Schülerinnen zu fragen, wer freiwillig schreiben beziehungsweise die Wese- Werte ablesen möchte, um die Schüler Schülerinnen nach- nicht nach Geschlecht eine Aufgabe zuzuweisen. Ja, das ist gut, die Idee. (.)
288	Vorlesen, Sonstiges*, Zustimmung AO + Begründung	Es wäre besser, die Aufgaben Aufschreiben, Werte Abzulesen durch zu tauschen, (...) ach zu rotieren, damit alle Schüler Schülerinnen eine Aufgabe wahrnehmen. (..) Ah ja das könnte klappen, könnte allerdings auch Zeit äh klauen, ich sage einfach mal ja. (.)
289	Vorlesen, Zustimmung AO	Du solltest darauf achten, dass- die Mädchen nicht für ihre Handschrift, sondern explizit für ihre physikalische Kompetenz zu loben, da du sonst das Bild vermittelst, Physik wäre eine Jungsfach. Ja. (.)
290	Vorlesen, Zustimmung AO + Begründung	Du sollst auf geschlechtsspezifische Zuschreibungen verzichten, da du so ein unangemessenes Verhältnis- Ja, einfach Geschlecht aus der Sprache rauszuhauen ist generell eine geile Idee, damit kannst du nichts falsch machen.
291	Vorlesen, Verweis auf frühere Stelle im Test	Du solltest grundlegend an deiner Einstellung gegenüber Mädchen und Jungen arbeiten, um dein stereotypisches Verhalten zu ändern. Ich finde die anderen Krit- äh Kritikpunkte fassen das sehr gut (unv.) zusammen bereits. #01:32:22-2#
292	Vorlesen, Verweis auf frühere Stelle im Test, Sonstiges	B1: Äh Feedback zu Robert zu seinem Umgang mit Messunsicherheiten geben. Ach ja richtig, ich habe ja zwei Kategorien ausgewählt. (.) Mist. (.)
293	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du sagst, du könntest exakt oder genau messen, da du so die Schülervorstellung bef- beförderst, man könnte großen- Größen ohne Messunsicherheiten bestimmen. Ja. (.)
294	Vorlesen, Zustimmung AO + Begründung, Überlegung zum Unterricht, Alternative, Überlegung zum Unterricht	Du gehst nicht auf die Messunsicherheiten ein. Das ist nicht gut, da diese eine wichtige Rolle in der Physik spielen. (.) Jein, technisch gesehen ja, Messunsicherheiten sind wichtig, aber Problem ist halt eben wie früh oder wie spät setzt man das ein. Ich würde damit erstmal noch ein bisschen warten, bis es dann tatsächliches Unterrichtsthema ist. Es sei denn, es war bereits ein Unterrichtsthema, dann äh ja sollte man da schon drauf achten. (.)

295	Vorlesen, Zustimmung AO + Begründung, Alternative	Es ist nicht gut, dass du nur einen der beiden Wagen wiegst, da die Schüler Schülerinnen nicht so sicher sein können, dass die Wagen wirklich gleich schwer sind. Nach dem Versuch mit dem Impuls? (.) Kann ich verstehen, die beiden Bälle sahen genau gleich schwer aus, waren es aber nicht. Laut Aussagen der beiden Schüler. Also eine kurze Beweisführung wäre ganz gut. (..)
296	Vorlesen, Sonstiges, Sonstiges*	Es ist nicht gut, dass du sagst, die äh die Wagen wären gleich schwer, da das auch wenn sie baugleich sind nicht stimmen muss. (..) Erneut gleiche Kritk wie bei dem Punkt vorher. Eine Menge Wiederholung in diesem Ding. (.)
297	Vorlesen, Ablehnung AO	Es ist nicht gut, dass du das Gewicht mit null Komma ein Kilo angibst, da du hier nicht berücksichtigst, welches die signa- signifikanten Nachkommastellen sind. Ach (.) komm, also das sehe ich jetzt nicht so, das geht jetzt ein bisschen zu weit. (.)
298	Vorlesen (2x), Zustimmung AO, Verweis auf frühere Stelle im Test	Es ist gut, dass du darauf verzichtest, auch auf Messunsicherheiten einzugehen, da die Kom- um die Komplexität des Experimentes überschaubar zu halten. (...) (unv.) gut, dass du darauf verzichtest, auf Messunsicherheiten, äh ja, ich meine, dass ist ziemlich genau ein (Tandem?) zu dem, was ich vorher schon gesagt habe. (.)
299	Vorlesen, Sonstiges, Ablehnung AO + Begründung, Zustimmung AO + Begründung	Es ist gut, dass du das Gewicht des Wagens mit null Komma eins angibst, um Fehler in der Berechnung des Impulses aufgrund vieler Nachkommastellen zu vermeiden. Ah hier haben wir jetzt eine Meinungsverschiedenheit. Ich benutze sehr gerne möglichst viele Nachkommastellen, um möglichst genau zu sein, andere benutz- äh runden sehr gerne auf ein oder zwei oder drei Nachkommastellen, um möglichst einfach zu arbeiten, also (.) hie- hier scheiden sich die Geister ein bisschen, ich würde sagen nein, manche würden sagen ja. (.)
300	Sonstiges	Weiter. #01:34:45-9#
301	Vorlesen	B1: Welche der unten stehenden Hinweise zu seinem Umgang mit Messunsicherheiten möchten Sie Robert für eine erneute Durchführung des Unterrichts.
302	Vorlesen, Überlegung zum Unterricht (2x), Sonstiges	Es ist gut, wie du mit Messunsicherheiten umgegangen bist. Du kannst dich- äh das kannst du erneut so machen. (.) Erneut Frage waren Messunsicherheiten bereits im Unterricht dran, oder waren (sie?) es halt eben nicht? (.) Ich weiß sowas nicht, ich weiß noch nicht mal, in welchem Bundesland der gute Mann ist. (.) Ich habe da absolut keine Ahnung von, also (.) wenn ja, dann ja, wenn nein, dann nein. (..) Quatsch, anders herum. Wenn ja, dann nein, wenn nein, dann ja. (..) So, sind wir jetzt auch alle äh grundlegend verwirrt, super, weiter. (.)

303	Vorlesen, Zustimmung AO	Es wäre besser, alle Nachkommastellen anzugeben, um die Labor- die die Laborwaage dir anzeigt, da du den Fokus auf genaue Messungen legst und die Laborwaage genauer als null Komma ein Kilogramm wiegen kann. (..) Wie vorhin bereits erwähnt, hier scheiden sich die Geister, ich sage ja, manche sagen nein und so hat jeder seine Meinung. (..)
304	Vorlesen, Überlegung zum Unterricht, Zustimmung AO	Es wäre besser, anstelle der Aussage, du könntest exakt messen, auch Messunsicherheiten einzubeziehen, da Messungen nie wirklich exakt sind. (.) Schon wieder dieses Ding, ich weiß nicht, ob die Messunsicherheiten schon kennen, ja oder nein. (...) Ah als Generalaussage sage ich einfach mal ja.
305	Vorlesen, Zustimmung AO	Es wäre besser, die beide Wagen zu wiegen, damit du den Schülerinnen ein korrektes Bild von genauem wissenschaftlichen Arbeiten vermittelst. Ja. (..) ja.
306	Vorlesen, Zustimmung AO + Begründung	Es wäre besser, beide Wagen zu wiegen, da die Wagen auch wenn sie baugleich (sind?) nicht gleich schwer sein müssen. Ja, Dichte ist immerhin ein Faktor. (..)
307	Vorlesen	Es wäre besser, beide Wagen zu wiegen, (.) da du im Vergleich der Massen auf Messunsicherheiten beziehungsweise Ungenauigkeiten eingehen kannst.
308	Vorlesen	Wäre besser, beide Wagen zu wiegen, damit die Schüler Schülerinnen s- äh selbst überzeugen können, dass die Wagen gleich schwer sind. #01:36:41-4#
309	Vorlesen	B1: Welches Feedback möchten Sie Robert zur Aktivierung der Schüler Schülerinnen geben?
310	Sonstiges	Okay, bevor ich diese Frage beantworte für die Dokumentierung eine kleine Aussage, ich habe mich- ich habe das zwar oft genug bereits gesagt, aber anfangs habe ich wirklich noch darauf fokussiert, jede Frage detailreich und genau zu beantworten und eine möglichst ehrliche Antwort zu geben. (.) Jetzt will ich einfach nur noch fertig werden. Also dieser Test Tex- Limesurvey oder wie auch immer das nennt ist definitiv zu lang für meinen Geschmack. (..) Äh (.) Betonung liegt auf mein Geschmack. (...)
311	Vorlesen	Welches Feedback möchten Sie Robert zur Aktivierung der Schüler Schülerinnen geben?
312	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du die Schüler Schülerinnen nur oberflächlich invol- involvierst, da die Schüler Schülerinnen trotz dieser Aktivierung nicht mitdenken müssen. (.) Stimmt. (.)
313	Vorlesen, Zustimmung AO	Es ist gut, dass du Schüler Schülerinnen für das Experiment einbeziehst, um sie f- im Unterricht zu aktiveren. Auch das stimmt. (.)

314	Vorlesen, Ablehnung AO, Verweis auf frühere Stelle im Test	Es ist gut, dass du das Experiment selbst durchführst, da du in der Durchführung im Vergleich zu den Schüler Schülerinnen dabei sehr sicher bist. (4s) Nein, da haben wir (dann?) einen Konflikt mit dem Aktivieren von der vorherigen Aussage. (..)
315	Vorlesen, Sonstiges*, Zustimmung AO	Es ist nicht gut, dass die Schüler Schülerinnen dir ohne Arbeitsauftrag zusehen, da die Schüler Schülerinnen so keinen Anreiz haben, aufmerksam zu äh zu sein. (..) Warte, da wurde kein Auf- Arbeitsauftrag gegeben? (.) Ich dachte, der wurde einfach nur aus Zeiteffizienz weggeschnitten. Ja, wenn es keinen gibt, dann definitiv. (.)
316	Vorlesen, Sonstiges*, Zustimmung AO	Es ist nicht gut, dass du die Wa- Masse des Wagens und insbesondere die Geschwindigkeit selbst bestimmst, da du den Schüler Schülerinnen vorher schon erklärt hast, wie sie das selber machen könnten. (.) Das hat er getan? War in dem Video gar nicht drin. Okay, angenommen, er hat es wirklich getan, dann ja. (..)
317	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du das Experiment so lehrerzentriert durchführst, da dies für die Schüler Schülerinnen langweilig ist. (.) Ja.
318	Vorlesen, Zustimmung AO + Begründung	Es ist nicht gut, dass Paula von der Tafel aus und die Schüler Schülerinnen (.) den hinteren Reihen den experimentellen Aufbau nicht gut erkennen können, da es für sie schwer ist, dir zu folgen. (.) Ja, besonders, wenn du mit dem Rücken zu ihnen vor den Leuten stehst. #01:38:52-1#
319	Vorlesen	B1: Welche der unten stehenden Hinweise zur Aktivierung der Schüler Schülerinnen möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben?
320	Vorlesen	Es wäre besser, die Schüler Schülerinnen zu fragen, wer freiwillig schreiben beziehungsweise die Werte ablesen möchte, um die Schüler Schülerinnen nicht nach Geschlecht nach eine Aufgabe zuzuweisen. (.)
321	Vorlesen, Sonstiges, Zustimmung AO + Begründung	Es wäre besser, die Schüler Schülerinnen zu fragen, ob sie den Ablauf verstanden oder noch Schwierigkeiten haben, um den Schüler Schülerinnen die Möglichkeit für Rückfragen zu geben. (...) Ach (unv., B1 murmelt, liest die AO vermutlich erneut). Ja, ja, ja, da lässt sich was definitiv zu sagen, wenn die Schüler Schülerinnen da Rückmeldung geben, lässt sich das leichter einschätzen, wer hat jetzt was wo wann verstanden und wer hat jetzt was wo wann nicht verstanden, (damit?) man (nochmal nachhaken?) kann. (.)
322	Vorlesen, Ablehnung AO	Du hast die Schüler Schülerinnen ausreichend aktiviert. (Ha?) Nein. (.)
323	Vorlesen, Zustimmung AO	Es wäre besser, Hypothesen der Schüler Schülerinnen abzufragen, um sie zur Beobachtung des Experiments zu motivieren. Ja, so als eine Art Ein- Einleitungsschritt, das wäre eine gute Idee.

324	Vorlesen, (Zustimmung AO) + Begründung, (Ablehnung AO) + Begründung, Überlegung zum Unterricht	Es wäre besser, den Schüler Schülerinnen Fragen zum Vorgehen während des Experiments oder zur Berechnung der Geschwindigkeit zu stellen, um sie zum Mitdenken anzuregen. (.) Mhm man könnte das äh zum einen so sehen, dass (es?) so äh dass es sehr wohl die zum Mitdenken aktiviert, allerdings auch andererseits könnte man das so sehen, dass es ja ein bisschen äh überfordert, nicht im Sinne von das sind zu schwierige Aufgaben, sondern das sind zu viele Aufgaben auf einmal, immerhin ist Konzentration eine begrenzte Ressource, egal was (.) das deutsche Schulministerium sagt. (..)
325	Vorlesen, Zustimmung AO	Es wäre besser, die Schüler Schülerinnen zum Experiment nach vorne kommen zu lassen, damit sie alles gut erkennen können. Ja, auch das ist eine gute Idee. #01:40:36-3#
326	Vorlesen	B1: Nachdem wir alle Messwerte aufgenommen habe, habe ich die Beobachtungen der Schüler Schülerinnen abgefragt. Um die Allgemeingültigkeit der Impulserhaltung nachzuweisen, habe ich sie dann in einem kurzen Exkurs theoretisch hergeleitet. Im Anschluss haben wir die Impulserhaltung am Experiment überprüft. (..) Aufgabe Überprüfung der Impulshaltung mit den Messwerten aus dem Experiment.
327	Sonstiges	Okay, weiter. #01:41:03-0#
328	-	[Video Einleitung Reflexion 06] #01:41:26-4#
329	Überlegung zum Unterricht	B1: Okay aus dieser Aussage lässt sich schließen, dass die Schüler Schülerinnen noch nicht eine Unterrichtseinheit zum Thema Messunsicherheiten und Messfehler hatten, (.) weil ansonsten hätten die einfach sagen können, ja, das ist halt ein Messfehler, das passiert, (.) aber da diese Aussage getätigt wurde, entweder erinnern sie sich nicht mehr daran oder diese Unterrichtseinheit ist nie passiert. #01:41:48-7#
330	Sonstiges	B1: Oh, noch ein Video, zwei in Folge? #01:41:54-2#
331	-	[Video Szene Reflexion 06] // [Weiterhin Video Szene Reflexion 06] #01:44:19-2#
332	Sonstiges (2x)	B1: Ah, Hiobsbotschaft Nummer zwei, mein Grafiktreiber ist mal wieder abgeschmiert. #01:42:58-5#
		B1: Excuse moi, was war denn das für eine Aussage am Ende? (..) Ihr Mädels. (.) Was? #01:44:27-3#
333	Zustimmung AO	B1: Umgang mit Schüler Schülerinnen, absolut. (.)
334	Sonstiges, Vorlesen	Ah ja, richtig, zur Dokumentierung. Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung, Lob oder Kritik, erhalten, um etwas für sich beziehungsweise seinen Unterricht zu lernen? (.)
335	Vorlesen	Umgang mit den Schüler Schülerinnen. #01:44:46-2#
336	Sonstiges	B1: Mal wieder genügend Rückmeldungen bekommen. #01:44:51-3#
337	Vorlesen	B1: Welches Feedback möchten Sie Robert geben?

338	Vorlesen, Zustimmung AO + Begründung, Sonstiges	Es ist nicht gut, dass du Mädels sagst, da das ein sexistischer Ausdruck ist. (...) Das ist jetzt ein bisschen außerhalb meiner Expertise, aber ich habe definitiv schonmal darü- äh Frauen darüber reden hören, dass der Ausdruck Mädels tatsächlich eher negative Konnotationen hat, also nehme ich einfach mal dieses Wissen von wann auch immer ich das gelernt habe und wende das hier drauf an und sage ja. (.)
339	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du die Mädchen und die Jungen so unterschiedlich behandelst, da das nicht wertschätzend gegenüber den Mädchen ist. Ja.
340	Vorlesen, Zustimmung AO + Begründung	Es ist nicht gut, dass du die Mädchen und Jungs Jungen in ihrem Können so unterschiedlich einschätzt, da du (.) so die Vorstellung förderst, Physik wäre ein Jungenfach. Auch das ja. Die- Diese (.) einfache, aber dennoch knackig und präzise schießende Aussage, ihr Mädels müsst euch das mal korrigieren, während die Jungs alles in Ordnung haben, ja, hilft definitiv, bei der Vorstellung.
341	Vorlesen, Sonstiges*	Du thematisierst die Einheit des Impulses nicht. (...) Was hat denn das mit dem Umgang der Schüler Schülerinnen zu tun? (..)
342	Vorlesen, Sonstiges*, Sonstiges	Äh da- das ist nicht gut, da dies (inkonsistent?) zur früheren Szene ist, in der du bei Paula auf die Notation der Einheit der Geschwindigkeit bestanden hast. Ich glaube, hier ist eine Aussage in der Kategorie verwechselt, weil das hat- das hat kaum bis gar nichts mit Schüler Schülerinnenumgang zu tun. (.) Okay, tun wir mal so als wäre (es?) (.)
343	Vorlesen, Sonstiges, Zustimmung AO, Alternative	Das ist nicht gut, da dies inkonsistent zur früheren Szene ist, (in der du bei Paula auf die Notation der?, gemurmelt) Einheit der Geschwindigkeit bestanden hast. (...) Ach, es geht um die Verwendung der Einheit, ah jetzt verstehe ich, was der Satz (.) soll. (..) aber später. (.) Äh das kann man auch danach machen, man muss nicht immer alles gleichzeitig machen. Ich sage einfach mal ja.
344	Vorlesen, Bewertung	Es ist gut, dass du die Mädchen dazu aufforderst, ihre Tabelle zu korrigieren, damit du sicher sein kannst, dass sie die gleichen Werte, wie die Jungen notiert haben. (.) Wegen der Aussage von einer Schülerin verdonnerst du eine gesamte äh Demographie dazu, ihre Sachen zu korrigieren, das ist eine dumme Idee. (.) E- egal welcher Kontext, das ist eine dumme Idee. (..)
345	Sonstiges	Weiter. #01:47:08-1#
346	Vorlesen	B1: Welche der unten stehenden Hinweise zu seinem Umgang mit den Schüler Schülerinnen möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben?
347	Vorlesen	Dein Umgang mit den Schüler Schülerinnen ist in Ordnung, du kannst ihn so beibehalten.

348	Vorlesen, Zustimmung AO	Du (möchtest?) die Mädchen stärker in den Unterricht einbeziehen, B- zum Beispiel berechnete Werte vorlesen lassen, um ihr physikalisches Selbstkon-zept zu stärken. Zum Beispiel. (.)
349	Vorlesen, Verweis auf frühere Stelle im Test, Zustimmung AO + Begründung	Du solltest auf geschlechtsspezifische Zuschreibungen verzichten, um die Schüler Schülerinnen nicht äh um die Schülerinnen nicht zu diskriminieren. Wie wir an einem früheren Punkt bereits gesagt, geschlechtsspezifische Zuschreibungen generell aus dem Wortschatz rauszuhauen ist die meiner Meinung nach einfachste und auch gleichzeitig sicherste Möglichkeit, also das ist generell eine gute Idee. (.)
350	Vorlesen, Zustimmung AO + Begründung	Du solltest von jedem jeder Schüler Schülerin je nur einen berech- berechneten Wert erfragen, da du so die berechneten Werte von mehr Schüler Schülerinnen kontrollieren kannst. (..) Ja, das kann man sehr wohl machen, allerdings würde ich das aus einem anderen Grund machen und zwar einfach nur äh um mehrere Schüler Schülerinnen und das will heißen so viele wie Möglich in den Unterricht mit einzubeziehen. (.)
351	Vorlesen, Zustimmung AO + Begründung, eigene Erfahrung, Ablehnung AO	Du solltest verstärkt schwächere Schüler Schülerinnen in den Unterricht einbinden, da wichtig ist, dass sie die Inhalte verstehen. Ja, ja, es ist theo- die Idee dahinter ist schon richtig, dieses Korr- äh Überprüfen, dass die Inhalte stimmen, aber als jemand, der früher selber ein schwacher Schüler war, ist- gibt es nichts schlimmeres, als vor der gesamten Klasse bloß gestellt zu werden und zu veröffentlichen nicht veröffentlichen, offenbaren, dass ich keinen Plan von gar nichts habe, also erneut (.) jein. (..)
352	Vorlesen, Sonstiges*, Ablehnung AO + Begründung	Du solltest die Schüler Schülerinnen unabhängig vom Geschlecht dazu auffordern, falsche Ergebnisse in der Tafel zu korrigieren, damit du am Ende alle Schüler Schülerinnen die richtigen Ergebnisse in der notiert haben. Was ist hier jetzt mit falsche Ergebnisse gemeint? Aussagen von den anderen Schülern Schülerinnen, die zu den eigenen Notizen nicht- weil wenn das passiert, dann äh wird es äh die ganze Situation nur in einem enden (.) zwar Chaos und das will niemand (.) also nein. (..)
353	Vorlesen, Zustimmung AO + Begründung	Du so- solltest hinterfragen, woher die geschlechts- spezifischen Zuschreibungen stammen, die du vertrittst, um sie zu überwinden. Ja, das ist ein generelles Ding, das man machen sollte, aber das hat eher weniger Bezug zum Unterricht, ich sage aber dennoch mal ja, weil sich selber zu hinterfragen ist immer eine gute Idee. Okay, vielleicht nicht immer, aber es ist eine gute Idee. #01:49:23-9#
354	Vorlesen	B1: Welches Feedback möchten Sie Robert zu seinem Umgang mit Maries Einwand beziehungsweise mit Messunsicherheiten geben? (.)

355	Vorlesen, Überlegung zum Unterricht	Die Reaktion auf M- Maries Einwand ist dir gut gelungen, da s- du so möglichst effizient zum zentralen Ergebnis der Impulserhaltung gelangst. (..) Äh wir haben jetzt ein äh Prioritätenproblem, was ist jetzt gerade (in?) dieser Situation wichtiger, dass Marie das wirklich versteht, äh was jetzt mit den Messfehlern äh auf sich hat oder dass alle verstehen, was jetzt die äh Impuls- äh Impulserhaltung bedeutet, das ist jetzt ein größeres Problem. (.) Je nach, was man höher als das höhere Problem setzt, fällt dann hier die Aussage ja oder nein zu. (.)
356	Vorlesen, Zustimmung AO, Alternative + Begründung (2x), Sonstiges*	Die Reaktion auf Maries Einwand ist dir gut gelungen, da du- da Rundungsfehler eine zentrale Fehlerquelle für den Unterschied- die Unterschiede der Ergebnisse sind. Ja, technisch gesehen hat diese Aussage schon recht, aber für meinen Geschmack hätte es ein bisschen ausführlicher, mehr mehr im Detail und vor allem generell mehr Erklärung passieren können, um wirklich sicher zu gehen, dass A Marie alles verstanden hat und B dass auch jeder andere das verstanden hat, der sich schlichtweg einfach nur nicht traut, zu fragen. (...) Ah (ach?) ja richtig, die Frage was kreuze ich hier an, warum nicht.
357	Vorlesen, Beschreibung, (Ablehnung AO) + Begründung	Die Reaktion auf Maries Einwand ist dir nicht gut gelungen, da du den Fokus von der experimentellen Bestätigung der Impulserhaltung hin zur Diskussion von leichten Ungenauigkeiten verschiebst. (..) Es ist ein- eine Neufokussierung, definitiv, aber es ist eine dem Kontext gegebenen äh (.) im gegebenen Kontext wichtige Diskussion, die geführt werden muss. (..)
358	Vorlesen, Zustimmung AO + Begründung, Sosntiges, Alternative + Begründung	Es ist nicht gut, dass du nur so knapp auf den Einwand von Marie einge eingehst, ich kreuze sie jetzt schonmal an, da (ich?) äh es definitiv meine Kritik ist, da die Begründung nicht ausreicht, um den Schüler Schülerinnen zu verdeutlichen, dass die Unterschiede aufgrund von Messunsicherheiten zustande kommen. Ja, Schüler ak- Schüler und Schülerinnen akzeptieren definitiv eine Menge Sachen, die man ihnen im Unterricht beibringt, aber es ist immernoch besser, dass sie diese Sachen auch verstehen und letzteres ist, was wir wollen, weswegen da definitv mehr Zeit drauf verwendet werden sollte. (..)
359	Vorlesen, Zustimmung AO, Verweis auf frühere Stelle im Test	Es ist nicht gut, dass du nur so knapp auf den Einwand von Marie eingehst, da so unklar bleibt, ob beziehungsweise wo die Schüler Schülerinnen Rundungsfehler gemacht haben. (.) Auch das kann man im Tandem mit der vorherigen Aussage kritisieren

360	Vorlesen, Bewertung + Begründung, Zustimmung AO + Begründung	Es ist nicht gut, dass du die Ungleichheit der Messwerte mit Rundungsfehlern begründest, da diese Begründung falsch ist. (..) Jein, jein, jein, jein, jein äh Rundungsfehler ist definitiv eine Fehlerquelle, es ist nur ein bisschen kurz gefasst, zu kurz gefasst zu dem Punkt, wo ei- wo einige wichtige Informationen ausgelassen werden. (...) Ich sage einfach mal ja, weil es ist doch ein Thema, das etwas länger und breiter äh besprochen werden sollte, als ein Nebensatz. (..)
361	Vorlesen, Ablehnung AO + Begründung	Es ist nicht gut, dass du die Ungleichheit der Messwerte mit Rundungsfehlern begründest, da du so den Schüler Schülerinnen die Schuld dafür gibst, dass die Ergebnisse nicht exakt gleich sind. Ich meine, das ist jetzt ein bisschen antagonistisch gedacht, dass jetzt eine Aussage zu Fehler gleich dazu führt, dass Schüler Schülerinnen an Selbstzweifel zerfallen, also ich würde sagen nein, (.) als Schüler generell hat man mit viel mehr emotionaler äh (..) Auswirkungen, Stress, emotional Trauma, mir fällt das deutsche Wort gerade nicht ein, zu tun, als nur eine Aussage von einem Lehrer über ein Thema und einen Rundungsfehler, also (.) Schüler sind da etwas härter im nehmen, also man erstmal denkt. (...) Oh äh Schüler und Schülerinnen, (die?) wollen wir ja nicht vergessen. #01:53:05-9#
362	Vorlesen	B1: Welche der unten stehenden Hinweise zu seinem Umgang mit Maries Einwand beziehungsweise mit Messunsicherheiten möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben?
363	Vorlesen, Ablehnung AO, Verweis auf frühere Stelle im Test	Deine Reaktion auf den Einwand von Marie ist angemessen und du kannst so er-, das kannst du erneut so machen. (..) Nein, die Gründe dafür habe ich vorhin bereits lang und breit erklärt. (.)
364	Vorlesen, Zustimmung AO, Überlegung zum Unterricht, Sonstiges	Du solltest den Schüler Schülerinnen angeben, auf wie viele Nachkommastellen sie runden sollen, ein oder zwei Stellen, damit du diese Situation umgehen kannst. Das wäre eine Möglichkeit, um diese Situation zu umgehen, allerdings ist dann (wieder?) die Frage, was willst du lieber, die Situation um- komplett umgehen oder die Situation voll voll front angreifen, damit die Schüler nämlich auch komplett verstehen, was da Sache ist. Erneut, je nachdem, wo die Prioritäten sind, setzt man hier jetzt das Kreuz, ja oder nein. (...)

365	Vorlesen (2x), Sonstiges (2X)	Es wäre besser, den Schüler Schülerinnen am Anfang mitzuteilen, dass Messunsicherheiten und Rundungsfehler nicht berücksichtigt werden, um diese Situation zu vermeiden. (...) Mein Hirn hat das nicht verarbeitet, nochmal. Es wäre besser, den Schüler Schülerinnen am Anfang mitzuteilen, dass Me- Messunsicherheiten und Rundungsfehler nicht berücksichtigt werden, um diese Situation zu vermeiden. (..) Entweder bin ich gerade dumm, oder dieser Satz macht keine Sinn für mich. (..)
366	Vorlesen, (Ablehnung AO) + Begründung (2x), Sonstiges	Es wäre besser, den Schüler Schülerinnen am Anfang mitzuteilen, dass Messunsicherheiten und Rundungsfehler nicht berücksichtigt werden- ja schön, wenn sie nicht berücksichtigt werden, das heißt aber noch lange nicht, dass sie nicht da sind. (.) Selbst, wenn man am Anfang sagt, ja, diese Fehler, die werden ignoriert, das ändert nichts an der Tatsache, dass äh da Unterschiede in der Ergebnissen (wird?) und somit Marie weiterhin ihre Kritik hat, also in diesem Punkt, egal, was du machst, ergibt (genau?) das selbe. (4s) Ja.
367	Vorlesen, Zustimmung AO + Begründung	Es wäre besser, ausführlicher zu begründen, wo die Rundungsfehler entstehen, damit die Schüler Schülerinnen deiner Erklärung besser folgen können. Das wäre eine Möglichkeit, dass Thema besser aufzuschlüsseln (.) entschuldigung und generell besser zu erklären. (.)
368	Vorlesen, Sonstiges*, Zustimmung AO, Alternative + Begründung	Du solltest bereits bei der Vorstellung des Experimentes über die vorhandene Reibung und Energieumwandlung beim Aufprall sprechen, damit den Schüler Schülerinnen in dieser Phase klar ist, dass die Ergebnisse aufgrund dieser Einflüsse nicht exakt gleich sein können. (.) Selbe Aussage, wie oft schon in diesem Survey-Text. Ja, aber später. Das letzte, was man will, ist die äh Schüler Schülerinnen mit Informationen überhäufen und das wäre zum Beispiel eine Sache. Gerade (frisch?) beim Impulsbegriff angekommen und jetzt soll man bereits über Energieumwandlung, Reibung, mögliche Wärmeentwicklung, mögliche Deformationsenergien und all diese anderen Sachen nachdenken? Schlichtweg zu viel, (.) man muss diese Sache Schritt für Schritt angreifen und nicht alles auf einmal. (.)
369	Vorlesen (2x), Sonstiges*, Zustimmung AO	Es wäre besser, zu thematisieren, dass es bei der Durchführung des Versuches Messunsicherheiten berücksi- Nochmal. Es wäre besser, zu thematisieren, dass es bei dieser bei der Durchführung des Versuches Messunsicherheiten berücksichtigt werden müssen, damit die Schüler Schülerinnen nachvollziehen können, wieso es zu Abweichungen kommt. (..) K- Ja, gute Idee.
370	Vorlesen, Zustimmung AO + Begründung	Du solltest das Thema Messunsicherheiten für dich noch einmal selber aufbereiten, damit du von solchen Fragen nicht vom Thema abgebracht wirst. (..) Ja, generell Vorbereitung ist eine gute Idee. (...)

371	Sonstiges, Sonstiges	(Oh?) mein Hals tut mir richtig weh mittlerweile von dem ganzen Reden. #01:56:35-0# B1: Ich meine, zu meiner Verteidigung (..) zwei Stunden später und wir haben noch nicht mal achtzig Prozent erreicht. (4s)
372	Vorlesen	Welches Feedback möchtest du Robert zu seinem äh Umgang mit der Einheit des Impulses geben? (..)
373	Sosntiges (2x)	Ah vielleicht wird es ein bisschen besser, wenn ich drauf drücke. (..) Nein, scheiß drauf, (hauen?) wir einfach weiter.
374	Vorlesen, Zustimmung AO, Alternative + Begründung	Es ist nicht gut, dass du nicht über die Einheit des Impulses sprichst, da die Einheit ein wichtiger Bestandteil der Angabe des Impulses ist. Ja, wenn man bereits mit solchen Einheits- äh spielereien anfängt, dann muss man das Spiel auch zu Ende spielen, weil ansonsten macht es schlichtweg keinen Sinn. (.)
375	Vorlesen, Sonstiges*, Sonstiges	Es ist nicht gut, dass du nicht über die Einheit des Impulses sprichst, da dies inkonsistent zur früheren Szene ist, in der du auf die Notation der Einheit der Geschwindigkeit be- Ist das nicht dasselbe, wie die vorherige Aussage? (.) Wenn nicht, dann ist wenigstens mein- meine Aussage dazu dieselbe. (.)
376	Vorlesen, Zustimmung AO	Es ist gut, dass du die Einheit des Impulses nicht so stark in den Fokus gerückt hast, (.) um die Schüler Schülerinnen nicht durch die neue Einheit zu verwirren. (...) Ja, gut, ich sage zwar immer ja, aber später, aber ein paar Sachen sollte man doch schon gleich am Anfang klären. (..)
377	Vorlesen, Zustimmung AO + Begründung, Alternative + Begründung	Es ist gut, dass du den Fokus auf die berechneten Werte und nicht auf die Einheit gelegt hast, da die gleichen Zahlen ausreichen, um die Impulserhaltung auf Grundlage des Experiments zu bestätigen. (..) Ach ja, es ist schon besser, wenn der Fokus mehr auf dem Wert an sich liegt, aber um den Wert in Kontext zu setzen, ist es besser, auch gleich eine Einheit zur Hand zu haben, dementsprechend wäre es besser, gleich die Einheit z- zusammen mit der Formel und dem Wert ganz am Anfang zu klären, (.) damit die Schüler auch was handfestes haben. (..)
378	Vorlesen, Alternative, Zustimmung AO	Es ist nicht gut, dass du die Schüler Schülerinnen nicht dazu aufforderst, die Einheiten anzugeben, da so unklar ist, ob die Impulse wirklich gleich groß sind. Ja, hätten die man die Einheiten gelernt. (..) Angenommen, in diesen Zwischencuts wurde die Einheit erklärt, dann ja. #01:58:36-2#
379	Sonstiges	B1: Achtzig Prozent, wuu. (..)
380	Vorlesen	Welche der unten stehenden Hinweise zu seinem Umgang mit der Einheit des Impulses möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben?
381	Vorlesen	Es ist in Ordnung, dass die Einheit des Impulses zunächst vernäch- dass du die Einheit des Impulses vernachlässigst, (.) das kannst du so erneut machen.

382	Vorlesen, Zustimmung AO	Du solltest die Einheit des Impulses angeben, da die Einheit ein wichtiger Bestandteil der Angabe des Impulses ist. Ja.
383	Vorlesen, Zustimmung AO	Du solltest darauf achten, dass die Schüler beim Vorlesen der Werte auch die Einheit angeben, (..) um in deinem eigenen Verhalten konsistent zu bleiben. Auch das. (.)
384	Vorlesen, Zustimmung AO + Begründung	Du solltest die Einheit des Impulses oben über die Spalte der Tabelle notieren, um die Schreibarbeit in jeder Zeile zu sparen. (.) Das ist eine Möglichkeit, aber Einheit muss angegeben werden, so oder so. (.)
385	Vorlesen, Zustimmung AO, Sonstiges	Du solltest die Schüler Schülerinnen dazu auffordern, die Einheit des Impulses zu nennen, damit die Schüler Schülerinnen sich an die Einheit des Impulses gewöhnen. Auch das, wow, komplette Sammlung, mhm nett. #01:59:44-2#
386	Sonstiges	B1: Und es geht weiter, meine Güte. (..)
387	Sonstiges	Nur mal so als Statistikfrage, wie lange haben die anderen etwa gebraucht, bis die da durch waren? #01:59:54-1#
388	-	I: Für das laute Denken? #02:00:01-4#
389	Sonstiges	B1: Für das gesamte Ding. #02:00:03-5#
390	-	I: Also das Sprechen macht den Test halt nochmal ein ganzes Stück länger, ne, also ohne das Sprechen ist es eine gute Stunde. #02:00:09-7#
391	Sonstiges	B1: Und mit? #02:00:11-9#
392	-	I: Ganz unterschiedlich, auf jeden Fall mehr. (.) Ich habe aber auch noch nicht viele Vergleichswerte, also- #02:00:20-8#
393	Sonstiges, Sonstiges	B1: Ja, ich meine ja nur, wir brechen jetzt schon die e- die dritte Stunde an. (..) Ah, ist jetzt eh zu spät für die Sachen, die ich noch machen wollte. #02:00:30-0#
394		[Video Einleitung Reflexion 07] // #02:00:43-4#
395	Überlegung zum Unterricht, Sonstiges	B1: Mhm (.) also wenn das so wie dein Merksatz zum zweiten Axiom ist, dann uh, bereite ich mich jetzt lieber schonmal auf die Fachsprache vor. (4s) Yay wir sind in der letzten Tabelle. (7s)
396	Vorlesen	(Reflexion?) null sieben. #02:01:03-3#
397	-	[Video Szene Reflexion 07] // [Weiterhin Video Szene Reflexion 07] // [Weiterhin Video Szene Reflexion 07] #02:02:38-3#
398	Sonstiges	B1: War da ein Videofehler gerade? #02:01:48-0# B1: Für einen einzelnen Frame war da gerade das ganze gesamte Bild so rot #02:01:56-9#
399	Sonstiges	B1: Abschreiben, das Ding das ein- jeder einzelne Lehrer macht, aber dennoch nie ein Schüler lieben lernt. #02:02:48-2#

400	Vorlesen (2x), Sonstiges	B1: Formuierung des Merksatzes zur Im- Ah ja, richtig, Dokumentierung. Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung, Lob und/oder Kritik, erhalten, um etwas für sich beziehungsweise seinen Unterricht zu lernen? (..)
401	Vorlesen	Formulierung des Merksatzes zur Impulserhaltung,
402	Vorlesen, Überlegung zum Unterricht, Ablehnung AO, Sonstiges	Interaktion mit den Schüler Schülerinnen. (4s) Ich meine, gab es Interaktion mit den Schüler Schülerinnen? (...) Ich sage mal nein und fokussiere mich auf das obere. #02:03:19-7#
403	Sonstiges	B1: Ja, ja, ja, schon genügend Rück- äh Rückmeldungen bekommen. #02:03:24-2#
404	Vorlesen	B1: Welches Feedback möchten Sie Robert zu seiner Formulierung des Merksatzes zum zur Impulserhaltung geben?
405	Vorlesen, Zustimmung AO + Begründung, Beschreibung	Die Formulierung des Merksatzes ist dir gut gelungen, da alle wichtigen Informationen korrekt enthalten sind. (..) Ja, ja, so kann man das definitiv sehen, es sind alle Informationen gegeben auch (ein?)- äh mit den Formeln gut zusammengefasst, (auch?) einmal die Formel für den Impuls an sich nochmal eingefügt mit diesem p gleich m mal v (..) und auch die Summe der beiden Impulse reingehend ist gleich die Summe der beiden rausgehend, auch das ist mit drin, ja, mir fällt jetzt spontan nichts ein, was da fehlt.
406	Vorlesen, Überlegung zum Unterricht, Ablehnung AO	Es ist gut, dass du im Merksatz auf das durchgeführte Experiment verweist, da die Schüler Schülerinnen den Merksatz direkt mit der Realität verbinden können. Das hat er getan? Kam mir jetzt überhaupt nicht so vor. (.) Ich sage einfach mal nein (unv.).
407	Vorlesen, Ablehnung AO	Die Formulierung des Merksatzes ist dir nicht gut gelungen, da sie für die Schüler Schülerinnen zu anspruchsvoll ist. (..) Ja, komm, bisschen mehr aushalten können die auch. (..)
408	Vorlesen, Zustimmung AO + Begründung	Es ist nicht gut, dass du deinen vorgefertigten Merksatz nutzt und die Merksätze der Schüler Schülerinnen nicht einbeziehst, dass du den Schüler Schülerinnen gegenüber nicht wertschätzend ist. So kann man das natürlich auch sehen, die Verwendung der Merksätze der Schüler Schülerinnen verfestigt das auch nochmal in der äh eigenen Erinnerung, weil das ist dann ein klarer Unterschied zwischen das wurde mir gegeben und das habe ich selber gemacht, das ist ein riesen Unterschied, das das hier, das lässt sich viel leichter merken. #02:04:55-5#
409	Vorlesen	B1: Okay, welche der unten stehenden Hinweise zu seiner Formulierung des Merksatzes zur Impulserhaltung möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben? (.)

410	Vorlesen, Zustimmung AO + Begründung	Die Formulierung des Merksatzes ist angemessen und du kannst sie erneut verwenden. (.) Ja, finde ich schon, es war alles mit dabei, Formulierung war auch ganz okay, war jetzt nicht so extrem, was die F- Fachwörter angeht, ja, ja, ich bin damit zufrieden.
411	Vorlesen, Sonstiges	Es wäre besser, Beispiel und Merksatz zu trennen, um die Komplexität des Merksatzes zu reduzieren. Warte, wo ist diese Aussage, wo ich mir nochmal den Merksatz angucken kann? (.) PowerPoint-Folie, da. (B1 öffnet Hilfe 'Kannst du mir die PowerPoint-Folie noch einmal zeigen?'). (..)
412	Vorlesen, Ablehnung AO + Begründung, Sonstiges	Impuls ist eine Erhaltungsgröße. In einem abgeschlossenen System ist die Summe der zu einer We- vor einer Wechselwirkung, zum Beispiel einem Stoß, gleich der Summe der Impulse nach der Wechselwirkung. (.) Ja, das ist jetzt, das sind (unv.) drei Wörter in dem gesamten Merksatz, die einfach nur auf das Beispiel fokussiert sind, das finde ich jetzt nicht so kritikwürdig. (...) Ah (einmal?) kurz neu hinsetzen.
413	Vorlesen, Zustimmung AO + Begründung, Überlegung zum Unterricht	Es wäre besser, zusätzlich zum langen Merksatz eine kurze Version zu formulieren, da sich die Schüler Schülerinnen eine kurze Version besser merken können. (.) Das ist schlichtweg eine generelle Wahrheit, äh kurze Sachen lassen sich leichter merken. (.) Ähm mir würde spontan nicht einfallen, wie man den Merksatz verkürzen könnte, ohne irgendwelche Informationen (zurück?) zu lassen. (.)
414	Vorlesen, Ablehnung AO + Begründung	Es wäre besser, den Merksatz nur in Bezug auf einen Stoßprozess zu formulieren, damit der Merksatz näher am besprochenen Experiment bleibt. Es wurde bereits ein Stoß als Beispiel in den Merksatz mit eingesetzt, also da finde ich jetzt nicht die Kritikwürdigkeit, äh da irgendwie was (zu ändern?).
415	Vorlesen, Überlegung zum Unterricht, Ablehnung AO + Begründung, Zustimmung AO + Begründung, Sonstiges	Es wäre besser, den Merksatz um die Einheit des Impulses zu ergänzen, da die Einheit stets zur Angabe einer physikalischen Größe gehört. (.) Womit wir wieder bei dem Problem wären, ob jetzt dieser physikalische Einheit für den Impuls bereits angegeben ist, ja oder nein, wenn nein, dann sollte er das, wenn ja (.) dann wurde es bereits und muss äh nicht unbedingt vielleicht nochmal wiederholt werden oder jetzt, wo ich drüber nachdenke, wäre es immer eine gute Idee, besonders am Anfang, Sachen so oft, wie möglich zu wiederholen, damit es (unv.) sofort gefestigt wird und nicht (das?) klassischer Fall von in einem Ohr rein, im anderen Ohr raus. (.) Äh wo war jetzt die Aussage, ich habe die komplett aus den Augen verloren.
416	Vorlesen, Zustimmung AO	Es wäre besser, den Merksatz um die Einheit des Impulses zu ergänzen, ja, das war die.

417	Vorlesen, Ablehnung AO	Du solltest den Merksatz schülernäher gestalten, zum Beispiel indem du eine Skizze ergänzt oder die Formulierung sprachlich vereinfachst, um die Schüler Schülerinnen bei Verständnis zu- beim Verständnis zu unterstützen. (...) Ja, man kann es mit einer Skizze erweitern, allerdings glaube ich äh schon, dass es in diesem Fall ganz ausreichend ist, den Merksatz so zu lassen, wie er gerade eben ist. (..)
418	Vorlesen, Zustimmung AO + Begründung, Verweis auf frühere Stelle im Test	Es wäre besser, eine Formulierung der Schüler Schülerinnen als Merksatz zu verwenden, da diese Formulierungen für sie besser verständlich sind. Das könnte man auch definitiv benutzen, eine Schüler Schülerin einfach äh selbst gemachte Sache, womit wir wieder bei meinem Punkt von vorhin machen äh gegeben vergleich mit selbstgemacht, der Selbstgemachte lässt sich leichter merken, weil es ist etwas, das hat man selber gemacht, man ist durch den Prozess gegangen, die falschen Sachen auszubügeln und die richtigen Sachen zu finden und allein dieser Prozess des richtig Findens äh setzt sich dann viel viel stärker und besser in den ei- ins eigene Gedächtnis ab. (...)
419	Sonstiges	Weiter. #02:08:24-3#
420	Vorlesen (2x), Sonstiges*	B1: Welches Feedback möchtest du Robert zu seiner Interaktion- ich habe den- (.) das habe ich nicht angekreuzt. (..) Ich wollte nur auf das (.) was, warum bin ich jetzt hier? Ich habe das nicht angekreuzt, ich will hier nichts zu sagen. (..) Technischer Fehler. (.) Äh welches Feedback möchtest du Robert zu seiner Interaktion mit den Schüler Schülerinnen geben?
421	Vorlesen, Zustimmung AO	Es ist gut, dass die Schüler Schülerinnen ihre Merksätze vorlesen, da du so erkennen kannst, ob sie die Impulserha- erhaltung umfassend verstanden haben. Ist eine gute Idee. (.)
422	Vorlesen, Zustimmung AO	Es ist gut, dass die Schüler Schülerinnen einen Merksatz formulieren sollen, um ein tief- gehendes Verständnis zu fördern. Ja. (..)
423	Vorlesen, Zustimmung AO + Begründung, Alternative	Es ist nicht gut, dass du den Schüler Schülerinnen keine Rückmeldung auf die Merksätze gibst, da die Schüler Schülerinnen so nicht einordnen können, ob ihre Merksätze angemessen formuliert sind. Das ist etwas definitiv kritikwürdi-ges, weil das könnte die Idee äh in den Kopf in die Köpfe der Schüler einsetzen, dass äh weil mein Merksatz ungleich Merksatz des Lehrers bedeutet, dass mein Merksatz falsch ist. (..) Also (..) da sollte man definitiv auch drauf eingehen. (.)

424	Vorlesen, Verweis auf frühere Stelle im Test, Zustimmung AO	Es ist nicht gut, dass du deinen vorgefertigten Merksatz nutzt und die Merksätze der Schüler Schülerinnen nicht einbesetzt, da du den Schüler Schülerinnen gegenüber nicht wertschätzend ist. (..) Den Punkt hatten wir doch schon, genau diesen Punkt, der war doch irgendwo anders auch schon. (4s) (Ja?), ich kreuze das einfach nochmal an.
425	Vorlesen, Überlegung zum Unterricht, Ablehnung AO + Begründung	Es ist nicht gut, dass du deine Formulierung als besser bezeichnest, (.) da du die Formulierungen der Schüler Schülerinnen so abwertest. Mhm da kann man jetzt äh das so oder so sehen, je nachdem, wie viel Selbstvertrauen und Resistenz man den Schülern und Schülerinnen (.) zuschreibt. Wenn man denen viel (.) pardon, wenn (man?) denen viel Selbstvertrauen zuschreibt, dann sollte man das jetzt nicht als Kritikpunkt ansehen, wenn man (denen?) viel zuschreibt, dann (..) äh, ne wenn man denen wenig (.) äh Selbstvertrauen zuschreibt, dann wäre das doch schon (eine?) größere Kritikpunkt. Ich vertraue da mal den Schüler Schülerinnen und sage, dass die in dem Punkt genügend Selbstvertrauen haben. (..)
426	Vorlesen, Zustimmung AO, Überlegungen zum Unterricht	Es ist nicht gut, dass nur zwei Schüler Schülerinnen ihre Merksätze vorgestellt haben, da du so nicht weißt, welchen Merksatz die restlichen Schüler Schülerinnen formuliert haben. Technisch gesehen ja, aber die Alternative ist sehr zeitaufwendig und zeitintensiv und nach so einem Lehrplan, angenommen, dieser gesamte Limesurvey-Kurs war eine Physikstunde (..) Junge, was für ein Zeitplan, da muss man erstmal hinterherkommen. (...)
427	Vorlesen, Überlegung zum Unterricht, Zustimmung AO + Begründung, Ablehnung AO + Begründung	Es ist nicht gut, dass nur zwei Jungen ihre Merksätze vorgestellt haben, da die Schüler Schülerinnen so lernen, dass du nur Jungen drannimmst. Ich erinnere mich jetzt gar nicht mehr, wie das Verhäl- Geschlechterverhältnis in der Klasse war. War das nicht so, dass tatsächlich mehr Mädchen in der Klasse sind als Jungen? (.) Wenn das der Fall ist, dann (.) würde ich sagen, ja, (.) weil das so meine Erinnerung ist, sage ich, kreuze ich hier mal an, wenn das nicht so der Fall ist und schlichweg mehr Jungs (.) in der Klasse sind also Mädchen, dann (.) glaube ich schon, dass es die meisten einfach zum äh dem Zufall anrechnen würden. (..)
428	Sonstiges	Weiter. #02:11:45-5#
429	Vorlesen	B1: Welche der unten stehenden Hinweise zu seiner Interaktion mit den Schüler Schülerinnen möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben? (..)

430	Vorlesen, Ablehnung AO + Begründung	Du solltest den Merksatz ausführlicher erklären beziehungsweise mit den Schüler Schülerinnen besprechen, da einige Aspekte des Merksatzes nicht selbsterklärend sind. (..) Das sehe ich jetzt tatsächlich nicht so, es war ziemlich ausführlich, besonders mit der Formel am Ende war es äh ziemlich gut erklärend, also den Punkt sehe ich nicht so.
431	Vorlesen, Verweis auf frühere Stelle im Test, (Zustimmung AO) + Begründung	Du solltest gemeinsam mit den Schüler Schülerinnen einen Merksatz erstellen, um sie stärker in den Unterricht einzubeziehen. (.) Ich wiederhol den Punkt jetzt nicht nochmal, aber kurz gesagt, das eigen gemachte lässt sich nicht äh besser merken. (.)
432	Vorlesen, Zustimmung AO, Ablehnung AO + Begründung	Es wäre besser, insgesamt mehr Schüler Schülerinnen ihre Merksätze vorstellen zu lassen, da du aus den Aussa-Antworten Rückschlüsse auf die Vorstellungen der Schüler Schülerinnen schließen kannst. (.) Erneut (.) ja, aber zeitaufwendig, also mach das lieber nicht. (..)
433	Vorlesen (3x), Zustimmung AO + Begründung, Beschreibung, Sonstiges (2x)	Du solltest die Schüler Schülerinnen sich melden lassen, wenn sie ihren Merksatz vorstellen wollen, da sich möglicherweise auch ein Mädchen gemeldet hätte. (...) Ich würde sagen ja, ja, das- diese Meldengeschichte gibt auch (.) gibt allen das Gefühl der Selbstbestimmung, der Autonomie, des ich bin hier, weil ich hier sein will, andererseits ist es eine (.) sehr gute Möglichkeit für Schülerinnen, strategisch ignoriert zu werden, in diesem Unterricht, also (..) das Autonomiegefühl der Schüler würde ich allerdings als höher einstufen, äh weswegen (...) wo ist die Zeile, ich habe die Au- Zeile aus den Augen verloren. Es wäre besser (unv., murmelt) nein, du warst das nicht, Du sollst die Schüler Schülerinnen, sich melden lassn (unv.)- ja, melden lassen, das war das.
434	Vorlesen, Sonstiges*, Zustimmung AO	Der Umgang mit den Schüler Schülerinnen war angemessen, so kannst du dich äh wieder verhalten. Uh schwierig, war das jetzt nur auf dieses eine Video bezogen oder auf die gesamte Stunde? Nur auf dieses eine Video? Ja, warum nicht. Auf die gesamte Stunde, (ah?) ne, lieber nicht. (..) Ähm (..) ich nehme einfach mal an, das ist nur auf dieses eine Video bezogen, weil es in der Kategorie nach diesem Video ist. (.)

435	Vorlesen, Überelgung zum Unterricht, Zustimmung AO + Begründung	Du solltest deinen Umgang mit den Schüler Schülerinnen überdenken. Da du sie und ihre Arbeit nicht lobst, kann dem Verhältnis der Schüler Schülerinnen zum Physikunterricht nachhaltig schaden. (...) Wie kann man das entweder so sehen, dass äh dem Punkt hier rechtgegeben wird und fehlende fehlender Lob am Selbstbewusstsein krazt, allerdings könnte man das auch so rum sehen, dass man definitiv auch mit Lob übertreiben kann, somit (der?) Gewichtung des einzelnen (.) Lobs äh wegnehmen kann. (...) Äh (.) ich sage einfach mal ja, (...) weil die emotionale Verbindung, weil Physik mag zwar ein logisches Fach zu s- äh Fach sein, aber Schüler sind nunmal Menschen und Menschen sind nunmal emotionale Wesen, (unv.) die emotionale Verbindung zu einem Punkt, einer Person, einem Thema, einem was auch immer gefestigt, gesichert und (..) mir fällt kein drittes Adjektiv ein äh gefestigt und gesichert werden soll. (..)
436	Vorlesen, Zustimmung AO + Begründung, Verweis auf frühere Stelle im Test	Du solltest darauf aufpassen, keine Rollenklischees in den Unterricht einzubringen, um die Schüler Schülerinnen nicht in diese Muster zu drängen. Ja, das ist eine generelle Sache, die hatten wir allerdings schon besprochen bei dem (.) Schienenexperiment? Ich nenne es einfach mal Schienenexperiment, äh da haben wir- habe ich da schon lang und breit drüber gesprochen.
437	Sonstiges	Weiter. #02:15:32-9#
438	Vorlesen, Sonstiges	B1: Hier geht es um den Unterricht insgesamt. Oh Junge. (.)
439	Vorlesen	Hier- Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung, Lob und/oder Kritik, erhalten, um sich f- etwas für sich beziehungsweise seinen Unterricht zu lernen? (..)
440	Vorlesen	Unterrichtsziele,
441	Vorlesen	Ablauf der Unterrichtsstunde,
442	Vorlesen	Auswahl,
443	Vorlesen	Umfang der Stundeninhalte,
444	Vorlesen	Medieneinsatz,
445	Vorlesen	Auftreten der Lehrkraft,
446	Vorlesen	Umgang mit Schüler Schülerinnen,
447	Vorlesen	Interaktion,
448	Vorlesen	fachliche Korrektheit.
449	Zustimmung AO, Verweis auf frühere Stelle im Test	Absolut Auftreten der Lehrerkraft, habe ich ja ganz an Anfang schon ein bisschen drüber gelästert. #02:16:05-3#
450	Sonstiges	B1: Natürlich hat er bereits genügend Rückmeldungen gekriegt. #02:16:11-1#

451	Sonstiges*, Sonstiges (3x)	B1: Inter- das überspringe ich jetzt einfach, das habe ich nicht ausgewählt, ich habe keine Ahnung, warum ich hier bin. (...) Gut, dann wähle ich halt eben das oberste aus, aber ich bin dennoch- will dennoch nicht sein, aber bin dennoch hier. #02:16:26-4# B1: Ja. (.) Weiter. #02:16:29-8#
452	Sonstiges, Sonstiges*	B1: Fachliche Korrektheit, auch das habe ich nicht ausgewählt. #02:16:33-3# B1: Warum bin ich hier? #02:16:36-0#
453	Sonstiges*, Sonstiges (2x)	B1: So und natürlich wurde das, worüber ich reden wollte übersprungen. (..) Was zum- (.) Ich würde diesen Test mal auf technische Korrektheit überprüfen, weil da scheint irgendwas im Argen zu sein. (.) Oder sollte das so? (4s) Okay, keine Antwort ist auch eine Antwort. #02:16:57-9#
454		[Video Prompt Schlnote] #02:17:03-8#
455	Sonstiges (2x), Beschreibung (3x), Bewertung (5x), Bewertung + Begründung, Alternative	B1: Uh uh schwierig. (..) Äh (..) das Experiment mit den Schienen war schon ziemlich kacke, die gesamte Aufführung (unv.) mit dem Rücken zum Schüler reden, das war schon etwas, wo ich puh, da hacke ich gerne drauf rum, der Umgang mit den Geschlechterrollen ist auch sehr kritisch, was für mich persönlich tatsächlich ein großer Aspekt (.) ist, weswegen das wiegt auch ziemlich schwer auf, das ges- generelle Auftreten, energielose, ja, bin ich halt eben hier, mache ich das Ding, das ging- das ging mir ordentlich gegen den Strich, muss ein bisschen Energie, bisschen Elan, bisschen Spaß, bisschen ah (schreit) mit dabei sein und (.) er hat nichts davon, er sieht selber schon gelangweilt aus und diese Emotion geht dann auch über auf die Schüler rüber, die dann auch gelangweilt sind (und das ist ja?) das letzte, was man will. (..) Aber es war fachlich korrecht- korrekt, die Merksätze waren okay bis auf den für den zweiten Axiom äh kurz und knackig, auch fachlich korrekt, (.) generelle Aufführung war gut, die Experimente wurden äh haben immer wieder die Schüler mit einbezogen, (.) im Großen und Ganzen würde ich drei minus vier plus sagen, äh (..) weil tief in mir noch ein Optimist und ein Gutmensch ist, sage ich drei (unv.) statt vier. (.) Weiter. #02:18:32-3#

O.2 Interview 2: LD2

Segment	Kodierung	Transkript
0	-	[Einverständniserklärung, Demographie] #00:07:26-4#
1	-	I: Genau, es geht darum, dass Ihnen während der Testdurchführung immer mal wieder Videos begegnen werden und im Anschluss werden dazu Fragen gestellt und ich bitte Sie dann eben logischer, ne, alle Gedanken während der Bearbeitung von diesen Fragen oder den Aufgaben immer laut auszusprechen, (.) so wie wir das bei der Übungsaufgabe auch gemacht haben, und hier jetzt auch die- also wenn sie Texte lesen oder Aufgabenstellungen lesen, die auch eben mit vorzulesen, dann kann ich besser nachvollziehen, wo Sie sind und man bleibt im Redefluss. Genau. Und wie immer gilt: Es geht nicht darum, zu erklären, sondern einfach laut auszusprechen, für sich selber laut zu denken. Wenn Sie aufhören, das zu tun, werde ich mich kurz einklinken und ansonsten ähm bin ich jetzt erstmal ruhig und sitze hier. #00:08:17-0#
2	Vorlesen (2x)	B2: Okay. (..) Okay, zur Situationsbeschreibung. Stellen Sie sich folgendes vor: Sie absolvieren zurzeit ein Praktikum an der Schule. Ihr Mitpraktikant Robert hat eine Doppelstunde seines Physikunterrichts aufgezeichnet und bittet Sie, ihm dazu eine Rückmeldung zu geben. Er zeigt Ihnen Ausschnitte seines Unterrichts und stellt Ihnen dazu Fragen. Um diese Fragen zu- um diese Fragen beantworten- ne, um diese Fragen zu können Sie Robert unter anderem darum bitten, Ihnen noch einmal das Unterrichtsvideo zu zeigen. Aber beachten Sie bitte die Zeit, denn Robert hat nur 70 min Zeit. Deswegen sollten Sie das Video maximal einmal- sollten sie Videos maximal einmal wiederholt werden. Schauen Sie sich also nur in dringenden Fällen das Video ein weiteres Mal an. Robert ist mit seinem Unterricht nicht zufrieden und bittet Sie um ein Feedback. #00:09:06-7#
3	-	[Video Lerngruppe] #00:09:32-5#
4	Überlegung zum Unterricht	B2: (...) Ist das wirklich so gut, wenn die immer ruhig sind? #00:09:38-1#
5	-	[Video Vorwissen] #00:11:53-9#
6	Sonstiges	B2: (..) Okay. #00:11:58-5#
7	-	[Video Unterrichtsziele, Stundenraster] #00:13:13-4#

8	Vorlesen, Sonstiges (4x), Beschreibung	B2: (7s) Schauen wir (jetzt?) das da [Anm: Stundenraster] einmal durch, (.) wie das grob abgelaufen ist. (.) Stellt Thema und Ablauf vor, erklärt Ablauf der Quizrunde, die beantworten Quizfragen, der moderiert Besprechung, Schüler beantworten die Fragen, (.) Vertiefung davon, Junge und Mädchen sprechen über das dritte Newtonsche Axiom, (.) welche Aussage richtig ist, welche nicht, (.) und die Erarbeitung (..) Thematik Impuls mit der Fragestellung, warum konnte die Titanic nicht bremsen, Lehrer leitet Experiment an, Schüler führen das Experiment durch, Lehrer sichert Erkenntnisse aus dem Experiment in einem Merksatz, Schüler schreiben Merksatz zur Sicherung ab. (..) Okay. (..) Erarbeitung und Sicherung, Versuchsaufbau, Versuchsdurchführung, Versuch, die Assistieren, leitet Gespräch zur Sammlung der Beobachtung, die nennen die Beobachtungen, okay, leitet Impulserhaltung aus den Newtonschen Axiomen her, okay, interessant, Aufgabe Überprüfung der Impulserhaltung mit den Messwerten aus dem Experiment, leitet Gespräch zur Besprechung, Schüler nennen Ergebnisse. (Lehrer?) sichert die Erkenntnisse aus dem Experiment in einem Merksatz. Okay. (..) Grober Verlauf also: Quiz, dieser kleine Comic, (...) und zwei kleine- zwei Experimente? Ein Experiment. Ja. #00:14:48-2#
9	Vorlesen, Überlegung zum Unterricht, Sonstiges	B2: Ich habe die Stunde mit der Vorstellung des Ablaufs begonnen. Als nächstes habe ich ein Quiz zur Wiederholung der Newtonschen Axiome angeleitet, welches die Schüler- innen in Partnerarbeit durchgeführt haben. Die Durchführung des Quiz habe ich zwar nicht gefilmt, dafür aber die Besprechung des Quiz mit den Schüler- innen, die ich dir jetzt zeige. Also der Teil Lehrer moderiert Besprechung der Quizfragen, Schülerinnen und Schüler beantworten die Quizfragen. Das war Vorstellung des Ablaufs, Quiz, Wiederholung. Okay. #00:15:20-4#
10	-	[Video Einleitung Reflexion 01] #00:15:39-7#
11	Sonstiges	B2: Ja. #00:15:41-7#
12	-	[Video Szene Reflexion 01] #00:17:44-3#
13	Vorlesen	B2: Okay. Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung, Lob und/oder Kritik erhalten, um etwas für sich bzw. seinen Unterricht zu lernen?
14	Vorlesen, Zustimmung AO (3x), Sonstiges (2x), Bewertung	Formulierung (Wortlaut)- (unv.) okay, Merksatz erstes Axiom, zweites, drittes Axiom (..) mhm ich überlege, wenn ich die Axiome nochmal durchgehe, (.) mhm habe ich zu allen was anzumerken? Beim ersten hätte ich was anzumerken, ja doch, definitiv, beim zweiten (.) was hätte ich da anzumerken? Es ging um die Formel, ja, Kleinigkeit, ähm, (.) braucht er das, um für sich zu lernen? Das Zweite war soweit noch das beste von denen. Das Dritte auf jeden Fall noch, da würde ich noch gerne was zu sagen.

15	Zustimmung AO, Verweis auf frühere Stelle im Test	Zur Gesprächsführung (.) ja, bezieht sich mit was ich zum ersten sagen wollte,
16	Sonstiges*, Bewertung, Zustmmung AO	Ergebnissicherung des Quiz, da geht es wahrscheinlich darum, (.) wie er das da an der Tafel präsentiert hat, das war soweit in Ordnung. (.) Ah, Kleinigkeit, komm, dann kann ich da auch was zu nennen.
17	Vorlesen	So hier können Sie zu bestimmten Fragen Antworten von Robert bekommen.
18	Vorlesen	Kann ich die Inforamtionen zu Lerngruppe noch einmal bekommen?
19	Vorlesen	Kannst du mir noch einmal sagen, welches Vorwissen die Schülerinnen und Schüler haben?
20	Vorlesen	Kannst du mir die Unterrichtsziele noch einmal nennen?
21	Vorlesen	Kannst du mir noch einmal das Stun- Studienras- Stundenraster zeigen?
22	Vorlesen	Kann ich den Unterrichtsausschnitt noch einmal sehen?
23	Vorlesen, Überlegung zum Unterricht	Kannst du mir die PowerPoint-Folien noch einmal zeigen. (.) Okay, soweit habe ich es noch im Kopf. #00:19:18-0#
24	-	I: Darf ich einmal ganz kurz dazwischen grätschen? #00:19:20-7#
25	Sonstiges	B2: Ja. #00:19:21-5#
26	-	I: Ähm, Sie haben jetzt grad bei so ein paar Sachen gesagt: Achso, ja, da- da wollte ich was zu sagen, oder da würde ich was zub sagen. Äh, wenn Sie mögen, können Sie das auch dann ruhig noch mit Aussprechen, also dann haben Sie ja wahrscheinlich was im Kopf, woran Sie denken. #00:19:33-8#
27	Sonstiges	B2: Achso, ich dachte, das kommt jetzt im Folgenden, deswegen habe ich es noch nicht gemacht. Aber ja, okay.
28	Sonstiges, Beschreibung, Überlegung zum Unterricht	Ähm, was hatte ich denn gesagt? Zur ersten, da äh er hatte sich vorher darüber beschwert- was heißt beschwert, ihm ist aufgefallen, dass die Schülerinnen und Schüler die [Anm: Merksätze] irgendwie nur auswendig können, aber damit nichts anfangen können und beim ersten hatte er gefragt, wie das- nur quasi gefragt, wie es benannt wurde, er hat die Antwort erhalten, wie es genannt wird und dann aber selber sofort erklärt, was das bedeutet und da würde ich mir nicht so sicher sein, ob Schülerinnen und Schüler äh- also offensichtlich hat ja keiner der Schülerinnen und Schüler sagen können, was das bedeutet, die haben nur den Namen genannt. Also das hängt so ein bisschen vielleicht zusammen mit dem 'die können das nur auswendig'.

29	Beschreibung, Bewertung (2x), Überlegung zum Unterricht (2x)	Ähm dann generell- (nochmal?) erstmal zum Dritten. Äh da kam am Ende das Ding mit Kraft gleich Gegenkraft. Bis dahin war das eigentlich alles gut, nur ähm gibt es ja diesen Stolperstein Kraft gleich Gegenkraft, dass die Schülerinnen und Schüler das dann irgendwie für ein Kräftegleichgewicht halten und dann würde sich ja nie ein Körper vorwärts bewegen, ähm da steckt ja diese Schülervorstellung irgendwie mit drinne. (.) Ähm die Formulierung fand ich an einigen Stellen, klar, die war richtig, die war wissenschaftlich, aber weiß ich nicht, wie sehr Schülerinnen und Schüler das verstehen würden. (.)
30	Zustimmung AO + Begründung, Überlegung zum Unterricht	Ähm ja und Ergebnissicherung des Quiz, das hatte ich glaube ich auch noch angekreuzt, äh ging es mir so ein bisschen darum, das war ja jetzt vorgefertigt (da?). Inwiefern hat denn das Quiz vorher Einfluss jetzt auf diese Ergebniszusammenfassung genommen? Gefühlt irgendwie gar nichts wahrscheinlich, das was ich so aus dem Unterricht erschließen kann. (..) #00:21:15-5#
31	Sonstiges, Vorlesen	B2: Okay. Dann zu Reflexion eins. Robert hat schon einige Rückmeldungen erhalten. Besonders wichtig ist ihm jetzt noch eine Rückmeldung zu seiner Formulierung des Merksatzes zum 3. Newtonschen Axiom. #00:21:29-4#
32	Vorlesen	B2: So, welches Feedback möchten sie Robert zu seiner Formulierung des Merksatzes zum dritten Axiom geben?
33	Vorlesen	Wählen Sie die Antwortoptionen aus, die Ihrer Einschätzung am nächsten kommt.
34	Vorlesen, Ablehnung AO	Die Formulierung des dritten Axioms ist gut. Ne.
35	Ablehnung AO + Begründung	Fachlich falsch auch nicht richtig, ist ja so gesehen fachlich richtig, da steckt nur der Stolperstein drinne.
36	Vorlesen, Zustimmung AO + Begründung	Die Formulierung des dritten Axioms ist für Schüler- innen nicht angemessen, da sie zu komplex ist, joa, das wäre es vielleicht, wie ich gesagt hatte mit der sprachlichen Formulierung.
37	Vorlesen, Überlegung zum Unterricht	Die Formulierung des dritten Axioms ist nicht gut, da die Vermischung von Beispiel und Merksatz die Schülerinnen verwirren kann. (.) Eventuell trifft das auf dieses Kraft-gleich-Gegenkraft-Argument zu.
38	Vorlesen, Zustimmung AO	Es ist nicht gut, dass nur die Kurzformel Kraft gleich Gegenkraft, aber nicht actio gleich reactio im Merksatz vorkommt. Diese Abkürzung ist nämlich etwas, was die Schüler- innen vielleicht kennen. Okay, vorerst wähle ich das aus, ich denke, das trifft es vorerst am besten.
39	Vorlesen, Ablehnung AO + Begründung	Es ist gut, dass du im Merksatz deutlich machst, dass die Kräfte an verschiedenen Körpern angreifen, da du so einer Verwechslung mit dem Kräftegleichgewicht vorbeugst. Ja, ne, finde ich glaube ich nicht so, weil nur dieses Kraft gleich Gegenkraft irgendwie so im Fokus steht.

40	Vorlesen, Sonstiges	Es ist nicht gut, dass die Kurzformal Kraft gleich Gegenkraft im Merksatz vorkommt, da Schüler- innen diese oft mit dem Kräftegleichgewicht assoziieren. (.) Ähm (.) ja, jetzt muss ich mich entscheiden zwischen dem, was ich angeklickt habe und dem, was da steht. (.)
41	Zustimmung AO + Begründung	Der Kräfte- Ja, das Kräftegleichgewicht ist schon das Ding, was verwirren könnte.
42	Überlegung zum Unterricht	Ob die die Abkürzung actio gl- stimmmt, die Abkürzung acio gleich reactio kennen die vielleicht überhaupt nicht,
43	Bewertung	das- die Verwechslung mit dem Kräftegleichgewicht ist für mich glaube ich das Kritische in dem Moment. #00:23:10-3#
44	Vorlesen (2x)	B2: So, welche der unten stehenden Hinweise zu seiner Formulierung des Merksatzes zum dritten Axiom möchten Sie Robert für eine erneute Durchführung des Unter- der Unterrichtsstunde geben? Wählen Sie die Antwortoptionen aus, die Ihrer Einschätzung am nächsten- (.)
45	Sonstiges	ah, ich darf mehrere auswählen, jetzt weiß ich gar nicht, ob ich das vorher auch durfte.
46	Voelesen, Ablehnung AO	Deine Formulierung des dritten Axioms ist gut so, so kannst du sie wieder verwenden. Nein.
47	Vorlesen, Ablehnung AO + Begründung	Es wäre besser, zuerst das Axiom ohne Beispiel zu notieren und anschließend ein Beispiel, um den Merksatz zu vereinfachen. Finde ich gar nicht so, er strickt das ja so um das Beispiel drum herum.
48	Vorlesen, Sonstiges	Es wäre besser, das Beispiel im Merksatz ausführlicher darzustellen, damit Schüler- innen die Anwendung des dritten Axioms verstehen können. Ah, vielleicht.
49	Vorlesen, Zustimmung AO, Überlegung zum Unterricht, Alternative	Es wäre sinnvoll das Axiom um eine Skizze zu ergänzen, damit es anschaulicher wird. (.) Kann man machen (.), wenn die Skizze gut ist. Wenn die Skizze die Pfeile zum Beispiel am gleichen Punkt zeigt, wird es wieder schwierig. An unterschiedlichen Punkten (.) ja, das- das würde ich sagen, doch.
50	Vorlesen, Ablehnung AO + Begründung	Es wäre besser, im Merksatz zu verdeutlichen, dass Kraft und Gegenkraft gleichzeitig wirken. Deine Formulierung anschließend übt Körper B eine Gegenkraft auf Körper A aus ist fachlich falsch. (.) Ah, ich glaube, das ist nicht so der Aufhänger, weiß ich nicht, ist mir jetzt auch nicht aufgefallen.
51	Vorlesen, Ablehnung AO + Begründung	Du solltest im Merksatz über den Kräften Vektorpfeile notieren, damit die Wirkrichtung der Kräfte deutlich wird. Ja, das würde in der Skizze schon deutlich werden, glaube ich.

52	Vorlesen, Zustimmung AO + Begründung	Du solltest im Merksatz verdeutlichen, dass Kraft und Gegenkraft in genau entgegengesetzte Richtungen wirken, damit die Schüler- innen die Bedeutung des Vorzeichens im Beispiel verstehen. Das würde ich auch sagen, (.) so wie ich das festgestellt habe, verstehen Schüler das Minus meistens nicht, woher das herkommt und warum das wieso da her kommt. #00:24:50-9#
53	-	[Video Einleitung Reflexion 02] #00:25:16-4#
54	Überlegung zum Unterricht	B2: So, zweite Phase, der Comic Junge und Mädchen sprechen über das dritte Newtonsche Axiom. #00:25:24-6#
55	-	[Video Szene Reflexion 02] #00:27:07-0#
56	Bewertung	B2: Okay. Schwieriges Beispiel.
57	Vorlesen	Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung, Lob und/oder Kritik erhalten, um etwas für sich beziehungsweise seinen Unterricht zu lernen? Gesprächsführung,
58	Vorlesen	Einsatz des Comics,
59	Vorlesen	Schülerfehlvorstellungen. (..)
60	Sonstiges	So, zu den drei Sachen:
61	Vorlesen, Zustimmung AO + Begründung, Beschreibung	Gesprächsführung ja, ein bisschen, weil er sofort da rein grätscht. Er stellt die Frage, kleiner zeigt in der gleichen Sekunde auf und er sagt sofort 'kommt schon, Leute'. Schreckt vielleicht sogar noch mehr ab.
62	Vorlesen, Zustimmung AO + Begründung	Einsatz des Comics, ne, finde ich gut, weil das genau so ein Grenzfall ist, irgendwie (.) zwischen erstem und drittem Axiom und im ersten Moment mich auch verwirrt hat
63	Überlegung zum Unterricht	und Schülerfehlvorstellungen weiß ich jetzt nicht, wie gut er sie aufgegriffen hat. Weiß ich noch nicht, wie gut er sie drauf hat. Vielleicht, ja, vielleicht. Einmal mit diesem Kraft gleich Gegenkraft, dass das ein zentraler Knackpunkt ist, dann versteht er vielleicht auch eher, wieso die Schüler damit so ein Problem haben, oder wie er es gesagt hat, wieso die so einen Quatsch geredet haben. #00:28:15-5#
64	Vorlesen	B2: Robert hat schon einige Rückmeldungen erhalten. Besonders wichtig ist ihm jetzt noch eine Rückmeldung zu den aufgetretenen Schülervorstellungen im Unterrichtsausschnitt. Okay. #00:28:25-0#
65	Vorlesen (2x)	B2: Wähle die entsprechenden Antwortoptionen aus, welche der folgenden Schülerfehlvorstellungen sind während des Unterrichtsausschnitts sichtbar geworden?
66	Vorlesen, Ablehnung AO	Es sind keine Schülerfehlvorstellungen- ne, stimmt nicht.
67	Vorlesen, Überlegung zum Unterricht, Ablehnung AO	Denken, dass Kräfte nur dann auftreten, wenn sich zwei Körper berühren. (.) Kommt das so richtig raus? (.) Ich glaube nicht.

68	Vorlesen, Ablehnung AO + Begründung	Die Schüler- innen schreiben dem Apfel eine Kraft als Eigenschaft zu. Nein, das kommt nicht so vor, sonst hätten sie ja gesagt, das Apfel hat eine innere Kraft, vor allem das innere oder irgendwie sowas.
69	Vorlesen, Zustimmung AO + Begründung	Die Schüler- innen denken, dass nur aktive- ja, ja, äh weil er gesagt hat, der Baum steht ja nur.
70	Vorlesen, Ablehnung AO + Begründung	Die Schüler- innen denken, dass einer Kraft immer eine Beschleunigung zugrunde liegen muss und schließen daher eine Kraftwirkung bei ruhenden Gegenständen aus. Das mit der Beschleunigung wurde nicht explizit gesagt, würde ich- würde ich ausschließen.
71	Vorlesen, Zustimmung AO + Begründung	Die Schüler- innen verwechseln Kraft und Gegenkraft mit den wirkenden Kräften- ja, das glaube ich ganz stark, ähm weil beide Aussagen irgendwie (.) irgendwie richtig interpretiert werden können, glaube ich. Also wenn man sagt, es geht um das Wechselwirkungsprinzip, dann ist nur eine richtig, aber generell, wenn man die nur ein bisschen durcheinander wirft, ist es schwierig, die auseinander zu dröseln in dem Moment.
72	Vorlesen (2x), Sonstiges* (3x), Überlegung zum Unterricht, Ablehnng AO + Begründung	Die Schüler- innen sehen nicht, dass der Apfel über den Baum mit der Erde verbunden ist und der Baum durch das Festhalten an der Erde eine Gegenkraft zum am Baum- am Baum ziehenden Apfel ausübt. Nochmal. Die Schüler- innen sehen nicht, dass der Apfel dem Baum- über den Baum mit der Erde verbunden ist und der Baum durch das Festhalten an der Erde eine Gegenkraft zum am Baum ziehenden Apfel ausübt (..). Sehen die das oder sehen die das nicht? Also der Baum hält den Apfel und hält sich selber an der Erde fest. Ne, das würde ich nicht sagen, das verwirrt mich selber ja irgendwie schon. Der Satz, der ist schwierig, vielleicht verstehe ich ihn auch gerade einfach nicht. #00:30:25-3#
73	Vorlesen (2x)	B2: Welche der unten stehenden Hinweise zum Umgang mit Schülerfehlvorstellungen möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben? Wählen Sie die Antwortoptionen aus, die Ihrer Einschätzung nach am nächsten kommen.
74	Vorlesen, Ablehnung AO	Dein Umgang mit den Schülerfehlvorstellungen- nein.
75	Vorlesen, Überlegung zum Unterricht	Du solltest den Prozess der Kraftwirkung thematisieren, um die Schülerfehlvorstellungen aus dem Weg zu räumen. Der Kraftwirkung? Weiß ich nicht, passt das (denn?) so? Was heißt denn Kraftwirkung überhaupt so richtig?
76	Vorlesen, Ablehnung AO + Begründung	Du solltest den Unterschied zwischen Kraft und Energie thematisieren, um zu verdeutlichen, warum die Schülerfehlvorstellungen falsch sind. Ne, die haben das nicht als Speichergröße thematisiert, das ist nicht das Ding.

77	Vorlesen, Zustimmung AO + Begründung, Überlegung zum Unterricht	Du solltest erklären/verdeutlichen, dass Kräfte auch dann wirken können, wenn man nichts, ja, aktiv tut. Dazu kannst du das Rollbrettexperiment verwenden: Zwei Schüler- innen stehe auf je einem Rollbrett und halten je ein Ende einer Schnut. Egal, ob beide Schüler- innen ziehen oder nur eine*r, treffen sich die Rollbretter in der Mitte. Das ist vielleicht eine gute Idee. Ob das mit dem Rollbrett so richtig passt, weiß ich zwar nicht, aber generell der Aspekt mit 'auch passive Körper können das ausüben'.
78	Vorlesen, Zustimmung AO + Begründung	Du solltest mit den Schüler- innen die verwendeten Begriffe und den Unterschied zwischen dem Kräftegleichgewicht und dem dritten Newtonschen Axiom klären, damit ihnen klar wird, warum die Vorstellungen falsch sind. Ja, Kräftegleichgewicht und drittes Newtonsches Axiom, das wurde so ein bisschen durcheinandergeworfen.
79	Vorlesen, Ablehnung AO + Begründung	Du solltest die Schüler- innen das Wechselwirkungsprinzip ausführlich mündlich oder schriftlich formulieren lassen, damit du siehst, was die Schüler- innen verstanden haben, beziehungsweise wo Fehlvorstellungen bestehen. Ne, glaube ich nicht, da laufen die Gefahr, dass sie nur aufschreiben, was sie auswendig gelernt haben und dann ist es vielleicht richtig auswendig gelernt #00:32:00-9#
80	Vorlesen	B2: Reflexion null drei. Bei der Besprechung des Comics hatte ich das Gefühl, dass die Schüler- innen einige Fehlvorstellungen haben. Deswegen habe ich die Schüler gebeten, Stellung zu einer Seite des Comics zu beziehen, also Argumente entweder für die Seite des Mädchens oder des Jungens zu sammeln. Diese Argumente haben wir dann besprochen und Anna hat fachlich korrekt erklärt, warum die Position des Mädchens die richtige ist. Anschließend hatten die Schüler- innen die Möglichkeit, offen gebliebene Fragen zum Comic zu stellen. Da die Schüler- innen hier keine Fragen hatten, habe ich den Unterricht forgeführt und zum Thema Impuls übergeleitet.
81	Überlegung zum Unterricht, Alternative	Da fällt mir direkt ein: Nur, weil jemand keine Fragen hat, heißt das noch nicht, dass er das verstanden hat. (.) Also oft sind die auch einfach schüchtern oder denken, 'jetzt haben wir es drei Mal durchgekaut, jetzt nochmal fragen ist irgendwie peinlich', weiß ich nicht, hätte ich an der Stelle vielleicht nochmal nachgehakt. #00:32:48-8#
82	-	[Video Einleitung Reflexion 03, Video Szene Reflexion 03] #00:34:48-9#
83	Vorlesen	B2: Zu welcher der folgenden Aspekte sollte Robert eine Rückmeldung, Lob und/oder Kritik erhalten, um etwas für sich beziehungsweise seinene Unterricht zu lernen.
84	Vorlesen, Ablehnung AO, Bewertung + Begründung	Einsatz des Beispiels Titanic, (.) ne, finde ich glaube ich ganz gut, das verdeutlicht das Ganze, also die Frage 'ja okay, wieso bremst das Schiff nicht einfach' (unv.) gesunken.

85	Vorlesen, Zustimmung AO + Begründung	Gesprächsführung, ja, weil er drängt die Leute dahin, dass sie vielleicht- also es kann sein, dadurch, dass er sofort eingeht auf die Schülerantworten, er möchte erstmal nur sammeln zu- zu so einem neuen Thema und ähm (.) äh er drängt die Leute ja so ein bisschen dahin, zu sagen, was er hören möchte und wenn man glaubt, man hat nicht das richtige raus, wird man wahrscheinlich auch nichts sagen. Also so findet man auch nicht raus, was die Schüler- innen vielleicht schon für Vorstellungen im Kopf haben, weil er sofort bewertet.
86	Vorlesen, Bewertung, Überlegung zum Unterricht	Äh fachliche Korrektheit und Fachsprache, würde ich so- ist in Ordnung, es heißt zwar noch Schwung und so weiter, aber Schwung wird ja später noch umbenannt und äh das ist nur, damit die Schüler- innen das jetzt verstehen.
87	Vorlesen, Zustimmung AO + Begründung, Alternative	Übergang der Unterrichtsphasen, ja, wirkte gekünstelt, so ein harter Cut, einfach, ja, jetzt machen wir weiter, nächstes Thema komplett was anderes, wir gehen von den Axiomen rüber zum Impuls und da (ist?) einfach nur ein Cut und fertig. Da- da muss man vielleicht anders planen, vielleicht dann genau eine Pause dazwischen legen, wenn es eine Doppelstunde ist oder genau vielleicht noch was dazu besprechen, aber nicht nur einfach so 'Okay, jetzt weiter'. #00:36:15-2#
88	Vorlesen, Sonstiges	B2: Robert hat schon einige Rückmeldungen erhalten. Besonders wichtig ist ihm jetzt noch eine Rückmeldung zur Gestaltung des Phasenwechsels (Anm: Hier fehlt im Text vor der Gestaltung des Phasenwechsels noch die Fachsprache; im Anschluss an dieses Interview ergänzt). Okay, gut, dass ich es angekreuzt hab. #00:36:25-5#
89	Vorlesen	B2: Welches Feedback möchten Sie Robert zur Nutzung der Fachsprache geben?
90	Vorlesen	Bitte wählen Sie die Antwortoptionen aus, die Ihrer Einschätzung am nächsten kommen.
91	Vorlesen, Ablehnung AO + Begründung, Alternative	Es ist nicht gut, dass Schüler- innen den Begriff Schwung nutzen, da die Bedeutung des Begriffs noch nicht geklärt ist und so unklar ist, was genau die Schüler- innen beschreiben. Ne, ich glaub, die verstehen schon, was richtig ist. Die verstehen schon, genau, er sagt ja mit dem- mit dem stabileren Schüler, der dich anrempelt oder der leichtere, ne, das ist schon fast genau das richtige, wir müssen es nur umbenennen, dann muss- das muss nur umgedeutet werden.
92	Vorlesen, Zustimmung AO	Es ist gut, dass Schüler- innen zunächst ihre Alltagssprache verwenden können, um zu erklären, warum es zur Kollision gekommen ist. So sinkt die Hemmschwelle für Schüler- innen, ihre Ideen zu äußern. Ja, das wär so richtig meine Sache.

93	Vorlesen, Zustimmung AO, Überlegung zum Unterricht	Es ist gut, zunächst mit dem alltagssprachlichen Begriff Schwung einzusteigen, da der Begriff Schwung ein tragfähiges Konzept für den Impulsbegriff aus dem Alltag ist. Ja und Schwung ist halt kein physikalischer Begriff, der ist nicht doppelt belegt, sowas wie Kraft.
94	Vorlesen, Ablehnung AO + Begründung	Die Formulierung stärker bremsen ist nicht gut gewählt, da auch Körper mit großem Impuls stärker bremsen könnten, um eine Kollision zu vermeiden. Ne, ne, das ist schon richtig. So ein LKW kann auch nicht stark bremsen, das verstehen die schon.
95	Vorlesen, Ablehnung AO	Es ist nicht gut, dass die Schüler- innen mit dem Begriff Schwung arbeiten- Nein, es ist gut, dass die das machen.
96	Vorlesen, Ablehnung AO, Sonstiges*	Ist nicht gut, dass die Schüler- innen- (.) Schüler- innen in der Benutzung des Begriffs Schwung verstärkst. So förderst du ihre Fehlvorstellungen, zum Beispiel dass beide Begriffe gleichgesetzt werden können und dass Körper mehr Kraft haben, wenn sie mehr Schwung besitzen. (.) Ne, glaube ich nicht. Wo- wo kommt denn hier jetzt noch die Kraft her? Hätte ich jetzt da gar nicht reingedeutet, die arbeiten- die sagen viel mit Geschwindigkeit und Masse, die zwei Sachen für den Impuls. Kraft sehe ich da jetzt nicht.
97	Vorlesen, Sonstiges, Überlegung zum Unterricht, eigene Erfahrung, Ablehnung AO	Es ist nicht gut, dass du so stockend redest und viele Füllwörter nutzt, das lenkt die Schüler- innen vom Inhalt ab. (..) Schwierig. (.) Wäre mir jetzt nicht aufgefallen, dass er super stockend redet, vielleicht ein bisschen, vielleicht ein bisschen Nervösität- (.) Nervosität, aber lenkt das die Schüler ab? Also ich weiß, wenn es richtig extrem wurde, dann- dann haben wir selber auch hinterher die 'Ähms' gezählt und uns nen kleinen Spaß drüber gemacht, aber äh ist das jetzt der Fall? Ich glaube nicht. #00:38:31-2#
98	Vorlesen	B2: So, welche der unten stehen Hinweise zur Nutzung der Fachsprache möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben?
99	Vorlesen	Wählen Sie die Antwortoptionen aus, die Ihrer Einschätzung am nächsten kommen.
100	Vorlesen, Ablehnung AO, Bewertung	Es wäre besser, die Begriffe Masse und Impuls gegebenenfalls mit der zugehörigen Formel einzuführen, ne, ich glaube, soweit als Einstieg ist das okay, wirklich.
101	Vorlesen, Ablehnung AO, Alternative	Es wäre besser, wenn die Schüler- innen- wenn du Schüler- innen dazu aufforderst, mehr Fachbegriffe zu nutzen, zum Beispiel den Massebegreiff, damit sie sich- nein, kann man selber später machen.
102	Vorlesen	Es wäre besser, den Impulsbegriff für den Schwung schon früher einzuführen, um zu verdeutlichen, um welches physikalisches Konzept es geht.

103	Vorlesen, Zustimmung AO, Sonstiges	Du solltest gegen die Schülervorstellung vorgehen, dass Körper Schwung haben, vergleichbar mit der Schülervorstellung Kraft haben, damit sich diese nicht verfestigt. Schwung haben, habe ich gar nicht so darauf geachtet, wenn das so stark rausgekommen ist, dann müsste man das machen. (..) Ich merke es mir mal.
104	Vorlesen	Die Verwendung der Begrifflichkeiten war angemessen, das kannst du wieder so durchführen.
105	Vorlesen, Zustimmung AO + Begründung	Du solltest dir in der Vorbereitung genauer überlegen, was du den Schüler- innen erzählen möchtest, damit du flüssiger und motivierender vortragen kannst. Ja, schadet eigentlich nie.
106	Sonstiges	Das [Anm: AO "Die Verwendung der Begrifflichkeiten war angemessen, das kannst du wieder so durchführen."] lassen wir auch mal angekreuzt.
107	Vorlesen, Ablehnung AO + Begründung	Es wäre besser die Schüler- innen zur Verwendung der Fachsprache aufzufordern, um zu vermeiden, dass sie die Konzepte Schwung und Impuls weiterhin verwechseln. (.) Im Einstieg glaube ich noch nicht. (.) Weil wenn man sie zur Verwendung auffordert, kriegen sie vielleicht Angst, was zu sagen, ne.
108	Ablehnung AO	Und das 'Schwung haben' [Anm: AO "Du solltest gegen die Schülervorstellung vorgehen, dass Körper 'Schwung haben' (vergleichbar"mit 'Kraft haben'), damit diese sich nicht verfestigt."], wäre es mir stark aufgefallen, würde ich es jetzt ankreuzen, aber ich- vielleicht habe ich auch nicht richtig zugehört. Ich meine, das wäre nicht so die Problematik gewesen in dem Moment. #00:40:10-3#
109	Vorlesen	B2: So, Robert zur Gestaltung des Phasenübergangs geben.
110	Sonstiges	Hier- Hier ist meine Meinung schon ein- schon deutlicher.
111	Vorlesen, Zustimmung AO	Der Phasenübergang ist dir nicht gut gelungen. Ja, zu abrupt, zu kurz, ja, genau.
112	Vorlesen, Zustimmung AO, Beschreibung, Überlegung zum Unterricht	Phasenübergang ist dir nicht gelungen, da die Inhalte zum Wechselwirkungsprinzip nicht ausreichend verstanden und gesichert wurden. Ja, er hat ja gesagt, (da?) wurde drüber geredet, da habe ich mich auch gefragt, wie- (.) bisschen fraglich. (.) Er hat gesagt, er hat das mit denen nochmal geklärt, aber die Folie war noch mal die gleiche, es gab kein neues Tafelbild, so wie das schien auf jeden Fall, kein neues Tafelbild, äh die gleiche Folie, wie vorher auch, also der aktive Unterricht nicht keinen Einfluss auf seine Planung, er zieht so durch, gefühlt.
113	Vorlesen, Ablehnung AO	Es ist unproblematisch- ne, ne, fin- finde ich doch, ist problematisch.
114	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du so einen abrupten Übergang zwischen den Themen gemacht hast, da der rote Faden der Stun- der Stunde, ja, glaube ich auch. (Es?) ist so ein künstlicher Cut.

115	Vorlesen, Bewertung + Begründung, Alternative	Es ist nicht gut, dass du sagst, das Thema ist komplett neu, da den Schüler- innen so das bisher Gelernte nicht mehr relevant erscheint. (unv.) Komplett neu ist schon in Ordnung, (..) weil es ja (auch?) ein neues Thema ist und irgendwie muss man einen Cut machen und man muss auch sagen, grenzt euch mal kurz ab von unseren jetzigen Konzepten und gehen wir über zum nächsten.
116	VorlesenÜberlegung zum Unterricht, Ablehnung AO + Begründung	Es ist nicht gut, dass du sagt, das Thema ist komplett neu, da es einen fachlichen Bezug zwischen Newtonschen Axiomen und dem Impulsbegriff gibt. Ja, stimmt, da gibt es einen, aber der verwirrt glaube ich mehr, das ist schon so ganz gut. #00:41:35-9#
117	Vorlesen (2x)	B2: Welche der unten stehen Hinweise zur Gestaltung des Phasenübergangs möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben? Wählen Sie die Antwortoptionen aus, die Ihrer Einschätzung am nächsten kommen.
118	Vorlesen, Ablehnung AO	Phasenübergang hast du gut gestaltet, ne.
119	Vorlesen, Ablehnung AO + Begründung	Es wäre besser, den Zusammenhang zwischen den Newtonschen Axiomen und dem Impu- Impulsbegriff zu verdeutlichen, da die Schüler- innen den Zusammenhang nicht unbedingt selbst erkennen können. (.) Noch nicht, es ist nur ein Einstieg.
120	Vorlesen, Zustimmung AO, Verweis auf frühere Stelle im Test	Es wäre besser, den Einstieg in den Impulsbegriff auf die nächste Stunde zu verschieben, um den Schüler- innen zunächst noch mehr Möglichkeiten zu geben, das Wechselwirkungsprinzip zu verstehen, ja,wie ich eben auch schonmal angedeutet hatte.
121	Vorlesen, Überlegung zum Unterricht	Es wäre besser, vor dem Einstieg in das neue Thema auch die Tafel zu wischen, damit die Schüler- innen sich voll auf das neue Thema konzentrieren können und nicht noch über das alte nachdenken. (..) Weiß ich nicht, denken Schüler so viel darüber nach, was noch an der Tafel steht? Ist das überhaupt ein signifikanter Effekt? Keine Ahnung.
122	Vorlesen, Zustimmung AO	Es wäre besser, einen glatteren Übergang zu gestalten, zum Beispiel durch eine kurze Verknüpfung (unv.), wie 'Wir finden jetzt etwas neues mit dem gelernten Wissen raus', ja, damit der rote Faden nachvollziehbar ist. Ja, das- das wär was gewesen, wär ne Möglichkeit.
123	Vorlesen, Überlegung zum Unterricht, Ablehnung AO	Wäre besser, den Übergang durch das Zeigen eines Experiments zu gestalten, (.) tja, was für ein Experiment? Ne, glaube ich nicht. #00:42:50-0#
124	-	[Video Einleitung Reflexion 04] #00:43:02-3#

125	Überlegung zum Unterricht, Vorlesen, Sonstiges	B2: So, wir sind in der Sicherungsphase. Er macht ein Experiment, führt das durch, sichert Erkenntnisse aus dem Experiment mit einem Merksatz, bin ich gespannt, wie er das macht, Schülerinnen und Schüler schreiben Merksatz zur Sicherung ab. #00:43:19-0#
126	-	[Video Szene Reflexion 04] // #00:43:28-0# #00:45:18-3#
127	Überlegung zum Unterricht	B2: Skizze bei dem Apfel hab ich tatsächlich übersehen, naja. #00:43:31-1#
128	Beschreibung	B2: (Naja?), okay, er schreibt etwas an und dann ist das (gesetzt?) (unv.) nicht geklärt, ob es verstanden wurde. #00:45:26-3#
129	Vorlesen	B2: Zu welcher der folgenden Aspekte sollte Robert eine Rückmeldung erhalten, um etwas für sich beziehungsweise seinen Unterricht zu lernen.
130	Vorlesen, Zustimmung AO + Begründung (2x)	Eingesetzte Beispiele, das mit dem Ball, (.) ja doch kleines Feedback, weil die haben das Experiment gemacht, aber erstmal müssen wir darauf vertrauen, dass die das ordentlich machen, dass sie das nicht gefährlich machen und irgendeinen Scheiß anstellen, ähm aber vor allem 'wirf ihn immer mit gleicher Geschwindigkeit' (.) schwierig. Weiß ich nicht, ob er das kann, müsste man gucken.
131	Vorlesen, Bewertung, Beschreibung	Fachsprache (.) äh Fachsprache soweit in Ordnung, doch, er macht ja den Übergang zum Impuls, hätte ich auch so- hätte ich ganz gut gefunden so.
132	Vorlesen, Bewertung + Begründung, Beschreibung, Alternative + Begründung	Gesprächsführung und Ergebnissicherung, die finde ich schwierig. Er sagt zum Beispiel wortwörtlich 'wir in der Physik'. Gefühlt erzeugt das meiner Meinung nach so eine kleine Ausgrenzung gegenüber jedem Schüler, der sich nicht mit der Physik verbunden fühlt, dessen Interesse das nicht ist. (Und?) wofür sollte ich etwas lernen, wofür sollt- warum sollte ich eine Sprechweise ähm einer Gemeinschaft annehmen, die mich nicht interessiert, der ich nicht zugehöre und der ich nicht zugehören will, das ist so, das ist schwierig, das 'wir' hätte ich nicht so gesagt. (Hätte ich?) eher gesagt 'man sagt in der Physik' oder sowas, das ist ein bisschen- dann dann grenzt man die Leute nicht direkt aus.
133	Vorlesen, Beschreibung (2x), Überlegung zum Unterricht (2x), Alternative + Begründung	Und Ergebnissicherung, ja okay, er hat gezeigt, dass es abhängig davon ist, aber sehen das die Schüler sofort, dass ich daraus die Formel ziehen kann? Und dann zu sagen 'ja, schreibt das jetzt ab, das ist richtig', weiß ich nicht, ich hätte noch gefragt, ob Schüler nicht irgendwas anderes noch sehen, ob die- wie die das schätzen oder dass die Schüler sagen, okay, was ist euch jetzt- die mal fragen: 'Was ist euch jetzt aufgefallen?' äh, 'wovon scheint das anscheinend abhängig zu sein? Ist das von was anderem noch abhängig', dass die ein bisschen was noch melden dürfen, sehr- sehr von ihm- auf ihn zentriert. #00:47:21-8#

134	Vorlesen	B2: Er hat schon einige Rückmeldung erhalten, besonders wichtig ist jetzt noch die Wahl des Experiments, zur Ergebnissicherung nach dem Experiment. #00:47:29-5#
135	Alternative + Begründung, Verweis auf frühere Stelle im Test	B2: Ja es- vorab, das Experiment: Ja, man hätte ein anderes nehmen können, ich habe ja zwei Sachen schon dazu gesagt mit- (.) ich kann mir vorstellen, dass die ein bisschen Scheiß machen und zu feste werfen oder sowas. Das ist stark abhängig davon, was zwei Schüler in dem Moment fühlen und was ist, wenn die was anderes fühlen? Das ist nicht hundert Prozent offensichtlich. (.)
136	Vorlesen, Zustimmung AO	So, das Experiment ist gut gewählt, da die beiden Schüler das Experiment durchführen den Impuls am eigenen Körper erfahren können, (.) ja, joa, Schüler können das selber machen, schon ganz gut.
137	Vorlesen, Beschreibung, Überlegung zum Unterricht	Das Experiment ist gut gewählt, da es den Einfluss von Masse und Geschwindigkeit auf den Impuls gut darstellt. Ja, er variiert die beiden Größen. (.) Ähm, (.) sieht man den direkten Einfluss (unv.) er sagt, die sind unterschiedlich schwer, ich weiß nicht, ob ein anderes Experiment das nicht besser machen könnte.
138	Virlesen, Zustimmung AO, Verweis auf frühere Stelle im Test, Beschreibung	Es ist nicht gut, dass nur über die Masse und die Geschwindigkeit gesprochen wurde, da unklar bleibt, ob der Impuls noch von anderen Größen, wie zum Beispiel der Form abhängen kann. Ja, was ich gesagt habe, die Schüler durften sich ja auch nicht melden.
139	Vorlesen, Sonstiges, Zustimmung AO + Begründung, Ablehnung AO + Begründung	Experiment ist nicht gut gewählt- also ich habe jetzt ist gut gewählt und ist nicht gut gewählt, weil irgendwie ist da (noch?) so ein Zwiespalt. Einerseits finde ich es ganz gut, weil sie es selber machen können und andererseits weiß ich nicht, wie gut die Erkenntnis daraus ist.
140	Vorlesen, Beschreibung	Nicht gut gewählt, da es den Zusammenhang beziehungsweise Abgrenzung des Impulses zur kinetischen Energie nicht veranschaulicht. (.) Stimmt, das tut es nicht, da findet kein Unterschied statt, (.) da hatten die Schüler aber bis jetzt kein Problem mit. Keiner von denen hat Energie gesagt.
141	Vorlesen, Zustimmung AO	Es ist gut, dass du die beiden Schüler in das Experiment einbezogen hast, so motivierst und aktivierst du sie. Ja, ist gut.
142	Vorlesen, Überlegung zum Unterricht	Es ist gut, dass du vorgibst, wer nach vorne kommen soll, so vermeidest du zu lange Organisations- Wartezeiten nach der Frage, wer will denn. Hat er es denn so vorgegeben? Ich weiß ja nicht.
143	Vorlesen, Alternative	Es ist gut, dass du nur zwei Schüler das Experiment durchführen lässt und nicht alle Schüler- innen, da du so Zeit sparst und das Experiment so einfach ist, dass es auch von den sitzenden Schüler- innen gut nachvollzogen werden kann. Ja, wenn es doch so einfach ist, dann lass es doch alle machen, (.) genug Bälle sind wahrscheinlich da. #00:49:31-3#

144	Vorlesen	B2: So, äh untenstehende Hinweise zur Wahl des Experiments Robert für eine erneute Durchführung geben.
145	Vorlesen	Experiment ist gut gewählt, kannst es erneut einsetzen.
146	Vorlesen, Ablehnung AO, Alternative + Begründung	Es wäre besser, statt des Experiments ein Video zu zeigen in welchem gezeigt wird, welchen Einfluss verschiedene Massen und Geschwindigkeiten auf den Impuls haben. Ne, Schüler können das doch selber erfahren, das ist recht anschaulich eigentlich.
147	Vorlesen, Zustimmung AO + Begründung	Es wäre besser ein anderes Experiment durchzuführen, das für alle Schülerinnen und Schüler (anschaulich?) ist, zum Beispiel eine Person, die auf einem Skateboard steht und einen Ball weg wirft oder das Werfen beider Bälle in den Sand zur Betrachtung der Krater. (.) Das ist sogar vielleicht noch besser, weil das ein bisschen unabhängiger davon ist, wie die Schüler werfen. Wenn sie circa gleich werfen. Der Krater ist deutlicher, als wenn der andere sagt, 'ja, ist schwerer, fühlt sich nach mehr Schwung an'.
148	Vorlesen, Zustimmung AO, Alternative	Es wäre besser, ein Stoßexperiment durchzuführen, da du so gut auf die Impulserhaltung überleiten kannst. Das habe ich auch überlegt mit zwei verschieden schweren ähm Blöcken auf- auf einer Bahn.
149	Vorlesen, Ablehnung AO + Begründung	Es wäre besser, ein Experiment zu wählen, zu dem eine Messreihe aufgenommen werden kann, ne, brauch glaube ich gar nicht, den Effekt- je nachdem, wie viel besprochen wurde vorher. Wenn es denen klar ist, dass proportionaler Zusammenhang so verdeutlicht werden kann, dann reicht das.
150	Vorlesen, Zustimmung AO + Begründung	Es wäre besser, wenn du das Experiment von allen Schüler- innen durchführen ließest, weil es den Schülerinnen gut tut, sich zwischendurch zu bewegen. Ja, das macht Physik auch ein bisschen aus.
151	Vorlesen, Ablehnung AO + Begründung, Bewertung	Du solltest darauf achten, dass die Schüler vorne so stehen, dass alle sitzenden Schüler erkennen können, wie leicht schwer es ist, den Ball zu fangen. Ne, das passte schon, die standen ja seitlich, man hat gesehen, wie sein Arm so nach hinten ging, das passte. #00:51:03-1#
152	Sonstiges	B2: So, zur Ergebnissicherung, ja da habe ich ein bisschen mehr.
153	Vorlesen, Zustimmung AO + Begründung	Es ist nicht gut, dass du Schüler- innen einen fertigen Merksatz vorgibst, da das reine Lesen und Abschreiben nicht verständnisförderlich ist. Ja, es ist sofort (unv.), wo das her kommt.
154	Vorlesen, Zustimmung AO, Bewertung (2x)	Es ist gut, dass Schüler- innen den Merksatz ins Heft übertragen, damit alle Schüler- innen den Merksatz als Sicherung haben. Ja, das stimmt, ist gut, dass man den hat (und?) wie man da hin kommt, ist das Problem.

155	Vorlesen, Zustimmung AO, Sonstiges	Es ist nicht gut, dass du die Ergebnisse des Experiments nicht sicherst, so ist unklar, was die Schüler- innen aus dem Ex- Experiment mitnehmen sollen. (.) Ja, stimmt, wäre gar nicht schlecht gewesen, (...) ja, würde ich auch noch ankreuzen.
156	Vorlesen, Beschreibung	Der Merksatz ist gut gelungen, da die Je-umso-Sätze die Zusammenhänge zwischen Masse, Geschwindigkeit und Impuls nachvollziehbar darstellen. Mhm, fällt so ein bisschen vom Himmel.
157	Vorlesen, Überlegung zum Unterricht	Es ist nicht gut, dass du den Impuls als skalare Größe einführst, so grenzt du den Impuls nicht ausreichend von der kinetischen Energie ab. Ist noch nicht vorgekommen das Problem, glaube ich.
158	Vorlesen, Ablehnung AO	Es ist gut, dass du den Begriff der Energie nicht erwähnst, so vermeidest du die Verwirrung bei den Schülerinnen, die verschiedene Fachbegriffe auf Grundlage ihrer Erfahrungen nicht trennen können. (.) Vielleicht, vielleicht, aber ich glaube, das ist nicht so das, was ich sagen wollte.
159	Vorlesen, Ablehnung AO, Alternative + Begründung	In deiner Sicherung fehlt eine Abgrenzung des Impulsbegriffs von der kinetischen Energie und Kraft, da die Begriffe den Schüler- innen zwar bekannt sind, aber Unterschiede für sie nicht ersichtlich sind. Ne, würde ich einfach im Anschluss irgendwann machen. Da ist das noch fehl am Platz, glaube ich. #00:52:27-2#
160	Vorlesen	B2: So unten stehenden Hinweise zur Ergebnissicherung für eine erneute Durchführung.
161	Vorlesen, Ablehnung AO	Experiment hast du gut gesichert, ne.
162	Vorlesen, Zustimmung AO	Es wäre besser, die Beobachtung des Experiments schriftlich an der Tafel zu sammeln, ja.
163	Vorlesen, Zustimmung AO, Verweis auf frühere Stelle im Test	Die Schüler- innen in die Sicherung mit einzubeziehen, zum Beispiel Merksatz formulieren zu lassen, damit sie stärker involviert, aktiviert- ja, mehr Schülermeldungen, was ich vorher sagte.
164	Vorlesen, Ablehnung AO + Begründung	Es wäre besser, zusätzlich zum Merksatz noch ein Beispiel zu zeigen, um auch die Formel darzustellen. (.) Weiß ich nicht, da war ja schon ein Beispiel.
165	Vorlesen, Ablehnung AO + Begründung, Überlegung zum Unterricht	Es wäre besser, in der Sicherung auch auf das Experiment einzugehen, zum Beispiel indem du auf der Folie zum Merksatz auch auf das Experiment eingehst oder Skizzen ergänzt, um die Inhalte stärker zu vernetzen. (.) Ne, glaube ich nicht, brauche ich da Skizzen? Noch nicht, nicht für- für einen Einstieg.
166	Vorlesen, Ablehnung AO, Verweis auf frühere Stelle des Tests	Es wäre besser, den Impuls als vektorielle Größe- ne, habe ich auch schon gesagt.

167	Vorlesen, Ablehnung AO, Alternative + Begründung	Es wäre besser, auch auf die kinetische Energie der Bälle einzugehen- nein. (.) Kein Vorwissen aktivieren, dann-dann kommt es zu Verwirrung, glaube ich. #00:53:25-0#
168	Vorlesen, Beschreibung (4x), Bewertung, Überlegung zum Unterricht (2x)	B2: Nachdem wir das Experiment mit den Tennisbällen abgeschlossen und die Schüler- innen den Merksatz zum Impuls abgeschrieben haben, habe ich zu einem Demonstrationsexperiment übergeleitet, an dem die Impulserhaltung- an dem ich die Impulserhaltung veranschaulichen möchte. Mit dem Demonstrationsexperiment möchte ich einen Auffahrunfall simulieren. Den Schülerinnen habe ich dann erklärt, wie der Auffahrunfall im Experiment dargestellt wird und wie die Messwertaufnahme funktioniert. Den experimentellen Aufbau und eine Messwerttabelle habe ich vor der Stunde vorbereitet. Hier habe ich dir zwei Fotos vom Aufbau und von der Tabelle mitgebracht. (.) Okay, scheinen zwei Wagen zu sein, (.) wahrscheinlich, so wie das aussieht, so- so eine Luftschwebebahn, damit die Reibung weg ist, okay. Unten jede Menge Kabel, hatte er das vorbereitet? Moment. Äh veranschaulichen möchte mit dem Demonstrationsexperiment, ich möchte simulieren, habe ich vor der Stunde vorbereitet. Okay, er hat das gemacht, alles klar. Ein bisschen kompliziert sonst. Messung eins, Messung zwei, drei, vier, zwei verschiedene Massen, zwei verschiedene Geschwindigkeiten, vor dem Stoß, nach dem Stoß. Ach, die pinnen dann zusammen? Ach, die haben diese- ah, okay, diese- diese Federspitzen. Okay, dann sind wir in der Erarbeitung und er führt den Versuch durch. Schauen wir mal. #00:54:46-5#
169	-	[Video Einleitung Reflexion 05] #00:54:54-8#
170	Sonstiges	B2: Kommt drauf an, wie es durchgeführt wurde. #00:54:59-9#
171	-	[Video Szene Reflexion 05] #00:57:25-0#
172	Sonstiges	B2: Oh, viel viel viel viel fällt mir ein. (lacht) So viel, dass ich schon fast zwischendurch stoppen wollte zum Sagen, aber erzähl ich alles jetzt. #00:57:35-2#
173	Vorlesen	B2: So, zu welchen der Aspekte sollte er Lob und oder Kritik erhalten?
174	Vorlesen, Zustimmung AO + Begründung	Äh Umgang mit den Schüler- innen ja, (.) das ist fast schon Sexismus eigentlich, finde ich, zu sagen, so Mädchen nach vorne, schöne Handschrift und Jungs, ihr könnt das besser ablesen, das fördert doch nur (.) das Klischee in der Physik, das hilft nicht.
175	Vorlesen, Zustimmung AO, Beschreibung, Alternative	Äh Notation der Messwerte, ja, sie sollte das aufschreiben, er hat ihr das vorgegeben, dann ja, vergiss die Einheiten nicht, die Einheit gibt er ihr auch noch vor, ähm ich würde vielleicht sagen, warum man die Einheit dazu schreiben sollte und warum ich zu dem Messwert komme.

176	Vorlesen, Bewertung + Begründung	Äh Auswertung- Durchführung und Auswertung des Experiments durch die Lehrkraft (.) na die Durchführung ist ja soweit okay, er macht das Ding an, er schiebt das an, (nochmal?), dass alle das sehen können, so ist in Ordnung.
177	Beschreibung (2x), Alternative, Alternative + Begründugn	Messunsicherheiten gar nicht thematisiert, nicht da, da wurde auch gesagt, äh dass man zum Beispiel- ich wa-ich wieg nur den einen, weil die sind ja Baugleich, macht man ja in der Wissenschaft nicht so, ich würde trotzdem beide wiegen und null Komma eins, das ist ja eine ziemlich ungenaue Waage, ähm ich würde trotzdem das Ganze als Unsicherheit mit aufnehmen, weil so denken die Leute, das Experiment ist äh ohne Fehler und ohne Unsicherheit, ist ja nicht der Fall, das vermittelt so ein leicht falsches Bild.
178	Vorlesen, Bewertung + Begründung, Alternative + Begründung	Wahl des Experiments, experimenteller Aufbau, das ist eigentlich ein gutes Experiment, glaube ich, das veranschaulicht das ganz gut, äh, da müsste man vielleicht zu sagen, äh warum mache ich jetzt das Gebläse an, was hat das für einen Sinn? (.) das wird nicht so ganz ersichtlich.
179	Vorlesen, Zustimmung AO, Beschreibung (2x)	Aktivierung, ja, einfach nur zu sagen so, schreib das jetzt auf, vergiss das nicht, äh aktiviert nicht so wirklich und zu sagen, wer (kann?) jetzt hier- die Jungs können das besser ablesen, da hat ja dann ein Mädchen keine Lust mehr drauf überhaupt, selbst wenn es vorher gewollt hätte, danach will es nicht mehr.
180	Vorlesen, Sonstiges*	Eingesetzte Beispiele (.) was denn für Beispiele? Weiß ich nicht. #00:59:22-7#
181	Vorlesen	B2: Robert hat schon einige Rückmeldungen erhalten. Besonders wichtig ist ihm jetzt noch eine Rückmeldung zum Umgang mit den Schüler- innen, dazu, wie er während des Experiments auf Messunsicherheiten eingegangen ist und zur Aktivierung. Mhm[ja]. #00:59:34-9#
182	Sonstiges	B2: Kann man viel zu sagen, glaube ich.
183	Vorlesen	Wählen Sie die Antwortoptionen aus, die Ihrer Einschätzung am nächsten kommen.
184	Vorlesen, Zustimmung AO + Begründung	Ist nicht gut, dass du sagst, dass Mädchen eine schöne Schrift haben und Jungen besser- ja, genau, die werden nur darauf reduziert.
185	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du die Mädchen nur für Fleißarbeit-ja, Stereotype gefördert, mhm[ja] genau was ich gedacht habe.
186	Vorlesen, Bewertung, Ablehnung AO + Begründung	Es ist gut, dass du die Mädchen für ihr Selbstkonzept lobst, um ihr Selbstkonzept im Physikunterricht zu stärken. Lob ist gut, aber nur dafür zu loben, dass die Handschrift so toll ist, das ist es nicht.

187	Vorlesen, Ablehnung AO, Alternative	Es ist nicht gut, dass du das Lob für die Handschrift der Mädchen mit einer inhaltlichen Kritik, Vergessen von Einheiten, kombinierst, da die Ursache hier nicht das Geschlecht ist. Kam mir jetzt nicht so vor. Ich hätte einfach generell anders reagiert auf die vergessene Einheit, ich hätte es am Ende vielleicht thematisiert oder oben drüber geschrieben m in Kilogramm und dann hätte ich gesagt 'okay, du hast die jetzt weggelassen, das darfst du nur, weil da in Kilogramm drüber steht'.
188	Vorlesen, Ablehnung AO, Beschreibung, Alternative	Ähm es ist nicht gut, dass du Schülerin schon vor dem Notieren des zweiten Messwerts erneut auf die Einheit hinweist, da du so das Bild festigst, Mädchen wären in Physik nicht kompetent. Ne, er wollte es ordentlich machen, aber hätte man am Ende besprechen können.
189	Vorlesen, Beschreibung	Es ist nicht gut, dass du der Schülerin an der Tafel direkt sagst, welche Einheit fehlt, da du ihr so unterstellst, sie wüsste das nicht. (.) Wurde ja nicht mal die Chance gelassen, zu zeigen, dass sie es weiß.
190	Vorlesen, Zustimmung AO, Alternative, Bewertung	Es ist in Ornung, dass du sagst, dass Paula aufgrund ihrer schönen Handschrift die Werte notieren soll, da es wichtig ist, dass alle Werte an der Tafel gut lesen können. Ja, ist wichtig, aber dann kann ich es auch mit dem PC machen, das muss dann nicht- (weiß ich nicht?) schwierig, ganz schwierig. #01:00:56-3#
191	Vorlesen	B2: Welche der unten stehenden Hinweise zum Umgang mit den Schüler- innen möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben?
192	Vorlesen, Ablehnung AO	Dein Umgang mit Mädchen und Jungen ist in Ordnung. Nein.
193	Vorlesen, Überlegung zum Unterricht, Alternative	Du solltest eine Geschichte um das Experiment herum konstruieren und zum Beispiel in Form eines Videos zeigen, um die Mädchen zu motivieren. (.) Eine Geschichte ist vielleicht nicht schlecht, aber um die Mädchen zu motivieren, weiß ich nicht. Also man könnte sagen, ja, wir simulieren jetzt einen Auffahrunfall, aber beim Auffahrunfall gibt es ja auch Reibung, irgendwie so ein- was hat das für eine Relevanz das Experiment, wo finde ich das und was kann ich damit sehen?
194	Vorlesen, Zustimmung AO	Es wäre besser, die Schüler- innen zu fragen, wer freiwillig schreiben beziehungsweise die Werte ablesen möchte, um den Schüler- innen nicht das Geschlecht einer Aufgabe zuzuweisen, ja. (.)
195	Vorlesen, Ablehnung AO, Alternative, Alternative + Begründung	Es wäre besser, die Aufgaben, Aufschreiben, Werte ablesen, durchzutauschen, damit alle Schüler- innen eine Aufgabe wahrnehmen, ne, kann man schon gleich lassen, man sollte nur gucken, dass nicht jede Stunde immer die gleichen Leute kommen, aber sonst hast du zu viel Gewusel.

196	Vorlesen, Zustimmung AO	Du solltest darauf achten, die Mädchen nicht für ihre Handschrift, sondern explizit für ihre physikalische Kompetenz zu loben, da du sonst das Bild vermittelst, Physik wäre ein Jungs- wäre ein Jungsfach, ja. (.) (Wird?) ein Problem.
197	Vorlesen, Zustimmung AO	Du solltest auf Geschlechtsspezifische Zuschreibungen verzichten, dass du so ein unangene- unangemessenes Verhältnis zwischen den Geschlechtern im Physikunterricht erzeugst, ja.
198	Vorlesen, Zustimmung AO, Alternative	Du solltest grundlegend an deiner Einstellung gegenüber Mädchen und Jungen arbeiten, um dein stereotypisches Verhalten zu ändern, ja. Vielleicht selber mal überdenken, was einem nicht auffällt. #01:02:24-7#
199	Vorlesen	B2: Welches Feedback möchten Sie Robert zu seinem Umgang mit Messunsicherheiten geben?
200	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du sagst, du könntest exakt oder genau messen, da du die Schülervorstellung förderst, man könnte die größ- ja, richtig.
201	Vorlesen, Zustimmung AO, Verweis auf frühere Stelle im Test	Du gehst nicht auf die Messunsicherheiten ein, das ist nicht gut, da diese eine wichtige Rolle spielen. Genau das hatte ich eben auch so im Kopf, oder hab ich auch sogar gesagt.
202	Vorlesen, Zustimmung AO + Begründung, Alternative + Begründung	Es ist nicht gut, dass du nur einen der beiden Wagen wiegst. Ja, das auch, das vermittelt (so?) ich mache eine Messung und die ist final. Stattdessen vielleicht nehme ich eine genauere Waage, die zwei oder drei Nachkommastellen macht und bilde (den?) Mittelwert, ich (mache?)- ich wiege das Ding fünf Mal, ich wiege beide Wagen fünf Mal, um zu zeigen, okay, wenn ich etwas messe, ist das nicht ein- ein- ein Wert, der von oben fällt, sondern den habe ich schon genau herausbekommen.
203	Vorlesen, Zustimmung AO, Alternative	Es ist nicht gut, dass du sagst, die Wagen wären gleich schwer, da das auch wenn sie baugleich sind nicht stimmen muss. Richtig, das muss (immer?) überprüft werden.
204	Vorlesen, Zustimmung AO, Sonstiges	Es ist nicht gut, dass du das Gewicht mit null Komma eins Kilogramm angibst, da du hier nicht berücksichtigst, welche die signifikanten Nachkommastellen sind. Auch das, einfach alles ankreuzen.
205	Vorlesen, Ablehnung AO	Es ist gut, dass du darauf verzichtest- ne, ne, muss drauf eingehen, egal, was jetzt hier als Begründung steht.
206	Vorlesen, Ablehnung AO, Alternative + Begründung	Es ist gut, dass du das Gewicht des Wagens mit null Komma eins Kilogramm angibst, um Fehler in der Berechnung des Impulses aufgrund vieler Nachkommastellen zu vermeiden. Ne, kann man ja schon machen und die Abweichung wird nicht riesig sein vermutlich. #01:03:44-7#
207	Vorlesen	B2: Welche der unten stehen Hinweise zu deinem Umgang mit den Messunsicherheiten möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben? (6s)

208	Vorlesen, Ablehnung AO + Begründung	Es ist gut, wie du mit Messunsicherheiten umgegangen bist, das kannst du erneut so machen. Nein (.) er hat es ja weg gelassen.
209	Vorlesen, Ablehnung AO, Sonstiges, Alternative	Es wäre besser, alle Nachkommastellen anzugeben, die die Laborwaage anzeigt, da du den Fokus auf genaue Messungen legst und die Laborwaage genauer als null Komma ein Kilogramm- (.) 'alle' Schwierig, bisher kommt es dem am nächsten, ich gucke mal, was noch kommt, also so zwei, drei hätte ich mir gewünscht, aber ich weiß nicht, wie viele sie hat, vermutlich bis auf Mikrogramm genau und das wäre natürlich zu viel. (.)
210	Vorlesen, Zustimmung AO	Es wäre besser, anstelle der Aussage, du könntest exakt messen, auch Messunsicherheiten einzubeziehen, ja.
211	Vorlesen, Zustimmung AO	Es wäre besser, beide Wagen zu wiegen, damit du den Schüler- innen ein korrektes Bild vom genauen wissenschaftlichen Arbeiten vermittelst, auch das.
212	Vorlesen, Zustimmung AO + Begründung	Es wäre besser, beide Wagen zu wiegen, da die Wagen auch wenn sie baugleich sind nicht gleich schwer sein müssen. Auch das, das gehört dazu zu dieser wissenschaftlichen Arbeit. Zum Beispiel. Ist das für mich so relevant, trifft es am ehesten zu? Nein, kommt hier so ein bisschen mit raus (B2 führt Mauszeiger über AO 2 und 3).
213	Vorlesen, Zustimmung AO, Sonstiges	Es wäre besser, beide Wagen zu wiegen, da du im Vergleich der Massen auf Messunsicherheiten bzw. Ungenauigkeiten eingehen kannst.
214	Vorlesen, Überlegung zum Unterricht	Es wäre besser, beide Wagen zu wiegen, damit die Schüler- innen sich selbst überzeugen können, dass die Wagen gleich schwer sind. Werden sie ja nicht sein, die werden ein kleines bisschen unterschiedlich sein. #01:05:11-2#
215	Vorlesen	B2: Welches Feedback möchten Sie Robert zur Aktivierung der Schüler- innen geben.
216	Vorlesen, Zustimmung AO, Beschreibung	Es ist nicht gut, dass du die Schüler- innen nur oberflächlich involvierst, da die Schüler- innen trotz dieser Aktivierung nicht mitdenken müssen. Ja, da wird nicht viel aktiviert.
217	Vorlesen, Zustimmung AO + Begründung	Es ist gut, dass du Schüler- innen in das Experiment einbeziehst. Das ist richtig, Einbezug ist wichtig und richtig, glaube ich auch, vor allem, weil die das mitmachen können, (.)
218	Alternative (2x)	aber äh nicht so (.) abgrenzend einbeziehen. Nicht sagen, Jungs jetzt das, Mädchen jetzt das und aber ich muss anschieben, damit es richtig bleibt. Wenn man den Jungs sagt, schieb ein bisschen schneller (und jetzt) schieb ein bisschen langsamer, das hätte auch funktioniert.
219	Vorlesen, Verweis auf frühere Stelle im Test, Ablehnung AO, Alternative	Es ist gut, dass du das Experiment selber durchführst- ah ja, genau was ich gerade gesagt habe, nein, hättest- hätten auch die Schüler machen können.

220	Vorlesen, Zustimmung AO + Begründung, Alternative, Beschreibung	Es ist nicht gut, dass die Schüler- innen dir ohne Arbeitsauftrag zusehen, da die Schüler- innen so keinen Anreiz haben, aufmerksam zu sein. Habe ich auch (im?) Video gehabt die Idee. Die wissen ja gar nicht, worauf sie achten sollen. Zumindest ein was fällt euch auf, was ist festzustellen, denkt an die Formel von gerade, womit hängt das zusammen. Kam nicht.
221	Vorlesen, Zustimmung AO + Alternative	Es ist nicht gut, dass du die Masse des Wagens und insbesondere die Geschwindigkeit selbst bestimmst, da du den Schüler- innen vorher schon erklärt hast, wie sie das selbst machen können. Ja, können sie auch selber machen unter Anleitung, unter Aufsicht, aber die können das selber machen.
222	Vorlesen, Überlegung zum Unterricht, Bewertung	Es ist nicht gut, dass du das Experiment so lehrerzentriert durchführst, da dies für die Schüler- innen langweilig ist. (..) Langweilig weiß ich nicht. Es war schon (.) was neues, besser als Tafelunterricht, sage ich mal. (.)
223	Vorlesen (2x), Sonstiges*, Zustimmung AO + Begründung, Alternative	Es ist nicht gut, dass Paula von der Tafel aus und die Schüler- innen aus den hinteren Reihen- (.) warte mal. Es ist nicht gut, dass Paula von der Tafel aus und die Schüler- innen aus den hinteren Reihen den experimentellen Aufbau nicht gut erkennen können, da es für sie schwer ist, dir zu folgen. Ja, Paula konnte nichts sehen, die war die leidtragende in der Situation und die anderen auch, man muss schon darauf achten, dass alle das- (.) das machen können. #01:07:02-4#
224	Vorlesen	B2: Welche der unten stehen Hinweise zur Aktivierung der Schüler- innen möchten Sie Robert für eine erneute Durchführung geben.
225	Vorlesen, Zustimmung AO	Es wäre besser, die Schüler- innen zu fragen, wer freiwillig schreiben beziehungsweise die Werte ablesen möchte, um den Schüler- innen nicht nach Geschlecht eine Aufgabe zuzuweisen. Ja, genau das.
226	Vorlesen, Zustimmung AO, Sonstiges	Es wäre besser, die Schüler- innen zu fragen, ob sie den Ablauf verstanden oder noch Schwierigkeiten haben, ja, auch das. Vielleicht- Wofür machen wir das überhaupt.
227	Ablehnung AO	Ausreichend aktiviert leider nicht.
228	Vorlesen, Überlegung zum Unterricht, Sonstiges	Es wäre besser, Hypothesen der Schüler- innen abzufragen, um sie zur Beobachtung des Experiments zu motivieren. (.) Da müsste ich nachdenken, (.) ähm geht nämlich darum, jetzt weiß ich im Impulsthema nicht auswendig, wie die Schülerfehlvorstellungen sind. Wenn es Fehlvorstellungen (.) gibt, wäre es ganz gut, sie im Vorhel- Vorhinein abzugreifen, wenn es keine gibt, erzeuge ich sie dadurch vielleicht. (..) Ich lasse es erstmal weg, weil ich mir unsicher bin in dem Bereich.

229	Vorlesen	Es wäre besser, den Schüler- innen Fragen zum Vorgehen während des Experiments oder zur Berechnung der Geschwindigkeit zu stellen, um sie zum Mitdenken anzuregen. Ja, er stellt sich einfach vor den Taschenrechner und sagt, ich rechne das eben aus, woher das kommt keine Ahnung, zumindest exemplarisch muss man das vielleicht mal machen auch mit einer Formel an der Tafel oder so. (.)
230	Vorlesen, Zustimmung AO	Es wäre besser, die Schüler- innen stärker in die Durchführung des Experiments einzubeziehen, zum Beispiel um Wagen anzustoßen, Werte abzulesen oder die Geschwindigkeit zu bestimmen. Ja, genau das, habe ich genauso gedacht.
231	Vorlesen, Zustimmung AO + Begründung, Verweis auf frühere Teststelle	Es wäre besser, die Schüler- innen zum Experiment nach vorne kommen zu lassen, damit sie alles gut erkennen können. (.) Wenn das nicht bereits möglich wäre, muss das so getan werden, ja, wie ich- wie es eben gesagt wurde, Paula kann das nicht sehen. #01:08:38-8#
232	Vorlesen, Beschreibung	B2: Nachdem wir alle Messwerte aufgenommen hatten, habe ich die Beobachtungen der Schüler- innen abgefragt. Um die Allgemeingültigkeit der Impulserhaltung nachzuweisen, habe ich sie dann in einem kurzen Exkurs theoretisch hergeleitet. Im Anschluss haben wir die Impulserhaltung am Experiment überprüft. Okay, Überprüfung der Impulserhaltung mit den Messwerten aus dem Experiment, Lehrer leitet Gespräch zur Besprechung, Schülerinnen und Schüler nennen Ergebnisse. Achso erst überprüfen die das in Partnerarbeit und dann ins Plenumsgespräch und machen dann daraus ein Plakat, okay. #01:09:09-1#
233	-	[Video Einleitung Reflexion 06, Video Szene Reflexion 06] // #01:12:09-4#
234	Beschreibung	B2: Da fällt mir direkt ein, jetzt kann ich sie sehen, ähm er hat keine Einheit genannt. Vorher war es ganz wichtig, jetzt ähm ich weiß nicht, wie die Einheit thematisiert wird, fällt mir nur direkt ein. #01:10:35-9#
235	Bewertung	B2: Krass, (unv.) nochmal so eine Mädelsklatsche hintendran ey. #01:12:16-5#
236	Vorlesen	B2: Zu welcher der folgenden Aspekte sollte Robert eine Rückmeldung erhalten.
237	Vorlesen, Zustimmung AO	Ja, Umgang mit- mit den Schüler- innen,
238	Vorlesen, Zustimmung AO	Messunsicherheiten auch, ja,
239	Vorlesen, Zustimmung AO + Begründung, Verweis auf frühere Teststelle	der Umgang mit den Einheiten, ja, Einheit, wie ich gesagt habe, keine Einheit, keine Einheit,
240	Sonstiges	alle- alle drei Sachen (lacht).

241	Sonstiges	Okay, äh sage ich auch nochmal kurz was zu.
242	Vorlesen, Bewertung + Begründung	Umgang mit den Schüler- innen, also die- die Mädchenklatsche zwischendrin mit Mädchen bitte Tabelle korrigieren geht gar nicht, geht nicht, überhaupt nicht, also das- das- das- also wenn ich mir vorstelle, das kriegt ein Elternteil hinterher mit, was der Lehrer sagt, das- das- allein, was das an Stress gibt und das ist einfach nicht richtig, das verfestigt das Bild (und?) ne. (.)
243	Vorlesen, Bewertung + Begründung, Alternative (2x), Beschreibung	Äh ja, Messunsicherheiten, Umgang mit dem Einwand von Marie. Der Einwand von ihr war ja super berechtigt, die sind ja überhaupt nicht gleich, das hätte von ihm schon kommen müssen meiner Meinung nach und da er aber die Messunsicherheiten überhaupt nicht thematisiert hat, ist er jetzt in so einer Zwickmühle, dass er damit dann auch nicht argumentieren kann. Notgedrungen hätte ich gesagt, dann für sie jetzt doch eben schnell ein und sag okay, die sind nicht unterschiedlich, weil wir haben folgende Unsicherheiten, hier gibt es was, da gibt es was, da gibt es was, die gibt es bei jedem Experiment.
244	Vorlesen, Alternative, Bewertung	Und Einheit des Impulses, vorher waren die Einheiten super wichtig, dann muss ich das stringent durchziehen, dann muss ich auch sagen okay, was hat denn das für eine Einheit und das ist ja eine zusammengesetzte Einheit, äh kann ich Einheiten so zusammensetzen, ist das schon bekannt, das muss ich einmal thematisieren. Einfach nur den Zahlenwert dahin zu schreiben, weil die Tabelle halt zu klein ist, so wie es auf dem Bild zu sehen war, ähm ne. #01:13:47-8#
245	Vorlesen, Sonstiges	B2: Robert hat zwar schon einige Rückmeldungen erhalten, würde aber trotzdem gern zu allen drei aufgeführten Aspekten noch eine Rückmeldung erhalten. Gut, dass ich alle drei angeklickt habe. #01:13:56-9#
246	Vorlesen	B2: Äh Robert zum Umgang mit Schülerinnen und Schülern.
247	Vorlesen (2x), Überlegung zum Unterricht (2x), Bewertung, Zustimmung AO	Es ist nicht gut, dass du Mädels sagst (..) weiß ich nicht, ist Mädels ein (.) was ist denn das für eine Klasse, steht das hier? Informationen zur Lerngruppe, zehnte Klasse (.) (B2 öffnet Hilfe, Beschreibung der Lerngruppe) so, zehnte Klasse an einer Gesamtschule, Klasse setzt sich aus vierzehn Mädchen und sieben Jungen, Schülerinnen und Schüler zeigen wenig Interesse, ja okay, geht ja nur um das Alter, in der zehnten Klasse Mädels (..) weiß ich nicht, man könnte auf jeden Fall- generell dass man da einen Gruppenbegriff für nutzen muss, das ist schon das Problem. Ja, ist ein Problem.
248	Vorlesen, Zustimmung AO	Ähm es ist nicht gut, dass du die Mädchen und Jungen so unterschiedlich behandelst, da das nicht wertschätzend gegenüber den Mädchen ist, auch das.

249	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du die Mädchen und Jungen in ihrem Können so unterschiedlich einschätzt, da du so die Vorstellung förderst- ja, das.
250	Vorlesen, Zustimmung AO, Beschreibung	Du thematisierst die Einheit des Impulses nicht, das ist nicht gut, da dies inkonsistent zur früheren Szene- (ja?) auch genauso gesagt. Bei ihr hat er drauf bestanden, ob das so ein Einheitending ist oder ob da ein Sexismusding ist.
251	Vorlesen, Ablehnung AO + Begründung, Alternative + Begründund	Es ist gut, dass du die Mädchen dazu aufforderst, ihre Tabelle zu korrigieren, damit du sicher sein kannst, dass sie die gleichen Werte wie die Jungen notiert haben. Ne, genau nicht, die sollten doch vorhin in Partnerarbeit kontrollieren, ob ihre Werte richtig sind, dann ist doch davon auszugehen, ähm oder selbst wenn würde ich davon ausgehen, Mädchen kriegen den Wert nicht hin, wenn das wirklich der Fall wäre, unwahrscheinlich, ähm dann müsste ich doch gemischte Gruppen machen, immer ein Mädchen und ein Junge, vielleicht auch, um dem Sexismusding entgegenzuwirken, dass auch Mädchen und Jungen super zusammenarbeiten können und nicht nur Jungen das drauf haben und dann hätte ich das Problem überhaupt nicht gehabt, wenn es denn überhaupt bestünde. Ähm ne ist nicht gut, dass er sie auffordert. (.) #01:15:40-4#
252	Vorlesen	B2: Welche der unten stehen Hinweise für seinen Umgang mit den Schülern möchten Sie Robert für eine erneute Durchführung mitgeben.
253	Vorlesen	Wählen Sie- ja Einschätzung am nächsten.
254	Vorlesen, Ablehnung AO	Dein Umgang mit den Schülerinnen ist in Ord- nein.
255	Vorlesen, Zustimmung AO + Begründung	Du solltest die Mädchen stärker in den Unterricht einbeziehen, zum Beispiel berechnete Werte vorlesen lassen mhm[ja], Ergebnissicherung, Zusammentragen war so ein Jungsding irgendwie.
256	Vorlesen, Zustimmung AO	Du solltest auf geschlechtsspezifische Zuschreibungen verzichten, um die Schülerinnen nicht zu diskriminieren. Ja.
257	Vorlesen, Zustimmung AO + Begründung	Du solltest von jedem Schüler- in nur einen berechneten Wert erfragen, da du so die berechneten Werte von mehreren Schüler- innen kontrollieren kannst. Mhm[Ja] auch wenn die Vorkontrolle bereits in Partnerarbeit erfolgt ist, weiß ich ja trotzdem nicht, ob es richtig ist, vielleicht haben die beide auch einen kleinen Murks raus. Vielleicht hat der eine gar nichts raus und der andere hilft ihm und dann habe ich nur einen Wert und vielleicht hat aber genau der, der gerechnet sich verrechnet oder was falsch eingetippt. (.)

258	Vorlesen, Zustimmung AO, Alternative	Du solltest verstärkt schwächere Schüler- innen in den Unterricht, da wichtig ist, dass sie die Inhalte verstehen. Ja, habe ich auch gedacht. Kurze Nachfrage, mal gucken, was habt ihr da raus, wie seid ihr darauf gekommen, wisst ihr, wie man das berechnen muss, vorab vielleicht schon.
259	Vorlesen, Zustimmung AO	Du solltest Schüler- innen unabhängig vom Geschlecht dazu auffordern, falsche Ergebnisse in der Tafel zu korrigieren, damit am Ende alle Schüler- innen die richtigen Ergebnisse in der Tafel notiert haben. (.) Ja, generell wenn jemand einen Fehler gemacht hat, dann egal wer, alle sollten korrigieren.
260	Vorlesen, Sonstiges*, Zustimmung AO + Begründung, Professionalisierung	Du solltest hinterfragen, woher die geschlechterspezifischen Zuschreibungen stammen, die du vertrittst, um sie zu überwinden. (..) Ah da geht es um Mädels und sowas, ne? Ja, schon, um als Lehrkraft besser zu werden. Wenn mir (da halt?) jemand sagt okay, du driftest da oft ab in so eine Schiene, die das Ganze nur noch viel schlimmer macht, äh ja, dann hinterfrage mal selber, woher das Bild kommt, nur so kannst du dem auf den Grund gehen. Ist jetzt nichts, was ich vorher schon im Kopf hatte, aber etwas, was auf jeden Fall getan werden sollte. #01:17:26-2#
261	Sonstiges	B2: Muss einmal schauen, was mein Akku sagt, der sieht nämlich- Okay achtunddreißig Prozent, der hält noch.
262	Vorlesen	Ähm welches Feedback möchten Sie Robert zum Umgang mit Maries Einwand beziehungsweise bezüglich- beziehungsweise mit Messunsicherheiten geben.
263	Vorlesen, Zustimmung AO + Begründung, Alternative + Begründung	Die Reaktion auf Maries Einwand ist dir gut gelungen- nein. Die Reaktion auf Maries Einwand ist dir gut gelungen- nein. Die Reaktion auf Maries Einwand ist dir nicht gut gelungen, da du den Fokus von der experimentellen Bestätigung hin zur Diskussion von leichten Ungenauigkeiten verschiebst. (.) Vielleicht, vielleicht, weil weil weil die Messunsicherheiten wurden vorher gar nicht thematisiert. Man müsste eigentlich sagen, dass in jedem Experiment, egal welches, immer leichte Verschiedenheiten auftauchen werden und wenn sie nicht auftauchen, macht man was falsch vermutlich oder misst nicht genau genug. Ja also man- man sagt dann, das ist generell nur ein Rundungsfehler, das hat mit dem Experiment nichts zu tun, das gibt ein falsches Bild.
264	Vorlesen, Zustimmung AO + Begründung	Es ist nicht gut, dass du nur so knapp auf den Einwand eingehst, da die Begründung nicht ausreicht, um den Schüler- innen zu verdeutlichen, dass die Unterschiede aufgrund von Messunsicherheiten zustande kommen, ja, es wird so abgesch- abgesch- abgeschweift einfach, keiner mehr darf was dazu sagen, ich habe das jetzt kurz erklärt und fertig (unv.).

265	Vorlesen, Sonstiges, Beschreibung, Ablehnung AO	Es ist nicht gut, dass du nur so knapp auf den Einwand von Marie eingehst, da so unklar bleibt, ob beziehungsweise wo die Schüler- innen Rundungsfehler gemacht haben. (.) Ob die Schüler selber welche gemacht haben. Die haben ja jetzt im Taschenrechner drei Nachkommastellen raus und immer vorgetragen. Ich glaube nicht, dass das ein Problem war.
266	Vorlesen, Überlegung zum Unterricht, Alternative	Es ist nicht gut, dass du die Ungleichheit der Messwerte mit Rundungsfehlern begründest, da diese Begründung falsch ist. (.) Ist die Begründung denn falsch? Ja und Nein. Erstens ist kein Experiment perfekt, erst recht nicht in Raumbedingungen, ähm man kann das schon mit Rundungsfehlern begründen, (das?) ist aber nicht der einzige Grund, so würde ich das anführen.
267	Vorlesen, Sonstiges* (2x), Überlegung zum Unterricht, Beschreibung (2x), Zustimmung AO + Begründung, Bewertung	Es ist nicht gut, dass du die Ungleichheit der Messwerte mit den Rundungsfehlern begründest, da du so den Schüler- innen die Schuld dafür gibst, dass die Ergebnisse nicht exakt gleich sind. (.) Ach Rundungsfehler der Kinder, ich dachte von der Einheit, jetzt- ach du Heilige. Jetzt versteh- jetzt, jetzt dämmert es mir auch. Achso, er sagt quasi, die- die Schüler haben sich ganz leicht verrechnet, irgendwo eine Einheit weggelassen, äh ne stimmt, ist genau richtig, das ja- da- da habe ich nicht so richtig drüber nachgedacht. Okay darf ich es jetzt ankreuzen, weil es ist ja nicht meine Einschätzung, aber doch, dadazu würde ich stehen im Nachhinein, man muss es ja mit Messunsicherheit begründen und nicht mit- Ich habe die ganze Zeit Rundungsfehler und Messunsicherheit durcheinander gehauen. (.) Ähm stimmt ja, das kommt nicht daher, dass sie sich verrechnet haben, sondern das kommt daher, dass nicht alles hundert Prozent genau gemessen werden kann und dann drei Nachkommastellen anzugeben ist ja in dem Moment richtig. #01:20:18-8#
268	Vorlesen	B2: Welche der unten stehenden Hinweise zu seinem Umgang mit Maries Einwand beziehungsweise Messunsicherheiten möchten Sie Robert für eine erneute Durchführung in der Unterrichtsstunde geben.
269	Vorlesen, Ablehnung AO	Deine Reaktion auf den Einwand ist angemessen- mhmh[nein].
270	Vorlesen, Zustimmung AO	Du solltest den Schüler- innen angeben, auf wie viele Nachkommastellen sie runden sollen, ja, das ist vorher auf jeden Fall zu sagen, rundet auf so und so viele Nachkommestellen. Vermutlich ist der Einfluss dahinter einfach nur mehr Aufwand und da wird nicht mehr viel passieren. (.)
271	Vorlesen, Ablehnung AO, Alternative + Begründung	Es wäre besser, den Schüler- innen am Anfang mitzuteilen, dass Messunsicherheiten und Rundungsfehler nicht berücksichtigt werden- nein, ist nicht richtig, die müssen berücksichtigt werden, sie gehören dazu zum Experiment, also die Messunsicherheiten, nicht die Rundungsfehler.

272	Vorlesen, Zustimmung AO, Alternative	Es wäre besser, ausführlicher zu begründen, wo die Rundungsfehler entstehen, damit die Schüler- innen deiner Erklärung besser folgen können. Wenn sie Rundungsfehler gemacht haben in Klasse zehn, was ja Thema von Klasse fünf, sechs ist, sollte man das auf jeden Fall thematisieren und fragen okay, wo kommt das jetzt her, was ist dein Wert, den du berechnet hast und was hast du aufgeschrieben. Das muss man auf jeden Fall nochmal ansprechen.
273	Vorlesen, Zustimmung AO, Beschreibung	Du solltest bereits bei der Vorstellung des Experiments über die vorhandene Reibung und Energieumwandlung beim Aufprall sprechen, damit den Schüler- innen in dieser Phase klar ist, dass die Ergebnisse aufg- aufgrund dieser Einflüsse nicht exakt gleich sein können. Ja, es wurde gesagt wir machen jetzt das Gebläse an, zum Beispiel, das gehört ja auch zu Reibung dazu, wir machen jetzt das Gebläse- warum, warum sollten wir das tun, das wurde (ja?) nicht thematisiert.
274	Vorlesen, Zustimmung AO, Beschreibung	Es wäre besser, dass es bei der Durchführung des Versuches Messunsicherheiten berücksichtigt werden müssen, damit die Schüler- innen nachvollziehen können, wieso es zu Abweichungen kommt, ja, wurde ja nicht gesagt, die sind einfach gleich schwer.
275	Vorlesen, Überlegung zum Unterricht (2x), Alternative + Begründung	Du solltest das Thema Messunsicherheiten für dich selbst noch einmal aufbereiten damit du von solchen Fragen nicht vom Thema abgebracht wirst. (.) Ich vermute, dass er das Thema verstanden hat, dass er ganz genau weiß, er hat das Studium durchlaufen und so weiter, Messunsicherheiten sind ein riesen Ding im Praktikum. Ähm er sollte es in dem Sinne nochmal für sich aufarbeiten äh im Umgang, wie ich das den Schüler- den Schüler und Schülerinnen vermittele, wie ich mit denen da umgehe, ja, damit das nicht so ein- so ein Thema ist, wo er nichts zu sagen kann, stimmt schon, er kann das bestimmt selber, aber denen das zu erklären, das ist- das ist das Problem. #01:22:28-0#
276	Vorlesen	B2: Welches Feedback möchten Sie Robert zu seinem Umgang mit der Einheit des Impulses geben?
277	Vorlesen, Zustimmung AO, Alternative + Begründung (2x)	Es ist nicht gut, dass du nicht über die Einheit des Impulses sprichst, da die Einheit ein wichtiger Bestandteil der Angabe ist. Ja, wenn ich vorher hundert Prozent drauf bestehe, muss ich zumindest thematisieren, wieso ich es gerade vielleicht noch weg lasse, wieso es gerade vielleicht noch schwierig ist, weil es halt zusammen gesetzt ist.
278	Vorlesen, Zustimmung AO + Begründung	Es ist nicht gut, dass du nicht über die Einheit des Impulses sprichst, da dies inkonsistent zur früheren Szene ist, in der du auf die Notation der Einheit der Geschwindigkeit bestanden hast. Ja, ich finde Konsistenz sehr wichtig.

279	Vorlesen, Zustimmung AO, Ablehnung AO, Verweis auf frühere Stelle im Test, Alternative	Es ist gut, dass du die Einheit des Impulses nicht so stark in den Fokus gerückt hast, um die Schüler- innen nicht durch die neue Einheit zu verwirren. Ja und nein, wie ich es vorher schon gesagt habe. Man kann- wenn man sie weglässt, dann sollte man definitiv begründen, warum habe ich sie gerade weggelassen und dann würde ich sagen, ja, euch ist jetzt- euch sollte auffallen, hier fehlt eine Einheit, die kommt gleich noch, hat folgenden Grund, ist- ist- wird ein bisschen komplizierter.
280	Vorlesen, Zustimmung AO + Begründung, Ablehnung AO + Begründung	Es ist gut, dass du den Fokus auf die berechneten Werte und nicht auf die Einheit gelegt hast, da die Zahlen ausreichen, um die Impulserhaltung auf Grundlage des Experiments zu bestätigen. (.) Ja und nein, ja die Zahlen reichen aus, um das zu sagen, aber (..) ein Ergebnis darf nie ohne Einheit da stehen, das ist für die Physik echt wichtig und das sollten die Schüler- innen lernen in dem Alter.
281	Vorlesen, Ablehnung AO + Begründung, Alternative (2x)	Es ist nicht gut, dass du die Schüler- innen nicht dazu aufforderst, die Einheiten anzugeben, da so unklar ist, ob die Impulse wirklich gleich groß sind. Ne, das- also schon, man sieht es an den Zahlenwerten, dass sie gleich groß sind, bei dem kleinen Experiment klappt das, äh er muss sie auch nicht unbedingt dazu auffordern, die Einheit anzugeben, er sollte es später nur thematisieren. So wäre mein Schluss dazu. #01:23:59-9#
282	Vorlesen	B2: Welche der unten stehenden Hinweise zu seinem Umgang mit der Einheit des Impulses möchten Sie Robert für eine erneute Durchführung geben?
283	Vorlesen, Zustimmung AO, Alternative (3x)	Es ist in Ordnung, dass du (die?) Einheit des Impulses zunächst vernachlässigst, das kannst du erneut so machen. Ja, vorerst vorerst vernachlässigen, wenn ich es vielleicht unmittelbar danach oder als Hausaufgabe oder im Einstieg zur nächsten Stunde, je nachdem, wie die Zeit ist, thematisiere. Wäre ich in der nächsten Stunde, würde ich ja sagen, okay guckt euch mal die Tabelle an, fällt euch da was auf, vielleicht auch als Hausaufgabe geben und dann darüber sprechen.
284	Vorlesen, Zustimmung AO, Alternative	Du solltest die Einheit angeben, da die Einheit ein wichtiger Bestandteil ist. Ja, eigentlich schon, im Nachgang finde ich.
285	Vorlesen, Alternative	Du solltest darauf achten, dass die Schüler beim Vorlesen der Werte auch die Einheit angeben, um in deinem eigenen Verhalten konsistent zu bleiben. Dann lieber vorher schon vielleicht kurz auf die Einheit verzichten und im Nachhinein sagen okay, haben wir jetzt gemacht- (.) auch schwierig. Beim Impuls vielleicht noch kurz weglassen, dann aber wie gesagt begründen, wieso ich es weglasse.

286	Vorlesen, Zustimmung AO, Alternative (2x)	Du solltest die Einheit des Impulses oben über die Spalte der Tabelle notieren, um die Schreibarbeit in jeder Zeile zu sparen, ja. Äh generell die Einheiten, Masse, Geschwindigkeit und so weiter, Einheit des Impulses generell jetzt dann am Ende in die Spalte der Tabelle schreiben, da ist ja auch schön noch Platz.
287	Vorlesen, Überlegung zum Unterricht	Du solltest die Schüler- innen dazu auffordern, die Einheit des Impulses zu nennen, damit die Schüler- innen sich an die Einheit des Impulses gewöhnen. Kommen sie von selber nicht drauf, würde ich jetzt einfach mal sagen, kommen die vielleicht selber nicht drauf. #01:25:28-5#
288	-	[Video Einleitung Reflexion 07] #01:25:41-3#
289	Vorlesen, Sonstiges	B2: So, Lehrer sichert Erkenntnisse aus dem Experiment in einem Merksatz, Schülerinnen und Schüler schreiben Merksatz zur Sicherung ab. (unv.) mal sehen. #01:25:50-5#
290	-	[Video Szene Reflexion 07] #01:27:24-8#
291	Beschreibung	B2: Joa, das Muster zieht sich so ein bisschen durch, habe ich das Gefühl, das relativ egal ist, was die Schülerinnen und Schüler sagen, weil alles schon vorgefertigt ist.
292	Zustimmung AO (2x)	Ja beides, zu beidem Feebdack.
293	Vorlesen, Beschreibung, Bewertung	Formulierung des Merksatzes ein rieseig langer Satz, (.) kompliziert, verschachtelt, alles, (.) schwierig, zu folgen schwierig und als Schüler wahrscheinlich noch schwieriger.
294	Vorlesen, Zustimmung AO + Begründung, Überlegung zum Unterricht, Beschreibung (2x), Alternative + Begründung	Interaktion auch das, nur zwei Jungs dran genommen, kein Mädchen dran genommen, obwohl es ja viel mehr Mädchen sind, ne, Moment (B2 öffnet Hilfe, Beschreibung der Lerngruppe), waren es nicht vierzehn Mädchen? Ja, vierzehn Mädchen und man nimmt zwei Jungs dran von den sieben, das äh dann nimmt man doch auch mindestens zwei, drei, vier Mädchen noch dran, damit das ungefähr ausgeglichen ist und ähm (die?) Interaktion, (.) äh zu sagen, ja richtiger wäre das stimmt vielleicht schon, für ein Lehrbuch wäre der andere richtiger oder beziehungsweise auch aus einem- äh aus einem Lehrbuch entnommen, aber dann ist ja die- die- die Mitarbeit der Schülerinnen und Schüler nichts wert in dem Moment, die hätten sagen können, was sie wollen, er hätte gesagt, richtiger ist das. #01:28:28-9#
295	Vorlesen, Sonstiges	B2: Robert hat zwar schon einige Rückmeldungen erhalten, würde aber trotzdem gern zu den beiden aufgeführten Aspekten noch eine Rückmeldung erhalten. Ja. #01:28:36-5#
296	Vorlesen	B2: Welches Feedback möchten Sie Robert zu seiner Formulierung des Merksatzes geben?

297	Vorlesen, Zustimmng AO, Bewertung	Die Formulierung ist dir gelungen, ja, alle wichtigen Informationen sind korrekt enthalten, (..) aber es ist zu kompliziert einfach.
298	Vorlesen, Ablehnung AO	Es ist gut, dass du im Merksatz auf das durchgeführte Experiment verweist, da die Schüler(*innen?)den Merksatz direkt mit der Realität verbinden könnten. (unv.) ja weiß ich nicht, nicht so relevant, glaube ich.
299	Vorlesen, Zustimmung AO + Begründung, Beschreibung	Die Formulierung des Merksatzes ist dir nicht gut gelungen, da sie für die Schüler- innen zu anspruchsvoll ist, ja, zehnte Klasse (.) glaube ich nicht, dass es- dass es ein verschachtelter Satz mit ganz viel Komma und Einschüben und alles Mögliche ähm ne.
300	Vorlesen, Zustimmung AO, Verweis auf frühere Stelle im Test	Es ist nicht gut, dass du deinen vorgefertigten Merksatz nutzt und die Merksätze der Schüler- innen nicht einbeziehst, da das den Schüler- innen gegenüber nicht wertschätzend ist, ja genau das, wie ich es vorher kurz gesagt hatte. #01:29:22-9#
301	Vorlesen	B2: Welchen der unten stehenden Hinweise zu seiner Formulierung des Merksatzes zur Impulserhaltung möchten Sie Robert für eine erneute Durchführung geben?
302	Vorlesen, Ablehnung AO + Begründung	Die Formulierung des Merksatzes ist angemessen, du kannst sie erneut verwenden. Nein, zu kompliziert.
303	Vorlesen, Zustimmung AO	Es wäre besser, Beispiel und Merksatz zu trennen, um die Komplexität des Merksatzes zu reduzieren. (.) Kann man schon noch so machen.
304	Vorlesen, Sonstiges, Ablehnung AO + Begründung, Alternative	Es wäre besser, zusätzlich zum langen Merksatz- entschuldigung, das war mein Hund im Hintergrund (lacht) ähm es wäre besser, zusätzlich zum langen Merksatz eine kurze Version zu formulieren, da sich die Schüler- innen eine kurze Version besser merken können. (.) Eine kurze prägnante Version, da zusätzlich nicht, dann wird es noch mehr. Generell etwas kürzer und vielleicht einen kleinen Satz, wo- wo es zusammenfassend ist.
305	Vorlesen, Ablehnung AO + Begründung	Es wäre besser, den Merksatz nur in Bezug auf einen Stoßprozess zu formulieren, damit der Merksatz näher am besprochenen Experiment bleibt. Jein, nein, nein, kein jein, nein. Äh (.) es ist ja eine Allgemeingültigkeit, die nicht nur auf das Experiment sich bezieht.
306	Vorlesen, Zustimmung AO, Alternative	Es wäre besser, den Merksatz um die Einheit des Impulses zu ergänzen, da die Einheit stets zur Angabe gehört. Ja, ein kurzer Nebeneinschub, die Einheit des Impulses ist und setzt sich wie folgt zusammen.
307	Vorlesen, Zustimmung AO	Du solltest den Merksatz schülernäher gestalten zum Beispiel indem du eine Skizze ergänzt oder die Formulierung sprachlich vereinfachst, um die Schüler- innen beim Verständnis zu unterstützen. Ja, definitiv das.

308	Vorlesen, Zustimmung AO, Alternative	Es wäre besser, eine Formulierung der Schüler als Merksatz zu verwenden, da diese Formulierungen für sie besser verständlich sind. Ja, vielleicht zu sagen okay, wir nehmen jetzt einen ersten, schreiben den auf und ergänzen den immer weiter mit Sachen, die andere Schüler auch noch gesagt haben, was vielleicht noch fehlt und wenn dann am Ende immer noch was Wichtiges fehlt, dann kann man sagen okay, ihr habt eine Kleinigkeit vernachlässigt, ist aber auch nicht schlimm, ihr seid ja auch Schüler, dafür ist der Lehrer schließlich auch da. #01:31:07-5#
309	Sonstiges	B2: So zur Interaktion.
310	Vorlesen, Bewertung (2x), Ablehnung AO	Es ist gut, dass die Schüler- innen ihre Merksätze vorlesen, da du so erkennen kannst, ob sie die Impulserhaltung umfassend verstanden haben. Ähm kann man es erkennen, ob es umfassend verstanden wurde, nicht, aber es ist auf jeden Fall gut, dass sie es vorlesen sollen, aber ich stimme dem nicht ganz zu, das kann ich nicht ankreuzen.
311	Vorlesen, Zustimmung AO + Begründung	Es ist gut, dass die Schüler- innen einen Merksatz formulieren sollen, um ein tiefergehendes Verständnis zu fördern. Ja, weil die drüber nachdenken müssen, was ist wichtig, was ist relevant, was- was muss da mit rein.
312	Vorlesen, Sonstiges, Bewertung, Alternative + Begründung	Es ist nicht gut, dass du den Schüler- innen keine Rückmeldung auf die Merksätze gibst, da die Schüler- innen so nicht einordnen können, ob ihre Merksätze angemessen formuliert sind. Ähm (..) schwierig, Feedback ist gut, dass man sagt- zumindest zu sagen, das ist eine gute Idee, das ist ein guter Einwand, das ist ein guter Winkel, aber nicht sagen ne, ist ein schlechter Merksatz, können wir nicht machen, weil sich dann nie wieder jemand traut, wenn er nicht gerade einen perfekten Merksatz hat. Also ich würde das auf jeden Fall akzeptieren und einordnen, aber das Negative immer ein bisschen außen vor lassen.
313	Vorlesen, Beschreibung	Es ist nicht gut, dass du deinen vorgefertigten Merksatz nutzt und die Schüler- innen- äh die Merksätze der Schüler- innen nicht einbeziehst, da das den Schüler- innen gegenüber nicht wertschätzend ist. Ja, wie ist der Einfluss auf den Unterricht, nicht vorhanden.
314	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du deine Formulierung als besser bezeichnest, ja, genau das.
315	Vorlesen, Zustimmung AO	Es ist nicht gut, dass nur zwei Schüler- innen ihre Merksätze vorgestellt haben, da du so nicht weißt, welche Merksätze die restlichen Schüler- innen formuliert haben. Ja, was ist mit dem ganzen Rest der Klasse, was ist mit den Mädchen?

316	Vorlesen, Zustimmung AO + Begründung, Beschreibung	Es ist nicht gut, dass (du?) nur zwei Jungen ihre Merksätze vorgestellt haben, da die Schüler- innen so lernen, dass du nur Jungs drannimmst. Ja, in fachlichen Inhalten scheint es da ein Muster zu geben, warum sollte man also noch mitmachen, wird ja nicht honoriert, wird wahrscheinlich auch nicht entsprechend benotet. #01:32:48-9#
317	Vorlesen	B2: Robert für eine erneute Durchführung der Unterrichtsstunde geben?
318	Vorlesen, Zustimmung AO, Alternative (2x)	Du solltest den Merksatz ausführlicher erklären beziehungsweise mit den Schüler- innen besprechen, da einige Aspekte des Merksatzes nicht selbsterklärend sind. Ja, wenn ich so einen komplizierten mache, würde ich ihn Satz für Satz durchgehen und selbst, wenn er einfacher ist, würde ich durchgehen, woher das kommt.
319	Vorlesen, Zustimmung AO, Verweis auf frühere Stelle im Test, Alternative	Du solltest gemeinsam mit den Schüler- innen einen Merksatz erstellen, ja, definitiv, wie ich gesagt habe mit dem ich nehme einen, der (unv.) einen als Vorlage und ergänze ihn um alles andere, was sie noch genannt haben.
320	Vorlesen, Zustimmung AO, Alternative	Es wäre besser, insgesamt mehr Schüler- innen ihre Merksätze vorstellen zu lassen, da du aus den Antworten Rückschlüsse auf die Vorstellung ziehen kannst. Auch das, deswegen ja nicht bewerten.
321	Vorlesen, Zustimmung AO	Du solltest die Schüler- innen sich melden lassen, wenn sie ihre Merksätze vorstellen wollen, da sich möglicherweise auch ein Mädchen gemeldet hätte, ja.
322	Vorlesen, Ablehnung AO + Begründung	Du solltest deinen- Ja, das (B2 zeigt auf Antwortoption 'Der Umgang mit den Schüler*innen war angemessen, so kannst du dich wieder verhalten.') lasse ich außen vor, das ist nicht angemessen.
323	Vorlesen, Zustimmung AO + Begründung	Du solltest deinen Umgang mit den Schüler- innen überdenken, da du sie und ihre Arbeit nicht lobst, kann dem Verhältnis der Schüler- innen zum Physikunterricht nachhaltig schaden. Ja, weil die denken, wa- warum soll ich was tun, wird nicht honoriert.
324	Vorlesen, Zustimmung AO, Beschreibung	Du solltest darauf aufpassen, keine Rollenklischees in den Unterricht einzubringen, ja, nur Jungs drangenommen und Mädchen nicht mal die Chance gelassen, selbst wenn sie gewollt hätten. #01:33:55-0#
325	Vorlesen	B2: So, hier geht es um den Unterricht insgesamt.
326	Vorlesen	Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung, Lob und/oder Kritik, erhalten, um etwas für sich beziehungsweise seinen Unterricht zu lernen?

327	Bewertung, Bewertung (2x) + Begründung (2x), Sonstiges (3x), Vorlesen,	Unterrichtsziele waren ja in Ordnung formuliert, war ja gut, dass die da ein Verständnis für kriegen. Haben wir die hier konkret nochmal? Ja (B2 öffnet Hilfe, Unterrichtsziele), ah jetzt kriege ich es nicht runtergezogen (B2 bewegt Fenster). Insgesamt möchte ich, dass die Schülerinnen und Schüler durch den Mechanikunterricht eine physikalische Sicht auf Bewegungsvorgänge in ihrem Alltag erlernen. In dieser Doppelstunde möchte ich insbesondere erreichen, dass die Schülerinnen und Schüler das Wechselwirkungsprinzip auf Beispiele anwenden können. Außerdem möchte ich, dass sie den Zusammenhang zwischen Impuls und- zwischen Impuls und Masse und Geschwindigkeit benennen können und das Prinzip des Impulserhalts auf ein Fahrbahnexperiment anwenden können. Zusätzlich möchte ich, dass die Schülerinnen und Schüler beim Experimentieren die Vorgänge beschreiben können und dabei Beobachtung und Erklärung voneinander trennen können. Gute Unterrichtsziele, grundlegend erstmal, vielleicht etwas viel, also es gleichzeitig zu schaffen, zwei Experimente durchzuführen, das dritte Axiom nochmal abgrenzen zu können mit Beispielen, generell aber, äh (.) achso Lob und oder Kritik, ja Lob, oh ich gebe ja quasi gerade Feedback, dann muss ich das auch ankreuzen. Ähm ja Unterrichtsziele sind ja gut gewählt, gute Ideen, vielleicht etwas viel, glaube ich.
328	Vorlesen, Bewertung + Begründung, Verweis auf frühere Stelle im Test	Ablauf der Stunde, ja ähm Kritik da schon weitestgehend genannt mit den harten Cuts.
329	Vorlesen, Bewertung (4x)	Ähm Auswahl, Umfang der Stundeninhalte, wie ich gerade sagte ähm eigentlich gut, das Ex- mit dem Experiment ist gut, generell das mit der Tabelle ist ja ganz gut, dass man die benutzt, die Umsetzung haperte dann etwas, ne also, aber mit Umfang, ich glaube, es könnte generell etwas viel sein.
330	Vorlesen, Bewertung, Beschreibung, Alternative	Medieneinsatz (..) neutral, ist in Ordnung, ähm kein Video gezeigt, vielleicht kann man ein Video zeigen, aber ein Video ist immer so- vielleicht was mit Impuls generell bei- bei- bei Autounfällen ganz prominent, aber eigentlich in Ordnung der Medieneinsatz.
331	Vorlesen, Beschreibung	Auftreten der Lehrkraft, ja zum Beispiel beim Rechnen so runtergeguckt, die Schülerinnen konnten dich nicht sehen, die Schüler und Schülerinnen konnten dich nicht sehen, ähm das Stottern, die leichte Unsicherheit, das Ignorieren der Schülermeldungen, also was heißt ignorieren ähm, dass (man das?) einfach so hinnimmt, was die sagen und nicht kommentiert oder manchmal soll man es ja auch nicht kommentieren, aber speziell ähm (.) dass es keinen Einfluss hat, was sie sagen.

332	Vorlesen, Zustimmung AO + Begründung	Umgang mit den Schüler- innen, ja, da diese Sexismusrichtung, die Wertschätzungsrichtung.
333	Vorlesen, Beschreibung, Bewertung, Bewertung + Begründung, Überlegung zum Unterricht, Sonstiges	Fachliche Korrektheit (.) ja, wenn ich- wenn ich sage, dass ich auch ein Lob ausspreche, dann ja, weil soweit richtig. Fachlich war alles weitestgehend korrekt, zumindest ist mir kein grober Fehler aufgefallen, habe ich aber auch nicht intensiv drüber nachgedacht, weil ich auf die anderen Sachen mehr fokussiert war. Aber so die Formel und der ganze Rest war soweit richtig, nur einmal äh die Einheit beim Impuls vergessen. #01:36:50-2#
334	Vorlesen	B2: Robert hat schon einige Rückmeldungen erhalten. Besonders wichtig ist ihm jetzt noch eine Rückmeldung zu seinem Umgang mit den Schüler- innen und der fachlichen Korrektheit. #01:36:59-6#
335	Vorlesen	B2: Welchen Feedback möchten Sie Robert zu seiner Interaktion mit den Schüler- innen geben?
336	Vorlesen, Zustimmung AO, Verweis auf frühere Stelle im Test	Es ist nicht gut, dass du die Beiträge von den Schülerinnen kaum berücksichtigst, ja, auch schon erwähnt.
337	Vorlesen, Bewertung, Alternative	Es ist nicht gut, dass du so lehrerzentriert unterrichtest, (.) äh da du so die Schüler- innen so nicht aktiv ein Verständnis für die Inhalte entwickeln. Das Lehrerzentrierte, also dieses lehrerzentrierte Lehrer-Schüler-Gespräch wo er sagt, was richtig und was falsch ist, ist eigentlich in Ordnung. Kann man glaube ich so machen, aber man muss mehr auf die Beiträge eingehen.
338	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du die Jungen und Mädchen unterschiedlich behandelst, ja.
339	Vorlesen, Zustimmung AO, Alternative	Es ist nicht gut, dass du Sicherungspha- ne. Es ist gut, dass du Sicherungsphasen eingebaut hast, damit du sichergehst, dass die Schüler- innen die zentralen Inhalte aus der Stunde mitnehmen. Wichtig, ich hätte die Sicherungsphasen nur leicht anders gestaltet.
340	Vorlesen, Überlegung zum Unterricht, Beschreibung, Sonstiges	So verschiedene Sozialformen, Einzelarbeit, Partnerarbeit, da du den Unterricht so abwechslungsreicher gestaltest. Verschiedene Sozialformen führen nicht zwangsläufig zu besserem Unterricht und ähm (.) sind jetzt auch in den Videos nicht so rausgekommen, deswegen würde ich dazu nichts sagen an der Stelle. Das stand zwar in den ähm in dem Verlaufsplan, aber da habe ich nicht ähm (..) kann ich nicht fundiert zu was sagen, glaube ich.

341	Vorlesen, Zustimmung AO, Beschreibung (2x), Alternative + Begründung	Es ist nicht gut, dass du im Unterricht nur von wenige Schüler- innen Beiträge anhörst, da die restlichen Schülerinnen so nicht aktiviert werden. Ja, es sind immer nur zwei, drei Wortmeldungen ähm (.) da kann mehr kommen, vor allem, weil das eine kleine Klasse ist, einundzwandig Leute, da kann man schon durchaus mehr Leute drannehmen, (.) dann macht man vielleicht weniger in der Stunde, aber es kommen mehr Wort- äh äußerungen dazu, mehr Interaktion.
342	Vorlesen, Zustimmung AO, Alternative	Deine Verständnisversicherung mit einer Frage, wie „Habt ihr noch Fragen?“ ist nicht zielführend, ja, die antworten darauf nicht, man muss gezielt nachfragen.
343	Vorlesen, Überlegung zum Unterricht, Professionalisierung, Beschreibung	Du thematisierst die auftretenden Schülervorstellungen nicht, das ist nicht gut, da sich diese so verfestigen. Kann man pauschal nicht sagen, äh kommt auf den Bereich der Physik an. Habe ich jetzt in der- im- im im Impulsverständnis nicht so drinne, bei der Kraft hat er sie ja schon ein bisschen thematisiert mit dem Comic. (..)
344	eigene Erfahrung, Überlegung zum Unterricht	Äh hier an der Stelle eigentlich auch ganz witzig, ähm wie wenn ich das mal- wenn ich da mal leicht ausschweifen darf. Wenn ich diesen Unterricht gegeben habe und ich dann frage, 'hast du das verstanden' und dann kommt ein unsicheres 'ja', das ist ungefähr eine ähnliche Frage, wie 'hast du noch Fragen' und dann frage ich aber auch oft hinterher 'ist das jetzt ein ja, damit ich die Klappe halte oder ist das ein ja, weil du es verstanden hast' und da gibt es hier vielleicht auch noch Probleme. #01:39:19-1#
345	Vorlesen	B2: Reflexion insgesamt. Welche der unten stehenden Hinweise zu seiner Interaktion mit den Schüler- innen möchten Sie Robert für eine erneute Durchführung geben?
346	Vorlesen, Ablehnung AO	Dein Umgang ist angemessen- nein.
347	Vorlesen, Zustimmung AO	Du solltest die Schüler- innen grundsätzlich stärker in den Unterricht einbeziehen, zum Beispiel in Diskussionen, bei der Durchführung von Experimenten oder der Formulierung von Merksätzen- ja.
348	Vorlesen, Zustimmung AO, Ablehnung AO + Begründung	Du solltest den Schüler- innen ausführlicher Feedback zu ihren Beiträgen geben, äh (.) ob ihre Aussagen korrekt sind. (.) Jein, mehr Rückmeldung generell ja, aber wenn man nur Antworten sammeln möchte, dann sollte man nicht direkt sagen, ob es richtig ist, dann findet man auch nicht raus, was vielleicht in den Köpfen steckt.
349	Vorlesen, Bewertung	Du solltest dir die Schülervorstellungen zum Thema vor dem Unterricht ansehen, damit du im Unterricht angemessen auf diese eingehen kannst. Ja, ging jetzt nicht so. Bei- mich- beim dritten Axiom vielleicht ein bisschen, beim Impuls fast gar nicht, eigentlich gar nicht.

350	Vorlesen,Überlegung zum Unterricht, Zustimmung AO	Du solltest dir grundsätzlich Gedanken darüber machen, wie, zum Beispiel mit welchen Aufgabenstellungen, du das Verständnis der Schüler- innen abprüfen kannst, um sicherzugehen, dass die Schüler- innen die Ziele der Stunde erreichen. Geht so wahrscheinlich bisschen in Richtung vielleicht mal eine Rechenaufgabe. (..) Weiß ich nicht. (.) Man müsste mehr prüfen, ob sie es verstanden haben, stimmt schon, ja Rechenaufgabe, das habe ich ja gerade selber dazu gesagt, stimmt schon, ja.
351	Vorlesen, Zustimmung AO	Du solltest deinen Umgang mit den Schüler- innen grundsätzlich überdenken. Dass du sie nicht lobst und wertschätzt, kann dem Verhältnis- ja, genau das.
352	Vorlesen, Zustimmung AO, Bewertung, Alternative	Es ist wichtig, dass du darauf aufpasst, keine Rollenklischees einzubringen, ja, darf nicht, darf überhaupt nicht, sollte (.) genau gegenteilig behandelt werden. #01:40:49-2#
353	Vorlesen	B2: So zur fachlichen Korrektheit.
354	Vorlesen, Zustimmung AO + Begründung, Bewertung	Den Unterricht hast du insgesamt gut durchgeführt, da er überwiegend fachlich korrekt war. (..) Ja, überwiegend fachlich korrekt, würde ich glaube ich unterschreiben. (.) Abgrenzung davon halt in den Rundungsfehlern, in den Einheiten, mal gucken, was hier noch so aufkommt. Es ist nicht-
355	Vorlesen, Ablehnung AO	ah du hast fachliche Fehler in der Unterrichtsstunde gemacht. Das ist nicht gut, da du den Schüler- innen so falsche Inhalte beibringst. Grundlegend war es ja richtig.
356	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du die Abweichungen beim experimentellen Nachweis der Impulserhaltung als Rundungsfehler bezeichnest, ja genau das.
357	Vorlesen, Zustimmung AO, Verweis auf frühere Stelle im Test	Es ist nicht gut, dass du die Einheit des Impulses nicht besprichst, auch das. Das ist das hier überwiegend, da kommt das überwiegend her, die beiden Sachen sind es halt nicht so.
358	Vorlesen, Professionalisierung, Zustimmung AO, Ablehnung AO	Du thematisierst die auftretenden Schülervorstellungen nicht. Das ist nicht gut, da sich diese so verfestigen. Kenne ich mich leider aus dem Kopf nicht gut genug aus, vielleicht, vermutlich, ist aber nicht so meine Position.
359	Vorlesen, Zustimmung AO, Alternative + Begründung	Es ist gut, dass du so großen Wert auf die Einheit der Geschwindigkeit legst, da die Einheit zentraler Bestandteil einer Größenangabe ist. (.) Wert auf die Einheit legen ja, aber dann stringent und nur die Einheit der Geschwindigkeit, die meistens ja bekannt ist, Kilometer pro Stunde, genau die kennen die Leute, oder Meter pro Sekunde, da muss man nicht so großen Wert drauf legen, da kommen sie von selber drauf.
360	Vorlesen, Zustimmung AO + Begründung	Es ist nicht gut, dass du bei der Durchführung des Demonstrationsexperimentes sagst, du würdest exakt messen- ja, das ist (diese?) eine Messung reicht, eine Messung gibt alles an, was gemessen (unv.), richtig, das ist genau das falsche Bild. #01:42:11-4#

361	Vorlesen	B2: Zur fachlichen Korrektheit für eine erneute Durchführung.
362	Vorlesen, Zustimmung AO, Alternative	Ja, du solltest auch bei der Besprechung der Impulse auf die Einheiten achten, um einheitlich vorzugehen. Am Ende, mitten drin nicht.
363	Vorlesen, Zustimmung AO + Begründung	Es wäre besser, die Schülervorstellungen, die im Unterricht auftreten, zum Beispiel dass ein Apfel Kraft „hat“, aufzugreifen, um den Schüler- innen die fachlich- ja, bei Kraft auf jeden Fall. Wenn sie auftreten, das ist so ein ganz starkes Ding, dass etwas Kraft hat, das sollte man auf jeden Fall aufgreifen.
364	Vorlesen, Zustimmung AO, Verweis auf frühere Stelel im Test	Du solltest dir die Schülervorstellungen zum Thema vor dem Unterricht angucken, damit du mit ihnen vertraut bist. Ja, vorher auch schon angekreuzt.
365	Vorlesen, Zustimmung AO	Es wäre besser, auch auf Mech- Messunsicherheiten beim Fahrbahn-Experiment einzugehen, Impulserhaltung trotz der leicht unterschiedlichen Werte- ja.
366	Vorlesen, Zustimmung AO, Alternative, Alternative + Begründung	Du solltest darauf achten, dass deine Merksätze fehlerfrei formuliert sind, damit die Schüler keine Fehler aus deinen Unterlagen übernehmen. Ja, generell, natürlich, mhm das umgeht man so ein bisschen, wenn man die Schüler dazu animiert, den Merksatz selber zu machen, natürlich muss man selber einen für sich selber skizzieren, ähm um das fachlich richtige im Kopf zu haben, definitiv.
367	Vorlesen, Sonstiges*, Bewertung	Du solltest stärker auf deine Wortwahl achten, um nicht ausversehen falsche Vorstellungen zu befördern. (.) Wüsste ich nicht, worauf sich das jetzt konkret bezieht, zum Beispiel das mit dem Schwung fand ich eigentlich in Ordnung. #01:43:21-0#
368		[Video Prompt Schulnote] #01:43:28-5#

369	Vorlesen, Beschreibung, Bewertung (3x), Bewertung + Begründung (4x), Sonstiges, Überlegung zum Unterricht	B2: Okay, ich gebe dir folgende Note, eins bis sechs. (..) Mhm eine eins ist es nicht, da lief zu viel falsch. Eine zwei ist es auch nicht, wenn ich die- die- die Wortlaute der Noten, also gut, würde ich den Unterricht als gut bezeichnen, nein, würde ich ihn als befriedigend oder ausreichend- ich tendiere mehr Richtung ausreichend. Jetzt frage ich mich, ist es denn noch ausreichend. (.) Ich würde sagen fast nicht, ich würde sogar sagen nicht, wenn ich so überlege, also ohne- ohne zu fies oder sowas zu werden, aber Gesamtnote des Unterrichts, das Sexismusproblem, das Klischeeproblem, (.) diese Vorfertigung ohne darauf einzugehen, im Endeffekt, wenn ich (darüber?) überlege haben die Schüler das verstanden und haben sie sich gut behandelt gefühlt und würden sie sich vielleicht auf die nächste Stunde freuen, um Interesse zu fördern, nein, nicht ausreichend. Ich vermute nicht, dass jeder das verstanden hat, also das kann auch nicht jeder verstehen, aber ich ver- vermute auch nicht, dass ein Großteil das verstanden hat, dieses aufdrücken mit- (.) mit dem vorgefertigten Merksatz, keine Berücksichtigung der Schüler, ist das ausreichend, kann das ausreichen, um den Schülern das ordentlich ausreichen? Nein, dann kann ich die Vier glaube ich nicht geben, ich glaube es wäre tatsächlich sogar eine Fünf vor dem Hintergrund, dass es nicht ausreichend ist und nicht fair ist. #01:44:49-1#

O.3 Interview 3: LD3

Segment	Kodierung	Transkript
0	-	[Einverständniserklärung, Demographie] #00:05:35-1#
1	-	I: Okay. #00:05:36-3#
2	Sonstiges	B3: (unv.) #00:05:37-0#
3	-	I: Genau, dann ist das quasi die Seite mit der der Test inhaltlich los geht ähm #00:05:43-3#
4	Sonstiges	B3: Mhm[Ja]. #00:05:43-6#
5	-	I: Genau ich würde nochmal diese Laut-Denken-Aufforderung, also quasi die die Aufgabe vorlesen // und danach schalte ich mich quasi stumm und bin stiller Beobachter und melde mich nur, wenn Sie eben aufhören zu sprechen. Ähm ja also es ist so, während der Testdurchführung werden Ihnen gleich immer wieder Videos begegnen und zu diesen Videos werden anschließend Fragen gestellt und ich bitte Sie, alle Gedanken während der Bearbeitung der ganzen Fragen beziehungsweise Aufgabenstellung immer laut auszusprechen und dann am besten auch, wenn Sie gerade einen Text oder eine Aufgabenstellung lesen, die laut vorzulesen. Es // geht nicht darum mir irgendwie Ihren Gedanken zu erklären, sondern ja sondern laut ähm einfach laut zu denken, also als wären Sie alleine in einem Raum, das können Sie sich dabei auch äh gut vorstellen, als würden Sie einfach zu sich selber sprechen. Genau und wie gesagt, ich würde mich jetzt stumm schalten und mich nur dann einschalten, wenn ich nichts von Ihnen höre (.) genau. #00:06:45-6#
6	Sonstiges (2x)	B3: Mhm[Ja]. #00:05:52-0#
		B3: Okay. #00:06:23-1#
7	Vorlesen	B3: Ähm äh stellen Sie sich folgendes vor, Sie absolvieren zurzeit ein Praktikum an der Schule. Ihr Mitpraktikant Robert hat eine Doppelstunde seines Physikunterrichts aufgezeichnet und bittet Sie, ihm dazu eine Rückmeldung zu geben. Er zeigt Ihnen Ausschnitte seines Unterrichts und stellt Ihnen dazu Fragen. Um diese Fragen zu beantworten können Sie Robert unter anderem darum bitten, Ihnen noch einmal das Unterrichtsvideo zu zeigen. Aber beachten Sie bitte die Zeit, (die-?) denn Robert hat nur siebzig Minuten Zeit. Deswegen sollen Sie die Videos maximal einmal wieder- deshalb sollten die Videos maximal einmal wiederholt werden. Schauen Sie sich also nur in dringenden Fällen die Videos ein weiteres Mal an. Robert ist mit seinem Unterricht nicht zufrieden und bittet Sie um ein Feedback. #00:07:27-8#
8	-	[Video Lerngruppe] #00:07:55-9#
9	-	I: Jetzt muss ich einmal mich ganz kurz doch noch einschalten. Ich gehe davon aus, dass Sie eben was gehört haben, ne, also das Video? #00:08:03-4#
10	Sonstiges	B3: Ja. #00:08:04-6#

11	-	I: Gut, ich habe das nämlich nicht gehört, wahrscheinlich weil ihr Geräteaudio nicht freigegeben ist. Äh // so lange das klappt, ist das in Ordnung, wenn das für Sie egal ist, könnten Sie das sonst vielleicht noch mit freigeben, das müsste unten linkts irgendwie so ein- #00:08:18-9#
12	Sonstiges (2x)	B3: Achso. #00:08:10-7#
		B3: Ja, müsste dann jetzt funktionieren. #00:08:22-4#
13	-	I: Super, dankeschön. Dann bin ich jetzt wieder ruhig. #00:08:25-4#
14	-	[Video Vorwissen, Video Unterrichtsziele, Video Stundenraster] #00:12:04-4#
15	Vorlesen	B3: Ich habe die Stunde mit der Vorstellung des Ablaufs begonnen. Als nächstes habe ich ein Quiz zur Wiederholung der Newtonschen Axiome angeleitet, welches die Schülerinnen in Partnerarbeit durchgeführt haben. Die Durchführung des Quiz selbst habe ich zwar nicht mitgefilmt, dafür aber die Besprechung des Quiz mit den Schülerinnen, die ich dir jetzt zeige. #00:12:25-3#
16	-	[Video Einleitung Reflexion 01, Video Szene Reflexion 01] #00:14:51-4#
17	Vorlesen, Zustimmung AO	B3: Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung erhalten, um etwas für sich beziehungsweise seinen Unterricht zu lernen? Zu den Formulierungen der Axiome. #00:15:10-3#
18	-	I: Ich grätsche nochmal ganz kurz rein, ich kann natürlich nicht in Ihren Kopf gucken, falls Sie (unv.) schon an was Konkretes gedacht haben, (.) dann wär das gut, das einfach mitzusagen quasi. Also wenn Sie jetzt eben- ich weiß es ja nicht, falls Sie einfach jetzt gedacht haben // Axiome sieht gut aus, nehme ich, dann- (passt?) das, aber genau.
19	Sontiges (2x)	B3: Mhm[Ja]. #00:15:29-1#
		B3: Okay. Merke ich mir.
20	Vorlesen	Äh (.) Robert hat schon einige Rückmeldungen erhalten. Besonders wichtig ist ihm jetzt noch eine Rückmeldung zu seiner Formulierung des Merksatzes zum dritten Newtonschen Axiom. #00:15:45-4#
21	Vorlesen	B3: Äh welches Feedback möchten Sie Robert zu seiner Formulierung des dritten Merksatzes geben? (...)
22	Ablehnung AO + Begründung	Die erste Antwort finde ich nicht so passend, da nicht deutlich wurde, in welche Richtung die Kräfte wirken. (4s)
23	Zustimmung AO + Begründung, Überlegung zum Unterricht (2x)	Die Formulierung könnte auch zu komplex gewesen sein, da sehr viele Fachbegrifffe genutzt wurden und es dann schwierig ist die Schüler äh dass die Schüler es verstehen und es dann eigentlich nur dazu führt, dass die Schüler es auswendig lernen und dann reproduzieren. (6s)

24	Vorlesen, Alternative + Begründung	Äh es ist nicht gut, dass nur die Kurzformel Kraft gleich Gegenkraft, aber nicht actio gleich reactio im Merksatz vorkommt. Diese Abkürzung ist nämlich etwas, was die Schülerinnen vielleicht kennen. (.) Jedoch kann durch die Kurzformel das äh Verständ- (vom?) Kräftegleichgewicht bei den Schülern auftauchen, weshalb ich das nicht in einem Merksatz (.) formulieren würde. (.)
25	Zustimmung AO	Äh es ist aber gut gelungen, dass in dem Merksatz deutlich wurde, dass die Kräfte an verschiedenen Körpern angreifen. (5s)
26	Zustimmung AO + Begründung	Und wie die letzte Antwort sagt, finde ich das (...) ja schwierig, dass Kraft gleich Gegenkraft im Merksatz vorkommt bezüglich des Kräftegleichgewichts. #00:17:14-0#
27	Vorlesen	B3: Welche der unten stehenden Hinweise zu seiner Formulierung des dritten Merksatzes zum dritten Axiom möchten Sie Robert für seine erneute Durchführung der Unterrichtsstunde geben?
28	Vorlesen, Ablehnung AO, Verweis auf frühere Stelle im Test	Deine Formulierung des dritten Axioms ist gut so, du kannst sie wieder so verwenden. Ähm nicht, wie die e-Frage vorher schon gezeigt hat.
29	Vorlesen, Zustimmung AO + Begründung (2x)	Es wäre besser, zuerst das Axiom ohne Beispiel zu notieren und anschließend ein Beispiel, um den Merksatz zu vereinfachen. Ja, da es so dann nicht zu Verwirrung im Merksatz kommen kann und sich die Schüler danach am Beispiel die Inhalte des Merksatzes nochmal verdeutlichen können. (.)
30	Vorlesen	Es wäre besser, das Beispiel im Merksatz ausführlicher darzustellen, damit die Schülerinnen die Anwendung vom dritten Axiom verstehen könnten. (.)
31	Vorlesen, Zustimmung AO + Begründung	Es wäre sinnvoll, das Axiom um eine Skizze zu ergänzen, damit es anschaulicher wird. (.) Ähm ja, da dadurch die Wirkung der Kräfte verdeutlicht werden könnte.
32	Vorlesen, Zustimmung AO	Es wäre besser, im Merksatz zu verdeutlichen, dass Kraft und Gegenkraft gleichzeitig wirken. Deine Formulierung anschließend übt Körper B eine Gegenkraft auf Körper A aus ist fachlich falsch. Finde ich sehr passend.
33	Vorlesen, Zustimmung AO + Begründung	Du solltest im Merksatz über den Kräften Vektorpfeile notieren, damit die Wirkrichtung der Kräfte deutlich wird. (.) Ja, damit da dann auch deutlich wird, dass Kräfte eine vektorielle Größe sind.
34	Vorlesen	Und du solltest im Merksatz verdeutlichen, dass Kraft und Gegenkraft in genau entgegengesetzte Richtungen wirken wirken, damit die Schülerinnen die Bedeutung des Vorzeichens im Beispiel verstehen. #00:18:40-6#
35	-	[Video Einleitung Reflexion 02, Video Szene Reflexion 02] #00:20:53-3#

36	Vorlesen	B3: Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung erhalten, um etwas für sich beziehungsweise seinen Unterricht zu lernen? (4s)
37	Vorlesen, (Zustimmung AO) + Begründung	Schülervorstellungen, da (.) die Schülervorstellungen wahrscheinlich zu der zu den Schwierigkeiten des Axioms führen
38	Vorlesen, (Zustimmung AO) + Begründung	und Gesprächsführung, da er sehr wenig auf die Aussagen der Schüler eingegangen ist. #00:21:16-2#
39	Vorlesen	B3: Robert hat schon einige Rückmeldungen erhalten. Besonders wichtig ist ihm jetzt noch eine Rückmeldung zu den aufgetretenen Schülervorstellungen im Unterrichtsausschnitt. #00:21:25-8#
40	Vorlesen	B3: Welche der folgenden Schülervorstellungen sind während des Unterrichtsausschnitts sichtbar geworden? (.)
41	Vorlesen, Überlegung zum Unterricht	Äh die Schülerinnen denken, dass Kräfte nur dann auftreten, wenn sich zwei Körper berühren, (..) habe ich jetzt nicht aktiv (.) gesehen.
42	Vorlesen	Die Schülerinnen schreiben dem äh Apfel eine Kraft als Eigensch- schaft zu.
43	Vorlesen, Beschreibung	Die Schülerinnen denken, dass nur aktive, sich bewegende Körper eine Kraft ausüben können, passive leisten Widerstand. Das wurde bei dem einen Schüler deutlich, der meinte der Kra- äh Baum steht da ja nur und kann keine Kraft ausüben. (.)
44	Vorlesen	Die Schülerinnen denken, dass einer Kraft immer eine Beschleunigung zugrunde liegen muss und schließen daher eine Kraftwirkung bei ruhenden Gegenständen aus.
45	Vorlesen	Und die Schülerinnen verwechseln Kraft und Gegenkraft mit den wirkenden Kräften, die einen Körper im Kräftegleichgewicht halten.
46	Vorlesen	Die Schülerinnen sehen nicht, dass der Apfel über den Baum mit der Erde verbunden ist und der Baum durch das Festhalten an der Erde eine Gegenkraft zu dem (.) zum am Baum ziehenden Apfel ausübt. #00:22:32-7#
47	Vorlesen	B4: Welche der unten stehenden Hinweise zum Übergang zum Umgang mit Schülerfehlvorstellungen möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben?
48	Vorlesen, Beschreibung	Dein Umgang mit den Schülervorstellungen ist angemessen, das kannst du wieder so machen. An sich ist er ja noch nicht aktiv auf Schülervorstellungen eingegangen.
49	Vorlesen, Zustimmung AO + Begründung	Du solltest den Prozess der Kraftwirkung thematisieren, um die Schülerfehlvorstellungen aus dem Weg zu räumen. (.) Ja, damit deutlich wird, dass auch ruhende Körper eine Kraft auswirken können.
50	Vorlesen	Du solltest den Unterschied zwischen Kraft und Energie thematisieren, um zu ver- um zu verdeutlichen, warum die Schülerfehlvorstellungen falsch sind.

51	Vorlesen	Du solltest erklären, verdeutlichen, dass Kräfte auch dann wirken können, wenn man nichts aktiv tut. Da kö- Dazu kannst du das Rollbrettexperiment verwenden: Zwei Schülerinnen stehen je ein Rollbrett und halten je ein Ende einer Schnur. Egal ob beide Schülerinnen ziehen oder nur einer eine, treffen sich die Rollbretter in der Mitte.
52	Vorlesen, Zustimmung AO + Begründung	Du solltest mit den Schülerinnen die verwendeten Begriffe und den Unterschied zwischen dem Kräftegleichgewicht und dem dritten Newtonschen Axiom klären, damit ihnen klar wird, warum ihre Vorstellungen falsch sind. (...) Ja, finde ich auf jeden Fall wichtig, dass man nochmal die Begriffe genau definiert (.) und den Unterschied verdeutlicht, (.) also den Unterschied zum Kräftegleichgewicht.
53	Vorlesen, Zustimmung AO, Alternative	Du solltest die Schülerinnen das Wechselwirkungsprinzip ausführlich mündlich oder schriftlich formulieren lassen, damit du siehst, was die Schülerinnen verstanden haben beziehungsweise Fehlvorstellungen bestehen. (.) Ja, aber da sollte er dann nochmal drauf achten, (.) dass der Merksatz nicht einfach auswendig gelernt wird, sondern die Schülerinnen, (.) die äh (eine?) eigene Wortwahl verwenden. #00:24:13-8#
54	Vorlesen	B3: Bei der Besprechung des Comics hatte ich das Gefühl, dass die Schülerinnen (einige?) Schülervorstellungen haben. Deswegen habe ich die Schülerinnen gebeten, Stellung zu einer Seite des Comics zu beziehen, also Argumente entweder für die Seite des Mädchens oder des Jungens zu sammeln. Diese Argumente haben wir dann besprochen und Anna hat fachlich korrekt erklärt, warum die Position des Mädchens die richtige ist. Anschließend hatten die Schülerinnen die Möglichkeit, offen gebliebene Fragen zum Comic zu stellen. Da die Schülerinnen hier keine Fragen hatten, habe ich den Unterricht fortgeführt und zum Thema Impuls übergeleitet. #00:24:47-0#
55	-	[Video Einleitung Reflexion 03, Video Szene Reflexion 03] #00:26:42-2#
56	Vorlesen	B3: Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung erhalten, um etwas für sich beziehungsweise seinen Unterricht zu lernen?
57	Vorlesen, (Zustimmung AO) + Begründung (2x)	Ähm (...) Fachsprache, und die Schülerantworten noch wenig Fachsprache beinhalteten. #00:27:08-8# da er noch nicht so intensiv auf die Fachsprache eingegangen ist
58	Vorlesen	B3: Robert hat schon einige Rückmeldungen erhalten. Besonders wichtig ist ihm jedoch eine eine Rückmeldung äh zur Nutzung der Fachsprache und zur Gestaltung des Phasenwechsels im Unterrichtsausschnitt. #00:27:20-0#
59	Vorlesen	B3: Welches Feedback möchten Sie Robert zur Nutzung der Fachsprache geben?

60	Vorlesen	Es ist nicht gut, dass die Schülerinnen den Begriff Schwung nutzen, da die Bedeutug des Begriffs noch nicht geklärt ist und so unklar ist, was die genau die Schülerinnen beschreiben.
61	Vorlesen, Zustimmung AO, Alternative (2x)	Es ist gut, dass die Schülerinnen zunächst ihre Alltagssprache verwenden können ,um zu erklären, warum es zur Kollision gekommen ist. So sinkt die Hemmschwelle für die Schülerinnen, ihre Ideen zu äußern. (.) Ja, jedoch sollte dann genau nochmal darauf eingegangen werden, was Schwung dann in der Physik bedeutet und dann sollte direkt der Fachbegriff eingeführt werden.
62	Vorlesen	Es ist gut, zunächst mit dem alltagssprachlichen Begriff Schwung einzusteigen, da der Begriff Schwung ein tragfähiges Konzept für den Impulsbegriff aus dem Alltag der Schülerinnen ist.
63	Vorlesen	Formulierung stärker bremsen ist nicht gut gewählt, da auch Körper mit großem Impuls, wie zum Beispiel die Titanic, stärker bremsen könnten, (.) um eine Kollision zu vermeiden. (.)
64	Vorlesen, Zustimmung AO, Alternative	Es ist nicht gut, dass die Schülerinnen mit dem Begriff Schwung arbeiten. So werden sie sich diesen Begriff merken und zukünftig nicht mit dem Impulsbegriff arbeiten. (...) Auch wieder ein guter Aspekt, jedoch (.) (ah?) ich glaube, wenn von Anfang an (.) dann genau die strikt gesagt (unv.) Schwung ist nur ein Begriff, den man aus dem Alltag kennt und dann direkt die Fachsprache eingeführt ist, könnte das funktionieren und wenn man dann halt nochmal klar den Begriff Schwung aus dem Unterricht verbannt in Anführungsstrichen.
65	Vorlesen	Es ist nicht gut, dass du die Schülerinnen in der Benutzung des Begriffs Schwung bestärkst. So förderst du ihre Fehlvorstellungen, zum Beispiel (..) dass beide Begriffe gleichgesetzt werden können und dass Körper mehr Kraft haben, wenn sie mehr Schwung besitzen. (...) #00:29:05-0#
66	-	I: Was denken Sie? #00:29:07-2#
67	Zustimmung AO + Begründung	B3: (...) Ja, dass das eigentlich ein guter Punkt ist, den ich bei der ersten Antwort noch nicht so (.) bedacht habe, da habe ich, glaube ich, nur (.) äh den Kontext alltägliche Welt beachtet und nur drüber nachgedacht, dass das die Motivation oder das Interesse der Schüler steigern könnte, aber dann halt trotzdem noch problematisch sein könnte, wenn sich die (.) Vorstellung dann bei den Schülern einbrennt in Anführungsstrichen. (6s)
68	Vorlesen, Ablehnung AO	Es ist nicht gut, dass du stockend- dass du so stockend redest und viele Füllwörter nutzt. Das lenkt die Schülerinnen vom Inhalt ab. Finde ich prinzipiell erstmal nicht, dass er sehr stockend im Video geredet hat. (..)

69	Bewertung, Alternative	Mhm ja wahrscheinlich ist es dann doch nicht so gut, dass direkt am Anfang der Begriff Schwung genutzt wurde, vielleicht solle er dann direkt den Begriff Impuls einführen. #00:30:10-3#
70	Vorlesen	B3: Welche der unten stehenden Hinweise zu der Nutzung von Fachsprache möchten Sie Robert für eine erneute Durchführung des der Unterrichtsstunde geben?
71	Vorlesen	Es wäre besser, die Begriffe Masse und Impuls gegebenenfalls mit der zugehörigen Formel einzuführen, damit die Schülerinnen die korrekte Fachsprache lernen.
72	Vorlesen	Es wäre besser, wenn du die Schülerinnen dazu aufforderst, mehr Fachbegriffe zu nutzen, zum Beispiel den Massebegriff, damit sie sich so dem Impulsbegriff annähern können. (.)
73	Vorlesen, Zustimmung AO + Begründung	Es wäre besser, den Impulsbegriff für den Begriff Schwung schon früher einzuführen, um zu verdeutlichen, um welches physikalische Konzept es geht. (..) Ja, damit dann auch nicht der Begriff Schwung sich zu (.) verinnerlicht bei den Schülern
74	Vorlesen, Zustimmung AO, Alternative	und du solltest gegen die Schülervorstellung vorgehen, dass Körper Schwung haben, vergleichbar mit der Schülervorstellung Kraft haben, damit diese nicht sich nicht verfestigt. Ähm ja aktiv auf die Schülervorstellungen einzugehen ist auf jeden Fall sehr sinnvoll, dann könnte man ja nochmal (.) auch das Beispiel mit Kraft haben im Unterricht anführen.
75	Vorlesen, Ablehnung AO + Begründung	Die Verwendung der Begrifflichkeiten war angemessen, das kannst du wieder so durchführen. (..) An sich wurden ja auch noch (kaum?/keine?) Fachbegriffe, deswegen würde ich nicht sagen, dass das angemessen war.
76	Vorlesen, Beschreibung	Du solltest dir in der Vorbereitung genauer überlegen, was du den Schülerinnen erzählen möchtest, damit du flüssiger und motivierender vortragen kannst. (..) Da er ja eigentlich bisher sich nur die Schülerantworten angehört hat, hat er ja noch nicht so viel selbst vorgetragen, sage ich mal.
77	Vorlesen, Zustimmung AO + Begründung (2x)	Und es wäre besser, die Schülerinnen zur Verwendung der Fachsprache aufzufordern, um zu vermeiden, dass sie die Konzepte Schwung und Impuls weiterhin verwechseln. Ja, damit sie sich dann direkt auch an den Begriff Impuls gewöhnen und den direkt richtig definieren können und es dann nicht zu falschen Vorstellungen kommt. #00:32:01-7#
78	Vorlesen	B3: Welches Feedback möchten Sie Robert zur Gestaltung des Phasenübergangs geben? (...)
79	Vorlesen	Der Phasenübergang ist dir nicht gut gelungen, da er zu abrupt war.
80	Vorlesen	Der Phasenübergang ist dir (.) nicht gut gelungen, da er zu kurz war.

81	Vorlesen, Zustimmung AO + Begründung	Der Phasenübergang ist die nicht gut gelungen, da die Inhalte zum Wechselwirkungsprinzip nicht ausreichend verstanden und gesichert wurden. (.) Ja, es gab ja noch viele Probleme zum dritten Newtonschen Axiom.
82	Vorlesen, Überlegung zum Unterricht	Es ist unproblematisch, dass der Phasenübergang abrupt war, da du mit dem Impulsbegriff wirklich ein komplett neues Thema beginnst. (..) Komplett neu ist das Thema ja nicht.
83	Vorlesen, Zustimmung AO + Begründung	Es ist nicht gut, dass du so einen abrupten Übergang zwischen den Themen gemacht hast, da der rote Faden der Stunde so für die Schülerinnen nicht erkennbar ist. (.) Ja, da könnten sich die Schüler ja direkt fragen, warum (.) kommt das Thema jetzt, warum brauchen wir das, was hat das mit dem Thema vorher zu tun. (.)
84	Vorlesen, Ablehnung AO + Begründung	Es ist nicht gut, dass du sagst, das Thema ist komplett neu, da den Schülerinnen so das bisher Gelernte nicht mehr relevant erscheint. (..) Nicht mehr relevant, würde ich nicht sagen, nur dass die Verbindungen dann einfach nicht (.) deutlich werden und nicht klar wird, warum jetzt das Thema in der Reihenfolge dran ist.
85	Vorlesen	Und es ist nicht gut, dass du sagst, das Thema ist komplett neu, da es einen fachlichen Bezug zwischen den Newtonschen Axiomen (.) und dem Impulsbegriff gibt. (4s)
86	Bewertung + Begründung	Und den Phasenübergang würde ich auch als sehr abrupt bezeichnen, da er ja dann direkt mit dem Beispiel zur Titanic angefangen hat und da nicht wirklich Bezüge zu den Themen vorher hatte. #00:33:41-0#
87	Vorlesen	B3: Welche der unten stehenden Hinweise zur Gestaltung des Phasenübergangs möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben? (..)
88	Vorlesen, Ablehnung AO	Den Phasenübergang hast du gut gestaltet, das kannst du so wieder machen. Paar Änderungen würde ich (ihm?) schon ans Herz legen.
89	Vorlesen, Zustimmung AO + Begründung	Es wäre besser, den Zusammenhang zwischen den Newtonschen Axiomen und dem Impulsbegriff zu verdeutlichen, da die Schülerinnen den Zusammenhang nicht unbedingt selbst erkennen können. (.) Ja, vor allem da es ähm vorhin noch die Probleme zu den (unv.) der Axiome gab, kann ich mir nicht vorstellen, dass da direkt die Bezüge zum Impulsbegriff schon deutlich wurden.
90	Vorlesen, Zustimmung AO + Begründung	Es wäre besser, den Einstieg in den Impulsbegriff auf die nächste Stunde zu verschieben, um den Schülerinnen zunächst noch mehr Möglichkeiten zu geben, (.) das Wechselwirkungsprinzip zu verstehen und anzuwenden. (..) Ja, ähm vor allem, wenn klar ist, gut es gibt noch zu den Themen Fragen, ist es, glaube ich, nicht unbedingt förderlich dann jetzt erstmal mit einem anderen Thema weiter zu machen und die Fragen dann noch einfach im Raum stehen zu lassen. (.)

91	Vorlesen, Ablehnung AO + Begründung	Es wäre besser, vor dem Einstieg in das neue Thema auch die Tafel zu wischen, damit die Schülerinnen sich voll auf das neue Thema konzentrieren können und nicht noch über das alte Thema nachdenken. (.) Ich denke nicht, dass ein Tafelanschrieb unbedingt so sehr ablenkend sein könnte, (.) vor allem, weil die Schülerinnen jetzt ja auch nicht so jung aussahen. (.)
92	Vorlesen	Und es wäre besser, einen glatteren Übergang zu gestalten, zum Beispiel durch eine kurze Verknüpfung,wie 'Wir finden jetzt etwas neues mit dem gelernten Wissen heraus', damit der rote Faden für- damit der rote Faden der Stunde für die Schülerinnen nachvollziehbar ist.
93	Vorlesen, Zustimmung AO + Begründung, Alternative	Es wäre besser, den Übergang durch das Zeigen eines Experiments zu gestalten, um den Bruch zwischen den Themen für die Schülerinnen weniger hart zu gestalten. (.) Ein Experiment könnte dann wiederum auch nochmal die Motivation oder das Interesse steigern, vor allem, wenn es dann vielleicht auch aus alltäglichen Kontexten aus der alltäglichen Lebenswelt kommt, (.) deswegen denke ich mal, dass das keine schlechte Idee wäre. #00:35:45-3#
94	-	[Video Einleitung Reflexion 04, Video Szene Reflexion 04] #00:37:54-8#
95	Vorlesen	B3: Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung erhalten, um etwas für sich beziehungsweise seinen Unterricht zu lernen? (..)
96	Vorlesen, Zustimmung AO, (Zustimmung AO) + Begründung, Bewertung, Alternative + Begründung	Eingesetzte Beispiele, da (.) das äh Experiment nicht sehr wissenschaftlich war. Gut, um, glaube ich, die Gedanken vom Impuls ein bisschen einzuführen, aber ich glaube, er hätte mehr deutlich machen sollte, (.) dass es jetzt kein wissenschaftliches Experiment ist, weil wir ja nicht die Masse oder die Geschwindigkeit vorher bestimmt haben. Was ja dann auch eigentlich die Wahl des Experiments ein bisschen ist.
97	Zustimmung AO	Und zur Ergebnissicherung. #00:38:37-2#
98	Vorlesen	B3: Robert hat schon einige Rückmeldungen erhalten. Besonders wichtig ist ihm jetzt noch eine Rückmeldung zur Wahl des Experiments und zur Ergebnissicherung nach dem Experiment. #00:38:46-7#
99	Vorlesen	B3: Welches Feedback möchten Sie Robert zur Wahl des Experiments geben?
100	Vorlesen	Das Experiment ist gut gewählt, da die beiden Schüler, die das Experiment durchführen, den Impuls am eigenen Körper erfahren können.

101	Vorlesen, Zustimmung AO, Alternative + Begründung (2x)	Das Experiment ist gut gewählt, da es den Einfluss von Masse und Geschwindigkeit auf den Impuls gut darstellt. (.) An sich ja, jedoch hätte ich es, glaube ich, besser gefunden, wenn er von Anfang auch sagt, gut je schneller der Ball ist, das bedeutet, er hat eine höhere Geschwindigkeit, schwerer bedeutet, dass die Masse größer ist, dass man die beiden Aspekte der Formel dann direkt schon im Experiment ein bisschen mehr verdeutlich hätte und nicht so sehr noch in der Alltagssprache geblieben wäre. (.)
102	Vorlesen, Zustimmung AO + Begründung	Es ist nicht gut, dass nur über die Masse und die Geschwindigkeit gesprochen wurde, da unklar bleibt, ob der Impuls noch von anderen Größen, wie zum Beispil der Form abhängen kann. (..) Mhm ja finde ich auf jeden Fall gut, weil er eigentlich nur die Formel angeschrieben hat und nicht nochmal genauer darauf eingegangen ist, was der Impuls überhaupt bedeutet.
103	Vorlesen, Überlegung zum Unterricht, Zustimmug AO, Ablehnung AO + Begründung	Das Experiment ist nicht gut gewählt, da es den Zusammenhang beziehungsweise die Abgrenzung des Impulses zur kinetischen Energie nicht veranschaulicht. (.) Mhm da müsste man das Experimente, glaube ich (.) (da?) kommt es auf die Einordnung des Experiments, glaube ich, gut an. Wenn er es als wissenschaftliches Experiment verkauft, dann ja, aber vielleicht um (.) so als ganz ganz kurzen Einstieg, (.) um vielleicht auch die Formel und die Abhängigkeiten von Masse und Geschwindigkeit zu verdeutlichen, finde ich an- finde ich es an sich gar nicht so schlecht, auch (weil?) die Schüler mit eingebunden werden. (.)
104	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du die beiden Schüler in das Experiment einbezogen hast. Es ist gut, dass du die beiden Schüler in das Experiment einbezogen hast. Ja
105	Vorlesen, Ablehnung AO	äh und es ist gut, dass du vorgibst, wer nach vorne kommen soll. (.) So vermeidest zu lange Organisations und Wartezeit nach der Frage 'Wer will denn'. (..) Ich denke nicht, dass es da zu sehr langen Organisations- oder Wartezeiten kommt.
106	Vorlesen, Zustimmung AO + Begründung	Und es ist gut, dass du nur zwei Schüler das Experiment durchführen lässt und nicht alle Schülerinnen, da du so Zeit sparst und das Experiment so einfach ist, dass auch von den sitzenden Schülerinnen gut nachvollzogen werden kann. (..) Ja, ich glaube, wenn er das von allen S- äh Schülern einmal- wenn er das mit allen Schülern einmal machen würde, würde es einfach zu viel Zeit in Anspruch nehmen und das Ergebnis des Experiments wäre dafür (ja?) nicht so wichtig, dass man so viel Zeit dafür in Anspruch nimmt. #00:41:21-1#
107	Vorlesen	B3: Welche der unten stehenden Hinweise zur Wahl des Experiments möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben? (..)

108	Vorlesen, Ablehnung AO	Das Experiment ist gut gewählt, du kannst es erneut einsetzen. (..) Ja, würde ich, glaube ich, verallgemeinert nicht sagen.
109	Vorlesen, Ablehnung AO + Begründung	Und es wäre besser, statt des Experiments ein Video zu zeigen, in welchem gezeigt wird, welchen Einfluss verschiedene Massen und Geschwindigkeiten auf den Impuls haben, (.) da das die Schülerinnen eher anspricht, als ein Experiment. (.) Muss ja nicht unbedingt sein, weil die Schüler ja durch das Experiment selbst (.) aktiviert werden und dann selbst handeln können, was ähm auch interessenfördernd sein kann.
110	Vorlesen, Zustimmung AO + Begründung	Es wäre besser, ein anderes Experiment durchzuführen, das für alle Schüler*innen anschaulich ist, zum Beispiel eine Person, die auf einem Skateboard steht und einen Ball wegwirft, oder das Werfen beider Bälle in den Sand Betrachtung der Krater. (...) Stimmt, das könnte auch gut sein, weil jetzt ja nur der eine Schüler, der den Ball zugeworfen bekommt überhaupt die (..) überhaupt merkt, was sich verändert und die anderen Schüler ja eigentlich zu zugucken.
111	Vorlesen, Bewertung, Alternative	Es wäre besser, ein Stoßexperiment durchzuführen, da du so gut auf die Impulserhaltung überleiten kannst. (..) Das fände ich, glaube ich, nicht direkt am Anfang des Ex- äh Unterrichts gut, sondern dann, wenn mehr über den Impuls gesprochen wurde und deutlich ist, was der Impuls ist und da kann man ja dann zu der äh Erhaltungsgröße überleiten.
112	Vorlesen, Zustimmung AO + Begründung (2x)	Es wäre besser, ein Experiment zu wählen, zu dem eine Messreihe aufgenommen werden kann, damit die Formel aus diesen objektiven (.) Messungen abgeleitet werden kann. (.) Auf jeden Fall, weil dann schon von Anfang an auch klar wird, gut, Masse und Geschwindigkeit spielt hier eine wichtige Rolle (.) und man dann direkt mehr schon auf die Fachsprache achtet und mehr in diesem physikalischen Kontext ist und nicht in diesem Alltagssprachlichen.
113	Vorlesen, Ablehnung AO	Es wäre besser, wenn du das Experiment von allen Schülern durchführen ließest, da es den Schülerinnen gut tut, sich zwischendurch äh zu bewegen. (..) Glaube ich jetzt nicht unbedingt.
114	Vorlesen, (Zustimmung AO) + Begründung	Du solltest darauf achten, dass die Schüler vorne so stehen, dass alle sitzenden Schüler erkennen können, (.) wie leicht, schwer es ist, den Ball zu fangen. (.) Könnte auf jeden Fall dann dazu beitragen, dass die Schüler, die sitzen (und?) nicht das Experiment (.) durchführen, dann ein bisschen mehr (.) nachvollziehen können, was jetzt hier das Ergebnis sein soll. #00:43:52-1#
115	Vorlesen	B3: Welches Feedback möchten Sie Robert zur Ergebnissicherung geben?

116	Vorlesen, Zustimmung AO + Begründung	Es ist nicht gut, dass du den Schülerinnen einen fertigen Merk- vorgibst, da das reine Lesen und Abschreiben nicht verstehens- verständnisfördern ist verständnisfördernd ist. Ähm ja, also nur durch das reine Aufschreiben, (.) ist ja nicht klar, gut, habe ich das (jetzt?) verstanden oder gibt es noch Fragen. (.)
117	Vorlesen, Zustimmung AO, Alternative + Begründung	Es ist gut, dass die Schülerinnen den Merksatz ins Heft übertragen, damit alle Schülerinnen den Merksatz als Sicherung notiert haben. (..) Das alle Schülerinnen den Merksatz haben, ist auf jeden Fall wichtig, aber dann sollte (natürlich?) noch mehr darüber gesprochen werden. Nur das im Heft Stehen- nur dadurch, dass es im Heft steht, haben die Schüler es ja nicht alle verstanden.
118	Vorlesen, Zustimmung AO, Alternative	Es ist nicht gut, dass du Er- die Ergebnisse des Experiments nicht sicherst, so ist unklar, was die Schülerinnen was die Schülerinnen aus dem Experiment mitnehmen sollten. (..) Ja, da wäre es ja dann angemessener, wenn man das Ergebnis vielleicht als Messreihe mit einer Messwerttabelle macht. (..)
119	Vorlesen, Zustimmung AO, Alternative + Begründung	Und der Merksatz ist gut gelungen, da die Je-umso-Sätze die Zusammenhänge von Masse und Geschwindigkei- Masse, Geschwindigkeit und Impuls nachvollziehbar darstellen. (4s) Das ist an sich nicht verkehrt, aber dann sollten, glaube ich, noch andere (.) Aspekte einbezogen werden, wie zum Beispiel eben schon gesagt die Form des Körpers, das könnte am Ende, glaube ich, noch zu Fragen führen.
120	Vorlesen, Zustimmung AO + Begründung	Es ist nicht gut, dass du den Impuls als skalare Größe einführst. So grenzt du den Impuls nicht ausreichend von der kinetischen Energie ab. (.) Ja, so wird ähm auch nicht deutlich, dass der Impuls noch eine Richtung haben kann, also viele Aspekte (.) werden hier gar nicht betrachtet.
121	Vorlesen, Alternative + Begründung	Es ist gut, dass du den Begriff der Energie nicht erwähnst. (.) So vermeidest du Verwirrung bei den Schülerinnen, die die verschiedenen Fachbegriffe auf Grundlage ihrer Erfahrung nicht trennen können. (..) An sich wäre es ja besser, wenn er den Begriff nochmal erwähnt und beide Begriffe klar voneinander abtrennt, damit halt nicht irgendwelche Verwirrungen oder Vermischungen dann in die Köpfe der Schüler gelangen.
122	Vorlesen, Zustimmung AO, Verweis auf frühere Stelle im Test	Und in deiner Sicherung fehlt eine Abgrenzung des Impulsbegriffs von der kinetischen Energie und der Kraft, da die Begriffe den Schülerinnen zwar bekannt sind, aber Unterschiede für sie nicht ersichtlich sind. Ja, das gehört ja eigentlich zu der Aussage da drüber. #00:46:17-9# (.)
123	Vorlesen	B3: Welche der unten stehenden Hinweise zur Ergebnissicherung möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben? (..)

124	Vorlesen, Ablehnung AO + Begründung	Die Ergebnisse des Experiments hast du gut gesichert, das kannst du wieder so machen. (...) Nicht wirklich, er hat ja nur die beiden je-mehr- (.) äh (.) Sätze (.) formuliert, aber so richtig der Bezug zum Experiment wurde ja auch nicht deutlich.
125	Vorlesen, Zustimmung AO	Es wäre besser, die Beobachtungen des Experiments schriftlich an der Tafel zu samm- zu sammeln, zum Beispiel in einer Tabelle), damit du im Verlauf der Stunde darauf zurückgreifen kannst. Auf jeden Fall.
126	Vorlesen, Alternative + Begründung	Es wäre besser, die Schülerinnen in die Sicherung mit einzubeziehen, zum Beispiel einen Merksatz formulieren zu lassen, damit sie stärker aktiviert und involviert werden. (..) Das könnte man ja nach dem Experiment machen, dass er alle Schüler dann dazu auffordert, einen Merksatz zu formulieren und die dann am Ende vorgestellt werden und man dann vielleicht nochmal zusammen einen allgemeinen Merksatz zusammen bastelt, dann könnte er ja auch mehr schauen, ob die Schüler(innen?) das- (..) die Inhalte zum Impuls überhaupt verstanden haben. (.)
127	Vorlesen, Zustimmung AO	Es wäre besser, zusätzlich zum Merksatz noch ein Beispiel zu zeigen, um auch die Formel anschaulich darzustellen. (..) Das könnte auf jeden Fall für das Verständnis helfen. (.)
128	Vorlesen, Zustimmung AO (2x) + Begründung (2x)	Es wäre besser, in der Sicherung auch auf das Experiment einzugehen, zum Beispiel indem du auf der Folie zum Merksatz auch auf das Experiment eingehst oder Skizzen ergänzt, um die Inhalte stärker zu vernetzen. (.) Ja, wenn man vorher schon ein Experiment gemacht hat, sollte ja in der weiteren Sicherung auch nochmal drauf eingegangen werden, damit die Schülerinnen und Schüler wissen, warum das überhaupt gemacht wurde und Skizzen (.) finde ich auf jeden Fall wichtig, (.) da man da dann auch nochmal gut verdeutlichen könnte, dass der Impuls ja (unv.) eine skalare Größe ist und eine Richtung besitzt. (.)
129	Vorlesen	Es wäre besser, den Impuls als vektorielle Größe einzuführen, um ihn von der kinetischen Energie abzugrenzen. (..)
130	Vorlesen, Zustimmung AO, Alternative	Es wäre besser, auch auf die kinetische Energie der Bälle einzugehen, um an das Vorwissen der Schüler anzuknüpfen. (..) Ja und dann, um vielleicht auch nochmal aktiv zu verdeutlichen, was der Unterschied-
131	Sonstiges	(oh?) (Anm: Verbindungsfehler) (5s) was habe ich getan? (12s) Äh (...) #00:48:57-0#
132	-	I: Ich sehe das jetzt auch gerade zum ersten Mal, ich weiß leider nicht, was das ist. #00:49:04-7#
133	Sonstiges	B3: Okay. #00:49:06-5#
134	-	I: Ich gucke einmal, ob ich den Test öffnen kann, dann können wir zumindest rausfinden, ob es- (...) ja, bei mir sieht das noch normal aus. #00:49:18-2#

135	Sonstiges	B3: Internet habe ich ja auch. #00:49:21-7#
136	-	I: Ja. #00:49:23-0#
137	Sonstiges	B3: Mhm okay. (10s) #00:49:35-4#
138	-	I: Das ist ja interessant. #00:49:39-1#
139	Sonstiges	B3: Das verwirrt mich ein wenig. #00:49:41-9#
140	-	I: Ähm (.) Sie hatten eben jetzt auf aktualisieren gedrückt, oder was war das? #00:49:49-7#
141	Sonstiges	B3: Äh ich weiß es nicht. (10s) Mhm, das ist doch doof. #00:50:06-2#
142	-	I: Oder ähm (...) können Sie oben in der URL-Zeile, also in der Adresszeile das einfach nochmal aufrufen und gucken, was dann passiert? #00:50:18-8#
143	Sonstiges	B3: Ja. #00:50:21-2#
144	-	I: (Ob?) Sie vielleicht wieder zum Anfang geleitet werden, (aber?) ne, ist immer das gleiche, ne? #00:50:25-4#
145	Sonstiges	B3: Ja mhm. #00:50:27-2#
146	-	I: Äh haben Sie hier auf Abbrechen oder auf Okay (.) äh gewählt. #00:50:31-4#
147	Sonstiges	B3: Äh ich hab- noch auf gar nichts. Soll ich mal auf Okay gehen? #00:50:34-5#
148	-	I: Vielleicht, (.) ich hätte es jetzt intuitiv gemacht, aber- #00:50:38-1#
149	Sonstiges	B3: Okay (.) ah. #00:50:40-4#
150	-	I: Okay. (.) Gut, dann haben wir es rausgefunden, falls es nochmal passiert. #00:50:44-8#
151	Vorlesen	B3: Äh welche der unten stehenden Hinweise zur Ergebnissicherung möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben? (.)
152	Vorlesen, Ablehnung AO	Ah (..) die Ergebnisse des Experiments hast du gut gesichert, das kannst du wieder so machen. (.) Würde ich nicht direkt sagen.
153	Vorlesen, Zustimmung AO, Alternative	Es wäre besser, die Beobachtungen des Experiments schriftlich an der Tafel zu sammeln, ja, (.) wie gesagt, in der Tabelle. (.)
154	Vorlesen, Zustimmung AO	Und ja, die Schülerinnen in die Sicherung mit einzubeziehen, (..)
155	Zustimmung AO	da noch ein Beispiel hinzuzufügen, auch vielleicht mit einer Skizze, damit die vektorielle Komponente deutlich wird. (..)
156	Zustimmung AO	Und dann nochmal die Begriffe kinet- also Energie und Impuls nochmal deutlicher voneinander abgrenzen. #00:51:34-9#

157	Vorlesen, Alternative	B3: Nachdem wir das Experiment mit den Tennisbällen abgeschlossen und die Schülerinnen den Merksatz zum Impuls abgeschrieben haben, habe ich zu einem Demonstrationsexperiment übergeleitet, an dem ich die Impulserhaltung veranschaulichen möchte: Mit dem Demonstrationsexperiment möchte ich einen Auffahrunfall simulieren. Den Schülerinnen habe ich dann erklärt, wie der Autounfall im Experiment nachgestellt wird und wie die Messwertaufnahme funktioniert. Den experimentellen Aufbau und eine Messwerttabelle habe ich vor der Stunde vorbereitet, hier habe ich dir zwei Fotos vom Aufbau und von der Tabelle mitgebracht. (6s) Hier (B3 betrachtet Tabelle) könnte er noch die Einheiten dazu schreiben, in welcher Einheit die Messung Mess- äh größen überhaupt Mess- werte überhaupt aufgenommen werden. #00:52:23-5#
158	-	[Video Einleitung Reflexion 05, Video Szene Reflexion 05] #00:55:02-7#
159	Vorlesen	B3: Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung erhalten, um etwas für sich bzw. seinen Unterricht zu lernen?
160	Zustimmung AO + Begründung	Umgang mit den Schülerinnen auf jeden Fall, da er ähm zwei Mal sch- geschlechterspezifische Stereotype da nochmal gefördert hat.
161	Vorlesen	Ähm Notation der Messwerte,
162	Vorlesen	Durchführung, Auswertung des Experiments durch die Lehrkraft,
163	Vorlesen	Messunsicherheiten,
164	Vorlesen	Wahl des Experiments,
165	Vorlesen	Aktivierung der Schülerinnen,
166	Vorlesen, Sonstiges	eingesetzte Beispiele (.) ähm. #00:55:36-8#
167	Vorlesen	B3: Robert hat schon einige Rückmeldungen erhalten. Besonders wichtig ist ihm jetzt noch eine Rückmeldung zum Umgang mit den Schülerinnen, dazu, wie er während des Experiments auf Messunsicherheiten eingegangen ist und zur Aktivierung der Schülerinnen. #00:55:49-9#
168	Vorlesen	B3: Welches Feedback möchten Sie Robert zu seinem Umgang mit den Schülerinnen geben?
169	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du sagst, dass Mädchen eine schöne Schrift haben und Jungen- Jungen bessere Werte ablesen können, da du die Mädchen so auf eine schöne Handschrift reduzierst. Auf jeden Fall.
170	Vorlesen, Zustimmung AO + Begründung	Ähm es ist nicht gut, dass du die Mädchen nur für Fleißarbeiten im Unterricht einbeziehst, da du so Stereotyp- Stereotype förderst. Ja, das könnte auch dazu führen, (.) dass die Mädchen dann denken, ja gut Physik kann ich eh nicht, ich habe nur eine schöne Schrift und sich dann auch nicht beteiligen.

171	Vorlesen, Beschreibung	Es ist gut, dass du die Mädchen für ihr Handeln im Unterricht lobst, um ihr Selbstkonzept im Physikunterricht zu stärken. (.) Hat er ja aktiv nicht gemacht, nur (für?) die schöne Schrift.
172	Vorlesen, Zustimmung AO + Begründung	Es ist nicht gut, dass du das Lob für die Handschrift von Mädchen mit einer inhaltlichen Kritik, vergessene Einheiten, kombinierst, da die Ursache hier nicht das Geschlecht ist. Ja, das kann ja noch weiter dazu führen, dass die Selbstsicherheit der Mädchen dadurch eher noch geschwächt wird und sie sich noch weniger trauen, in Physik irgendwie sich zu beteiligen.
173	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du (die?) Schülerin schon vor dem Notieren des zweiten Messwertes erneut auf die Einheit hinweist, da (.) du so das Bild festigst, Mädchen wären in Physik nicht kompetent. (.) Ja.
174	Vorlesen, Beschreibung	Und es ist nicht gut, dass du der Schülerin an der Tafel direkt sagst, welche Einheit fehlt, da du ihr so unterstellst, sie wüsste das nicht. (.) Und dadurch hat er ihr auch keine Zeit gegeben, da selbst nochmal drüber nachzudenken.
175	Vorlesen	Und es ist in Ordnung, dass du sagst, dass Paula aufgrund ihrer schönen Handschrift die Werte notieren soll, da es wichtig ist, dass alle die Werte an der Tafel gut lesen können. #00:57:30-3#
176	Vorlesen	B3: Welche der unten stehenden Hinweise zum Umgang mit den Schülerinnen möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben?
177	Vorlesen, Ablehnung AO + Begründung	Dein Umgang mit den Mädchen und Jungen ist in Ordnung so, da Paula wirklich eine gut lesbare Handschrift hat. (..) Der Umgang war ja eben nicht in Ordnung, (da er?) hier nochmal Klischees gefestigt hat.
178	Vorlesen, Zustimmung AO + Begründung	Du solltest eine Geschichte um das Experiment herum konstruieren und zum Beispiel in Form eines Videos zeigen, um die Mädchen zu motivieren. (..) Wenn er da bestimmte alltägliche Kontexte verwendet, kann das auf jeden Fall förderlich sein, um das Interesse zu steigern.
179	Vorlesen, Ablehnung AO, Alternative	Und es wäre besser, die Schülerinnen zu fragen, wer freiwillig schreiben beziehungsweise die Werte ablesen möchte, um den Schülerinnen nicht nach Geschlecht eine Aufgabe zuzuweisen. (4s) Ja nicht zwingend, er könnte ja auch einfach irgendeinen Schüler oder Schülerin dran nehmen, aber dann muss er n- ja nicht unbedingt herausstellen, du schreibst jetzt, weil du ein Mädchen bist, deswegen musst du das können. (..)
180	Vorlesen, (Ablehnung AO) + Begründung	Es wäre besser, die Aufgaben Aufschreiben, Werte Ablesen durch zu tauschen, damit alle Schülerinnen eine Aufgabe wahrnehmen. Das könnte, glaube ich, eher zu mehr Stress und einfach dazu führen, dass sich das Experiment und die Durchführung in die Länge zieht. (.)

181	Vorlesen, Zustimmung AO	Du solltest darauf achten, die Mädchen nicht für ihre Handschrift, sondern explizit für ihre physikalischen Kompetenzen zu loben, da du sonst das Bild vermittelst, Physik wäre ein Jungen- Jungsfach. (.) Mhm ja.
182	Vorlesen, (Zustimmung AO) + Begründung	Und du solltest auf geschlechtsspezifische Zuschreibungen verzichten, da du so ein unangemessenes Verhältnis zwischen den Geschlechtern im Physikunterricht erzeugst. Und dadurch wird ja auch wieder vermittelt, Physik wäre nur ein Jungsfach.
183	Vorlesen	Und du solltest grundlegend an deiner Einstellung gegenüber Mädchen und Jungen arbeiten, um dein stereotypisches Verhalten zu ändern. #00:59:19-5#
184	Vorlesen	B3: Welches Feedback möchten Sie Robert zu seinem Umgang mit Messunsicherheiten geben? (..)
185	Vorlesen, Zustimmung AO, Alternative	Es ist nicht gut, dass du sagst, du könntest exakt oder genau messen, da du so die Schülervorstellung beförderst, man könnte Größen ohne Messunsicherheiten bestimmen. (.) Ja, da sollte er direkt sagen, dass jede Größe eigentlich von Unsicherheiten behaftet ist.
186	Vorlesen, Zustimmung AO, Überlegung zum Unterricht	Du gehst nicht auf die Messunsicherheiten ein. Das ist nicht gut, da diese eine wichtige Rolle in der Physik spielen. (..) Ja, mir ist auf jeden Fall nicht aufgefallen, dass er darauf eingeht. (.)
187	Vorlesen, Ablehnung AO, Alternative	Es ist nicht gut, da du nur einen der beiden Wagen (.) wiegst, da die Schülerinnen so nicht sicher sein können, dass die Wagen wirklich gleich schwer sind. (..) Mhm ja, könnte man machen, aber ich glaube, dass es schon reicht, wenn er nur einen der Wagen wiegt. Er hätte ja auch noch einen Schüler vielleicht nach vorne nehmen können, der sich beide Wagen anguckt (..)
188	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du sagst, die Wagen wären gleich schwer, da das auch wenn sie baugleich sind nicht stimmen muss. (...) Auch wieder wahr.
189	Vorlesen	Es ist nicht gut, dass du das Gewicht mit null Komma ein Kilogramm angibst, da du hier nicht berücksichtigst, welches die signifikanten Nachkommastellen sind. (...)
190	Vorlesen, Alternative	Äh es ist gut, dass du darauf verzichtest, auch auf Messunsicherheiten einzugehen, um die Komplexitä- Komplexität des Experiments überschaubar zu halten. (..) Mhm (..) er muss die Messunsicherheiten ja nicht mit einbeziehen, aber er sollte auf jeden Fall erwähnen, dass es halt Messunsicherheiten gibt und alle Größen davon belastet sind, aber dann könnte er sagen, dass wir die jetzt nicht bei der Auswertung mit berücksichtigen, aber erwähnen sollte er die auf jeden Fall. (.)

191	Vorlesen, Ablehnung AO, Alternative + Begründung	Äh es ist gut, dass du das Gewicht des Wagens mit null Komma einem Kilogramm angibst, um Fehler in der Berechnung des Impulses aufgrund vieler Nachkommastellen zu vermeiden. (..) Ne, ich glaube, da fänd ich es besser, wenn er die Werte ein bisschen genauer angibt, damit (er?) dann die Schüler auch nochmal sehen, dass (man?) nie perfekte Messwerte rauskriegt, sondern halt immer Unsicherheiten berücksichtigen muss, oder es zu Rundungsfehlern kommen kann. #01:01:28-1#
192	Vorlesen	B3: Welche der unten stehenden Hinweise zu seinem Umgang mit Messunsicherheiten möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben? (.)
193	Vorlesen, Beschreibung	Es ist gut, wie du mit Messunsicherheiten umgegangen bist. Er ist ja gar nicht mit Messunsicherheiten umgegangen.
194	Vorlesen, Zustimmung AO	Es wäre besser, alle Nachkommastellen anzugeben, die die Laborwaage dir anzeigt, da du den Fokus auf genaue Messungen legst und die Laborwaage genauer, als null Komma ein Kilogramm wiegen kann. Ja, wenn er explizit sagt, dass er genau messen will, sollte er das auf jeden Fall mit (einbeziehen?).
195	Vorlesen, Zustimmung AO + Begründung	Es wäre besser, anstelle der Aussage, du könntest exakt messen, auch Messunsicherheiten einzubeziehen, da Messungen nie wirklich exakt sind. (..) Äh ja, dass dann auch nicht die Schüler die Vorstellung haben, dass man exakt messen könnte und man immer perfekte Messwerte (da?) rauskriegt.
196	Vorlesne	Es wäre besser, beide Wagen zu wiegen, damit du den Schülerinnen ein korrektes Bild von genauem wissenschaftlichen Arbeiten vermittelst. (...)
197	Vorlesne	Es wäre besser, beide Wagen zu wiegen, da die Wagen auch wenn sie baugleich sind nicht gleich schwer sind. (..)
198	Vorlesen, Zustimmung AO + Begründung	Äh es wäre besser, beide Wagen zu wiegen, da du im Vergleich der Massen auf Messunsicherheiten beziehungsweise Ungenau- igkeiten eingehen kannst. (..) Stimmt, das ist eigentlich ein gutes Beispiel, um das zu verdeutlichen, (.) dass es halt Messunsicherheiten gibt.
199	Vorlesen, Zustimmung AO + Begründung	Es wäre besser, beide Wagen zu wiegen, damit die Schülerinnen sich selbst überzeugen können, dass die Wagen gleich schwer sind. (..) Und stimmt, eigentlich nochmal wissenschaftliches Arbeiten zu verdeutlichen, wäre ja auch wichtig. #01:03:03-7#
200	Vorlesen	B3: Welches Feedback möchten Sie Robert zur Aktivierung der Schülerinnen geben? (..)
201	Vorlesen	Es ist nicht gut, dass du die Schülerinnen nur oberflächlich involvierst, da die Schülerinnen trotz dieser Aktivierung nicht mitdenken müssen. (.)

202	Vorlesen, Beschreibung, Alternative	Es ist gut, dass du Schülerinnen in das Experiment einbeziehst, um sie im Unterricht zu aktiveren. (..) Aber er hat sie ja nicht- er hat sie ja nur teilweise mit einbezogen, er könnte da vielleicht noch selbst für die Durchführung Schüler nach vorne bitten.
203	Vorlesen, Ablehnung AO + Begründung	Es ist gut, dass du das Experiment selbst durchführst, da du in der Durchführung im Vergleich zu den Schülerinnen dabei sehr sicher bist. (..) Ich denke nicht, dass es riesen Unterschiede gegeben hätte, (da es jetzt?) ja auch kein so komplexes Experiment ist. (.)
204	Vorlesen, Zustimmung AO + Begründung	Es ist nicht gut, dass die Schülerinnen dir ohne Arbeitsauftrag zusehen, da die Schülerinnen so keinen Anreiz haben, aufmerksam zu sein. (..) Ja, so wissen sie ja auch nicht aktiv, was sie tun sollen oder worauf sie achten sollen. (.)
205	Vorlesen, Zustimmung AO + Begründung	Es ist nicht gut, dass du die Masse des Wagens insbesondere (.) die Geschwindigkeit selbst bestimmst, da du den Schülerinnen vorher schon erklärt hast, wie sie das selbst machen könnten. (.) Ja, so könnte er die Schüler mit einbeziehen und hier dann auch direkt sehen, ob sie das vorher Erklärte verstanden haben. (..)
206	Vorlesen, Sonstiges*, Überlegung zum Unterricht	Es ist nicht gut, dass du das Experiment so lehrerzentriert durchführst, da dies für die Schülerinnen langweilig ist. (..) Finde ich schwierig ohne Kontext was dazu (zu?) sagen, da ja lehrerzentriert nicht unbedingt schlecht sein muss. (..)
207	Vorlesen, Überlegung zum Unterricht, Beschreibung	Es ist nicht gut, dass Paula von der Tafel aus und die Schülerinnen aus den hinteren Reihen den experimentellen Aufbau (.) nicht gut erkennen können, da es für sie schwer ist, dir zu folgen. (4s) (Ja?) ich glaube, dass man von hinten das schon gut erkennen kann, es war ja relativ groß (...) mhm (4s). #01:05:04-9#
208	-	I: Was denken Sie gerade? #01:05:06-4#
209	Sonstiges, Zustimmung AO, Alternative	B3: (.) Äh ich lese mir nochmal die oberen Antworten durch, (..) ja, dass er die Schülerinnen auf jeden Fall einbezieht ist gut, (..) aber es könnte halt mehr Einbezug sein. #01:05:20-8#
210	Vorlesen	B3: Welche der unten stehenden Hinweise zur Aktivierung der Schülerinnen möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben?
211	Vorlesen, Alternative, Ablehnung AO	Es wäre besser, die Schülerinnen zu fragen, wer freiwillig schreiben beziehungsweise die Werte ablesen möchte, um den Schülerinnen nicht nach Geschlecht eine Aufgabe zuzuweisen. (.) Solange er ja nicht aktiv dann sagt, okay du machst- sagst, okay du machst das jetzt wegen deines Geschlechts, (.) äh muss das, glaube ich, nicht unbedingt freiwillig sein.

212	Vorlesen, Zustimmung AO, Beschreibung	Es wäre besser, die Schülerinnen zu fragen, ob sie den Ablauf verstanden oder noch Schwierigkeiten haben, um den Schülerinnen die Möglichkeit für Rückfragen zu geben. (.) Stimmt, da ist er ja nicht drauf eingegangen, sondern (.) hat sofort mit der Durchführung angefangen.
213	Vorlesen, Alternative	Du hast die Schülerinnen ausreichend aktiviert. Das kannst du so wieder machen. (.) (Naja?), es könnte auf jeden Fall mehr Schüleraktivierung geben.
214	Vorlesen, Zustimmung AO + Begründung (2x), Beschreibung	Es wäre besser, Hypothesen der Schülerinnen abzufragen, um sie zur Beobachtung des Experiments zu motivieren. (.) Stimmt, dann hätte man auch mehr Schüleraktivierung und die Schüler werden dann zum Mitdenken angeregt, so sehen sie ja nur vorne, dass der Lehrer eben was vorführt in Anführungsstrichen und wissen ja nicht so wirklich, worauf sie achten müssen.
215	Vorlesen, Zustimmung AO + Begründung	Es wäre besser, den Schülerinnen Fragen zum Vorgehen während des Experiments oder zur Berechnung der Geschwindigkeit zu stellen, (.) um sie zum Mitdenken anzuregen. Ja, dann ist die Gefahr auch nicht so groß, dass die Schüler währenddessen einfach abschalten und denken, (.) gut, der Lehrer macht das vorne schon.
216	Vorlesen, Beschreibung	Es wäre besser, die Schülerinnen stärker in die Durchführung des Experiments einzubeziehen, zum Beispiel um Wagen anzustoßen, Werte abzulesen oder die Geschwindigkeit zu bestimmen, (.) ähm (.) Werte abzulesen, dazu wurden ja dann Schüler aktiviert und nach vorne gebeten.
217	Vorlesen, Zustimmung AO	Es wäre besser, die Schülerinnen zum Experiment nach vorn kommen zu lassen, damit sie alles gut erkennen können. (.) Ja. #01:07:21-7#
218	Vorlesen	B3: Nachdem wir alle Messwerte aufgenommen hatten, habe ich die Beobachtungen der Schülerinnen abgefragt. Um die Allgemeingültigkeit der Impulserhaltung nachzuweisen, habe ich sie dann in einem kurzen Exkurs theoretisch hergeleitet. Im Anschluss haben wir die Impulserhaltung am Experiment überprüft. #01:07:42-0#
219	-	[Video Einleitung Reflexion 06, Video Szene Reflexion 06] #01:10:33-9#
220	Vorlesen	B3: Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung erhalten, um etwas für sich beziehungsweise seinen Unterricht zu lernen?
221	Zustimmung AO + Begründung	(unv.) auf jeden Fall wieder Umgang mit den Schülerinnen, da er (hier?) wieder Klischees verfestigt.
222	Zustimmung AO + Begründung	Und Messunsicherheiten, weil ja dann am Ende das Problem mit den nicht perfekten Messwerten deutlich wird und die Schüler dann die äh Impulserhaltung nicht wirklich nachvollziehen können. #01:11:01-5#

223	Vorlesen	B3: Robert hat sch- zwar schon einige Rückmeldungen erhalten, würde aber trotzdem gerne zu allen drei aufgeführten Aspekten noch eine Rückmeldung erhalten. #01:11:10-5#
224	Vorlesen	B3: Welches Feedback möchten Sie Robert zu Umgang mit den Schülerinnen geben?
225	Vorlesen, Ablehnung AO, Alternative	Es ist nicht gut, dass du Mädels sagst, da das ein sexistischer Ausdruck ist. (..) Wäre mir jetzt neu, aber er sollte halt nicht unbedingt nur sagen, gut, ihr Mädels habt das jetzt falsch gemacht.
226	Vorlesen, Zustimmung AO, Beschreibung (2x)	Es ist nicht gut, dass du die Mädchen und Jungen so unterschiedlich behandelst, da das nicht wertschätzend gegenüber den Mädchen ist. Ja, hier ist er ja gar nicht drauf eingegangen oder hat überhaupt keine Möglichkeiten, dass die Mädchen das auch richtig gemacht haben könnten. (.)
227	Vorlesen, Zustimmung AO, Verweis auf frühere Stelle im Test	Es ist nicht gut, dass du die Mädchen und Jungen in ihrem Können so unterschiedlich einschätzt, da du so die Vorstellung förderst, Physik wäre ein Jungsfach. Ja, das wurde ja auch schon vorher deutlich.
228	Vorlesen, (Zustimmung AO) + Begründung	Und du thematisierst die Einheit des Impulses nicht. (.) Das ist nicht gut, da dies inkon- inkonsistent zur früheren Szene ist, in der du bei Paula auf die Notation der Einheit der Geschwindigkeit bestanden hast. Und es ist halt einfach relevant, neue physikalische Begriffe einzuführen und die Einheiten mit zu thematisieren, da sie ja dazu gehören.
229	Vorlesen, Alternative + Begründung	Und es ist gut, dass du die Mädchen dazu aufforderst, ihre Tabelle zu korrigieren, damit du sicher sein (kannst?), dass sie die gleichen Werte, wie die Jungen notiert haben. (.) Er sollte es einfach allgemein sagen, alle Schüler sollen ihre Tabelle korrigieren, damit alle am Ende die gleichen Werte haben. Müssen ja nicht unbedingt nur die Mädels sein. #01:12:34-5#
230	Vorlesen	B3: Welche der unten stehenden Hinweise zu seinem Umgang mit den Schülerinnen möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben? (..)
231	Ablehnung AO + Begründung	Dein Umgang mit den Schülern war ja nicht in Ordnung, weil er hier wieder Klischees verfestigt.
232	Vorlesen, Zustimmung AO + Begründung, Alterantive	Du solltest die Mädchen stärker in den Unterricht einbeziehen, zum Beis- zum Beispiel berechnete Werte vorlesen lassen, um ihr physikalisches Selbstkonzept zu stärken. Ja, aber das kann er dann wiederum ja nur auch fördern, wenn er nicht diese Klischees mit einbezieht, (.) sonst, wenn er nur seinen Umgang ja weiter pflegt, ist es ja eigentlich klar, dass die Mädchen dadurch weniger selbstbewusst werden und sich nicht aktiv beteiligen. (..)

233	Vorlesen, (Zustimmung AO) + Begründung	Du solltest auf geschlechtsspezifische Zuschreibungen verzichten, um die Schülerinnen nicht zu diskriminieren. (.) Und halt wieder nicht diese Klischees zu verfestigen. (.)
234	Vorlesen, Ablehnung AO + Begründung, Alternative	Du solltest von jedem Schüler Schülerin je nur einen berechneten Wert erfragen, da du so die berechneten Werte von mehr Schülerinnen kontrollieren kannst. (..) Naja, ich glaube, jeder Schüler wird zeitlich wieder ein bisschen schwierig, aber dann mehr Schüler mit einzubeziehen kann auf jeden Fall nicht schaden.
235	Vorlesen, Alternative (2x), Zustimmung AO, Ablehnung AO	Du solltest verstärkt schwächere Schüler in den Unterricht einbinden, da wichtig ist, dass sie die Inhalte verstehen. (..) Solange er dann nicht sagt, okay, du bist ein schwacher Schüler, kannst du bitte deine Werte vorlesen, (..) ja, (unv.) aber irgendwie finde ich die Aussage nicht so passend. Er sollte halt einfach darauf achten, dass alle Schüler mit einbezogen werden und alle die Inhalte verstehen. (..)
236	Vorlesen, Zustimmung AO	Du solltest die Schülerinnen unabhängig vom Geschlecht dazu auffordern, falsche Ergebnisse in der Tafel zu korrigieren in der Tabelle zu korrigieren, damit am Ende alle Schüler die richtigen Ergebnisse in der Tabelle notiert haben. Auf jeden Fall.
237	Vorlesen	Und du solltest hinterfragen, woher die geschlechtsspezifischen Zuschreibungen stammen, die du vertrittst, um sie zu überwinden. #01:14:35-7#
238	Vorlesen	B3: Welches Feedback möchten Sie Robert zu seinem Umgang mit Maries Einwand bzw. mit Messunsicherheiten geben?
239	Vorlesen, Ablehnung AO + Begründung	Die Reaktion auf Maries Einwand ist dir gut gelungen, da du so möglichst effizient zum zentralen Ergebnis der Impulserhaltung gelangst. (.) Aber ich denke nicht, dass er so- dass Marie es dann trotzdem verstanden hat, weil er ja nicht wirklich viel auf ihren Einwand eingegangen ist und sie einfach nur schnell abfrühstücken wollte.
240	Vorlesen, Beschreibung	Die Reaktion auf Maries Einwand ist dir gut gelungen, da Rundungsfehler eine zentrale Fehlerquelle für die Unterschiede der Ergebnisse sind. (.) Aber er ist ja nicht auf weitere Fehlerquellen eingegangen, (es?) sind nicht immer nur Rundungsfehler. (.)
241	Vorlesen	Die Ei- Reaktion auf Maries Einwand ist dir nicht gut gelungen, da du den Fokus von der experimentellen Bestätigung der Impulserhaltung hin zur Diskussion von leichten Ungenauigkeiten verschiebst. (..) Äh da du den Fokus von der experimentellen Bestätig- (.)

242	Vorlesen, Zustimmung AO + Begründung, Beschreibung	Äh es ist nicht gut, dass du nur so knapp auf den Einwand von Marie eingehst, da die Begründung nicht ausreicht, um den Schülerinnen zu verdeutlichen, dass die Unterschiede aufgrund von Messunsicherheiten zustande kommen. Ja, ich denke nicht, dass sie es so kurz (dann?) verstanden hat und den Begriff Messunsicherheiten hat er ja überhaupt gar nicht gebraucht.
243	Vorlesen, Zustimmung AO, Alternative	Es ist nicht gut, dass du nur so knapp auf den Einwand von Marie eingehst, da so unklar bleibt, (.) ob beziehungsweise wo die Schüler*innen Rundungsfehler gemacht haben. (.) Ja, da hätte er noch detaillierter drauf eingehen können. (.)
244	Vorlesen, Alternative	Es ist nicht gut, dass du die Ungleich-heit der Messwerte mit Rundungsfehlern begründest, da diese Begründung falsch ist. Er sollte auf jeden Fall auch noch andere Begründungen mit einbeziehen.
245	Vorlesen, Ablehnung AO, Alternative	Es ist nicht gut, dass die Ungleichheit der Messwerte mit Rundungsfehlern begründest, da du so den Schülerinnen die Schuld dafür gibst, dass die Er- äh Ergebnisse nicht exakt gleich sind. (.) Ich denke nicht, dass da irgendwie ein Schuldaspekt bei (.) deutlich wird oder das überhaupt eine Rolle spielt, er sollte halt einfach nochmal (.) detaillierter auf den Einwand eingehen und das ein bisschen besser erläutern. #01:16:36-8#
246	Vorlesen, Ablehnung AO + Begründung	B3: Welche der unten stehenden Hinweise zu seinem Umgang mit Maries Einwand beziehungsweise mit Messunsicherheiten möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben? (..) Ähm ich fand die Reaktion nicht angemessen, da sie halt nur sehr kurz war. (.)
247	Vorlesen, Zustimmung AO + Begründung	Du solltest den Schülerinnen angeben, auf wie viele Nachkommastellen sie runden sollten. Ja, damit könnte er solche Probleme ein bisschen minimieren.
248	Vorlesen	Es wäre besser, den Schülerinnen am Anfang mitzuteilen, dass Messunsicherheiten und Rundungsfehler nicht berücksichtigt werden, um diese Situation (.) zu vermeiden. (.)
249	Vorlesen, Zustimmung AO	Es wäre besser, ausführlicher zu begründen, wo die Rundungsfehler entstehen, damit die Schülerinnen deiner Erklärung besser folgen können. Ausführlicher ist auf jeden Fall nochmal ein guter Punkt, den er berücksichtigen sollte. (.)

250	Vorlesen, Zustimmung AO, Alternative + Begründung (2x)	Du solltest bereits bei der Vorstellung des Experiments über die vorhandene Reibung und Energieumwandlung beim Aufprall sprechen, damit den Schülerinnen in dieser Phase klar ist, dass die Ergebnisse aufgrund dieser Einflüsse nicht exakt gleich sein können. (..) Ja, da sollte er auf jeden Fall von Anfang an drauf eingehen, wenn er das erwähnt hätte, (.) hätte sich, glaube ich, Maries Einwand erübrigt am Ende und er hätte ja schon von vornherein fragen können, welche Einflüsse wohl dazu führen können, dass man keine exakt gleichen Messwerte (.) bekommt, dann hätte er die Schüler ja auch nochmal zum denken beziehungsweise nachdenken anregen können.
251	Vorlesen, Zustimmung AO	Es wäre besser, zu thematisieren, dass es bei der Durchführung des Versuches Messunsicherheiten (...) äh zu berücksichtigen gibt, damit die Schülerinnen nachvollziehen können, wieso es zu Abweichungen kommt. Ja, das hätte er auf jeden Fall vorher erwähnen sollen.
252	Vorlesen	Und du solltest das Thema Messunsicherheiten für dich selbst noch einmal aufbereiten, damit du von solchen Fragen nicht vom Thema abgebracht wirst. #01:18:24-8#
253	-	I: Was denken Sie da bei dem letzten Punkt zum Beispiel? #01:18:29-6#
254	Zustimmung AO + Begründung	B3: Äh ja, da hätte er das auf jeden Fall vorher nochmal ein bisschen mehr für sich ins Gedächtnis rufen sollen, weil er so, glaube ich, ein bisschen überfordert war mit dem Einwand und da gar nicht drüber nachgedacht hat und so dann, glaube ich, von seinem eigentlichen Plan schnell abgekommen ist. (7s) Ja. #01:18:55-3#
255	Vorlesen	B3: Äh welches Feedback möchten Sie Robert zu seinem Umgang mit der Einheit des Impulses geben? (.)
256	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du nicht über die Einheit des Impulses sprichst, da die Einheit ein wichtiger Bestandteil der Angabe des Impulses ist. (.) Ja.
257	Vorlesen, Zustimmung AO + Begründung	Und es ist nicht gut, dass du nicht über die Einheit des Impulses sprichst, da dies inkonsistent zur früheren Szene ist, (.) in der du auf die Notation der Einheit der Geschwindigkeit bestanden hast. (Ja es?) ist halt generell wichtig, alle Werte immer mit Einheiten anzugeben.
258	Vorlesen, Zustimmung AO + Begründung	Es ist gut, dass du die Einheit des Impulses nicht so stark in den Fokus gerückt hast, um die Schülerinnen nicht durch die neue Einheit zu verwirren. (.) Die Einheit gehört halt einfach zum Impuls dazu, also es ist wahrscheinlich eher verwirrender, wenn sie dann in irgendwelchen Büchern die Einheit lesen, wenn sie vorher gar nicht thematisiert wurde. (.)

259	Vorlesen, Ablehnung AO + Begründung	Es ist gut, dass du den Fokus auf die berechneten Werte und nicht auf die Einheit gelegt hast, da die gleichen Zahlen ausreichen, um die Impulserhaltung auf Grundlage des Experiments zu bestätigen. (..) Die Einheit gehört halt immernoch zum Impuls dazu, deswegen denke ich nicht, dass es da gut ist, dass die Einheit weggelassen wird.
260	Vorlesen, Ablehnung AO + Begründung, Überlegung zum Unterricht	Es ist nicht gut, dass du die Schülerinnen nicht dazu aufforderst, die Einheiten anzugeben, da so unklar ist, ob die Impulse wirklich gleich groß sind. (9s) Ja, (.) ne, also es wird ja trotzdem deutlich, dass die Zahlen ja gleich sind, dann könnten sie auch, glaube ich, ohne Einheit nachvollziehen, dass die Impulse gleich groß sind, (.) aber die Einheit gehört halt einfach zum Impuls. #01:20:34-2#
261	Vorlesen	B3: Welche der unten stehenden Hinweise zu seinem Umgang mit der Einheit des Impulses möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben?
262	Vorlesen, Ablehnug AO + Begründung	Es ist in Ordnung, dass du die Einheit des Impulses zunächst vernachlässigst, das kannst du erneut so machen. (.) Ne, da es dann am Ende wahrscheinlich eher zu Verwirrung kommt, wenn dann auf einmal trotzdem Einheiten auftauchen.
263	Vorlesen	Du solltest die Einheit des Impulses angeben, da die Einheit ein wichtiger Bestandteil der Angabe des Impulses ist. (.)
264	Vorlesen, (Zustimmung AO) + Begründung	Du solltest darauf achten, dass die Schüler beim Vorlesen der Werte auch die Einheit angeben, um in deinem eigenen Verhalten konsistent zu bleiben. Und dadurch lernen die Schüler ja auch von Anfang an, Werte werden halt immer mit Einheiten angegeben, (.) brauchen sie auch für Klausuren oder Tests.
265	Vorlesen, Zustimmung AO, Alternative	Du solltest die Einheit des Impulses oben über die Spalte der Tabelle notieren, um die Schreibarbeit in jeder Zeile zu sparen. (..) Wenn das dann nicht am Ende zu Verwirrung führt, könnte das helfen, aber ich glaube nicht, dass es so viel mehr Schreibarbeit ist, wenn man jetzt noch jeweils die Einheit dazu schreibt, das waren ja (jetzt?) auch nicht so viele Messwerte.
266	Vorlesen, Zustimmung AO	Und du solltest die Schülerinnen dazu auffordern, die Einheit des Impulses zu nennen, damit die Schülerinnen sich an die Einheit des Impulses gewöhnen. Ja. #01:21:50-4#
267	-	[Video Einleitung Reflexion 07, Video Szene Reflexion 07] #01:23:45-4#
268	Vorlesen	B3: Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung erhalten, um etwas für sich beziehungsweise seinen Unterricht zu erlernen?

269	Vorlesen, (Zustimmung AO) + Begründung	Interaktion mit den Schülerinnen, da er ja durch auch dieses Wort 'besser' die Arbeit der Schüler ein bisschen herabwürdigt
270	Zustimmung AO	und Formulierung des Merksatzes. #01:24:07-4#
271	-	I: Warum? #01:24:09-1#
272	Vorlsen	B3: Robert hat s- (5s)
273	[Segment 270] + Begründung	weil er da auch immernoch nicht die Einheiten mit berücksichtigt hat und auch immer (sehr?) nur auf die Formel eingegangen ist und noch nicht so richtig deutlich wurde, was der Impuls ist. (.)
274	Vorlesen	Äh Robert hat zwar schon einige Rückmeldungen erhalten, würde aber trotzdem gern zu den beiden aufgeführten Aspekten noch eine Rückmeldung erhalten. #01:24:32-3#
275	Vorlesen	B3: Welches Feedback möchten Sie Robert zu seiner Formulierung des Merksatzes zur Impulserhaltung geben?
276	Vorlesen, (Ablehnung AO) + Begründung	Die Formulierung des Merksatzes ist dir gut gelungen, da alle wichtigen Informationen korrekt enthalten sind. (.) Aber die Einheiten ja auch noch nicht deutlich wurden.
277	Vorlesen, Beschreibung	Es ist gut, dass du im Merksatz auf das durchgeführte Experiment verweist, da die Schülerinnen den Merksatz direkt mit der Realität verbinden könnten. (.) (Also?) richtig das Experiment hatte er ja eigentlich nicht erwähnt. (.)
278	Vorlesen, Zustimmung AO + Begründung	Die Formulierung des Merksatzes ist dir nicht gut gelungen, da sie für die Schülerinnen zu anspruchsvoll ist. Kann ich mir gut vorstellen, da ja die Schüler noch nicht in ihren Merksätzen viele Fachbegriffe benutzt haben, sondern sich nur auf die Formel eigentlich äh fokussiert haben (.) und das könnte da zu Verständnisproblemen führen, wenn dann so viele Fachbegriffe in dem Merksatz des Lehrers sind.
279	Vorlesen, Zustimmung AO + Begründung, Alternative	Es ist nicht gut, dass du deine- n vorgefertigten Merksatz nutzt und die Merksätze der Schülerinnen nicht einbeziehst, da das den Schülern gegenüber nicht wertschätzend ist. Ja, auch durch diese Aussage mit mein Merksatz ist besser, wertet er sie auch ein bisschen her- herab, da hätte er eher mit den Schülern gemeinsam nochmal einen neuen Merksatz erarbeiten können. #01:25:43-9#
280	Vorlesen, Bewertung	B3: Welche der unten stehenden Hinweise zu seiner Formulierung des Merksatzes zur Impulserhaltung möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben? (.) Ähm, ja wie gesagt, war die Formulierung noch nicht angemessen.
281	Vorlesen, Ablehnung AO + Begründung	Es wäre besser, Beispiel und Merksatz zu trennen, um die Komplexität des Merksatzes zu reduzieren. (.) Glaube ich nicht, ein Beispiel kann ja zu mehr Verständnis (unv.) dazu beitragen.

282	Vorlesen, (Ablehnung AO) + Begründung, Bewertung	Es wäre besser, zusätzlich zum langen Merksatz eine kurze Version zu formulieren, (..) da sich die Schülerinnen eine kurze Version besser merken können. Aber wenn in der kurzen Version dann die Hälfte fehlt, ist es ja auch nicht (unbedingt?) förderlich und zu lang fand ich den Merksatz jetzt nicht.
283	Vorlesen, Ablehnung AO + Begründung	Es wäre besser, den Merksatz nur in Bezug auf einen Stoßprozess zu formulieren, damit der Merksatz näher am besprochenen Experiment bleibt. (.) Aber ein Merksatz sollte ja eher allgemein formuliert sein, (.) deshalb fänd ich das nicht so passend.
284	Vorlesen, Zustimmung AO, Beschreibung	Es wäre besser, den Merksatz um die Einheit des Impulses zu ergänzen, da die Einheit stets zur Angabe einer physikalischen Größe gehört. Ja, die hat er ja bis jetzt immer noch nicht erwähnt.
285	Vorlesen, Zustimmung AO (2x)	Du solltest den Merksatz schülernäher gestalten, zum Beispiel indem du eine Skizze ergänzt oder die Formulierung sprachlich vereinfachst, um die Schülerinnen beim Verständnis zu unterstützen. (...) Äh eine Skizze kann auf jeden Fall förderlich sein, die Formulierung sprachlich vereinfachen (.) auf (jeden?) kann gut sein, solange es (nachher?) noch fachlich korrekt ist und trotzdem die wichtigsten Fachbegriffe genannt werden.
286	Vorlsen, Zustimmung AO, Bewertung, Alternative + Begründung	Es wäre besser, eine Formulierung der Schüler- innen als Merksatz zu verwenden, da diese Formulierungen für sie besser verständlich sind. (...) Ja, also ich glaube nicht, die f- Merksätze, die jetzt vorgetragen wurden, ich glaube, es wäre besser, wenn er dann mit den Schülerinnen zusammen einen Merksatz (.) dann erarbeitet, weil jetzt in den Merksätzen, die vorgetragen wurden, ja nur eigentlich die Formel dargestellt wurde. #01:27:44-4#
287	Vorlesen	B3: Welches Feedback möchten Sie Robert zu seiner Interaktion mit den Schülerinnen geben?
288	Vorlesen, Zustimmung AO, Alternative	Es ist gut, dass die Schülerinnen ihre Merksätze vorlesen, da du so erkennen kannst, ob sie dir (.) Impulserhaltung umfassend ob sie die Impulserhaltung verstanden haben. Ja, ist auf jeden Fall gut, dass er die Antworten der Schüler mit einbezieht, (.) äh jedoch hätte er da mehr drauf (eingehen?) können.
289	Vorlesen, Zustimmung AO + Begründung	Es ist gut, dass die Schülerinnen einen Merksatz formulieren sollen, um ein tiefergehendes Verständnis zu fördern. (.) Ja, so merken die Schüler ja selbst, ob sie es verstanden haben, oder nicht. (.)
290	Vorlesen, Zustimmung AO, Alternative	Es ist nicht gut, dass du den Schülerinnen keine Rückmeldung auf die Merksätze gibst, da die Schülerinnen so nicht einordnen können, ob ihre Merksätze angemessen formuliert sind. (.) Ja, also er hätte (unv.) mehr auf deren Antworten eingehen können.

291	Vorlesen, Zustimmung AO, Verweis auf frühere Stelle im Test	Es ist nicht gut, dass du deinen vorgefertigten Merksatz nutzt und die Merksätze der Schüler nicht einbeziehst, da das den Schülerinnen gegenüber nicht wertschätzend ist. Ja, das gehört ähm ein bisschen zu der Aussage davor und die Aussage danach, die spiegelt ja den Inhalt eigentlich auch wieder ungefähr wieder.
292	Vorlesen	Es ist nicht gut, dass du deine Formulierung als besser bezeichnest, da du die Formulierungen der Schüler so abwertest, was ja auch wieder nicht wertschätzend ist. (.)
293	Vorlesen, Zustimmung AO, Alternative (2x)	Und es ist nicht gut, dass nur zwei Schüler ihre Merksätze vorgestellt haben, da du so nicht weißt, welche Merksätze die restlichen Schülerinnen formuliert haben. Ja, hier hätte er auf jeden Fall noch mehr Beispiele einbeziehen können und dann vielleicht eine Mischung aus den Merksätzen dann (.) nochmal zu einem neuen Merksatz dann zusammenfassen können.
294	Vorlesen, Zustimmung AO, Alternative	Äh und es ist nicht gut, dass nur zwei Jungen ihre Merksätze vorgestellt haben, da die Schülerinnen so lernen, dass nur Ju- dass nur du nur Jungen drannimmst. Ja, er hätte auf jeden Fall drauf achten können, dass er ei- Mädchen und Jungen dran nimmt. #01:29:33-5#
295	Vorlesen	B3: Welche der unten stehenden Hinweise zu seiner Interaktion mit den Schülerinnen möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben? (.)
296	Vorlesen, Zustimmung AO + Begründung (2x)	Du solltest den Merksatz ausführlicher erklären beziehungsweise mit den Schülerinnen besprechen, da einige Aspekte des Merksatzes nicht selbsterklärend sind. (.) Ja, dadurch, dass er ja viele Fachbegriffe benutzt, wenn er jetzt eigentlich nur sagen würde, dass der Merksatz abgeschrieben werden soll, weiß er auch nicht, ob der Merksatz überhaupt verstanden wurde.
297	Vorlesen, Zustimmung AO + Begründung (2x)	Du solltest gemeinsam mit den Schülerinnen einen Merksatz erstellen, um sie stärker in den Unterricht einzubeziehen. Ja, dadurch sieht er dann auch nochmal, wie groß jetzt das Verständnis überhaupt war und die Schüler sind angeregt, mitzudenken. (..)
298	Vorlesen, Zustimmung AO + Begründung	Es wäre besser, insgesamt mehr Schülerinnen ihre Merksätze vorstellen zu lassen, da du aus den Antworten Rückschlüsse auf die Vorstellungen der Schülerinnen ziehen kannst. Ja, äh da er ja dann auch nochmal sieht, ob das Experiment verstanden wurde. (.)
299	Vorlesen	Du solltest die Schülerinnen sich melden lassen, wenn sie ihren Merksatz vorstellen wollen, da sich möglicherweise auch ein Mädchen (.) äh gemeldet hätte. (.)
300	Vorlesen	Äh der Umgang mit den Schülerinnen war angemessen, so kannst du dich wieder verhalten. (...)

301	Vorlesen, Zustimmung AO + Begründung	Äh du solltest deinen Umgang mit den Schülerinnen überdenken. Dass du sie und ihre Arbeit nicht lobst, kann das kann dem Verhältnis der Schülerinnen zum Physikunterricht nach- kann dem Verhältnis der Schülerinnen zum Physikunterricht nachhaltig schaden. Ja, wenn sie sich das- äh der Gedanke dann verfestigt, dass ihre Arbeit gar nicht gewertschätzt ist wird, könnte die Motivation weiter runter gehen und die Schüler würden sich weniger im Unterricht beteiligen.
302	Vorlesen, Zustimmung AO, Alternative	Du solltest darauf aufpassen, keine Rollenklischees in den Unterricht einzube- einzubringen, um die Schülerinnen nicht in diese Muster zu drängen. (.) Ja, da könnte er (dann?) gegen angehen, indem er dann auch mehr Mädchen in den Unterricht einbezieht und dann deren Ergebnisse auch vorgestellt werden, die er dann aber wiederum wertschätzen sollte #01:31:33-9#
303	Vorlesen	B3: Äh hier geht es um den Unterricht insgesamt.
304	Vorlesen	Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung erhalten, um etwas für sich beziehungsweise seinen Unterricht zu lernen? (...)
305	Vorlesen, (Zustimmung AO) + Begründung	Auftreten der Lehrkraft, da ja oft deutlich wurde, dass er noch Klischees (..) ähm (...) im Unterricht verfestigt hat
306	Zustimmung AO + Begründung	und auch noch gar nicht auf die äh jetzt hier im Beispiel Merksätze der Schüler eingegangen ist, deshalb auch noch Umgang mit Schülerinnen und Interaktion. #01:32:10-9#
307	Vorlesen	B3: Äh Robert hat schon einige Rückmeldungen erhalten. Besonders wichtig ist ihm jetzt noch eine Rückmeldung zu seinem Umgang mit den Schülerinnen und der fachlichen Korrektheit in der ganzen Stunde. #01:32:21-4#
308	Vorlesen	B3: Welches Feedback möchten Sie Robert zu seiner Interaktion mit den Schülerinnen geben?
309	Vorlesen, Zstimmung AO + Begründung	Es ist nicht gut, dass du die Beiträge von Schülerinnen kaum berücksichtigt, da dies demotivierend für die Schülerinnen ist. (.) Ja und dann kann ja die Frage aufkommen, warum die Schüler das überhaupt machen, wenn deren Arbeit gar nicht mit einbezogen oder gewürdigt wird.
310	Vorlesen, Zustimmung AO + Begründung, Überlegung zum Unterricht	Es ist nicht gut, dass du so lehrerzentriert unterrichtest, da die Schülerinnen so nicht aktiv ein Verständnis für die Inhalte entwickeln. (..) Ja, also mehr Schüleraktivität (kann?) auf jeden Fall förderlich sein. (.) Obwohl lehrerzentriert nicht automatisch schlecht bedeutet, kommt ja immer auch noch auf das Verhältnis an, vielleicht in den anderen Unterrichtsstunden.
311	Vorlesen	Es ist nicht gut, dass du die Jungen und Mädchen unterschiedlich behandelt, da- behandelst, da das demotivierend ist. (.)

312	Vorlesen, Alternative	Es ist gut, dass du Sicherungsphasen in deinen Unterricht eingebaut hast, damit du sichergehst, dass die Schülerinnen die zentralen Inhalte aus der Stunde mitnehmen. (.) Jedoch sollte er da nochmal genauer darau- auf das Verständnis pochen und die Schüler das nicht nur abschreiben lassen. (.)
313	Vorlesen, Zustimmung AO + Begründung	Es ist gut, dass du verschiedene Sozialformen im Unterricht einplanst, Einzelarbeit, Partnerarbeit, da du den Unterricht so abwechslungsreicher gestaltest. (..) Ja, das kann auf jeden Fall helfen, um auch die Motivation zu steigern.
314	Vorlesen, Zustimmung AO, Alternative (2x)	Es ist nicht gut, dass du im Unterricht nur von wenigen Schülerinnen Beiträge anhörst, da die restlichen Schüler so nicht aktiviert werden. (.) Ja, da könnte er auf jeden Fall (auf?) mehr Beiträge (erwarten?) und dann auch wieder nochmal den Geschlechteraspekt berücksichtigen und mehr Mädchen äh antworten lassen. (.)
315	Vorlesen, zustimmung AO, Alternative	Deine Verständnis- äh versicherung mit einer Frage, wie 'Habt ihr noch Fragen?' ist nicht zielführend, da Schüler*innen auf diese Frage eher nicht antworten. Äh ja, da könnte er dann nochmal aktiv ähm das umgehen und eher nochmal mit den Schülerinnen die Inhalte diskutieren. (.)
316	Vorlesen, Zustimmung AO, Beschreibung, Alternative	Du thematisierst die auftretenden Schülervorstellungen nicht. Das ist nicht gut, da sich diese so verfestigen. Ja, das wurde ja am Anfang vor allem bei den Newtonschen Axiomen deutlich, da hätte er nochmal aktiver auf Schülervorstellungen eingehen können. #01:34:33-7#
317	Vorlesen, Ablehnung AO	B3: Welche der unten stehenden Hinweise zu seiner Interaktion mit den Schülerinnen möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben? (.) Äh wie gesagt, sein Umgang mit den Schülerinnen war eben nicht so angemessen. (.)
318	Vorlesen	Es solltest die Schüler mehr in den Unterricht mit einbeziehen (.) und (ähm?) dadurch dann die Inhalte gemeinsam zu erarbeiten, was dann wiederum dazu führt, dass die Schüler mehr nachdenken und aktiviert werden und er dann auch mehr sieht, ob die Schüler die Inhalte jetzt verstanden haben. (..)
319	Vorlesen, Zustimmung AO, Beschreibung	Du solltest den Schülerinnen ausführlicher Feedback zu ihren Beiträgen geben, damit die Schülerinnen einordnen können, ob ihre Aussagen korrekt sind. (.) Ja, das hat man jetzt ja zum Beispiel bei den Merksätzen gesehen.

320	Vorlesen, Zustimmung AO + Begründung	Du solltest dir die Schülervorstellungen zum Thema vor dem Unterricht ansehen, damit du im Unterricht angemessen auf diese eingehen kannst, um das Verständnis der Schülerinnen zu fördern. (.) (unv.) auf jeden Fall helfen, wenn er die Schüler- Schülervorstellung vorher schon ähm im Kopf hat und dann nicht im Unterricht, sage ich mal, von den Schülervorstellungen (ein bisschen?) überrumpelt wird und das gar nicht in seinen Plan passt. (.)
321	Vorlesen	Du solltest dir grundsätzlich Gedanken darüber machen, wie, zum Beispiel mit welchen Aufgabenstellungen, du das Verständnis der Schülerinnen abprüfen kannst, um sicherzugehen, dass die Schülerinnen die Ziele der Stunde erreichen. (4s)
322	Vorlesen, Zustimmung AO, Beschreibung, Alternative + Begründung	Äh du solltest dir grundsätzlich Gedanken darüber machen, mit welchen Aufgabenstellungen du das- (.) ähm ja, also er könnte auf jeden Fall mehr Aufgabenstellungen mit einbeziehen, (.) da muss ich jetzt an das reine Abschreiben denken, das ja nicht unbedingt zielfördernd ist, da hätte er dann eher nochmal eine Aufgabe stellen können, dass die Schüler nochmal Merksätze formulieren (.) sollen, um das Ziel der Stunde dann zu festigen. (.)
323	Vorlesen	Du solltest deinen Umgang mit den Schülerinnen grundsätzlich überdenken. Dass du sie nicht lobst und wertschätzt, kann dem Verhältnis der Schülerinnen zum Physikunterricht nachhaltig schaden. (.)
324	Vorlesen, Zustimmung AO + Begründung	Und es ist wichtig, dass du darauf aufpasst, keine Rollenklischees in den Unterricht einzubringen, um die Schülerinnen nicht in diese Muster zu drängen. Ja, da hatteäh also die Rollenklischees wurden hier ja sehr sehr oft verdeutlicht. #01:36:51-4#
325	Vorlesen	B3: Welches Feedback möchten Sie Robert zur fachlichen Korrektheit der Stunde geben? (...)
326	Vorlesen, (Ablehnung AO) + Begründung	Deinen Unterricht hast du insgesamt gut durchgeführt, da er überwiegend fachlich korrekt war. Aber er hat ja oft wichtige Aspekte vergessen, wie Messunsicherheiten oder äh die Einheit des Impulses.
327	Vorlesen	Du hast fachliche Fehler in der Unterrichtsstunde gemacht. Das ist nicht gut, da du den Schülerinnen so fachliche Inhalte beibringst.
328	Vorlesen, Alternative	Es ist nicht gut, dass du die Abweichungen beim experimentellen Nachweis der Impulserhaltung als Rundungsfehler darstellst, (.) das ist fachlich falsch. Da hätte er auf jeden Fall nochmal auf Messunsicherheiten eingehen sollen.
329	Vorlesen	Ist nicht gut, dass du die Einheit des Impulses nicht besprichst, da die Einheit ein wichtiger Bestandteil der Angabe eines Impulses ist. (.)

330	Vorlesen, Beschreibung	Und du thematisierst die auftretenden Schülervorstellungen nicht. Das ist nicht gut, da sich diese so verfestigen. (.) Das hat man ja vor allem am Anfang gesehen bei den Newtonschen Axiomen, wo er dann aber mit dem neuen Thema weitergemacht hat und sich nicht nochmal aktiv mit den Schülervorstellungen auseinander gesetzt hat.
331	Vorlesen, Zustimmung AO, Alternative	Es ist gut, dass du so großen Wert auf die Einheit der Geschwindigkeit in der Tabelle legst, da die Einheit zentraler Bestandteil einer Größenangabe ist. (..) Ja, an sich nicht falsch, aber er sollte dann halt auf jede Einheit Wert legen und nicht nur bei bestimmten Größen.
332	Vorlesen, Zustimmung AO, Alternative, Alternative + Begründung	Es ist nicht gut, dass du bei der Durchführung des Demotr- Demonstrationsexperimentes sagst, du würdest exakt messen, da du den Schülerinnen das Bild vermittelst, du könntest Messunsicherheiten vermeiden. (.) Äh ja also da sollte er auf (seine?) Wortwahl achten, da man ja nicht exakt messen kann, da vor dem Experiment hätte er nochmal gut das Thema Messunsicherheiten (.) aufgreifen können. #01:38:37-8#
333	Vorlesen	B3: Welche der unten stehenden Hinweise zur fachlichen Korrektheit der Stunde möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben? (..)
334	Alternative + Begründung (2x)	Wie gesagt sollte er hier auf die Einheit achten, um, ja, einmal einheitlich vorzugehen, dass jede Größe immer mit einer Einheit behaftet ist, (..) aber auch um die Schüler dann nicht danach zu verwirren, wenn dann in Büchern oder so von Einheiten gesprochen wird. (.)
335	Vorlesen, Zustimmung AO, Alternative	Es wäre besser, die Schülervorstellungen, die im Unterricht auftreten, zum Beispiel dass ein Apfel Kraft hat, aufzugreifen, um den Schülerinnen die fachlich korrekte Sichtweise näher zu bringen. (.) Ja, da sollte (er?) auf jeden Fall Schülervorstellungen thematisieren und die dann vielleicht einfach gegenüberstellen, (unv.) zum Beispiel auch den sprachlichen Aspekt, dass man alltagssprachlich das, also Kraft haben sagen kann, aber dann im Kontext des Unterrichts eine andere Sprache verwendet.
336	Vorlesen, Zustimmung AO + Begründung	Du solltest dir die Schülervorstellungen zum Thema vor dem Unterricht angucken, damit du mit ihnen vertraut bist und Ideen entwickeln kannst, wie du ihnen begegnen kannst. (.) Ja, auf jeden Fall besser, als wenn man dann spontan dann mit den Schülervorstellungen ähm (.) konfrontiert wird. (.)
337	Vorlesen, (Zustimmung AO) + Begründung	Es wäre besser, auch auf Messunsicherheiten beim Fahrbahn-Experiment einzugehen, damit die Schüler(innen?) die Impulserhaltung trotz der leicht unterschiedlichen Werte akzeptieren. Dadurch hätte er dem Einwand von Marie vorbeugen können. (.)

338	Vorlesen, Zustimmung AO	Du solltest darauf achten, dass deine Merksätze fehlerfrei formuliert sind, damit die Schülerinnen keine Fehler aus deinen Unterlagen übernehmen. (...) Ja, da würde ich auch sagen, dass die Merksätze ausführlicher sein sollten.
339	Vorlesen	Und du solltest stärker auf deine Wortwahl achten, um nicht ausversehen falsche Vorstellungen zu befördern. #01:40:24-3#
340	-	[Video Prompt Schulnote] #01:40:30-9#
341	Bewertung, Bewertung + Begründung, Alternative	B3: (7s) So ein-e gute Drei, da auf jeden Fall gute Aspekte da waren, aber vor allem der Umgang mit den Schülern noch verbesserungswürdig ist (mit?) in Bezug auf die Rollenklischees und noch mehr auf die Arbeit der Schüler eingegangen werden könnten. #01:40:55-4#

O.4 Interview 4: LD4

Das Transkript zu Interview 4 wird zweigeteilt dargestellt: Für die ersten 32 Segmente liegt eine zweite Kodierung vor, hier werden beide Kodierungen zum Transkript notiert. In einer zweiten Tabelle wird der Rest des Transkripts mit Kodierung aufgeführt.

Segment	Kodierung	2. Kodierung	Transkript
0	-		[Aufgabenstellung (fälschlicherweise vor Einverständniserklärung), Einverständniserklärung] #00:01:35-8#
1	-		I: Und jetzt starten wir inhaltlich mit dem Test. Genau, der Test sieht so aus, dass immer wieder Videos kommen und zu diesen Videos soll dann- werden danach anschließend Fragen gestellt und wie eben schon gesagt, ich bitte dich, dass du dann während der Bearbeitung dieser Fragen beziehungsweise Aufgabenstellung alle deine Gedanken einfach immer laut aussprichst und wenn du gerade was liest dann eben das auch mit laut vorzulesen, um in diesem Redefluss zu bleiben. Es geht nicht darum, mir das irgendwie zu erklären, sondern quasi diese Vorstellung, äh du wärst alleine in einem Raum und würdest zu dir selber sprechen und genau, falls du aufhörst zu sprechen, dann, genau, kann ich nicht in deinen Kopf gucken, ich weiß nicht, warum das passiert, aber ich würde dann einfach kurz einwerfen irgendwie 'bitte weiter sprechen' oder 'was denkst du' ähm genau ja einfach so so als Reminder. Und ansonsten würde ich mich jetzt stumm schalten und bin ja quasi nicht mehr da. #00:02:31-6#

2	Vorlesen	Vorlesen	B4: Okay. (..) Okay, also Sch- erste Situationsbeschreibung. Sie sind äh absolvieren zurzeit ein Praktikum an der Schule. Ihr Mitpraktikant Robert hat eine Doppelstunde seines Physikunterrichts aufgezeichnet und bittet Sie, (.) ihm dazu eine Rückmeldung zu geben. Er zeigt Ihnen Ausschnitte seines Unterrichts und stellt Ihnen dazu Fragen. Um diese Fragen zu beantworten zu können (.) können Sie Robert unter anderem darum bitten, Ihnen noch einmal das Unterrichtsvideo zu zeigen. Aber beachten Sie bitte die Zeit, denn Robert hat nur Siebzig Minuten Zeit. Deswegen sollten die Videos maximal einmal wiederholt werden. Schauen Sie also- schauen Sie sich also nur in dringenden Fällen die Videos ein weiteres Mal an. Robert (.) äh ist mit seinem Unterricht nicht zufrieden und bittet Sie um ein Feedback. #00:03:24-2#
3	-		[Video Beschreibung Lerngruppe, Video Vorwissen, Video Unterrichtsziele, Video Stundenraster (Anm: Vor Beginn des Videos schaut B4 kurz Stundenraster an)] #00:08:11-0# B4: (liest Stundenraster) #00:09:04-8#
4	Vorlesen	Vorlesen	B4: Ich habe die Stunde mit der Vorstellung des Ablaufs begonnen. Als nächstes habe ich ein Quiz zur Wiederholung der Newtonschen Axiome ange- leitet, welches die Schülerinnen in Sch- Partnerarbeit durchgeführt haben. Die Durchführung des Quiz selbst habe ich zwar nicht gefilmt, dafür aber die Besprechung des Quiz mit den Schülerinnen, die ich dir jetzt zeige. #00:09:30-2#
5	-	-	[Video Einleitung Reflexion 01, Video Szene Reflexion 01] #00:12:02-1#

6	Vorlesen	Vorlesen	B4: Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung, Lob und/oder Kritik, erhalten, um was für (.) um etwas für sich beziehungsweise seinen Unterricht zu lernen? (.)
7	Vorlesen	Vorlesen	Formulierung Wortlaut des Merksatz zum ersten äh Axiom,
8	Vorlesen	Vorlesen	Formulierung Wortlaut zum zweiten Axiom. (.)
9	Sonstiges	Sonstiges	Okay ähm also der erste Merksatz (6s)
10	Bewertung, Alternative	Zustimmung AO, Bewertung, Alternative	also beim (.) dritten- beim- bei der Formulierung zum Merksatz zum dritten Axiom fand ich es in dem Merksatz etwas störend, dass direkt in dem Merksatz ein Beispiel schon genannt wurde, sondern ich hätte das wahrscheinlich eher getrennt gemacht, dass man halt ähm auf jeden Fall das Beispiel dahinter äh (.) macht, ja. (...)
11	Überlegung zum Unterricht	Überlegung zum Unterricht	Okay und wir war nochmal genau der erste das erste Axiom formuliert, ähm (...) ich bin grad (13s) ah, ja. (16s) (liest Hilfestellung, öffnet Hilfe 'Kann ich den Unterrichtsausschnitt noch einmal sehen?'). #00:13:57-6#
12	-		[Video Szene Reflexion 01 aus Hilfestellung, B4 überspringt Teile] #00:15:11-8#

13	Beschreibung, Überlegung zum Unterricht (3x), Zustimmung AO + Begründung, Alternative	Beschreibung, Überlegung zum Unterricht (3x), Zustimmung AO, Alternative+Begründung	B4: Okay und ähm (..) zu dem Thema Ergebnissicherung des Quizes (...) also es war wieder ein sehr äh lehrermoderierter Vorgang, das heißt er wollte ja von den Schülern immer genau diese Antworten (...) hören. (.) Ich weiß nicht, ob das so das Geschickteste ist, um- also ich bin grad am überlegen, ob das das Geschickteste ist, um so ein ähm Quiz z- (..) auf jeden Fall zu sichern, aber auf jeden Fall fand ich die Gesprächsführung (4s) mhm (4s) also wenn man bedenkt, dass es eine sehr ruhige Lerngruppe ist, muss er wahrscheinlich, wenn sich keiner freiwillig meldet oder so, einfach Leute dran nehmen, aber ähm (..) ich frage mich jetzt gerade auch in diesem Moment, die haben jetzt ja das Quiz nochmal komplett selbst, also in der Gruppe, wie es scheint, besprochen, (...) ob das so geschickt ist. (6s) Mhm (8s) also ich würde definitiv am Thema Gesprächsführung was nochmal sagen, weil ähm, er hat zwar jetzt immer nur nach dem Namen gefragt und dann halt nochmal (denen?) seinen gleichen Merksatz, den die ja wahrscheinlich schon vorher besprochen hatten, nochmal genauso vorgelesen und (.) es (.) war vielleicht sogar schon so in die Richtung doppelte Sicherung und so, das heißt da würde ich auf jeden Fall noch was ändern. (5s)
14	Sonstiges	Sonstiges	Ja, (.) so. (6s) Noch was? #00:17:29-1#
15	Vorlesen	Vorlesen	B4: Robert hat schon einige Rückmeldungen erhalten. Besonders wichtig ist ihm jetzt noch eine Rückmeldung zu seiner Formulierung des Merksatzes zum dritten Newtonschen Axiom. #00:17:42-6#
16	Vorlesen	Vorlesen	B4: Welches Feedback möchten Sie Robert zu seiner Formulierung des Merksatzes zum dritten Axiom geben?

17	Vorlesen, Ablehnung AO + Begründung, Zustimmung AO	Vorlesen, Ablehnung AO + Begründung, Zustimmung AO	Ähm, deine Formulierung des dritten Axioms ist gut, da alles wichtige enthalten ist. Ach, nein. (..) Äh also auf jeden Fall nicht, weil ich fand die auf jeden Fall schon sehr sehr komplex formuliert (.) ähm, das heißt ich wäre wahrscheinlich auf (..) genau, es ist auf jeden Fall dieser dritte Punkt, also- äh der vierte,
18	Vorlesen, Zustimmung AO, Alternative	Vorlesen, Zustimmung AO, Alternative	die Formulierung des dritten Axioms ist nicht gut, da die Vermischung von Beispiel und Merksatz die Schüler verwirren kann, das find ich auf jeden Fall, er kann das mit dem Beispiel finde ich sehr gut, (.) ähm aber auf jeden Fall ähm (.) nicht in dem Merksatz direkt, sondern eher als Beispiel noch danach.
19	Vorlesen, Zustimmung AO, Überlegung zum Unterricht, Alternative + Begründung	Vorlesen, Zustimmung AO, Überlegung zum Unterricht, Alternative + Begründung	Es ist nicht gut, dass nur die Kurzformel Kraft gleich Gegenkraft, aber nicht actio gleich reactio im Merksatz vorkommt. Diese Abkürzung ist nämlich etwas, was die Schülerinnen vielleicht äh (.) kennen. (.) Ja, also das (.) könnte man auf jeden Fall sagen, dass man ähm (.) dieses actio gleich reactio jetzt äh reinbringt, aber da ist auch jetzt finde ich so ein bisschen die Frage, (.) ob alle Schüler das schon kennen, aber das auf eine Aktion Reaktion (.) eine Reaktion folgt, ist vielleicht für die Schüler auch einfach sehr viel verständlicher, als das andere, das heißt das würde ich auf jeden Fall mit rein nehmen.

20	Vorlesen, Zustimmung AO, Überlegung zum Unterricht, Bewertung, Alternative (2x)	Vorlesen, Zustimmung AO, Überlegung zum Unterricht, Alternative	Es ist gut, dass du im Merksatz deutlich machst, dass die Kräfte verschiedenen- an verschiedenen Körpern angreifen, da du so einer Verwechslung mit äh dem Kräftegleichgewicht vorbeugst. (...) Äh es ist auf jeden Fall gut, das klar zu machen, (.) ich weiß jetzt- bin jetzt gerade am überlegen, ob es mich in der Formulierung, wie er sie hatte, hat es mich auf jeden Fall (.) mit dem Beispiel drin erstmal noch verwirrt, (.) also man (.) könnte (..) das auch- also man kann das auf jeden Fall (.) äh mit rein bringen, aber eigentlich, wenn man das vorher, dass äh Kräfte immer an einem Körper angreifen und so, (.) finde ich, muss das jetzt nicht in den Merksatz zum dritten äh Axiom mit hinein. (7s)
21	Vorlesen, Ablehnung AO	Vorlesen, Ablehnung AO	So, dann hatte ich deine Formulierung des dritten Axioms ist fachlich falsch, würde ich jetzt nicht sagen.
22	Vorlesen, Bewertung, (Ablehnung AO) + Begründung	Vorlesen, Bewertung, Ablehnung AO + Begründung	Äh die Formulierung des dritten Axioms ist für Schüler- rinnen nicht angemessen, da sie zu komplex ist. Also von der Sprache her würde ich das jetzt auch- also fand ich (die?) jetzt nicht komplex. Wie gesagt, ich fand, die Komplex- heit kam halt durch diese Vermischung von Beispiel und Merksatz.
23	Sonstiges, Vorlesen, Zustimmung AO + Begründung, Alternative + Begründung	Sonstiges, Vorlesen, Zustimmung AO + Begründung	Und jetzt noch das letzte, es ist nicht gut, dass die Kurzformel Kraft gleich Gegenkraft im Merksatz vorkommt, da Schülerinnen diese oft mit dem Kräfte- äh (.) äh Kräftegleichgewicht assozi- assoziieren. Also das würde ich auf jeden Fall sagen, ich finde die Beschreibung über actio gleich reaction eingänglicher für die Schüler und auch eine Sache, die sich Schüler sehr sehr gut merken können, das heißt ich würde da eher mit den Begriffen arbeiten, als mit Kraft gleich Gegenkraft. (...)
24	Sonstiges	Sonstiges	So, habe ich alles? (..) Ja. #00:21:42-5#

25	Vorlesen	Vorlesen	B4: Welche der unten stehenden Hinweise zu seiner Formulierung des Merksatzes zum dritten Axiom möchten Sie Robert für eine erneute Durchführung des Unterrichts eine erneute Durchführung der Unterrichtsstunde geben?
26	Vorlesen, Ablehnung AO + Begründung	Vorlesen, Ablehnung AO + Begründung	Deine Formulierung des dritten Axioms ist gut so, du kannst sie wieder so verwenden. Also würde ich auf jeden Fall nicht sagen, da mir da jetzt zu viele Sachen nicht passen.
27	Vorlesen, Zustimmung AO + Begründung	Vorlesen, Zustimmung AO + Begründung	Es wäre besser, zuerst das Axiom ohne Beispiel zu not- notieren und anschließend ein Beispiel, um den Merksatz zu vereinfachen. Auf jeden Fall, ich finde ja, also das macht es ja (.) sehr deutlich, ähm wenn du (.) in einen Merksatz noch Zahlen oder eine Formel mit drin hast, ist der immer viel komplizierter, als wenn du es einmal hinschreibst und danach erst dir ein Beispiel und die Formel und so anguckst. (.)
28	Vorlesen, Zustimmung AO, Alternative	Vorlesen, Zustimmung AO, Alternative	Es wäre besser, das Beispiel im Merksatz ausführlicher darzustellen, damit die Schülerinnen (.) die Anwendung vom dritten (.) Axiom verstehen können. Also ich finde, das ähm das Beispiel mehr auszuführen ist auf jeden Fall eine gute Idee, aber das gehört meiner Meinung nach nicht so in den Merksatz mit rein, sondern es ist für mich halt wirklich eher ein Punkt, den man nach dem Merksatz auch auf jeden Fall ausführlich machen sollte, aber halt nicht im Merksatz. (...)

29	Vorlesen, Zustimmung AO, Überlegung zum Unterricht	Vorlesen, Zustimmung AO, Überlegung zum Unterricht	Es wäre sinnvoll, das Axiom um Skiz- eine Skizze zu ergänzen, damit es anschaulicher wird. (..) Ja, also das kann man auf jeden Fall, ich bin jetzt gerade nur am überlegen, ob ähm die Sache über eine Skizze für die Schüler (ah?) viel deutlicher wird, weil was willst du da machen, du malst zwei Körper hin ähm wo dann Kräftepfeile (wieder?) zu sehen sind. Mhm (.) das ist schwer finde ich gerade. (..)
30	Vorlesen, Überlegung zum Unterricht	Vorlesen, Überlegung zum Unterricht	Es wäre besser, im Merksatz zu v-verdeutlichen, dass Kraft und Gegenkraft gleichzeitig wirken. Deine Formu-lierung anschließend übt Körper B eine Gegenkraft auf Körper A aus ist fachlich falsch. (...) Ja, (..) mhm (.) kann man natürlich sagen, (..) (ja?) mhm (unv.) da ist jetzt die Überlegung, ob das für die (...) also das das anschließend in dem Satz, (.) das- das stimmt, das könnte auch noch ein Problem sein, aber ich bin gerade am überlegen, wie man das für die Schüler sonst fachlich gut formulieren kann (..) oder so. (4s)
31	Vorlesen, Zustimmung AO, Überlegung zum Unterricht	Vorlesen, Zustimmung AO, Überlegung zum Unterricht	Du solltest im Merksatz über den Krä- äh über den Kräften Vektorpfeile notieren, damit die Wirkungäh Wirkrichtung der Kräfte deutlich wird. (.) Ja, gut also das finde ich ist etwas- wenn man ähm bis jetzt immer Kräfte mit Vektorpfeilen geschrieben hat, dann sollte man das auch weiter so machen, (.) also wenn er schon die Richtung von äh Kräften und so eingeführt hat und die mit Vektorpfeilen beschrieben hat, dann sollte man das definitiv auch weiterhin machen. (.)

32	Vorlesen, (Zustimmung AO) + Begründung, Überlegung zum Unterricht Sonstiges, Alternative + Begründung	Vorlesen, Zustimmung AO + Begründung, Überlegung zum Unterricht Alternative, Sonstiges	Du solltest im Merksatz verdeutlichen, dass Kraft und Gegenkraft in genau entgegengesetzte Richtungen wirken, damit die Schülerinnen die Bedeutung des Vorzeichens im Beispiel verstehen. (5s) Mhm (also?) das ist ja eigentlich der entscheidende Punkt (..) ähm von (..) dem (.) äh von dem Satz, äh dass die beiden entgegengesetzt und gleich groß (.) sind. (4s) Okay mhm ich bin jetzt gerade am überlegen, ob das so richtig ist. (3s) Also es (6s) mhm (6s) also das auf jeden Fall mit diesem anschließend Körper A, ich glaube, solche Formulierungen sollte man definitiv wenn- in so einem Merksatz vermeiden, vor allen Dingen, wenn sie halt zu physikalischen Problemen führen können, also das ist auf jeden Fall sinnvoll. (6s) (Ach?) (2s) ich bin jetzt gerade am überlegen nochmal wegen dem letzten mit (7s) mhm (6s) mhm (...) (so?). #00:27:21-4#

Segment	Kodierung	Transkript
33	-	[Video Einleitung Reflexion 02, Video Szene Reflexion 02] #00:29:43-8#
34	Vorlesen	LD4: Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung erhalten, um etwas für sich beziehungsweise seinen Unterricht zu lernen? (.)
35	Vorlesen	Gesprächs- (also?) der erste Punkt Gesprächsführung,
36	Vorlesen	Einsatz des Comics
37	Vorlesen, Beschreibung, Alternative (2x), Zustimmung AO	und Schülerfehlvorstellungen. Mhm (...) also ich fand jetzt in diesem äh (...) in disem Videoclip hat man ja auf jeden Fall bei den Schülern gesehen, dass die f- also noch nicht die kor- äh korrekte physikalische- also nicht alle die korrekte physikalische Erklärung hatten, das heißt (.) an dem Punkt müsste man sich auf jeden Fall (..) mit den Schülerfehlvorstellungen noch mal beschäftigen (.) genau und ähm halt auch am besten schauen, wie er das für eine erneute Durchführung so einer Stunde besser einbinden kann, ähm (.) (unv.) irgendwie mit dem Comic halt schon vorher das das machen (.) kann, oder so. Also das würde ich auf jeden Fall. (..)

38	Vorlesen, Ablehnung AO, Bewertung (2x) + Begründung (2x)	Mhm also Einsatz des Comics (4s) ich finde es ehrlich gesagt nicht, also ich fand den Comic jetzt (unv.) in Ordnung, als es ist eine gute Diskussionsgrundlage, obwohl für die Schüler wird der Comic natürlich auch (..) ähm (.) sch- schwierig sein ähm zu erklären, weil das (auf?) jetzt eine andere Situation und dieses ähm Axiom jetzt auf den Comic zu übertragen mhm (9s)
39	Beschreibung, Bewertung	äh (.) zum Thema der Gesprächsführung, also er hat ja jetzt, wenn ich das richtig gesehen habe, einfach erstmal nur moderiert, das heißt nacheinander (.) äh dran genommen und ähm war halt eher als Moderator tätig. Das also (.) fand ich jetzt erstmal also er hat die Antworten noch nicht bewertet, (.) sondern einfach nur erstmal so stehen lassen, (..) also zum Sammeln fand ich das gut von der Gesprächsführung her (.) jetzt bin ich- ich bin halt, also das würde ich auf jeden Fall nicht (unv.) (...)
40	Sonstiges, Bewertung, Alternative + Begründung	aber auf jeden Fall beim Comic, (.) mhm (..) da bin ich jetzt gerade am überlegen, ob man dazu was sagen könnte (...) mhm (9s) naja, okay. (..) Ich weiß nicht, also ich fand den Comic schon gut, aber ich glaube, ich würde den nicht so am Anfang (be?)nutzen, weil er dann doch schon schwieriger ist, sondern eher dann als ein (...) später nochmals als Anwendung, also (unv.) danach. #00:33:31-6#
41	Vorlesen	LD4: Robert hat schon einige Rückmeldungen erhalten. Besonders wichtig ist ihm jetzt noch eine Rückmeldung zu den aufgetretenen Schülervorstellungen im Unterrichtsausschnitt. #00:33:45-0#
42	Vorlesen	LD4: Welche der folgenden Schülerfehlvorstellungen sind während des Unterrichtsausschnitts sichtbar geworden?
43	Ablehnung AO + Begründung	Also es sind auf jeden Fall Schülerfehlvorstellungen sichtbar geworden, das heißt das ist schonmal (.) ähm (..) falsch.
44	Vorlesen, Zustimmung AO + Begründung	Ähm die Schülerinnen denken, dass Kräfte nur dann auftreten, wenn sich zwei (..) äh (.) Körper berühren, also das ist auf jeden Fall aufgetreten, das war ja mit ähm dass der Apfel auch am Baum hängen muss und so. (...)
45	Vorlesen, Überlegung zum Unterricht	Die Sch- Schüler schrieben dem Apfel eine Kraft als Eigenschaft (.) zu. (.) Ich bin mal grad am- wenn ich gerade nochmal da drüber nachdenke, ich (.) ich weiß jetzt gerade nicht, ob die gesagt haben, dass der (4s) ähm- da muss ich (gleich?) nochmal nachschauen, ob der Apfel jetzt wirklich die Kraft hat.
46	Vorlesen, Sonstiges	Äh die Schülerinnen denken, dass nur aktive, das heißt sich bewegende Körper, eine Kraft ausüben können, passive leisten nur Widerstand. (.) Ja, ähm okay, weiß ich jetzt gerade auch nicht.

47	Vorlesen, Ablehnung AO	Die Schülerinnen denken, dass einer Kraft immer eine Beschleunigung zugrunde liegen muss äh und schließen daher eine Kraftwirkung bei ruhenden Gegenständen aus. (.) Nein. (4s)
48	Vorlesen	Die Schülerinnen verwechseln Kraft und Gegenkraft mit den wirkenden Kräften, die einen Körper im Kräfte- (.) äh gleichgewicht halten. (.)
49	Vorlesen	Und die Schülerinnen sehen nicht, dass der Apfel über den Baum mit der- über den- über den Baum mit der Erde verbunden ist und der Baum durch das Festhalten an der Erde eine Gegenkraft am Baum- (..) zum am Baum ziehenden Apfel ausübt. (7s)
50	Sonstiges	Okay, ich muss jetzt gerade einmal nochmal (..) schauen. [LD4 öffnet Hilfe 'Kannst du mir den Comic noch einmal zeigen?'] (unv.) (..) [LD4 öffnet Hilfe 'Kann ich den Unterrichtsausschnitt noch einmal sehen?'] #00:36:10-6#
51	-	[Video Szene Reflexion 02 aus Hilfestellung, LD4 überspringt Teile] #00:37:19-6#
52	Beschreibung	LD4: Also man hat bei dem Letzten vor allen Dingen diese Schülervorstellung gesehen, dass der Apfel eine Kraft als Eigenschaft hat, äh (.) das hat man genau bei der letzten Schüleräußerung (.) auf jeden Fall nochmal gehört.
53	Vorlesen, Ablehnung AO	Die Schülerinnen denken, dass nur aktive, das heißt sich bewegende Körper eine Kraft ausüben (.) (unv.), passive leisten nur Widerstand, ich- die habe ich jetzt gerade nicht so gesehen.
54	Vorlesen	Die Schülerinnen denken, dass eine Kraft immer (.) eine Beschleunigung zugrunde liegen muss äh und schließen daher eine Kraftwirkung bei ruhenden Gegenständen aus. (4s)
55	Vorlesen	Die Schülerinnen verwechseln Kraft und Gegenkraft mit wirkenden Kräften, die einen Körper im (.) Kräftegleichgewicht halten. (...)
56	Vorlesen, Zustimmung AO	Äh die Schüler*innen sehen nicht, dass der Apfel über den Baum mit der Erde verbunden ist und der Baum durch das Festhalten an der Erde eine äh (..) Gegenkraft zum am Baum ziehenden Apfel ausübt. Das tun sie, also nicht alle auf jeden Fall. (...)
57	Sonstiges	Genau (4s) so. #00:38:44-6#
58	Vorlesen	LD4: Welche der unten stehenden Hinweise zum Umgang mit Schülervorstellungen möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben?
59	Vorlesen, Ablehnung AO + Begründung, Beschreibung	Ähm dein Umgang mit Schülervorstellungen ist angemessen, (.) das kannst du wieder so machen. (.) Äh nein also er geht ja zumindest hier jetzt noch- oder ist wahrscheinlich noch gar nicht darauf eingegangen, weil sonst müssten die ja, wenn man es auf jeden Fall gut macht, nicht noch so präsent sein, wie bei ihm.

60	Vorlesen, Zustimmung AO	Du solltest den Prozess der Kraft-wirkung thematisieren, um die Schülervorstellungen aus dem Weg zu räumen. Mhm (6s) das würde ich auf jeden Fall sagen, dass er das machen sollte.
61	Vorlesen, Überlegung zum Unterricht, Ablehnung AO	Du solltest den Unterschied zwischen Kraft und Energie thematisieren, um zu verdeutlichen, warum die Schülervorstellungen falsch sind. Ich wüsste jetzt gerade nicht, was er ähm (...) vielleicht das mit warum- aber ich weiß (da?) jetzt gerade nicht, was er da so wirklich machen will. (.) Vielleicht das ähm achso, (..) vielleicht über die ähm (.) dass die Kraft halt als die Eigenschaft vom Apfel gesehen wird, aber ich finde es eher nicht.
62	Vorlesen, Zustimmung AO	Du solltest erklären, verdeutlichen, dass Kräfte auch dann wirken können, wenn man nichts aktiv tut. Dazu kannst du das Rollbrettexperiment verwenden, zwei Schülerinnen stehen auf je einem Rollbrett und halten je ein Ende einer Schnur. Egal ob beide Schülerinnen ziehen oder nur einer, tre- äh treffen sich (.) beide Rollbretter in der Mitte. (4s) Mhm ja. (..)
63	Vorlesen	Du solltest äh mit den Schülerinnen die verwendeten Begriffe und den Unterschied zwischen dem Kräftegleichgewicht und dem dritten Newtonschen Axiom (er?)klären, damit ihnen klar wird, warum ihre Vorstellungen f-falsch sind. (5s)
64	Vorlesen, Ablehnung AO + Begründung	Du solltest (den?) Schüler- du solltest die Schüler das Wechselprinzip ausführlich mündlich oder schriftlich formulieren lassen, damit du siehst, was die Schülerinnen verstanden haben bzw. wo ähm Fehlvorstellungen liegen. (4s) Also ich weiß nicht, über das Formulieren würde ich das eigentlich nicht so tun, (..) weil ähm beim Formulieren würden die wahrscheinlich einfach nur das nochmal sagen, was sie mit dem Merksatz gelernt haben mit dem auswendig- von wegen auswendig lernen, ich glaube da bringt das an der Stelle gar nicht so viel.
65	Zustimmung AO	Ähm (.) also auf jeden Fall sollte man die- den Aspekt, dass die Kräfte (.) wirken, auch wenn man nichts äh aktiv tut (...) machen.
66	Vorlesen, Zustimmung AO	Und du solltest mit den Schülerinnen die verwendeten Begriffe und den Unterschied zwischen Kräftegleichgewicht und dem dritten Newtonschen Axiom erklären, damit ihnen klar wird, warum ihre Vorstellungen falsch (4s) sind. (.) Ja, auf jeden Fall auch. So. #00:42:15-9#

67	Vorlesen, Sonstiges	LD4: Bei der Besprechung des Comics hatte ich das Gefühl, dass die Schüler eine Schülervorstellung haben. Deswegen habe ich die Schüler gebeten, Stellung zu einer Seite des Comics zu beziehen, also Argumente entweder für eine äh für die Seite des Mädchens oder des Jungens zu sammeln. Diese Argumente haben wir dann besprochen und Anna hat fachlich korrekt erklärt, warum die Position des Mädchens die richtige ist. Anschließend haben- anschließend hatten die Schüler die Möglichkeit, offen gebliebene Frage zu- Fragen zum Comic zu stellen. Da die Schüler hier keine Fragen hatten, habe ich äh den Unterricht fortgeführt äh zum Thema Impuls übergeleitet. (...) Okay. #00:43:03-3#
68	-	[Video Einleitung Reflexion 03, Video Szene Reflexion 03] #00:45:13-3#
69	Vorlesen	LD4: Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung erhalten, um etwas für sich beziehungsweise seinen Unterricht zu lernen? (.)
70	Vorlesen, Bewertung	Also (zum?) Beispiels Titanic (.) um auf die (.) um auf die Impulserhaltung zu kommen, (.) ja gut, kann man vielleicht machen, aber so wie er es gemacht hat, fand ich es jetzt nicht wirklich gut.
71	Bewertung	Gesprächsführung fand ich in Ordnung, also dazu (.) habe ich jetzt keine Kritik.
72	Vorlesen, Beschreibung, Ablehnung AO	Ähm (.) fachliche Korrektheit (..) und Fachsprache, ja okay also (...) äh die Schüler und so kennen ja den Begriff von Impuls und so noch nicht, das heißt, dass die das haben (.) ähm Schwung (denn?) ja okay. (.) Mhm (.) finde ich jetzt, ist ja, glaub- an der Stelle noch gar nicht so wichtig,
73	Bewertung	aber der Übergang zwischen den Unterrichtsphasen, wenn er wirklich von dem Beispiel direkt einfach so übergegangen ist, sollte er da noch was dran tun. (..)
74	Sonstiges, Ablehnung AO + Begründung	Ja (.) also fachliche Korrektheit, Fachsprache, ich bin jetzt gerade am überlegen, aber eigentlich ist das jetzt ja das Einführungsbeispiel und es ging da darum, dass die Schüler ihre Vermutung und so ähm geben, das heißt (.) eigentlich finde ich, muss man- glaube ich, ist das da noch nicht so wichtig, dass er auf eine richtige Fachsprache und fachliche Korrektheit achtet. #00:47:23-9#
75	Vorlesen	LD4: Robert hat schon einige Rückmeldungen erhalten. Besonders wichtig ist ihm jetzt noch die Rückmeldung zur- zu der Fachsprache und zur Gestaltung des Phasenwechsels im Unterrichts- äh (fortschritt?). #00:47:38-8#
76	Vorlesen	LD4: Welches Feedback möchten Sie Robert zu zu der Fachsprache geben?

77	Vorlesen, Sonstiges*	Ähm es ist nicht gut, dass Schülerinnen den Begriff Schwung nutzen, da die Bedeutung des Begriffs noch nicht ge- (.) noch nicht (.) geklärt ist und so unklar ist, wo genau die Schülerinnen (noch?) Fragen- erstmal schön mit Rechtschreibfehler noch. (..)
78	Vorlesen	Mhm äh es ist gut, dass die Schülerinnen (ihm?) zunächst ihre Alltagssprache verwenden können, um zu erklären, warum es zur Kollision gekommen ist, so sinkt die Hemmschwelle für die Schülerinnen, ihre Ideen zu äußern.
79	Vorlesen	Es ist gut, zunächst mit dem alltagssprachlichen Begriff Schwung einzusteigen, da der Begriff Schwung ein tragfähiges Konzept für die äh für den Impulsbegriff aus dem Alltag für Schülerinnen ist.
80	Vorlesen	Die Formulierung stärker bremsen ist nicht gut gewählt, da auch Körper mit großem Impuls wie zum Beispiel die Titanic stärker bremsen können, um eine- könnten um eine (.) Kollision zu vermeiden. (..)
81	Vorlesen	Es ist nicht gut, dass die Schülerinnen den Beg- Begriff Schwung arbeiten, so werden sie den Begriff merken und zukünftig nicht mehr mit dem Impulsbegriff arbeiten.
82	Vorlesen, Zustimmung AO	Es ist nicht gut, dass die Schülerinnen in der Benutzung des Begriffs Schwung bestätigst, so förderst du die Vorstellung, dass äh dass beide Begriffe gleich gesetzt werden können und dass ein Körper mehr Kraft- (.) und das Körper mehr Kraft haben, wenn sie mehr Schwung besitzen. Also das ist auf jeden Fall so. (..)
83	Vorlesen	Es ist nicht gut, dass du stockend redest und viele Füllwörter nutzt, das lenkt die Schüler vom Inhalt ab.
84	Zustimmung AO + Begründung, Beschreibung, Alternative + Begründung	Mhm (4s) ich glaube, das ist eigentlich für die Schüler, also (weil?) man von der Schülerseite betrachtet, ist es eigentlich gut, dass die erstmal ihre Vermutung äußern könnt- können, warum es zur ähm Kollision gekommen ist, dass es halt auch wirklich diese Hemmschwelle Hemmschwelle senkt, aber es ist auf jeden Fall- dann hätte ich halt auf jeden Fall das gesagt, dass er sie dann nicht benutzen sollte. (...) Also es ist auf jeden Fall nicht gut, (unv.) nachher weiter benutzt (.) den Begriff (5s) mhm (5s)
85	Überlegung zum Unterricht	oh darauf (Anm: LD4 bewegt die Maus über AO 'Es ist nicht gut, dass du so stockend redest und viele Füllwörter nutzt. Das lenkt die Schüler*innen vom Inhalt ab') habe ich jetzt ehrlich gesagt gar nicht so geachtet mhm (5s) also (.) okay. #00:50:43-5#
86	Vorlesen	LD4: Welche der unten stehen Hinweise zur Nutzung von Fachsprech- sprache möchten ja Sie Robert für eine erneute Durchführung des Unterrichts geben.
87	Vorlesen	Es wäre besser, den Begriff Masse und Impuls gegebenenfalls mit zugehörigen Formeln einzuführen, damit die Schüler die ko- äh konkre- äh die korrekte Fachsprache lernen. (...)

88	Vorlesen	Es wäre besser, wenn du die Schüler auffordern würdest, mehr Fachbegriffe zu nutzen, zum Beispiel den Massebegriff, damit sie sich so dem Impulsbegriff annähern können.
89	Vorlesen	Es wäre besser, den Impulsbegriff für den Begriff Schwung schon früher einzuführen, um zu verdeutlichen, um welches physikalische Konzept es geht.
90	Vorlesen, Zustimmung AO	Du solltest gegen die Schüler-äh-vor- Schülervorstellung vorgehen, dass Körper Schwung haben, vergleichbar äh (.) äh genau, vergleichbar mit der Schülervorstellung Kraft haben, damit diese sich nicht festigt. Also das sehe ich auf jeden Fall so, dass wenn man über die Alltagsvorstellungen kommt, dass man auf jeden Fall äh gegen die Schülervorstellungen, die da entstehen können, äh, arbeitet. (...)
91	Vorlesen, Ablehnung AO	Also die Verwendung der Begrifflichkeiten war angemessen, du kannst- äh das kannst du wieder so durchführen, nein.
92	Vorlesen, Zustimmung AO	Du solltest dir äh in der Vorbereitung genauer überlegen, was du den Schülerinnen erzählen möchtest, damit du flüssiger und motivierender vortragen kannst. Das ist eigentlich immer sch- achso, ja (6s) (unv.) also das ist eigentlich immer gut, dass man sich vorher vorbereitet und äh überlegt, was man sagen will, (..) ja.
93	Vorlesen, Ablehnung AO, Zustimmung AO, Alternative	Es wäre besser, die Schülerinnen zur Verwendung der Fachsprache aufzufordern, um zu vermeiden, dass sie das Konzept Schwung und Impuls weiterhin verwechseln (.) mhm (6s) also ich weiß jetzt nicht, ob es geschickt wäre, den Begriff Masse und Impuls vorher einzuführen, ähm das kann ich glaube ich ist eher nicht so geschickt. Ähm (6s) das man also auf jeden Fall diesen Effekt, dass man die Schülerinnen auffordert, Fach- äh Fachbegriffe zu nutzen, das kann man machen, also das ist auf jeden Fall glaube ich eine gute Sache, aber dann halt immer noch, ich würde sagen, man muss halt immer noch drauf achten, dass das die dann nicht zu sehr demotiviert (.) oder ach ja. (...) Ja. #00:53:52-2#
94	Vorlesen	LD4: Welches Feedback möchten Sie Robert zur Gestaltung der äh des Phasenübergangs geben.
95	Vorlesen, Zustimmung AO	Der Phasenübergang ist dir nicht gut gelungen, er war zu abrupt, ja, würde ich definitiv sagen.
96	Vorlesen, Ablehnung AO	Äh der Phasenübergang ist dir nicht gut gelungen, da er kurz- da er zu kurz war, naja (...) mhm er war eigentlich einfach abrupt

97	Vorlesen, Zustimmung AO, Alternative + Begründung	und der Wechsel- Der Phasenübergang (.) wo bin ich, ist dir nicht gelungen äh nicht gut gelungen, da die Inhalte zum Wechselwirkungsprinzip nicht ausreichend verstanden und gesichert wurden. (Also?) das würde ich definitiv sagen, vor allen Dingen, da er bei dem äh bei der anderen Anwendungsaufgabe gemerkt hat, dass es da noch totaäh da viele Probleme gibt, hätte ich da dann nochmal eine Aufgabe zur Sicherung eigentlich gemacht.
98	Vorlesen, Zustimmung AO	Es ist unproblematisch, dass der Phasenübergang abrupt war, da du mit dem Impulsbegriff wirklich ein komplett neues Thema beginnst. (...) Ja, (das?) ist natürlich auch eine Sache.
99	Vorlesen	Es ist nicht gut, dass du so äh einen abrupten Übergang zwischen den- den Themen gemacht hast, da der rote Faden der Stunde für die Schüler nicht erkennbar ist. (4s)
100	Vorlesen, Sonstiges, Bewertung + Begründung, Überlegung zum Unterricht	Äh es ist nicht gut, dass das äh dass du sagst, dass das Thema komplett neu, da die Schüler so das bisher Gelernte nicht mehr äh das bisher Gelernte nicht mehr relevant erscheint. (...) Also ich bin da jetzt gerade am überlegen, ähm so dieser abrupte Themenübergang, ich glaube, dass das wichtigste, was ich finde, was dadurch verloren gegangen ist, ist halt wirklich dieser rote Faden, weil (.) so ist-sind das halt eigentlich zwei einzelne Stundenziele auch oder erstmal ich (meine?) das ist jetzt- man müsste- ich müsste jetzt eigentlich gerade nochmal nachgucken, was er als Stundenziel formuliert hat, aber (.) es ist auf jeden Fall nicht so gut.
101	Vorlesen, Sonstiges, Zustimmung AO + Begründung	Es ist nicht gut, dass du sagst, das Thema ist komplett neu, da es einen fachlichen Bezug zwischen den Newtonschen- Newtonschen Axiomen und dem Impulsbegriff gibt. (.) Ja, (...) da bin ich jetzt gerade am überlegen, klar gibt es den Bezug, (.) aber also das (er?) es (so?) sagt, dass es komplett neu ist, ist auf jeden Fall nicht gut, (..) aber ich glaube eher, weil ähm das für die Schüler das bisher (...) Gelernte nicht mehr relevant ist, weil Schüler ja gerne eigentlich vergessen, (.) wenn man was Neues macht, was sie vorher gelernt hatten. #00:56:51-1#
102	Vorlesen	LD4: Welche der unten stehenden Hinweise zur Gestaltung des Phasenübergangs möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben?
103	Vorlesen, Ablehnung AO	Den Phasen- Den Phasenübergang hast du gut gestaltet, das kannst du wieder so machen. Nein.
104	Vorlesen, Zustimmung AO	Es wäre besser, den Zusammenhang zwischen den Newtonschen Axiomen und dem Impulsbegriff zu verdeutlichen, da die Schüler den Zusammenhang nicht unbedingt selbst erkennen können. (Also?) das ist auf jeden Fall gut. (.)

105	Vorlesen, Zustimmung AO + Begründung	Also es wäre besser, den Einstieg zur Impulsbe- in den Impulsbegriff auf die nächste Stunde zu verschieben, um den Schülern zunächst noch mehr Möglichkeiten zu geben, das Wechselwirkungsprinzip zu verstehen und anzuwenden. Das hätte ich jetzt auf jeden Fall einmal gesagt, weil ähm die ja doch schon echt noch Probleme mit diesem Comic hatten. (.)
106	Vorlesen, Zustimmung AO	Es wäre besser, vor dem Ein-stieg in das neue Thema auch die Tafel zu wischen, damit die Schüler sich äh voll auf das neue Thema konzentrieren können und nicht noch über das alte Thema nachdenken. Ja.
107	Vorlesen, Zustimmung AO	Es wäre besser, einen glatteren Übergang zu gestalten, zum Beispiel durch eine kurze Verknüpfung, wie wir finden jetzt etwas neues mit dem gelernten Wissen heraus, damit der roten Fa- der rote Faden (..) der Stunde für die Schüler- innen nachvollziehbar ist. Also wenn man es auf jeden Fall in der gleichen Stunde machen will, würde ich das auf jeden Fall- ist das auf jeden Fall besser, (.) als ähm (...) als einfach so (..) ähm (4s) ist das auf jeden- also wenn man das in der Stunde machen will, ist das besser, äh eine Verknüpfung zu haben, als auf einmal zu sagen, das ist etwas komplett Neues. (.)
108	Vorlesen, Zustimmung AO	Es wäre besser, den Übergang durch das Zeigen eines Experiments (.) zu gestalten, um den Bruch zwischen den Themen für die Schüler-innen weniger hart zu gestalten. Ja okay, mhm (10s) ja. #00:59:11-3#
109	-	[Video Einleitung Reflexion 04, Video Reflexion Szene 04] #01:01:29-4#
110	Vorlesen, Sonstiges	LD4: Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung erhalten, ja. (7s)
111	Zustimmung AO + Begründung	Also auf jeden Fall hier zum Thema Fachsprache, weil er benutzt jetzt ja auch schon wieder das ähm (.) äh die Vorstellung vom Schwung und (...) äh genau.
112	Vorlesen, Überlegung zum Unterricht	Dann ähm zum eingesetzten Beispiel, (.) ja gut also das mit dem (4s) Ball zu werfen, ist (..) ich bin gerade am überlegen, ob es da eine bessere Wahl gibt.
113	Beschreibung, Bewertung + Begründung, Zustimmung AO + Begründung	Es war halt es ist halt sehr, finde ich, sehr (...) subjektiv, aber das ist eher die Wahl des Experiments hier, weil er hat jetzt ja einfach nur die zwei Personen das machen lassen und so befragt, ja und was habt ihr jetzt so gemerkt und (.) das ist halt (.) finde ich, also kein richtiges Experiment, um halt die Ergebnisse, auf die er gekommen ist, wirklich drauf zu kommen, finde ich, ist das auf jeden Fall (.) keine gute Wahl ähm des Experiments, um auch einfach ein Beispiel zu haben, (...) kann man es wahrscheinlich machen,

114	Beschreibung, Bewertung	aber (.) ähm (4s) zum Thema Ergebnissicherung, ja okay, er hat einmal die Ergebnisse (.) (unv.) (.) geschrieben, aber was- da hat er auch (einmal?) nur gesagt, ja okay das- der Impuls ist über die Formel definiert und dann, wann er sich verändert, aber das fand- also wenn (er?) so den äh Impuls einführt, dann fehlt da eigentlich was. (.) So. #01:03:41-3#
115	Vorlesen	LD4: Robert hat schon e- äh einige Rückmeldungen (.) erhalten. Besonders wichtig ist ihm jetzt äh noch eine Rückmeldung zur Wahl des Experiments und zur Ergebnissicherung nach dem Experiment. #01:03:54-2#
116	Vorlesen	LD4: Welches Feedback möchten Sie Robert zur Wahl des Experiments geben?
117	Vorlesen, Sonstiges	Das Experiment- ah das Experiment ist gut gewählt, da die beiden Schüler, äh die das Experiment durchführen, den Impuls am eigenen Körper erfahren können. Naja.
118	Vorlesen, Ablehnung AO	Das Experiment ist gut gewählt, da es den Einfluss von Masse und Geschwindigkeit auf den Im-puls gut darstellt. Naja, bin ich eher (.) nicht der Meinung.
119	Vorlesen	Es ist nicht gut, da es nur über die Masse und die Geschwindigkeit gesprochen wurde, da un- da unklar bleibt, ob der Impuls noch von einer äh von anderen Größen, zum Beispiel wie der Form abhängen kann.
120	Vorlesen, Zustimmung AO + Begründung	Das Experiment ist nicht gut gewählt, da äh es den Zusammenhang beziehungsweise die Abgrenzung des Impulses zur kinetischen Energie nicht veranschaulicht. (7s) (Also?) es ist auf jeden Fall, dass man nur äh zwei runde Körper be- äh be- (.) das nur mit zwei runden Körpern durchgeführt hat und so wirklich noch andere Größen ähm es von anderen Größen abhängen könnte. (4s)
121	Sonstiges, Vorlesen	Und dann noch (unv.). Es- äh es ist gut, dass du beiden Schüler das- in das Experiment einbezogen hast, so motivierst und aktivierst du sie.
122	Vorlesen	Es ist gut, dass du vorgehst- dass du vorgibst, wer nach vorne kommen soll, so vermeidest zu lange Organisations-, Wartezeiten nach der Frage 'wer will denn'.
123	Vorlesen, Ablehnung AO + Begründung	Es ist gut, dass du nur zwei Schüler das Experiment durchführen lässt und nicht alle Schüler, da du so Zeit sparst und das Experiment so einfach ist, dass es auch von den sitzenden Schülern aus gut nachvollzogen werden kann. Also da würde ich eigentlich definitiv beim letzten nicht zustimmen, weil (.) die anderen sagen ja nur, dass in Anführungszeichen der Schwung weniger wird, also da stimme ich überhaupt nicht mit überein, (.)
124	Vorlesen, Zustimmung AO + Begründung	ähm (...) das mit dem Wartezeiten, wer will denn, ja okay, ich finde, das hängt immer davon ab, wie man das sonst- also wenn man es- der Reihe dann nach alle mal experimentieren lässt oder nach vorne kommen lässt, dann ist es wahrscheinlich fair und dann kann man das ähm (...) auf jeden Fall machen.

125	Zustimmung AO	Aber das auf jeden Fall (.) die Schüler mit eingebunden werden ist (.) gut. (...) Mhm ja (...) (also?). (6s)
126	Sonstiges	Habe ich alles? #01:06:47-4#
127	Vorlesen, Sonstiges	LD4: Welchen der unten st- ja okay.
128	Vorlesen, Ablehnung AO	Ähm das Experiment ist gut gewählt, du kannst es erneut einsetzen. Nein.
129	Vorlesen	Es wäre besser, äh statt des Experiments ein Video zu zeigen, in welchem gezeigt wird, welchen Einfluss verschiedene Massen und Geschwindigkeiten auf den Impuls haben, da die Schüler eher anspricht- da das die Schüler eher anspricht als ein Experiment.
130	Vorlesen, Zustimmung AO + Begründung	Es wäre besser, ein anderes Experiment durchzuführen, äh das für alle Schüler anschaulich ist, zum Beispiel eine Person, die auf einem Skateboard steht und einen Ball wegwirft, oder das Werfen beider (.) beider Bälle in den Sand zur Betrachtung der Krater. Also das würde ich auf jeden Fall sagen, weil das meinte ich ja mit das ist auf jeden Fall nicht einsichtig, (.) was für die anderen Schüler, die die haben ja nur die Erfahrung der zwei, die Experimentieren.
131	Vorlesen, Zustimmung AO	Es wäre besser, ein (.) Stoßexperiment durchzuführen, da du (.) so auf die Impulserhaltung überleiten kannst. Ja.
132	Vorlesen	Es wäre besser, ein Experiment zu wählen, zu dem eine Messreihe aufgenommen werden kann, damit du aus dieser Formel diesen objektiven Messungen- aus diesen objekt- objektiven Messungen abgeleitet werden kann.
133	Vorlesen	Es wäre besser, wenn das Experiment von allen Schülern durchführen lässt, da es den Schülern gut tut, sich zwischendurch zu bewegen.
134	Vorlesen	Du solltest darauf achten, dass die Schüler vorne stehen (.) dass die Schüler vorne so stehen, dass alle sitzenden Schüler erkennen könnten, wie leicht, schwer der Ball (..) wie leicht, schwer es ist, den Ball zu fangen. (.)
135	Zustimmung AO	Mhm also ich würde auf jeden Fall sagen, dass ein anderes Experiment gut ist, ich würde sogar wahrscheinlich nicht das mit den (..) Krater doch schon zur Einführung machen,
136	Überlegung zum Unterricht	oder (.) ich bin gerade am überlegen oder ob das mit der Messreihe- (aber da?) bei der Messreihe (...) das (unv.) (...) (unv.) da muss man gucken, ich bin gerade am überlegen, ob das so einfach umsetzbar wäre (.) mhm (unv.) (14s) #01:09:23-2#
137	-	I: Was überlegst du gerade? #01:09:25-6#
138	Sonstiges	LD4: Mhm? #01:09:26-8#
139	-	I: Was denkst du? #01:09:29-0#
140	Sonstiges	LD4: Ach, (entschuldigung?)

141	Sonstiges, Überlegung zum Unterricht (3x)	ähm ja ich bin gerade eigentlich nur- ach (.) ich finde die Frage hier gerade echt schwer, weil ach es ist (so?) ein richtiges Einführungsexperiment für äh den Impuls zu finden, das ist kommt total drauf an, ich bin da gerade am überlegen, ob das bei diesem Kr- Betrachtung der Krater auch (.) Fehlvorstellungen kommen kann, aber beim Stoßexperiment (...) mhm also das ist auf jeden Fall für die Impulserhaltung gut, aber ich weiß nicht, ob das schon von Anfang an so gut ist, um auch den Impuls einzuführen (...) ähm (.) naja. (5s)
142	Sonstiges	Also es (7s) (unv.) erstmal weiter, oder? (...) Ja. #01:10:41-3#
143	Vorlesen	LD4: Ähm (welches?) Feedback möchten Sie Robert zur Ergebnissicherung geben?
144	Vorlesen	Es ist nicht gut, dass- es ist nicht gut, dass du den Schülern einen fertigen Merksatz vorgibst, da das da- da das reine Lesen und Abschreiben nicht verständnisfördernd ist. (...)
145	Vorlesen	Es ist gut, dass die Schüler den Merksatz ins Heft übertragen, damit alle Schüler den Merksatz als Sicherung notiert haben.
146	Vorlesen	Es ist nicht gut, dass du die Ergebnisse des Experiments nicht sicherst, so ist unklar, was die Schüler aus dem Experiment mitnehmen sollten.
147	Vorlesen	Der Merksatz ist gut gelungen, da du je-umso-Sätze die Zusammenhänge von Masse, Geschwindigkeit und Impuls nach-vollziehbar darstellst.
148	Vorlesen	Es ist nicht gut, dass äh du den Impuls als skalare Größe einführst. So grenzt du den Impuls nicht ausreichend von der kinetischen Energie ab.
149	Vorlesen	Es ist gut, dass du den Begriff der Energie nicht erwähnst, so vermeidest du die Verwirrung bei den Schülerinnen, die (.) die die verschiedenen Fachbegriffe auf Grundlage ihrer Erfahrung nicht trennen können.
150	Zustimmung AO + Begründung, Bewertung, Alternative	Also es ist auf jeden Fall gut, so einen Merksatz als Sicherung in das Heft zu übertragen, damit man den halt auch wirklich hat und nachschauen kann, wie wir jetzt nochmal wirklich genau den Impuls haben, aber (.) dieses fertige- der fertige Merksatz ist (...) also ein fertiger Merksatz ist okay, finde ich, aber du musst den halt eigentlich mit den Schülern zusammen herleiten, dass du dann den fertigen Merksatz nachher hast, wenn du mit den Schülern besprochen hast, welche Teile da drin stehen, da spricht eigentlich meiner Meinung nach nichts gegen.
151	Zustimmung AO	Ähm (..) die Ergebnisse von deinem Experiment (.) oder so sollt- (.) sollten auf jeden Fall (.) gesichert wer-den, (4s) weil so- ja also das würde ich auf jeden Fall erstmal auch sichern, (.)

152	Zustimmung AO	ähm (...) und eigentlich wenn man nachher noch auf Impulserhaltung und so eingeht, (dann?) ist es halt (.) auch echt schlecht
153	Zustimmung AO	oder (.) ähm wenn man den Impuls halt wirklich nur als skalare Größe einführt (.) und nichts von der Richtung (.) mit sagt. (7s) Ja. #01:13:23-4#
154	Vorlesen	LD4: Welchen der unten stehenden Hinweise zur Ergebnissicherung möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben?
155	Vorlesen, Ablehnung AO	Die Ergebnisse des Experiments hast du gut gesichert, das kannst du wieder so machen. Nein.
156	Vorlesen, Zustimmung AO	Es wäre besser, die Beobachtungen des Experiments schriftlich an der Tafel zu sammeln, damit du im Verlauf der Stunde darauf zugreifen kannst. Also ich würde immer sagen, dass auf Ex- äh die Ergebnisse von Experimenten schriftlich gesammelt werden. (5s)
157	Vorlesen, Zustimmung AO + Begründung, Sonstiges	Es wäre besser, die Schüler in die Sicherung mit einzubeziehen, zum Beispiel einen Merksatz formulieren zu lassen, damit sie stärker aktiviert und involviert werden. (..) Ähm (4s) ja auf jeden Fall zumindest (6s) ah da bin ich jetzt gerade am überlegen, weil einen fertigen Merksatz zu haben, schadet glaube ich eigentlich nicht, (denn?) weil wenn die Formulierung st- stimmt, aber die Schüler mit einzubeziehen, was in dem Merksatz stehen sollte, (...) ist glaube ich ganz gut eigentlich (unv.). (.)
158	Vorlesen, Zustimmung AO	Es wäre besser, za- zusätzlich zum Merksatz noch ein Beispiel zu zeigen, um auch die Formel anschaulich darzustellen. (.) Ja, obwohl das (...) mhm (.) ich glaube, das ist jetzt gerade erstmal nicht so das wichtigste mit.
159	Vorlesen, Alternative	Es wäre besser, in der Sicherung auf das Experiment einzugehen, zum Beispiel indem du auf der Folie zum Merksatz auch auch auf das Experiment eingehst oder Skizzen ergänzt, um die Inhalte zu- stärker zu vernetzen. Ich finde, das sollte man eher auch mit (an der?) Tafel und mit (.) nicht vorher vorbereitet machen.
160	Vorlesen, Zustimmung AO	Es wäre besser, den Impuls als vektorielle Größe einzuführen, um ihn von der kinetischen Energie abzugrenzen. Ja. (...)
161	Vorlesen, Ablehnung AO	Es wäre besser, auch auf die kinetische Energie der Bälle einzugehen, um das Vorwissen der Schülerinnen anzuknüpfen. Naja, finde ich jetzt gerade glaube ich in dem Moment nicht ganz so wich- tig. #01:15:54-4#

162	Vorlesen	LD4: Nachdem wir das Experiment mit den Tennisbällen abgeschlossen und die Schüler den Merksatz zum Impuls abgeschrieben haben, habe ich zu Demo- habe ich zu einem Demonstrationsexperiment übergeleitet, an dem ich die Impulserhaltung veranschaulichen möchte. Mit dem Demonstrationsexperiment möchte ich ich einen Auffahrunfall simulieren. Den Schülerinnen haben ich dann erklärt, wie der Autounfall im Experiment nachgestellt wird und wie die Messwertaufnahme funktioniert. Den experimentellen Aufbau und eine Messwerttabelle habe ich vor der Stunde vorbereitet, hier habe ich dir zwei Fotos vom Aufbau und der Tabelle mitgebracht. (.) Ah (.) mhm (12s). #01:16:50-9#
163	-	[Video Einleitung Reflexion 05, Video Szene Reflexion 05] #01:19:34-1#
164	Sonstiges	LD4: Ach du Scheiße. (.)
165	Vorlesen	Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung erhalten, um etwas für seinen Unterricht zu (sehen?)?
166	Vorlesen, Zustimmung AO + Begründung	Umgang mit Schülerinnen und Schülern definitiv, das ist (war?) war ja dieses typische Mädchen können was- äh Mädchen Schrift, Jungs experimentieren, sehr schön. (4s)
167	Vorlesen, Zustimmung AO + Begründung	Äh (.) Durchführung, Auswertung des Experiments durch die Lehrkraft, ähm ja (du?) ja dieses dann, ja mhm ich berechne denn jetzt je gerade mal schnell die (.) Geschwindigkeit. (.)
168	Vorlesen, Zustimmung AO + Begründung, Alternative	Notation der Messwerte, (.) ja okay, er hat nochmal sehr auf die Einheiten beharrt, die könnte man natürlich auch einfach vorne in den Klammern (.) zur Masse oder so dazuschreiben, aber (.) ja.
169	Vorlesen, Zustimmung AO + Begründung, Überlegung zum Unterricht	Mhm Wahl des Experiments, experimenteller Aufbau, würde ich s- (..) d- (4s) definitiv noch was sagen, weil ich weiß nicht, ob (.) jetzt (das?) in diesem Unterrichtsausschnitt, zumindest hat er halt nicht nochmal gesagt, dass man da halt diese zwei Lichtschranken hat, wo die Geschwindigkeit gemessen wird o- oder bestimmt werden kann.
170	Vorlesen, Zustimmung AO, Beschreibung, Sonstiges	Aktivierung der (.) Schüler, ja okay, er hat jetzt zwei Schüler mit eingebunden, mhm ähm also ich glaub aber, dass diese (.) drei Punkte für mich das wichtigste sind
171	Zustimmung AO + Begründung	und natürlich von wegen Messunsicherheiten, ähm darüber kann man eigentlich bei jedem physikalischen Experiment, wenn man das mit den Schülern durchführt, einmal reden, wo denn Messunsicherheiten entstehen können. #01:21:24-0#

172	Vorlesen	LD4: Robert hat schon einige Rückmeldungen erhalten. Besonders wichtig ist ihm jetzt noch der Umgang mit den Schülerinnen, dazu, wie er das währ- während des Experiments auf Messunsicherhei- Messunsicherheiten eingeht ist und die Aktivierung der Schüler. #01:21:41-3#
173	Vorlesen	LD4: Äh welches Feedback möchten Sie zum Umgang mit den Schülern geben?
174	Vorlesen, Zustimmung AO + Begründung	Es ist nicht gut, dass du äh sagst, Mädchen eine schöne Schrift haben und Jungen besser Werte ablesen können, da du die Mädchen so auf eine äh (.) schöne Hand-schrift reduzierst. Definitiv, das ist dieses typische (.) ähm ja. (...)
175	Vorlesen	Und das ist auch definitiv nicht gut, dass du die Mädchen nur für die fürs Schreiben und also (unv.) das ist halt wirklich dieses Stereotyp fördern.
176	Vorlesen, Zustimmung AO + Ablehnung AO + Begründung	Es ist gut, dass du die Mädchen für eine Ha- ihr Handeln im Unterricht lobst, um ihr Selbstkonzept in dem Physikunterricht zu stärken. Ja, aber nicht, wenn es nichts mit Physik zu tun hat, sondern nur mit Aufschreiben. (.)
177	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du das Lob für die Handschrift vom Mädchen mit einer inhaltlichen Kritik kombinierst, da du die Ursache hier hier nicht das Geschlecht ist. (4s) Mhm ja.
178	Überlegung zum Unterricht	Und ich bin jetzt gerade sowieso bei den Einheiten am überlegen, ob er es überhaupt gesagt hat, als er das gewogen hat, ob er das überhaupt gesagt hat. (.)
179	Vorlesen	Es ist nicht gut, dass du die Schülerin schon vor der Notierung des zweiten Messwertes erneut auf die Einheit hinweist, da du so das Bild festigst, Mädchen wären in Physik nicht kompetent. (.)
180	Vorlesen, Sonstiges	Es ist nicht gut, dass du der Schülerin an der Tafel direkt sagst, welche Einheit fehlt, da du ihr so unterstellst, dass sie das nicht wüsste. Ja, ich weiß nicht, bin ich gerade am überlegen.
181	Vorlesen, Ablehnung AO + Begründung	Es ist in Ordnung, dass du sagst, dass Paula aufgrund ihrer schönen Handschrift die Werte notieren soll, da es wichtig ist, (.) dass alle die Werte an der Tafel gut lesen können. Äh nein, definitiv nicht, da kann- gibt es vielleicht auch Jungs oder (unv.) einfach. (.) Ja. (.)

182	Sonstiges, Bewertung, Zustimmung AO + Begründung	Mhm ich bin jetzt gerade bei dem (10s) also ich bin jetzt gerade bei dem Punkt mit dem ähm vor dem Notieren des zweiten Messwerts, ob das wirklich daran liegen kann, also ob man da wirklich das (Bild?) (...) davon machen könnte, dass man die äh Mädchen nicht kompetent sind, ich glaube, dass- es vermittelt halt eher also- es ist definitiv nicht gut, dass er sie schon darauf hinweist, aber ich würde auch eher sagen, dass es das- ähm (.) äh (nicht?) das (nicht?)- nicht, dass die Schülerin nicht kompetenz ist, sondern dass du es ihr dann einfach- ja okay, dass du es ihr nicht mehr zutraust, also definitiv nicht Kompetenz, ja gut. #01:24:19-8#
183	Vorlesen	LD4: Welche der unter st- steh- Umgang mit Schülern möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben?
184	Vorlesen, Ablehnung AO	Dein Umgang mit den Mädchen und Jungen ist in Ordnung, da- nein, definitiv nicht. (.)
185	Vorlesen, Ablehnung AO + Begründung	Du solltest äh solltest eine Geschichte um das Experiment (.) Experiment herum konstruieren, zum Beispiel in Form eines Videos, um die äh Mädchen zu motivieren. (.) Mich stört bei der Antwort grad das halt um die Mädchen zu motivieren, wenn um alle Schüler zu motivieren, aber nicht, um die Mädchen zu motivieren.
186	Vorlesen, Zustimmung AO	Es wäre besser, die Schüler (äh?) zu fragen, wer freiwillig schreiben will, die Werte ablesen möchte, um die Schüler nicht nach Geschlecht eine Aufgabe zuzuweisen. Definitiv.
187	Vorlesen	Es wäre besser, die Aufgaben Aufschreiben, Werte Ablesen durch zu tauschen, damit alle Schüler eine Aufgabe wahrnehmen. (5s)
188	Vorlesen, Zustimmung AO	Es- du solltest darauf achten, die Mädchen nicht für ihre Handschrift, sondern explizit für ihre physikalische Kompetenz (.) äh (.) zu loben, da du sonst das Bi- da du sonst das Bild vermittelst, Physik wäre ein ein reines Jungsfach. Ja, (.) aber auf-
189	Vorlesen, Sonstiges	Du solltest auf geschlechterspezifische Zu- äh schreibungen verzichten, da du (.) so ein unangemessenes Ver- da du so ein unangemessenes Verhältnis zwischen den Geschlechtern im Physik- (.) unterricht äh (...) erzeugst, also (..)
190	Vorlesen, Zustimmung AO, Alternative	du solltest grundlegend an deiner Einstellung gegenüber Mädchen und Jungen arbeiten, um dein stereotypisches Verhalten zu ändern. Definitiv, da bin ich völlig dabei. (4s) Aber ich glau- also ich würde halt eigentlich sagen, dass es wirklich gut wäre, wenn man die (..) also das durchzutauschen jetzt auf diese eine (.) Stunde bezogen nein, aber so allgemein bezogen definitiv, aber ich glaube, wenn man das jetzt nur auf die eine Stunde (.) bezieht, dann würde das meiner Meinung nach zu viel Zeit kosten (.) ähm (..)

191	Vorlesen, Zustimmung AO	und (.) man sollte (auf?) ja auf die geschlechterspezifischen Zuschreibungen verzichten. (..) Ja. (..) Okay. #01:26:51-8#
192	Vorlesen	LD4: Welches Umgang mit Messunsicherheiten geben.
193	Vorlesen, Zustimmung AO	Mhm es ist nicht gut, dass du sagst, du könntest exakt oder genau messen, da du die Schülervorstellung beförderst, man könnte (äh?) Größen ohne Unsicherheiten bestimmen. (..) (Ja?) das ist eigentlich vor allen Dingen (.) äh.
194	Vorlesen, Sonstiges*, Zustimmung AO (2x)	Du gehst nicht auf die Unsicherheiten ein. Das ist nicht gut, da diese eine wichtige Rolle in der Physik spielen. Also die beiden (LD4 zeigt auf AO 'Du gehst nicht auf die Messunsicherheiten ein. Das ist nicht gut, da diese eine wichtige Rolle in der Physik spielen.' und 'Es ist nicht gut, dass du nur einen der beiden Wagen wiegst, da die Schüler*innen so nicht sicher sein können, dass die Wagen wirklich gleich schwer sind.') finde ich hängen halt stark damit zusammen und das sind auch meiner Meinung nach ziemlich wichtige (.) Punkte wenn es um Experimentieren geht.
195	Vorlesen, Sonstiges, Zustimmung AO + Begründung	Es ist nicht gut, dass du von äh dass du nur einer der beiden Wagen wiegst, da die Schüler so nicht sicher sein können, dass die Wagen wirklich (.) gleich schwer sind. (...) Ich (bin da?) jetzt gerade am überlegen, also meiner Meinung- also ich würde es halt einfach machen, auch wenn (.) sieht, dass die halt genau gleich schwer sind, es ist auch was für die Schüler (dann?) einfach (..) ähm einfacher.
196	Vorlesen, Sonstiges, Zustimmung AO	Es ist nicht gut, dass du sagst, die Wagen wären gleich schwer, da das auch wenn sie baugleich sind nicht stimmen muss. (5s) (Auch wenn?) da bin ich jetzt gerade am überlegen, also (5s) also eigentlich bin ich eher bei dem Ersten.
197	Vorlesen	Es ist nicht gut, dass du das Gewicht mit null Komma eins Kilo angibst, da du hier nicht berücksichtigst, welche signifikanten Nachkomm- äh welches die signifikanten Nachkommastellen sind.
198	Vorlesen, Ablehnung AO	Es ist gut, dass du darauf verzichtest, auch auf Messunsicherheiten einzugehen, um die Komplexi- äh Komplexität des Experiments überschaubar zu halten. (.) Nein.

199	Vorlesen, Ablehnung AO + Begründung (3x)	Es ist gut, dass du das Gewicht des Wagens mit null Komma ein Kilo angibst, um äh (.) (Fehler?) in der Berechnung des Impulses aufgrund vieler Nachkommastellen zu vermei- vermeiden. Also (..) eigentlich ist es ja wirklich so, dass bei so einem kleinen Objekt mit null Komma eins Ki- das Gewicht mit null Komma eins Kilo (anzugeben?) ist halt eigentlich echt (..) ungenau, weil man halt wirklich mit der Waage genauer messen könnte und auch das Gewicht (.) genauer angeben kann und das (sind?) definitiv nicht alle signifikanten Kommastellen. (4s) Ja. #01:29:23-2#
200	Vorlesen	LD4: Welche der unten stehenden äh (.) Hinweise zum Umgang mit Messunsicherheiten möchten Sie Robert (.) mitgeben?
201	Vorlesen, Ablehnung AO	Es ist gut, wie äh du mit Messunsicher- unsicherheiten umgegangen bist, das kannst du erneut so machen. Nein, er hat sie ja nur ignoriert.
202	Vorlesen, Ablehnung AO	Es wäre besser, alle Nachkommastellen anzugeben, (.) die die Laborwaage dir anzeigt, da du den Fokus auf genaue Messungen legst und die Laborwaage genauer, als null Komma eins wie- äh Kilogramm wiegen kann. (.) Ja, alle finde ich jetzt auch schwer, signifkant-
203	Vorlesen, Zustimmung AO	es wäre besser (.) äh, anstelle der Aussage, du könntest exakt messen, auch Messunsi- Messunsicherheiten einzubeziehen, da Messungen nicht exakt sind. (6s) Ähm (.) Ja.
204	Vorlesen, Zustimmung AO + Begründung	Es wäre besser, beide Wagen zu wiegen, damit du den Schülern ein konkretes Bild von genauem wissenschaftlichen Arbeiten vermittelst. (Also?) das finde ich jetzt echt ein wichtiger Punkt, weil (.) nur weil sie (baugleich?) sind, ist es ja auch (4s) ja. (.)
205	Vorlesen, Zustimmung AO	Es wäre besser, beide Wagen zu- beide Wagen zu wiegen, da du im Vergleich der Massen auf (unv.) Messunsicherheiten beziehungsweise Ungenauigkeiten eingehen kannst. Ja.
206	Vorlesen, Zustimmung AO + Begründung, Alternative + Begründung	Es wäre besser, beide Wagen zu wiegen, damit die Schülerinnen sich selbst überzeugen können, dass die Wagen gleich schwer sind. Also (6s) also ich würde eigentlich auch noch das sagen, dass wenn sie baugleich sind, sie nicht unbedingt gleich schwer sein müssen, (.) kann ja immer sein und ähm (..) ich würde auch (.) nicht (unv.) wahrscheinlich nicht alle (Anm: gemeint sind vmtl. Nachkommastellen) obwohl eigentlich (4s) also es ist auf jeden Fall besser, ähm ne- das dieses exakt messen stömuss auf jeden Fall eigentlich weg, weil man kann ja nie exakt messen se- messen selbst alle ja Werte sind mit Nachkomma- mit Fehlern angegeben.
207	Sonstiges	Äh ja, So. #01:31:40-5#
208	Vorlesen	LD4: Ähm (.) Aktivierung der Schülerinnen geben.

209	Vorlesen, Zustimmung AO + Begründung, Alternative	Es ist nicht gut, dass du die Schülerinnen nur oberflächlich involvierst, da Schüler trotz dieser Aktivierung nicht mitdenken müssen. (..) Es- also das stimmt bei diesem Experiment wirklich, weil ähm er berechnet ja selbst die Geschwindigkeit einmal ganz schnell kurz selbst, (.) da könnte man sie ja zumindest mit einbinden.
210	Vorlesen, Zustimmung AO, Beschreibung	Es ist gut, dass du Schüler (.) in das Experiment einbeziehst, um sie im Unterricht zu aktiveren. (.) Ja, also er aktiviert ja eigentlich nur den einen, der dann nachher die Werte ablesen kann meiner Meinung nach (so wirklich?).
211	Vorlesen, Ablehnung AO	Es ist gut, dass du das Experiment selbst durchführst, da du in der Durchführung im Vergleich zu den Schülern dabei sehr sicher bist. Ja, aber Sicherheit muss nicht unbedingt sein, also es ist nicht gut.
212	Vorlesen, Zustimmung AO	Es ist nicht gut, dass der- dass du die Schüler ohne äh dir ohne Arbeits- äh (.) dir ohne Arbeitsauftrag zusehen, da Schüler so keinen Anreiz haben, aufmerksam zu sein. (.) Ähm ja.
213	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du die Masse des Wagens und insbesondere die Geschwindigkeit selbst bestimmst, da du die Schüler vorher schon erklärt hast, wie sie es selbst machen könnten. Da stimme ich- also das ist eigentlich meine- also da stimme ich eigentlich völlig zu.
214	Vorlesen, Zustimmung AO + Begründung	Äh es ist nicht gut, dass du das Experiment so lehrerzentriert durchführst, da den Schülern langweilig ist. (.) Das (würde?) ich eigentlich auch sagen, also (das?) die machen ja nichts.
215	Vorlesen, Bewertung, Bewertung + Begründung	Es ist nicht gut, dass Paula von der Tafel aus (..) von der Tafel aus und die Schülerinnen aus den hinteren Reihen den experimenta- experimentellen Aufbau nicht gut erkennen können, da es für sie schwer ist, dir zu folgen. Mhm (4s) mhm (7s) na (.) also eigentlich fehlt, finde ich, auch erstmal der Arbeitsauftrag und auch erstmal eigentlich diese Beobachtung, was bei dem Zusammenstoß passiert und so, das ist ja direkt dann eigentlich schon die Auswertung, (.) also da fehlt halt wirklich erstmal dieser Arbeitsauftrag auch noch.
216	Vorlesen, Zustimmung AO, Alternative	Es ist nicht gut, dass Paula von der Tafel aus- ja mhm (...) (das?) also das von der Tafel aus das man den experimen- das Experiment nicht gut sehen kann, (.) da muss er halt echt drauf achten. So. #01:34:21-4#
217	Vorlesen	LD4: Welche der unten stehenden Hinweise zur Aktivierung der Schülerinnen möchten Sie Robert für eine erneute (.) Durchführung der Unterrichts-stunde geben? (..)

218	Vorlesen, Zustimmung AO, Sonstiges	Äh es wäre besser, die Schülerinnen zu fragen, wer freiwillig schreiben beziehungsweise Werte ablesen möchte, um die Schüler nicht nach dem Gesch- äh nach dem Geschlecht eine Aufgabe zuzuweisen. Definitiv, immer. (D?)iese Geschlechterrollen (.) im Physikunterricht.
219	Vorlesen	Es wäre besser, die Schülerinnen zu fragen, ob sie den Ablauf verstanden oder noch Schwierigkeiten haben, um Schüler die Möglichkeit zu- für Rückfragen zu geben. (.)
220	Vorlesen, Ablehnung AO	Du hast die Schüler ausreichend aktiviert. Das kannst du wieder so machen, nein, hat er definitiv nicht.
221	Vorlesen, Zustimmung AO	Es wäre besser, Hypothesen der Schülerinnen abzufragen, um sie zur Beobachtung des Experiments zu (.) motivieren. Also das würde ich eigentlich immer machen, Hypothesenbildung (.) vor dem Experiment.
222	Vorlesen, Zustimmung AO + Begründung	Es wäre besser, den Schülerinnen Fragen zum Vorgehen während des Experiments oder zur Berechnung der Geschwindigkeit zu stellen, um sie zum Mitdenken anzuregen, ja, also das kann- gerade die Geschwindigkeit können die ja selbst berechnen.
223	Vorlesen	Es wäre besser, die Schüler stärker in die Durchführung des Experiments einzubinden, zum Beispiel um Wagen anzustoßen, Werte abzulesen oder die Geschwindigkeit zu bestimmen, (.) um die Schüler aktiv (.) einzubinden.
224	Vorlesen, Zustimmung AO (2x) + Begründung (2x)	Es wäre besser, die Schüler zum Experiment nach vorn kommen zu lassen, damit sie alles gut erkennen können. (.) Also das mit dem einbinden auf jeden Fall, es ist meistens dann doch besser, wenn (.) gerade bei so einem Experiment, wo das jemand durchführen kann, wenn (es?) (gerade?) einfach ein anderer durch-führt und auch das nach vorne Kommen, damit halt oder zumindest so, dass alle gut sehen können ist eigentlich wichtig. (8s)
225	Sonstiges	Mhm weiter. #01:36:14-6#
226	Vorlesen	LD4: Nachdem wir alle Messwerte aufgenommen haben, habe ich die Beobachtungen der Schülerinnen abgefragt. Um die Allgemeingültigkeit der Impulserhaltung nachzuweisen, habe ich sie dann in einem kurzen Exkurs theoretisch hergeleitet. Im Anschluss haben wir die Impulserhaltung am Experiment überprüft. #01:36:33-5#
227	-	[Video Einleitung Reflexion 06, Video Szene Reflexion 06] #01:39:27-8#
228	Sonstiges	LD4: Ach du Scheiße. (4s)
229	Vorlesen	Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung erhalten, um etwas für sich beziehungsweise seinen Unterricht zu lernen?
230	Vorlesen, Zustimmung AO	Umgang mit Schülerinnen vor allen Dingen.
231	Vorlesen, Zustimmung AO	Messunsicherheit (.) Umgang mit dem Einwand von Marie auch, finde ich auch, dass er da komplett das nur auf eine Rundung geschoben hat, passt ja eigentlich nicht

232	Vorlesen, Zustimmung AO + Begründung	und Einheit des Impulses, das ist natürlich auch eine sehr wichtige Sache, vor allen Dingen, da er vorher bei der Geschwindigkeit in der Tabelle ja so drauf beharrt hat, dass da Einheiten stehen (.) und sie dann jetzt zumindest wie ich es gesehen habe nicht (.) also mit aufgeschrieben hat. #01:40:16-5#
233	Vorlesen, Sonstiges	LD4: Robert hat zwar schon einige Rückmeldungen erhalten, würde aber trotzdem gern zu allen drei aufgeführten Aspekten noch eine Rückmeldung erhalten. (.) Yay. #01:40:26-8#
234	Vorlesen	LD4: Ähm welches Feedback möchten Sie Robert zu Umgang mit Schülerinnen geben?
235	Vorlesen, Zustimmung AO	Es ist gut, dass du Mädels sagst, da das- (.) es ist nicht gut, dass du Mädels sagst, da das ein sexistischer Ausdruck ist. Das fand ich eher das geringste Problem dabei, a- auch.
236	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du Mädchen und Jungs so unterschiedlich behandelst, da das nicht wertschätzend geüber- äh wertschätzend gegenüber den Mädchen ist. Ja, vor a- auf jeden Fall. (.)
237	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du Mädchen und Jungen äh in ihrem Können so unterschiedlich einschätzt, da du so die Vorstellung forderst, Physik wäre ein Jungsfach. Definitiv, das hat er (ja?) sowas von gefordert. (.)
238	Vorlesen, Zustimmung AO	Du thematisierst die Einheit des Impulses nicht. Das ist nicht gut, da dies inkosiste- inkonsistent zur früheren Szene ist, in der du Paula auf die Notation der Einheit der Geschwindigkeits bestanden hat- äh bestanden hast. Das meinte ich ja auch schon, dass das echt (un-?) inkonsequent ist.
239	Vorlesen, Alternative (2x)	Und es ist gut, dass du Mädchen dazu aufforderst, ihre Tabelle zu korrigieren, damit du sicher sein kannst, (lacht, Anm: LD4 lacht häufig leicht, hier ist es notiert, da im Vergleich stark gelacht wurde.) dass sie die gleichen Werte haben, wie Jungs. Ach, man kann alle darauf auffordern, die Tabelle zu korrigieren und nicht nut die Mädels und das mit dem sexistischen Ausdruck, ja gut, ähm (.) ja, sollte man (lacht) sollte er auch korrigieren. (.) Ey. #01:41:54-5#
240	Vorlesen	LD4: Welchen der unten stehenden Hinweise zu seinem Umgang mit den Schülerinnen möchten Sie Robert äh für eine erneute Durchführung der Unterrichtsstunde geben?
241	Vorlesen, Ablehnung AO	Dein Umgang mit Schülerinnen ist in Ordnung, du kannst ihn so behalten- beibehalten. Definitiv nicht.
242	Vorlesen, Alternative	Du solltest die Mädchen stärker in den Unterricht mit einbeziehen, zum Beispiel berechnete Werte vorlesen lassen, um ihr physikalisches Selbstkonzept zu stärken. (.) (Ein?) (ah?) ich finde eher das (unv.) eigentlich alle-

243	Vorlesen, Zustimmung AO	Du solltest geschlechtsspezifische Zuschreibungen verzichten, um die Schülerinnen nicht zu demis- äh diskriminieren. Auf jeden Fall. (..)
244	Vorlesen, Ablehnung AO, Alternative	Du solltest vo jedem Schüler Schülerin je nur einen berechneten Wert erfragen, da du so (.) die berechneten Werte von mehreren Schülern (.) kontrollieren kannst. (.) Ja gut, äh es oder zumindest (.) vielleicht dann sich nachher nochmal angucken, wo Unterschiede sind oder ob ich jetzt genau jeden Schüler nur einen Wert- man könnte nat- das finde ich jetzt nicht so wichtig.
245	Vorlesen, Zustimmung AO	Du äh solltest verstärkt schwächere Schüler in den Unterricht einbinden, da wichtig ist, äh dass sie die Inhalte verstehen. Ja, eigentlich schon.
246	Vorlesen, Zustimmung AO, Alternative	Du solltest Schüler unabhängig vom Geschlecht dazu auffordern, falsche Ergebnisse in der Tabelle zu korrigieren, damit am Ende alle Schüler die richtigen Ergebnisse in der Tabelle notiert haben. (..) Eigentlich sollte- ist ja das das Wichtigste, zumindest dann zu gucken, wo nachher bei den (unv.) Werten noch Unterschiede bestehen.
247	Vorlesen, Zustimmung AO + Begründung	Du solltest hinterfragen, (.) woher die geschlechterspezifischen Zuschreibungen stammen, die du vertrittst, um sie zu überwinden. Könnte er jedenfalls ähm auf jeden Fall machen (.) und wenn man das jetzt auf den dann sollte man- er sollte natürlich auch die Mädchen stärker in den Unbe- äh Unterricht mit einbeziehen, eigentlich bei ihm ist es die Mädchen stärker mit einbeziehen, eigentlich würde man ja sagen, ähm (..) gleichbe- gleichbe- (.) gleichberechtigt. #01:44:02-3#
248	Vorlesen	LD4: Welches Feedback möchten Sie zu Robert zu seinem Umgang mit Maries Einwand beziehungsweise mit der Messunsicherheit geben?
249	Vorlesen, Ablehnung AO + Begründung	Die Reaktion auf Maries Einwand ist dir gut gelungen, da du so möglichst effizient zum zentralen Ergebnis der Impulserhaltung gelangst. Nein, definitiv nicht, er hat es ähm auf Rundungsfehler geschoben. (.)
250	Vorlesen, Ablehnung AO	Die Reaktion auf Maries Einfall äh Einwand ist dir gut gelungen, da Rundungsfehler ein zentraler Fehlerquelle für die Unterschiede sin- für die Unterschiede der Ergebnisse sind. Auch nicht.
251	Vorlesen, Ablehnung AO + Begründung	Die Reaktion auf Maries Einwand ist dir nicht gut gelungen, da du den Fokus von der experimentellen Bestätigung der Impuls- erhaltung hin zur Diskussion von leichten Ungenauigkeiten verschiebst. (.) Ich finde die Diskussion (als?) wichtig und diese also ist das auch nicht das.

252	Vorlesen, Zustimmung AO, Alternative	Es ist nicht gut, dass du nu- nur so knapp auf den Einwand von Marie eingehst, da die Begründung nicht ausreicht, um Schüler zu verdeutlichen, dass die Unterschiede aufgrund von Messunsicherheiten zustande kommt. (.) Also das finde ich halt eher wichtig, wenn schon diese Frage kommt, dann kann man auch wirklich (.) (im Team?) überlegen, woher Messunsicherheiten entstehen.
253	Vorlesen, Zustimmung AO, Sonstiges	Es ist nicht gut, dass du so nur so knapp auf den Einwand von Marie eingehst, da so unklar bleibt, äh ob beziehungsweise wo Schüler Rundungsfehler gemacht haben. (.) Äh das ist es ja auch, (also?) wenn du es auf Rundungsfehler (..) ah ne, das ist der nächste.
254	Vorlesen, (Ablehnung AO) + Begründung	Es ist nicht gut, dass du die Ungleich- Ungleichheit der Messwerte auf Rundungsfehler begründest, da diese Begründung falsch ist. Es sind ja Messunsicherheiten und keine Rundungsfehler. (.)
255	Vorlesen, Zustimmung AO + Begründung	Es ist nicht gut, dass du die Ungleichheit der Messwerte mit Rundungsfehlern begründest, da da du so den Schülern die Schuld dafür gibst, dass das Experiment nicht exakt gleich sind. Also das könnte man halt auch sehen, dass sie dann eher die Schuld bei sich suchen, als zu überlegen, wo beim Experiment Unsicherheiten ähm sind. #01:45:57-0#
256	Vorlesen	LD4: Welche unten stehenden Hinweise zu seinem Umgang mit Maries Einwand beziehungsweise Messunsicherheiten möchten Sie Robert für eine erneute- (4s).
257	Vorlesen, Ablehnugn AO	Die Reaktion auf Marie äh auf den Einwand von Marie ist angemessen, du kannst äh das kannst du erneut so machen. Definitiv nicht.
258	Vorlesen, Verweis auf frühere Stelle im Test	Du solltest den Schülern angeben, auf wie viele Nachkommastellen sie runden sollen, (.) ein oder zwei Stellen, damit du die in- diese Situation umgehen kannst. (Ja dann?) wären wir wieder bei der Sache mit den sigfi- signifikanten Stellen. (.)
259	Vorlesen, Ablehnung AO	Es wäre besser, den Schülern am Anfang mitzuteilen, dass Messunsicherheiten (ähm) und Rundungsfehler nicht berücksichtigt werden, um diese Situation zu vermeiden. Es wäre- ah ja, ne, eigentlich mhm (.).
260	Vorlesen, Sonstiges	Es wäre besser, ausführlich zu begründen, wo die Rundungsfehler entstehen, damit die Schüler deiner Erklärung besser folgen können. (.) Du- (.) (das ist ein?) na. (.)
261	Vorlesen, Alternative	Du solltest (.) bereits bei der Vorstellung des Experiments über die vorhandene Reibung und Energieumwandlung beim Aufprall sprechen, damit die Schülerinnen in dieser Phase klar ist, dass äh dass Ergebnisse aufgrund dieser Einflüsse nicht nicht exakt (.) sein können. (..) Ich (würde?) das eher dann thematisieren, wenn es durch kommt.

262	Vorlesen	Es wäre besser, zu thematisieren, dass es bei der Durchführung des Versuches Messunsicherheiten (..) des Versuches Messunsicherheiten berücksichtigt werden müssen, damit die Schülerinnen und Schüler äh nachvollziehen können, wo (.) äh wieso die Abwei- wieso es zu dieser Abweichung kommt. (4s)
263	Vorlesen, Zustimmung AO + Begründung, Sonstiges (2x), Alternative	Du solltest das Thema Messunsicherheiten für dich selbst noch einmal aufarbeiten, da du solche Fragen nicht vomdamit du von solchen Fragen nicht vom Thema abgebracht wirst. Also eigentlich definitiv, weil (.) es ist ja eigentlich eine Chance, vor allen Dingen, wenn sie einen darauf ansprechen, dass man darüber nochmal redet. (.) Ähm (...) ich gucke mal gerade nochmal durch äh das eine (..) ich weiß nicht, eigentlich finde ich, dass man- (.) aber mit der Energieumwandlung und Reibung, (.) ich finde aber ich würde das nicht vorher machen, ich würde da vielleicht drüber sprechen, wenn jetzt wirklich es zu dieser Situation kommt. (.) Also lass ich es, glaube ich. #01:48:23-0#
264	Vorlesen	LD4: Welches Feedback möchten Sie Robert zum Umgang mit der Einheit des Impulses geben?
265	Vorlesen, Zustimmung AO + Begründung	Es ist nicht gut, dass du n- es ist nicht gut, dass du nicht über die Einheit des Impulses sprichst, da die Einheit ein wichtiger äh Bestandteil der der Angabe des Impulses ist und vor allen Dingen, da er so auf Einheiten besteht. (.)
266	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du nicht über die Einheit des Impulses sprichst, da dies inkonsequen- inkonsiste- inkonsistent zu früheren Szenen ist, in der du auf die Notation der Einheit der Geschwindigkeit bestanden hast. Das stimmt halt auch wirklich so.
267	Vorlesen, Ablehnung AO	Es ist gut, dass (du?) die Einheit des Impulses nicht so stark in den Fokus gerückt hast, um die Schüler nicht durch äh die neue Einheit zu verwirren. Nein, finde ich jetzt nicht so, es ist halt-
268	Vorlesen, Ablehnung AO + Begründung	Es ist gut, dass du den Fokus auf die berechneten Werte und nicht auf die Einheit gelegt hast, da die gleichen Zahlen ausreichen, um die Impulserhaltung auf Grundlage des Experiments zu bestätigen. Nein, weil das könnte ja auch sein, dass ein Schüler auf einmal nicht mit Kilogramm, sondern Gramm oder so gerechnet hat.
269	Vorlesen, Zustimmung AO + Begründung, Verweis auf frühere Stelle im Test	Es ist nicht gut, dass du die Schüler (.) ähm nicht auf- es ist nicht gut, dass du die Schüler nicht dazu aufforderst, die Einheiten anzugeben, da so unklar ist, ob sie ob die Impulse wirklich gleich groß sind. Das ist es halt, was ich meinte, man könnte natürlich auch (.) eine Masse in Kilogramm, die andere in Gramm haben und trotz- und dann ganz verschiedene Impulse haben. #01:49:54-1#

270	Vorlesen	LD4: Welche der unten stehenden Hinweise zu seinem Umgang (.) äh mit der Einheit des Impulses möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben?
271	Vorlesen	Es ist in Ordnung, dass du die Einheit des Impulses zunächst (.) vernachlässigst, das kannst du erneut so machen.
272	Vorlesen, Ablehnung AO + Begründung	Du solltest die Einheit des Impuls- also würde ich eher nicht sagen, wenn er vor allem so auf Einheiten beharrt und so.
273	Vorlesen, Zustimmung AO, Alternative	Du solltest die Einheit des Impulses angeben, da die Einheit ein (.) wich- (.) tiger Bestandteil (.) der Angabe des Impulses ist. (unv.) also würde ich eigentlich schon sagen, wenn nicht im Merksatz, dass man sich das danach einmal gemeinsam überlegt.
274	Vorlesen, Zustimmung AO	Du solltest darauf achten, dass die Schüler beim Vorlesen der Werte auch die Einheit angeben, um (.) äh dein- (.) um in deinem eigenen Verhalten konsistent zu bleiben. Ja, definitiv.
275	Vorlesen, Zustimmung AO Ablehnung AO + Begründung, Alternative	Du solltest die Einheit des Impulses oben über die Spalte der Tabelle notieren, um die Schr- äh Schreibbarkeit in jeder- um Schreibbarkeit (.) ne, Schreibarbeit in jeder Zeile zu sparen. Ja, aber dann müsste man es auch konsistent durch alle äh Spalten machen und da er es bei den anderen nicht gemacht hat, würde ich es auch nicht sagen.
276	Vorlesen, Zustimmung AO + Begründung	Du solltest die Schülerinnen dazu auffordern, die Einheit des Impulses zu nennen, damit die Schüler sich an die Einheit des Impulses gewöhnen. (.) Finde ich eigentlich richtig. Vor allem, dass sie wissen halt, dass vor allen Dingen bei physikalischen Größen halt die Einheit dazugehört. #01:51:25-9#
277	-	[Video Einleitung Reflexion 07, Video Szene Reflexion 07] #01:53:23-8#
278	Vorlesen, Zustimmung AO, Alternative	LD4: Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung erhalten, um etwas von für sich beziehungsweise seinen Unterricht zu lernen? (4s) Ähm also auf jeden Fall zur Interaktion mit den Schülern, er (.) wenn er schon so fragt, was die machen wollen, dann sollten die auch in der Gruppe zusammen einen (.) Satz formulieren und die Formulierung des Merk- also das finde ich halt dieses (...) das ist äh da halt wenn er schon so fragt, dann soll er (unv.) die Schülerantworten auch mit einbeziehen. (4s) #01:54:02-9#
279	Sonstiges	LD4: Robert war (unv.) (...) ja. #01:54:08-6#
280	Vorlesen	LD4: Welches Feedback möchten Sie Robert zu seiner Formulierung des Merksatzes geben?
281	Vorlesen	Die Formulierung des Merksatzes ist dir gut gelungen, da alle wichtigen Informationen korrekt enthalten- äh (...).

282	Vorlesen, Zustimmung AO, (Ablehnung AO) + Begründung, Sonstiges	Es ist gut, dass du im Merksatz auf das durchgeführte Experiment verweist, da die Schülerinnen den Merksatz direkt mit der Realität verbinden können. Ja, aber dann (passt?) man das jetzt wieder nur für diesen einen Fall und keinen allgemeingültigen. (.) Ach.
283	Vorlesen, Zustimmung AO	Die Formulierung des Merksatzes ist dir nicht gut gelungen, da sie für die Schülerinnen zu anspruchsvoll ist. Ja okay mhm.
284	Vorlesen, Zustimmung AO	Es ist nicht gut, (.) dass du deinen vorgefertigten Merksatz nutzt und die Merksätze der Schülerinnen nicht einbeziehst, da die Schüler äh gegenüber nicht wertschätzend ist. Also das ist es halt, das finde ich ist mit das wichtigste, dass (.) wenn man schon nach der Meinung fragt, die dann auch gemeinsam erarbeitet. (...) Ja. #01:55:12-0#
285	Vorlesen	LD4: Welches de- äh der unten stehenden Hinweise zu seiner Formulierung des Merksatzes äh zur Impulserhaltung möchten Sie Robert äh für eine erneute Durchführung geben?
286	Vorlesen, Ablehnung AO	Die Formulierung des Merksatzes ist angemessen, du kannst sie erneut so verwen- äh verwenden. Nein.
287	Vorlesen, Zustimmung AO	Es wäre besser, Beispiele und Merksatz zu trennen, um die Komplexität des Merksatzes zu reduzieren. Das ist wieder das trennen auf jeden Fall.
288	Vorlesen, Alternative	Es wäre besser, zusätzlich zum langen Merksatz eine kurze Version zu formulieren, da sich Schülerinnen eine kurze Version besser merken können. Das ist ja eigentlich schon auch wenn man das trennt, wird die ja auch schon viel kürzer.
289	Vorlesen, Ablehnung AO	Mhm es wäre besser, den äh Merksatz nur in Bezug auf einen Stoßprozess zu for- entschuldigung (.) zu formulieren, damit der Merksatz näher am besprochenen Experiment bleibt. (Nein?) (unv.).
290	Vorlesen, Alternative	Es wäre besser, den Merksatz um die Einheit des Impulses zu ergänzen, da die Einheit stets zu den Angaben der physikalischen Größe gehört. (.) Ich f- (..) (na?) ich finde, es gehört eher zur Definitions- die Einheit gehört eher zur Definition vom Impuls und nicht äh erst zu dem Erhaltungssatz (wozu?) er jetzt hier gerade den Merksatz gemacht hat.
291	Vorlesen	Äh du solltest den Merksatz äh schülernäher gestalten, zum Beispiel indem du eine Skizze ergänzt oder die Formulierung sprachlich vereinfachst, um den Schülerinnen beim Verständnis zu hel- äh zu unterstützen.

292	Vorlesen, Ablehnung AO + Begründung, Überlegung zum Unterricht (2x), Ablehnung AO (2x) + Begründung (2x), Sonstiges	Es wäre besser, eine Formulierung der Schülerinnen zu verwenden, da diese Formulierungen für sie besser verständlich sind. Also (.) eine Formulierung der Schüler, da ist es schwer, dass man es halt auch physikalisch korrekt aufschreibt, das würde ich jetzt nicht so wirklich sagen. (...) Mhm (7s) ich bin (hier?) jetzt gerade am überlegen mit dem (.) Merksatz schülernäher gestalten ähm mit der Skizze, (.) ob das wirklich beim Verständnis helfen könnte, ob er das machen sollte, oder wie das damit aussieht. (...) Mhm (4s) (ach?) ich weiß- da bin ich mir jetzt gerade unsicher (.) mhm, aber ich bin jetzt (auch?) beim Stoßprozess, (.) also es macht es auf jeden Fall einfacher, das erst einmal auf einen Stoßprozess und auf ein Zweikörperproblem, also zwei Körper äh zu vereinfachen (...) mhm, aber ich weiß nicht, (...) (naja?) eigentlich ist er dann (.) nicht mehr so allgemein, der könnte dann die Schüler (Anm: denken?) dass das nur bei zwei Körpern oder so erhalten bleiben muss. (..) Ich glaube, ich werde es so lassen. #01:58:04-3#
293	Vorlesen	LD4: Welches Feedback möchten Sie Robert zu seiner Interaktion mit den Schülern geben?
294	Vorlesen, Zustimmung AO	Es ist gut, dass die Schüler ihre Merksätze vorlesen, da du so erkennen kannst, ob äh sie dir Impulserhaltung umfassen- äh umfassend verstanden haben. Äh ja (.) also definitiv, aber (.)
295	Vorlesen, Zustimmung AO, Bewertung + Begründung	Es ist gut, dass die Schülerinnen einen Merksatz formulieren sollen, um ein tieferes Verständnis zu fördern. (Ja?), aber das ich finde das mit den zwei Merksätzen immer schwierig wegen der doppelten Sicherung. (.)
296	Vorlesen, Alternative	Es ist nicht gut, dass die Schülerinnen keine Rückmeldung auf ihre Merksätze gibst, da die Schüler so nicht einordnen können, ob ihre Merksätze angemessen formuliert sind. Also ich glaube, wenn man halt wirklich das macht mit eigenen Merksätzen, dann sollte man (es?) individuellere Rückmeldungen geben.
297	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du deinen vorgefertigten Merksatz nutzt und die Merksätze der Schüler nicht einbeziehst, da die Schüler gegenüber nicht wertschätzend ist. Finde ich (immer?).
298	Vorlesen, Zustimmung AO + Begründung	Ähm es ist nicht gut, dass du deine Formulierung als besser bezeichnest, da du die Formulierungen der Schüler so abwertest. (.) (Eigentlich?) auch (unv.) etwas als besser zu bezeichnen ist sowieso immer blöd.
299	Vorlesen	Es ist nicht gut, dass nur zwei Schüler Merksätze vorgestellt haben, da du so nicht weißt, welche Merksätze die restlichen Schüler formuliert haben. (..)
300	Vorlesen, Verweis auf frühere Stelle im Test, Bewertung	Und es ist nicht gut, dass zwei Jungen dran nimmst, aber das ist ja wieder das gleiche Problem wie vorher, was vielleicht eher ein Problem mit seinem Unterricht ist. (...)

301	Sonstiges	Mhm ja, ich glaube aber so passt das. #01:59:53-3#
302	Vorlesen	LD4: Welche der unter stehenden Hinweise zu seiner äh Interaktion mit den Schülern möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben?
303	Vorlesen, Sonstiges	Äh (.) du solltest den Merksatz ausführlicher erklären beziehungsweise mit den Schülern besprechen, da einige Aspekte des Merksatzes nicht selbsterklärend sind. (.) Ich also (unv.) ich (bin?/finde?) halt eher-
304	Vorlesen	Du solltest gemeinsam mit den Schülern einen Merksatz erstellen, um sie stärker mit in den Unterricht einbin- einzubeziehen. (...)
305	Vorlesen, Zustimmung AO, (Ablehnung AO) + Begründung, Sonstiges	Es wäre besser, insgesamt mehr Schüler ihre Merksätze vorstellen zu lassen, da du aus den Antworten Rückschlüsse auf die Vorstellungen der Schüler ziehen kannst. Ja, aber dann hört man den Gleichen auch so häufig. (.) (Ah?) es ist, ich finde das schwer.
306	Vorlesen, Zustimmung AO	Du solltest Schüler, die sich melden- sich melden lassen, wenn sie ihre Merksätze vorstellen wollen, da sich möglicherweise auch ein Mädchen gemeldet hätten. Ja.
307	Vorlesen, Ablehnung AO	Der Umgang mit den Schülern war angemessen, so kannst du (.) dich verhalten. Fand ich jetzt nicht so.
308	Vorlesen, Zustimmung AO	Du solltest deinen Umgang mit den Schülern überdenken. Dass du (4s) ähm ja. Du solltest deinen Übergang (.) ähm Umgang mit den Schülern überdenken. Dass du sie ihre Arbeit nicht lobst, kann dem Verhältnis Schüler zum Physikunterricht nachhaltig schaden. Ja. (.)
309	Zustimmung AO, Verweis auf frühere Stelle im Test	Und auch das Rollenklieschee (...) definitiv, das ist ja wieder das gleiche. #02:01:32-7#
310	Vorlesen	LD4: Reflexion Gesamt.
311	Vorlesen	Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung, Lob, erhalten, um etwas (an?) sich äh für sich beziehungsweise seinen Unterricht zu lernen? (..)
312	Zustimmung AO	Eigentlich zu den Unterrichtszielen,
313	Zustimmung AO + Begründung	zu dem Ablauf auch auf jeden Fall, der war ja nicht gut (.)
314	Zustimmung AO + Begründung	und die Auswahl und Umfang der Stundeninhalte, da gab es ja, (ich?) (...) ich fand das mit der (unv.) (.) mit den zwei Aspekten im Unterricht auch echt schwer
315	Zustimmung AO + Begründung	und zum Medieneinsatz, ja okay mit seinen vorgefertigten (.) Fol- obwohl Medieneinsatz, vorgefertige Folien,
316	Zustimmung AO	ja, also die Interaktion auf jeden Fall,
317	Zustimmung AO	fachliche Korrektheit auch auf (..) jeden Fall.
318	Sonstiges, Bewertung (2x)	(Ob?) mhm ich bin bei den Unterrichtszielen gerade im (unv.) am überlegen, weil eigentlich passt es ja mit den Newtonschen Axiomen wiederholen und dann ähm (.) eigentlich Impulserhaltung, Unterrichtsziel war ja eigentlich klar gegeben, das (.) passte,

319	Zustimmung AO + Begründung	aber zum Auftreten als Lehrkraft definitiv nochmal was (...) von wegen (.) Klischees und ja. #02:02:52-0#
320	Vorlesen	LD4: Robert hat sich schon einen hat schon eigene einige Rückmeldungen erhalten. Besonders wichtig ist ist ihm jetzt noch eine Rückmeldung zu seinem Umgang mit den Schülerinnen und der fachlichen Korrektheit im ganzen Stunde. #02:03:05-7#
321	Vorlesen	LD4: Welches Feedback möchten Sie Robert zu seiner Interaktion mit Schülerinnen geben?
322	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du Beiträge von Schülerinnen kaum berücksichtigst, da dies demotivierend für die Schüler ist. Ja, vor allen Dingen bei den Merksätzen.
323	Vorlesen, Überlegung zum Unterricht, Beschreibung	Es ist nicht gut, dass du so lehrerzentriert unterrichtest, da äh Schüler so nicht aktiv ei- (.) ein Verständnis für die Inhalte entwickeln. Bin ich jetzt gerade am Überlegen mhm (.) was hat er nicht Lehrerzentriertes gemacht? Also er hat ja immer die Merksätze lehrerzentriert gemacht, äh das Experiment war lehrerzentriert, okay es war eigentlich alles lehrerzentriert, ja gut.
324	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du die Jungen und Mädchen unterschiedlich behandelst, da dies demotivierend ist. Vor allen Dingen für die Mädchen, ja.
325	Vorlesen, Zustimmung AO	Es ist gut, dass du äh Sicherungsphasen im Unter- in deinen Unterricht eingebaut hast, damit du sicher- äh (.) äh gehst, dass die Schülerinnen die zentrale In- halte aus der Stunde mitnehmen. Ja, schon.
326	Vorlesen, Überlegung zum Unterricht, Beschreibung	Es ist gut, dass du verschiedene Sozialformen im Unterricht einplanst, Einzel-, Partnerarbeit, (.) da du den- (.) es ist- ah Unter- abw- Unterricht so abwechs- abwechslungsreicher gestaltest. Ich bin gerade am überlegen, hatte er überhaupt Partnerarbeit und so wirklich drinne, (.) vielleicht das Auswerten der Experimente oder ach ja, stimmt, bei (.) ähm auswerten, stimmt da stand Partnerarbeit mhm (..) bin ich gerade am überlegen.
327	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du im Unterricht nur von wenige Schülern Beiträge anhörst, da restliche Schüler so nicht aktiv werden. (.) Ja. (.)
328	Vorlesen, Zustimmung AO	Deine Verständnissicherung mit einer Frage, wie 'Habt ihr noch Fragen?' ist nicht zielführend, da Schüler auf diese Frage eher nicht antworten. Das stimmt schonmal auf jeden Fall.
329	Vorlesen, Zustimmung AO	Du thematisierst die auftretenden Schülervorstellungen nicht, das ist nicht gut, da sich diese so verfestigen. Ja vor allen Dingen am Anfang mit dem Schwung und so. (..)
330	Zustimmung AO, Bewertung	Ähm mhm (.) ich weiß nicht, ob das äh also es ist gut, dass er verschiedene Sozialformen einbaut, aber ich fand äh ich finde das bei ihm noch nicht so wirklich gelungen (..) mhm, aber eigentlich ist das schon gut. (.)
331	Sonstiges	So (.) (unv.) weiter. #02:05:32-9#

332	Vorlesen	LD4: Ähm welche der unten stehenden Hinweise zu seiner Interaktion mit den Schülern möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben?
333	Vorlesen, Ablehnung AO	Dein Umgang mit den Schülerinnen ist angemessen- nein, definitiv nicht.
334	Vorlesen, Zustimmung AO	Du solltest die Schüler grundsätzlich stärker in den Unterricht einbeziehen, zum Beispiel in Diskussionen, bei der Durchführung von Experimenten oder äh der Formulierung von Merksätzen, um die Impulserhaltung gemeinsam mit den Schülern zu erarbeiten. Jap, das finde ich ein sehr wichtiges Thema. (.)
335	Vorlesen, Zustimmung AO	Äh du solltest den Schülern ausführliches Feedback zu seinen äh zu ihren Beiträgen geben, damit die Schüler einordnen können, ob ihre Aussage korrekt ist. Also gerade, wenn man die Merksätze bespricht, finde ich das schon echt wichtig, dass man den Schülern nochmal ein Feedback gibt, ob sie jetzt das richtig haben oder nicht.
336	Vorlesen, Zustimmung AO	Du solltest die Schülervorstellungen zum Thema vor dem Unterricht ansehen, damit du im Unterricht angemessen auf diese eingehen kannst, um das Verständnis der Schüler zu fördern. Das ist ein sehr wichtiges Thema. (.)
337	Vorlesen, Sonstiges, Zustimmung AO	Äh du solltest dir grundsätzlich Gedanken darüber machen, wie, zum Beispiel mit welchen Aufgabenstellungen, du das Verhältnis der Schüler- innen abprüfen kannst, um sicherzust- gehen, dass die Schüler die Ziele der Stunde erreichen. (...) (Ah?) bin ich jetzt gerade am überlegen mhm, aber eigentlich ist das auch wichtig.
338	Vorlesen, Zustimmung AO + Begründung	Du solltest deinen Umgang mit der mit den Schülern grundsätzlich überdenken. Dass du sie nicht lobst und wertschätzt, kann dem Ver- hältnis Schüler zum Physikunterricht nachhaltig schaden. Ja und auch einfach das äh Benachteiligen der Mädchen vor allen Dingen auch.
339	Vorlesen, Zustimmung AO	Es ist wichtig, dass du darauf aufpasst, keine Rollenklischees in den Unterricht mit einzubringen, um die Schüler in die Muster zu drängen. Ja, definitiv mit den Mädchen und der schönen Handschrift, yay. #02:07:29-2#
340	Vorlesen	LD4: Welches Feedback möchten Sie Robert zur fachlichen Korrektheit der Stunde geben?
341	Vorlesen	Deinen Unterricht hast du insgesamt gut durchgeführt, da er überwiegend fachlich korrekt war.
342	Vorlesen, Sonstiges	Du hast fachliche Fehler in den Unterrichtsstunden gemacht. Das ist nicht gut, da du den Schülern so falsche Inhalte beibringst. (.) Ach ich finde das gerade echt schwer.
343	Vorlesen, Zustimmung AO + Begründung	Es ist nicht gut, dass du die Abweichungen beim experimentellen Nachweis der Impulserhaltung als Rundungsfehler darstellst. Das ist fachlich falsch. Das ganz wichtiger Aspekt, vor allem mit dem physikalischen Arbeiten im Hintergrund. (.)

344	Vorlesen, Zustimmung AO + Begründung	Es ist nicht gut, dass du äh die Einheit des Impulses nicht berücksichtigst, da Einheiten ein wichtiger Bestandteil des- der Angabe des Impulses ist. Auch wichtig, vor allen Dingen, wenn er schon die Formel nennt, ist dann auch der Schritt zur Einheit (ja?) nicht mehr groß.
345	Vorlesen, Zustimmung AO + Begründung	Du thematisierst die auftretenden Schülervorstellungen nicht. Das ist nicht gut, da sich diese so verfestigen. (.) Auch ein wichtiger Punkt, hat man ja bei dem Kraftbe- also bei der Kraft noch gesehen.
346	Vorlesen, Zustimmung AO, Alternative	Es ist gut, dass du äh so großen Wert auf die Ein- heit der Geschwindigkeit in der Tafel legst, da die Einheit zentraler Bestandteil der Größenangabe ist. Stimmt schon, aber damit sollte er halt auch wirklich konsequent sein.
347	Vorlesen, Zustimmung AO	Mhm es ist nicht gut, dass du bei der Durchführung des Demonstrationsexperimentes sagst, du würdest exakt messen, da du die Schüler (.) äh das Bild vermittelst, du könntest Messunsicherheiten vermeiden. Ja (...) genau, so. #02:09:03-6#
348	Vorlesen	LD4: Welche der unten stehenden Hinweise zur fachlichen Korrektheit der Stunde möchten Sie mit Robert für eine erneute Durchführung der Unterrichtsstunde geben?
349	Vorlesen, Zustimmung AO	Äh du solltest au- du solltest auch bei der Besprechung der Impuls auf die Einheiten achten (.) achten, um einheitlich vorzugehen. Ja.
350	Vorlesen, Zustimmung AO, Alternative	Es wäre besser, die Schülervorstellungen, ähm die im Unterricht auftreten, zum Beispiel dass du äh zum Beispiel, dass der Apfel Kraft hat, aufzugreifen, um den Schülern die fachliche korrekte Sichtweise näher zu bringen. Ja, oder die auch vor allen Dingen, ich weiß nicht, auch zu widerlegen.
351	Vorlesen, Zustimmung AO	Du solltest dir die Schülervorstellungen zum Thema vor dem Unterricht angucken, damit du mit ihnen vertraut bist und Ideen entwickeln kannst, wie du ihnen begegnen kannst. Ja, auch (.) wichtiger Punkt.
352	Vorlesen, Zustimmung AO	Es wäre besser, auch auf Messunsicherheiten beim Fahrbahn-Experiment einzugehen, da die Schüler die Impulserhaltung trotz der leicht unterschiedlichen Werte akzeptieren. (.) Ich glaube, es ist halt (..) echt wichtig.
353	Vorlesen, Zustimmung AO	Du solltest darauf achten, dass deine Merksätze fehlerfrei formuliert sind, damit die Schülerinnen keine Fehler aus deinen Unterlagen übernehmen. (..) Ja, eigentlich auch
354	Vorlesen, Zustimmung AO, Sonstiges	und er muss auch mit dem Schwung vor allem auf die Wortwahl achten. Ja ach (.) (..) ach Mist. #02:10:31-1#
355	-	[Video Promt Schulnote] #02:10:38-5#

356	Bewertung (2x) + Begründung (2x), Sonstiges (3x)	LD4: Mhm (.) ich glaube, ich wäre eher bei einer (mhm?) (.) Fünf so weil mich vor allen Dingen auch das mit den Geschlechterrollen und (.) ähm der Einbindung der Schüler, Eingehen auf Schülerantworten und so echt gestört hat. (.) Also die fachliche Korrektheit, (äh?) ne Vier, weil ah ich weiß es nicht, Vier oder Fünf ha mich hat das jetzt echt (doch?) gestört, aber fachlich war, (.) naja (doch?), mache ich. #02:11:13-9#

O.5 Interview 5: LD5

Segment	Kodierung	2. Kodierung	Transkript
0	-	-	[Einverständniserklärung, Demographie (Anm: An letzte Schulnoten nicht genau erinnert, sondern geschätzt), Hinweis auf Aufgabenstellung] #00:04:08-5#
1	-	-	I: Also es ist so, dass während der Testdurchführung Ihnen gleich immer mal wieder Videos begegnen äh und im Anschluss an diese Videos werden Fragen gestellt, die Sie beantworten und ich bitte Sie eben dann alle Gedanken während der Bearbeitung dieser Fragen beziehungsweise Aufgabenstellung immer laut auszusprechen und wenn Sie gerade etwas lesen, dann auch diese Texte oder die Aufgabenstellung mit vorzulesen. Ähm es geht wie gesagt nicht darum, jemandem etwas zu erklären, sondern Sie können sich vorstellen, Sie wären alleine in einem Raum, würden zu sich selber sprechen. Ähm genau, ich schalte mich gleich stumm, es sei denn, Sie äh unterbrechen Ihren Redefluss, dann melde ich mich mal kurz mit so einer Aufforderung, wie ich es vorhin schon angesprochen hatte, also mit einem kurzen 'bitte weitersprechen' oder so. #00:04:54-8#
2	Sonstiges	Sonstiges	B5: Auch während des Videos, wenn da Gedanken kommen, die auch immer formulieren, also alles, was man denkt. #00:04:59-9#
3	-	-	I: Genau, alles, was kommt. #00:05:01-6#
4	-	-	[Video Lerngruppe, Video Vorwissen] // #00:07:50-8#
5	Beschreibung	Beschreibung	B5: Das klingt recht lang #00:07:06-0#
6	-	-	[Video Unterrichtsziele, Video Stundenraster] #00:09:02-0#

7	Vorlesen, Bewertung	Vorlesen, Beschreibung, Bewertung	B5: Lehrer erklärt Ablauf der Quizrunde, Schüler beantworten Quiz als Einstieg. Dann (ein?) Bild zum Plenumsgespräch (.) (unv.) Erarbeitungs- Sicherungsphase mit dem Impuls und mit einem Kontext (.) und anschließend Versuch dazu, das klingt ja erstmal sehr sinnvoll. #00:09:32-4#
8	Vorlesen	Vorlesen	B5: Ich habe die Stunde mit der Vorstellung des Ablaufs begonnen. Als nächstes habe ich ein Quiz zur Wiederholung der Newtonschen Axiome angeleitet, welches die Schüler- innen in Partnerarbeit durchgeführt haben. Die Durchführung des Quiz selbst habe ich zwar nicht gefilmt, dafür aber die Besprechung des Quiz mit den Schüler*innen, die ich dir jetzt zeige. #00:09:52-4#
9	-	-	[Video Einleitung Reflexion 01, Video Szene Reflexion 01] // [Weiterhin Video Szene Reflexion 01] // #00:12:15-7#
10	Überlegung zum Unterricht, Alternative	Überlegungen zum Unterricht, Alternative	B5: Wieso stellt er das jetzt alleine vor und nicht mit dem Schüler? #00:10:52-0#
11	Beschreibung	Überlegungen zum Unterricht	B5: (Er?) stelle ja eine Nachfrage (unv.) und (unv.) nicht wirklich #00:11:36-6#
12	Vorlesen	Vorlesen	B5: Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung, Lob und/oder Kritik, erhalten, um etwas für sich bzw. seinen Unterricht zu lernen?
13	Zustimmung AO + Begürndung	Zustimmung AO+Begründung	Seine Formulierungen waren (beim?) Zweiten recht mathematisch, deswegen würde ich dazu eine Rückmeldung geben
14	Zustimmung AO	Zustimmung AO	und die Ergebnissicherung,
15	Zustimmung AO + Begründung	Zustimmung AO+Begründung	die Gesprächsführung ist mir auch aufgefallen, da sie (.) erstmal sehr wenig fragend war. #00:12:39-2#

16	Vorlesen	Vorlesen	B5: Robert hat schon einige Rückmeldungen erhalten. Besonders wichtig ist ihm jetzt noch eine Rückmeldung zu seiner Formulierung des Merksatz- satzes zum dritten Newton- Newtonschen Axiom. #00:12:49-5#
17	Vorlesen	Vorlesen	B5: Welches Feedback möchten Sie Robert zu seiner Formulierung des Merksatzes zum dritten Axiom geben? Wählen Sie die Antwortoption aus, die Ihrer Einschätzung am nächsten kommen. Bitte wählen Sie die zutreffenden Antworten aus.
18	Vorlesen	Vorlesen	Deine Formulierung des dritten Axioms ist gut, da alles wichtige enthalten ist.
19	Vorlesen, Ablehnung AO + Begründung	Vorlesen, Ablehnung AO + Begründung	Da- Deine Formulierung des dritten Axioms ist fachlich falsch. Das würde ich nicht zustimmen, an sich ist ja der Grund erkannt bei ihm.
20	Vorlesen, Überlegung zum Unterricht	Vorlesen, Überlegung zum Unterricht	Die Formulierung des dritten Axioms ist für die Schüler- innen nicht angemessen, da sie zu komplex ist. Ich glaube, dass dritte Newtonsche Axiom an sich komplex ist und deswegen es recht lang sein muss.
21	Vorlesen	Vorlesen	Die Formulierung des dritten Axioms ist nicht gut, da die Vermischung von Beispiel und Merksatz die Schüler- innen verwirren kann.
22	Vorlesen, Ablehnung AO + Begründung	Vorlesen, Ablehnung AO + Begründung	Es ist nicht gut, dass du- dass nur die Kurzformel Kraft gleich Gegenkraft, aber nicht actio gleich reactio im Mer- Diese Abkürzung ist nämlich etwas, was die Schüler- innen vielleicht kennen. Dem würde ich nicht zustimmen, weil das andere Latein und das (eine?) Deutsch ist, das stimmt- wird es noch schwerer eher.
23	Vorlesen	Zustimmung AO + Begründung	Ja, die Formulierung allein Kraft gleich Gegenkraft könnte ein bisschen Ver- äh (unv.) unsichern, denke ich.

24	Vorlesen, Bewertung	Vorlesen, Bewertung	Es ist gut, dass du im Merksatz deutlich machst, dass die Kräfte an verschiedenen Körpern angreifen, da du so einer Verwechslung mit dem Kräftegleichgewicht vorbeugst. Fand ich der (Vorlage?) sehr gut gelungen.
25	Vorlesen, Zustimmung AO	Vorlesen, Zustimmung AO	Es ist nicht gut, dass die Kurzformel Kraft gleich Gegenkraft im Merksatz vorkommt, da Schüler diese oft mit dem Kräftegleichgewicht assoziieren. Das ist genau das Problem. #00:14:13-6#
26	Vorlesen	Vorlesen	B5: Welche der unten stehenden Hinweise zu seiner Formulierung des Merksatzes zum dritten Axiom möchten Sie Robert für eine erneute Durchführung der Unterrichtseinheit- stunde geben? Wählen Sie die Antwortoptionen aus, die Ihrer Einschätzung am nächsten kommen.
27	Vorlesen, Ablehnung AO	Vorlesen, Ablehnung AO, Sonstiges	Deine Formulierung des dritten Axioms ist gut so, du kannst sie wieder so verwenden. Ich habe da schon widersprochen.
28	Vorlesen	Vorlesen	Es wäre besser, zuerst das Axiom ohne Beispiel zu notieren und anschließend ein Beispiel, um den Merksatz zu vereinfachen.
29	Vorlesen	Vorlesen	Es wäre besser, das Beispiel im Merksatz ausführlicher Darzustellen, damit die Schüler- innen die Anwendung vom dritten Axiom verstehen können.
30	Vorlesen, Zustimmung AO + Begründung	Vorlesen, Zustimmung AO + Begründung	Es wäre sinnvoll, das Axiom um eine Skizze zu ergänzen, damit es anschaulicher wird. Das würde ich auf jeden Fall erwähnen, eine Skizze ist sehr lernförderlich.
31	Vorlesen, Ablehnung AO	Vorlesen, Ablehnung AO	Es wäre besser, im Merksatz zu verdeutlichen, dass Kraft und Gegenkraft gleichzeitig wirken. Deine Formulierung anschließend übt Körper B eine Gegenkraft auf Körper A aus ist fachlich falsch. Da würde ich nicht zustimmen.

32	Vorlesen, Überlegung zum Unterricht, Ablehnung AO + Begründung	Vorlesen, Ablehnung AO + Begründung	Du solltest im Merksatz über den Kräften Vektorpfeile notieren, damit die Wirkrichtung der Kräfte deutlich wird. Da würde ich jetzt schwer überlegen, weil an sich ist eine Mechanik mit Kräften gut äh mit Vektorpfeilen auch sinnvoll, eine zweidimensionale Mechanik, es könnte die Schüler (aber?) beim dritten Axiom noch mehr verwirren, dadurch dass das dritte Axiom an sich schon sehr komplex ist, würde ich das jetzt nicht ankreuzen.
33	Vorlesen, Zustimmung AO + Begründung	Vorlesen, Zustimmung AO + Begründung	Du solltest im Merksatz verdeutlichen, dass Kraft und Gegenkraft in genau entgegengesetzte Richtungen wirken, damit die Schüler- innen die Bedeutung des Vorzeichens im Beispiel verstehen. Dem würde ich genau zustimmen, damit der Unterschied (auch?) zwischen Kraft und Gegenkraft noch deutlicher wird und nicht der- das Problem zum Kräftegleichgewicht her- erzeugt wird. #00:15:49-2#
34	-	-	[Video Einleitung Reflexion 02, Video Szene Reflexion 02] // [Weiterhin Video Szene Reflexion 02] // #00:17:56-8#
35	Bewertung, Alternative	Überlegung zum Unterrcht, Alternative	B5: Auf die Frage (bringt?) das relativ wenig, er sollte eher Ansatzpunkte geben (unv.) (um den Schülern zu helfen?) #00:17:13-1#
36	Überlegung zum Unterricht	Überlegung zum Unterricht	B5: Da muss er aufpassen mit Schülervorstellungen (unv.) #00:17:53-2#
37	Vorlesen	Vorlesen	B5: Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung, Lob und/oder Kritik, erhalten, um etwas für sich beziehungsweise seinen Unterricht zu lernen?
38	Zustimmung AO + Begründung	Zustimmung AO + Begründung	Die Schülerfehlvorstellungen, (.) weil die am Ende vor allem deutlich werden. #00:18:13-8#

39	Vorlesen	Vorlesen	B5: Robert hat schon einige Rückmeldungen erhalten. Besonders wichtig ist ihm jetzt noch eine Rückmeldung zu den aufgetretenen Schülervorstellungen im Unterrichtsausschnitt. #00:18:23-0#
40	Vorlesen	Vorlesen	B5: Welche der folgenden Schüler-Schülerfehlvorstellungen sind während der Unterrichtsphase sichtbar geworden?
41	Beschreibung	Beschreibung	Es sind auf jeden Fall Schülerfehlvorstellungen sichtbar geworden.
42	Zustimmung AO + Begründung	Zustimmung AO + Begründung	Die Kraft als etwas, was ein Körper hat, wurde ja einmal erwähnt, (.) als Eigenschaft, was jetzt hier auch erkennbar ist [Anm: B5 kreuzt AO 3 an].
43	Vorlesen	Vorlesen	Die Schüler denken, dass nur aktive Körper eine Kraft ausüben können, passive leisten nur Widerstand.
44	Vorlesen	Vorlesen	Die Schüler denken, dass eine Kraft-einer Kraft immer eine Beschleunigung zugrunde liegen muss und schließen daher eine Kraftwirkung bei ruhenden Gegenständen aus.
45	Vorlesen, Zustimmung AO + Begründung	Vorlesen, Zustimmung AO + Begründung	Die Schüler verwechseln Kraft und Gegenkraft mit den wirkenden Kräften, das würde ich genau sagen, weil die Schüler sagen immer von Kraft und Gegenkraft der Apfel wird ja auch angezogen und die Gegenkraft (.) vom Apfel auf die Erde ist eigentlich nicht relevant in diesem Fall, glaube ich.
46	Vorlesen	Vorlesen	Die Schüler- innen ziehen- sehen nicht, dass der Apfel über den Baum mit der Erde verbunden ist und der Baum durch das Festhalten (.) an der Erde eine Gegenkraft zum am Baum ziehenden Apfel ausübt.
47	Vorlesen, Sonstiges* (2x)	Sonstiges	Die Schüler- innen sehen nicht, der Apfel ist am Baum verbunden, (.) der Baum hält den Apfel fest und übt deswegen die Gegenkraft (.) zum am Baum- das macht für mich wenig Sinn, deswegen würde ich bei diesen beiden Punkten bleiben. #00:19:36-2#

48	Vorlesen	Vorlesen	B5: Welcher der unten stehenden Hinweise zum Umgang mit Schülerfehlvorstellungen möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben? Wählen Sie die Antwortoption aus, die Ihrer Einschätzung am nächsten kommt.
49	Ablehnung AO + Begründung	Ablehnung AO + Begründung	Er geht ja (mit?) Schülerfehlvorstellungen wenig um, das heißt er bezieht sich darauf gar nicht, deswegen ist der nicht angemessen.
50	Vorlesen, Zustimmung AO + Begründung	Vorlesen, Zustimmung AO + Begründung	Er sollte den Prozess der- der Unterschied zwischen Kraft und Energie, da ist es, damit Kraft keine Eigenschaft ist, die ein Körper haben und die Schülerfehlvorstellungen falsch s- nicht korrekt sind und die (um-?) äh verändert werden können (.).
51	Vorlesen, Ablehnung AO	Vorlesen, Ablehnung AO	Trifft auch dann, wenn man nichts tut, (.) das würde ich nicht sagen.
52	Vorlesen, Zustimmung AO + Begründung	Vorlesen, Zustimmung AO + Begründung	Du solltest Schülern die ver- verwendeten Begriffe, den Unterschied zwischen Kräftegleich- das ist genau richtig, (weil?) auch im dritten Newtonschen das relativ vermischt wird. (.)
53	Vorlesen, Zustimmung AO + Begründung	Vorlesen, Zustimmung AO + Begründung	Das Wechselwirkungsprinzip ausführlich mündlich oder schriftlich formulieren lassen, damit du siehst, was die Schüler*innen verstandendas wäre gut, um die Schülerfehlvorstellung genau einschätzen zu können. #00:20:37-9#

54	Vorlesen, Überlegung zum Unterricht, Alternative	Vorlesen, Überlegung zum Unterricht, Alternative	B5: Bei der Besprechung des Comics hatte ich das Gefühl, dass die Schüler- innen einige Vorstellungen-Schülervorstellungen haben. Deswegen habe ich die Schüler- innen gebeten, Stellung zu einer Seite des Comics zu beziehen, also Argumente entweder für die Seite des Mädchens oder des Jungens zu sammeln. Diese Argumente haben wir dann besprochen und Anna hat fachlich korrekt erklärt, warum die Position des Mädchens die richtige ist. Anschließend hatten die Schüler- innen die Möglichkeit, offen gebliebene Fragen zum Comic zu stellen. Da die Schüler- innen hier keine Fragen hatten, habe ich den Unterricht fortgeführt und zum Thema Impuls übergeleitet. (Und?) ich glaube schon, dass die Schüler Fragen haben, sie trauen sich vielleicht eher nicht, die zu stellen, da hättest du vielleicht noch ein bisschen mehr reingehen sollen in (dem?) Unterricht. #00:21:22-1#
55	-	-	[Video Einleitung Reflexion 03, Video Szene Reflexion 03] // #00:23:13-6#
56	Überlegung zum Unterricht	Überlegung zum Unterricht	B5: (unv.) #00:22:38-8#
57	Vorlesen	Vorlesen	B5: Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung, Lob und/oder Kritik, erhalten, um etwas für sich beziehungsweise seinen Unterricht zu lernen?
58	Vorlesen, Zustimmung AO + Begründung, Alternative	Vorlesen, Zustimmung AO + Begründung, Alternative	Den Einsatz des Beispiels, da fand ich die Wahl des Beispiels passend, aber (auf?) die Erklärung hätte vielleicht durch die Schüler passieren können oder ähnliches, deswegen möchte ich ihm da Rückmeldung geben. (.)
59	Zustimmung AO + Begründung (2x)	Zustimmung AO + Begründung (2x)	Und die Gesprächsführung, weil er gut auf die Schülervor- Schüler eingegangen ist und die Schülervorstellungen als Anknüpfpunkt genommen hat. #00:23:40-1#

60	Vorlesen	Vorlesen	B5: Robert hat schon einige Rückmeldungen erhalten. Besonders wichtig ist ihm jetzt noch eine Rückmeldung zur Nutzung der Fachsprache und zur Gestaltung des Phasenwechsels im Unterrichtsausschnitt. #00:23:50-0#
61	Vorlesen	Vorlesen	B5: Welches Feedback möchten Sie Robert zur Nutzung der Fachsprache geben?
62	Vorlesen, Bewertung + Begründung	Vorlesen, Ablehnung AO, Bewertung + Begründung	Es ist nicht gut, dass die Schülerinnen den Begriff Schwung nutzen, da die Bedeutug des Begriffs noch nciht geklärt ist (.) (unv.) und so unklar ist, was genau die Schüler(*innen?) beschreiben. Das würde jetzt- ich finde es gerade gut, dass der Begriff Schwung genutzt wird, weil dadurch die Schülervorstellungen an die Schülervorstellungen angegknüpft wird und er den Begriff Schwung (unv.) auch umdenken kann bei den Schülern die Vorstellung dazu beim Begriff Impuls. (.)
63	Vorlesen, Zustimmung AO, Alternative + Begründung	Vorlesen, Zustimmung AO Begründung	Es ist gut, dass die Schüler- innen zunächst ihre Alltagssprache verwenden können um zu- um zu- um zu erklären, warum es zur Kollision gekommen ist. So sinkt die Hemmschwelle für- genau das ist es. Die Schülervorstellungen- die Schüler sollen ihre Vorstellungen äußern und Fehler sind im Unterricht gut.
64	Vorlesen, Zustimmung AO + Begründung	Vorlesen, Zustimmung AO + Begründung	Es ist gut, zunächst mit dem alltagssprachlichen Begriff Schwung einzusteigen, da der Begriff Schwung ein tragfähiges Konzept- das würde ich auch sagen, es ist genau das anknüpfen an den Schülervorstellungen.
65	Vorlesen, Sonstiges*	Vorlesen, Ablehnung AO, Sonstiges	Die Formulierung stärker bremsen ist nicht gut gewählt, da auch Körper mit großem Impuls, wie zum Beispiel die Titanic, stärker bremsen könnten, um eine Kollision zu vermeiden. (..) Da weiß ich jetzt nicht ganz, was er damit meint und deswegen kreuze ich es auch nicht an.

66	Vorlesen, Ablehnung AO + Begründung, Alternative	Vorlesen, Ablehnung AO + Begründung, Alternative	Es ist nicht gut, dass die Schülerinnen mit dem Begriff Schwung arbeiten. So werden sie sich merken und zukünftig nicht mit dem Impulsbegriff arbeiten. Das finde ich gerade nicht, sie können mit dem Begriff Schwung arbeiten, nur muss der einfach nur umdefiniert werden für die Schüler.
67	Vorlesen, Ablehnung AO + Begründung	Vorlesen, Ablehnung AO + Begründung	Es ist nicht gut, dass die Schülerinnen in der Benutzung des Begriffs Schwung bestärkst. Das finde ich immer noch nicht, weil der Begriff Schwung für mich ein anschlussfähiges Konzept ist.
68	Vorlesen, Zustimmung AO, Bewertung + Begründung	Vorlesen, Zustimmung AO, Bewertung + Begründung	Es ist nicht gut, dass du so stockend redest und viele Füllwörter nutzt, das lenkt die Schüler vom Inhalt ab. Das finde ich teilweise, ich finde, dass du ein bisschen viele Pausen machst und dadurch die Schüler das Gefühl haben, dass ihre Antwort vielleicht falsch sein könnte und du sie verunsicherst. #00:25:36-7#
69	Vorlesen, Alternative	Vorlesen, Überlegung zum Unterricht	B5: Welche der unten stehenden Hinweise zu der Nutzung von Fachsprache möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben? Ja, (unv.) vor allem auf die Begriffe Schwung, dass er den behalten soll und die Fachbegriffe ja erst später einführen sollte vielleicht auch eher.
70	Vorlesen, Ablehnung AO	Vorlesen, Ablehnung AO	Den Impuls für den Schwung schon etwas f- für Schwung- ne, eben nicht.
71	Vorlesen, Zustimmung AO + Begründung	Vorlesen, Zustimmung AO + Begründung	Du solltest- Du solltest gegen die Schülervorstellung vorgehen, vergleichbar (unv.) Schwung haben, ist genau richtig, (weil er?) hat keinen Impuls, sondern ein Impuls wird ausgeübt.
72	Vorlesen	Vorlesen	Die Verwendung der Begrifflichkeiten war angemessen, das kannst du wieder so durchführen.
73	Vorlesen, Zustimmung AO, Alternative	Vorlesen, Zustimmung AO, Alternative	Genauer überlegen, was Schülern erzählen möchtest, das finde ich auch und er sollte auch genuer darauf sich Gedanken machen, wie er auf die Schüler- antworten eingeht.

74	Vorlesen, (Ablehnung AO) + Begründung	Vorlesen, (Ablehnung AO) + Begründung	Es wäre besser, Schüler- innen zur Verwendung der Fachsprache aufzufordern, (.) Konzepte Schwung und Impuls weiterhin verwechseln. Ich fand nicht, dass er es verwechselt ist. #00:26:30-9#
75	Vorlesen	Vorlesen	B5: Welches Feedback möchten Sie Robert zur Gestaltung des Phasenübergangs geben? (.)
76	Vorlesen	Vorlesen	Der Phasenübergang ist dir- ist dir nicht gut gelungen, da er zu abrupt war.
77	Vorlesen	Vorlesen	Der Phasenübergang ist dir nicht gut gelungen, da er zu kurz war. (.)
78	Vorlesen	Vorlesen	Inhalte zum Wechselwirkungsprinzip nicht ausreichend verstanden und gesichert wurden.
79	Vorlesen	Vorlesen	Es ist unproblematisch, dass der Phasenübergang abrupt war, da du mit dem Begriff Impulsbegriff wirklich ein komplett neues Thema beginnst.
80	Vorlesen	Vorlesen	Es ist nicht gut, dass du so einen abrupten Übergang zwischen den Themen gemacht hast, da der rote Faden der Stunde für so für die Schüler- innen nicht erkennbar ist.
81	Vorlesen, Zustimmung AO + Begründung	Vorlesen, Zustimmung AO + Begründung	Es ist nicht gut, dass du sagst, das Thema ist komplett neu, da den Schüler- innen so das bisher nicht mehr relevant erscheint. Das finde ich genau richtig, die sollten das bisher Gelernte immernoch auf den Begriff beziehen und somit Kraft und Impuls in Relation setzen oder in Beziehung setzen.
82	Vorlesen, Zustimmung AO	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du sagst, das ist komplett neu, da es einen fachlichen Bezug zwischen- und genau das ist es auch noch. #00:27:24-3#
83	Vorlesen	Vorlesen	B5: Welche der unten stehenden Hinweise zur Gestaltung des Phasenübergangs möchten Sie Robert für eine erneute Durchführung desder Unterrichtsstunde geben?
84	Zustimmung AO	Vorlesen, Zustimmung AO	Er sollte mehr die Unterrichtsphasen aneinander anknüpfen zwischen dem Zusammenhang, das finde ich gut [Anm: B5 kreuzt AO 2 an].

85	Vorlesen, (Ablehnung AO) + Begründung, Alternative	Vorlesen, Ablehnung AO + Begründung, Alternative	Der Einstieg in den Impuls auf die nächste Stunde zu verschieben, ja das Verschieben ändert jetzt auch nicht viel, man sollte eher den Zusammenhang dazwischen formulieren, dann kann man das auch in einer Stunde trotzdem machen.
86	Vorlesen, Zustimmung AO + Begründung, Alternative	Vorlesen, Zustimmung AO + Begründung, Alternative	Auch die Tafel zu wischen, damit die Schüler sich voll auf das neue Thema (unv.)- das finde ich auch, dass das neue Thema, das alte Thema damit (verwischt wird?) also weggewischt werden sollte, das neue Thema jetzt der Einstieg ist, außer anschlussfähige Konzepte, da könnte (man-?) man jetzt vielleicht noch ein bisschen Eingrenzung machen, ein bisschen mehr Antwortmöglichkeiten dazu.
87	Vorlesen, Bewertung	Vorlesen, Zustimmung AO + Begründung	Es wäre besser, einen glatteren Übergang zu gestalten, zum Beispiel durch eine kurze Verknüpfung,wie 'wir finden jetzt etwas neues mit dem gelernten Wissen heraus', damit der rote Faden- (.) der rote Faden fehlt gerade der Zusammenhang zwischen Impuls und Kraft.
88	Vorlesen, (Ablehnung AO) + Begründung	Vorlesen, Ablehnung AO + Begründung	Durch das Zeigen eines Experiments zu gestalten, ich finde einen Kontext auch genauso gut, wie ein Experiment und ist auch genauso lehrfähig. #00:28:31-6#
89	-	-	[Video Einleitung Reflexion 04, Video Szene Reflexion 04] // [Weiterhin Video Szene Reflexion 04] // [Weiterhin Video Szene Reflexion 04] // #00:30:40-3#
90	Beschreibung	Überlegung zum Unterricht	B5: (Und?) jetzt fangen die beiden (unv.) eigentlich. #00:29:25-8#
91	Alternative + Begründung	Alternative + Begründung	B5: Warum macht er das jetzt nicht selber wieder, ich glaube, es sollte eher der andere wieder werfen, damit er den Schwung vergleichen kann. #00:29:47-7#
92	Bewertung	Bewertung, Überlegung zum Unterricht	B5: Genau richtig, die Schülervorstellung an- (unv.) anzuschließen (unv.) Konzept (unv.). #00:30:25-3#

93	Bewertung	Alternative	B5: Abschreiben alleine wird denen jetzt wenig bringen, denke ich. #00:30:44-0#
94	Vorlesen	Vorlesen	B5: Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung, Lob und/oder Kritik, erhalten, um etwas für sich bezibeziehungsweise seinen Unterricht zu lernen?
95	Vorlesen	Vorlesen	Eingesetzte Beispiele,
96	Vorlesen	Vorlesen	Fachsprache,
97	Vorlesen	Vorlesen	Gesprächsführung,
98	Vorlesen	Vorlesen	Wahl des Experiments,
99	Vorlesen	Vorlesen	Ergebnissicherung.
100	Zustimmung AO + Begründung	Zustimmung AO + Begründung	Ich würde Wahl des Experiments und Ergebnissicherung nehmen, weil alleine Abschreiben finde ich nicht tragfähig. #00:31:05-3#
101	Vorlesen	Vorlesen	B5: Robert hat schon einige Rückmeldungen erhalten. Besonders wichtig ist ihm jetzt noch eine Rückmeldung zur Wahl des Experiments und zur Ergebnissicherung nach dem Experiment. #00:31:14-2#
102	Vorlesen	Vorlesen	B5: Welches Feedback möchten Sie Robert zur Wahl des Experiments geben?
103	Zustimmung AO + Begründung, Bewertung	Zustimmung AO, Bewertung	An sich war das Experiment gut gewählt, die Durchführung würde ich gerne ein bisschen kritisieren. Äh deswegen das Experiment ist gut gewählt, da die beiden Schüler durchführen, Impuls erfahren können, das finde ich gut,
104	Zustimmung AO, Alternative	Zustimmung AO, Alternative	es stellt auch den Einfluss von Masse und Geschwindigkeit den Unterschied gut dar, aber da hätte er die Schüler vielleicht anders wechseln sollen. (.)
105	Vorlesen, Ablehnung AO + Begründung	Vorlesen, Ablehnung AO + Begründung	Nur über Masse und die Geschwindigkeit, von anderen Größen, das finde ich jetzt gar nicht, weil er erstmal den Impuls einführen möchte.
106	Bewertung	Ablehnung AO	(Und?) das Experiment finde ich gut gewählt.
107	Zustimmung AO	Zustimmung AO+Begründung	(Verhalten?) des Einbeziehen von den Schülern das ist immer gut, dass die Schüler Experimente

108	Alternative, Sonstiges	Alternative, Sonstiges	auch an sich könnte das Experiment glaube ich auch noch noch offener für die Schüler (allein?) durchführen, das hätte ich jetzt gerne noch als Ankreuz- hier oder als Hinweis.
109	Vorlesen, Zustimmung AO, Alternative	Vorlesen, Ablehnung AO, Alternative	Dass du vorgibst, wer nach vorne kommen soll, (.) wer will denn, da weiß ich jetzt nicht ganz, warum, aber es ist schon gut so, dass er mal welche vorgibt, da sollte er gucken, dass er das ein bisschen abwechselt.
110	Vorlesen, Überlegung zum Unterricht	Zustimmung AO + Begründung, Sonstiges	Es ist gut, dass du nur zwei Schüler das Ex- durchführen lässt und nicht alle Schüler, es spart Zeit, das muss man immer gucken, (wie er?) das einbetten will, dazu fehlen mir jetzt ein bisschen die Hinweise zum weiteren Unterricht oder zur Zeit, wie viel Stunden Unterricht er dafür noch hat für diese Einheit. #00:32:24-0#
111	Vorlesen	Vorlesen	B5: Welche der unten stehenden Hinweise zur Wahl des Experiments möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben?
112	Vorlesen, Zustimmung AO	Vorlesen, Zustimmung AO	Das Experiment ist gut gewählt, das finde ich auf jeden Fall.
113	Vorlesen, Ablehnung AO + Begründung	Vorlesen, Ablehnung AO + Begründung	Es wäre besser, (.) ein Video finde ich gar nicht, (beim?) Video erfahren die Schüler das nicht. (.)
114	Vorlesen, (Ablehnung AO) + Begründung, Alternative	Vorlesen, Ablehnung AO + Begründung, Alternative	Durchzuführen, dass für alle Schüler anschaulich ist. Ich fand es schon sehr anschaulich, dass die Schüler beschrieben haben, man könnte mit mehr Tennisbällen, dass jeder Schüler das Experiment durchführt.
115	Vorlesen, (Zustimmung AO) + Begründung, Bewertung	Vorlesen, Ablehnung AO + Begründung, Alternative	Es wäre besser, ein Stoßexperiment- (.) da könnte man den Zusammenhang darstellen, aber an sich finde ich das jetzt immer trotzdem noch so gut.
116	Vorlesen, Ablehnung AO + Begründung	Vorlesen, Ablehnung AO + Begründung	Zu wählen, zu dem eine Messreihe- ich finde, eine Messreihe und Formeln sind nicht unbedingt immer sinnvoll, auch objektive Experimente sind manchmal gut oder auch quali- qualitative Experimente. (.)

117	Vorlesen, Zustimmung AO + Begründung	Vorlesen, Zustimmung AO, Alternative	Von allen Schülern, das finde ich, weil das kann man mit Tennisbällen auch sehr gut umsetzen, ich meine da- er braucht nur Tennisbälle, mehr braucht er ja nicht. (.)
118	Vorlesen, (Ablehnung AO) + Begründung	Vorlesen, Ablehnung AO + Begründung	Dass die Schüler, die vorne stehen, dass alle sitzenden erkennen können, ich finde schon, dass sie es erkennen konnten
119	Überlegungen zum Unterricht, Bewertung	Überlegungen zum Unterricht, Bewertung	und die Schüler müssen da auch mal ein bisschen ihre Gedanken äußern, das finde ich schon (genau?) gut. #00:33:27-0#
120	Vorlesen	Vorlesen	B5: Welches Feedback möchten Sie Robert zur Ergebnissicherung geben?
121	Bewertung, Alternative	Bewertung, Alternative	Der fertige Merksatz, finde ich, ist eine Katastrophe. Ein- die Schüler sollten den vielleicht auch mitkonstruieren können den Merksatz
122	Sonstiges, Bewertung	Bewertung, Alternative	oder die Einwirkung von Geschwindigkeit und Masse, (.) (gerade?) die Sicherung notieren, die Notation davon ist gut, dass sie es aufschreiben,
123	Bewertung, Alternative, Sonstiges	Alternative, Sonstiges	aber halt nicht, dass sie den Merksatz einfach abschreiben, da hätte ich gern noch eine Option, dass sie was aufschreiben auf jeden Fall, aber nicht den Merksatz abschreiben. (.)
124	Vorlesen, Ablehnung AO + Begründung	Vorlesen, Ablehnung AO + Begründung	Es ist nicht gut, dass du die Ergebnisse des Experiments nicht sicherst, das finde ich nicht schlimm, weil es eher ein qualitatives Gespräch im Unterricht ist.
125	Vorlesen, Zustimmung AO, Beschreibung	Vorlesen, Zustimmung AO, Beschreibung	Der Merksatz ist gut gelungen, ich finde ja, aber er ist halt von ihm konstruiert, trotzdem ist er halt gut gelungen.
126	Vorlesen, Zustimmung AO	Vorlesen, Zustimmung AO + Begründung	Es ist nicht gut, dass du für den Impuls als skalare Größe- (.) das ist genau das Problem von der Kinematik.
127	Vorlesen, Zustimmung AO	Vorlesen, Zustimmung AO	Dass du den Begriff der Energie nicht erwähnst, das finde ich gerade gut. (.)

128	Vorlesen, Ablehnung AO + Begründung, Alternative	Vorlesen, Ablehnung AO + Begründung, Alternative	In deiner Sicherung fehlt eine Abgrenzung des Impulsbegriffs von der kinetischen Energie, das finde ich nicht unbedingt nötig, dass er das jetzt unbedingt noch abgrenzt, dann würde er glaube ich nur noch die Schüler verwirren, das kann er ja später machen. #00:34:35-5#
129	Vorlesen	Vorlesen	B5: Welche der unten stehenden Hinweise zur Ergebnissicherung möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben?
130	Vorlesen, Ablehnung AO	Vorlesen, Ablehnung AO	Das Ergebnis des Experiments hast du gut gesichert, das finde ich gar nicht.
131	Vorlesen, Ablehnung AO	Vorlesen, Ablehnung AO + Begründung	Es wäre gut, dass die Beobachtungen schriftlich, das finde ich bei qualitativen Fall (unv.) auch nicht unbedingt nötig.
132	Vorlesen, Zustimmung AO, Beschreibung, Alternative	Vorlesen, Zustimmung AO, Alternative	In die Sicherung einzubeziehen auf jeden Fall. Die Schüler haben einen Merksatz bekommen, der sollte vielleicht mit den Schülern besprochen sein.
133	Ablehnung AO + Begründung, Alternative, Überlebung zum Unterricht	Ablehnung AO + Begründung, Alternative, Überlebung zum Unterricht	Noch Beispiele finde ich auch nicht wichtig, das war- das war ein Beispiel, ich finde dieses Beispiel können die Schüler auch alle durchprobieren auch selber, ein Beispiel kommt wahrscheinlich auch noch später.
134	Vorlesen, Zustimmung AO, Alternative	Zustimmung AO, Alternative	Auf das Experiment einzugehen, (.) ja das- das fände ich, könnte man bei der Sicherung mit den Schülern gemeinsam machen, dass man das Experiment dazu auch noch bezieht.
135	Zustimmung AO	Vorlesen, Zustimmung AO	Als vektorielle Größe auf jeden Fall,
136	Ablehnung AO	Ablehnung AO	der Rest finde ich nicht. #00:35:25-6#

137	Vorlesen (2x), Beschreibung (2x), Alternative (2x), Bewertung, Überlegung zum Unterricht, Sonstiges	Vorlesen, Beschreibung, Alternative, Vorlesen, Beschreibung, Überlegungen zum Unterricht, Bewertung, Alternative, sonstiges	B5: Nachdem wir das Experiment mit den Tennisbällen abgeschlossen und die Schüler- innen den Merksatz zum Impuls abgeschrieben haben, habe ich zu einem Demonstrationsexperiment übergeleitet, an dem ich die Impulserhaltung veranschaulichen möchte. Er hat jetzt ein Schülerexperiment mit zwei Schülern und dann (das?) Exp- Demonstrationsexperiment. Wieso jetzt nicht ein Schülerexperiment zum Beispiel? Mit dem Demonstrationsexperiment möchte ich einen Auffahrunfall simulieren. Den Schülerinnen habe ich dann erklärt, wie der Auffahr- Autounfall im Experiment nachgestellt wird und wie die Messwertaufnahme funktioniert. Den experimentellen Aufbau und eine Messwerttabelle habe ich vor der Stunde vorbereitet, hier habe ich dir zwei Fotos vom Aufbau und von der Tabelle mitgebracht. Im Aufbau sieht man zwei Wagen, die fahren wahrscheinlich gegeneinander, denke ich mal und dann soll es ein Impuls- Stoßexperiment darstellen. Ist das jetzt ein gekoppelter Stoß oder ein elastischer Stoß, das kann ich nicht ganz erkennen, aber das wird er dann ja wahrscheinlich klar abgrenzen. Die Messwerttabelle sieht erstmal sehr sinnvoll aus. Die Größen sollte man vielleicht noch mit einer Einheit einführen oder in der Tabelle auf die Einheiten achten, da bin ich jetzt gleich mal gespannt, wie er das durchführt. #00:36:40-0#
138	-	-	[Video Einleitung Reflexion 05, Video Szene Reflexion 05] // [Weiterhin Video Szene Reflexion 05] // [Weiterhin Video Szene Reflexion 05] // [Weiterhin Video Szene Reflexion 05] // [Weiterhin Video Szene Reflexion 05] // [Weiterhin Video Szene Reflexion 05] // [Weiterhin Video Szene Reflexion 05] #00:39:20-4#

139	Überlegung zum Unterricht, Alternative	Überlegung zum Unterricht, Alternative	B5: (Ich?) finde jetzt ein bisschen nicht physikalisch tragfähig auch für weitere Experimente (unv.) anschlusfähig für die Schüler, er sollte beide messen. #00:37:20-7#
140	Beschreibung, Bewertung, Alternative	Beschreibung, Bewertung, Alternative	B5: Gut, das ist jetzt sehr vorurteilbehaftet, das ist ziemlich schwierig, das so zu formulieren, weil gerade der Unterricht an Mädchen orientiert sein sollte, damit die nicht das Bild von der Physik haben, dass sie nur schreiben sondern auch in der Physik (das?) Durchführen. #00:37:42-9#
141	Bewertung, Alternative	Bewertung, Alternative	B5: Es ist gut, dass er auf die Einheit hinweist (.) er hätte aber die Schüler fragen sollen, was falsch ist. #00:37:58-0#
142	Alternative	Alternative	B5: Wieso rechnet er die jetzt aus und la- lässt die nicht die Schüler machen? (unv.) die Schüler(- innen?) #00:38:31-9#
143	Beschreibung , Alternative + Begründung	Beschreibung, Alternative, Beschreibung	B5: Die Gesprächsführung ist jetzt sehr harsch zu den Schülern, da könnte er ein bisschen mit den Schülern arbeiten, (weil?) eigentlich (arbeitet?) er jetzt alleine und die Schüler gucken (unv.) nur zu. #00:38:49-4#
144	Alternative, Überlegung zum Unterricht	Alternative, Überlegung zum Unterricht	B5: Wieso gibt er jetzt alle Messdaten vor und hilft- lässt die Schüler es nicht durchführen. (.) Ich glaube, dass er jetzt gerade die Kontrolle über die Klasse verlieren könnte. #00:39:07-9#
145	Zustimmung AO	Zustimmung AO	B5: Da würde ich ganz gern den Umgang von ihm besprechen
146	Zustimmung AO	Zustimmung AO	und auch die Aktivierung der Schüler.
147	Vorlesen	Vorlesen	Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung, Lob und/oder Kritik, erhalten, um etwas für sich bzw. seinen Unterricht zu lernen? #00:39:36-8#

148	Vorlesen	Vorlesen	B5: Robert hat schon einige Rückmeldungen erhalten. Besonders wichtig ist ihm jetzt noch eine Rückmeldung zum Umgang mit den Schüler- innen, dazu, wie er während des Experiments auf Messunsicherheiten eingegangen ist und zur Aktivierung der Schüler- innen. #00:39:50-5#
149	Vorlesen	Vorlesen	B5: Welches Feedback möchten Sie Robert zu seinem Umgang mit den Schüler- innen geben?
150	Bewertung, Bewertung + Begründung	Bewertung, Bewertung + Begründung	Der Umgang mit der Schülerin ist gerade nicht gut. Da ist die Reduktion und das finde ich ganz schlecht.
151	Vorlesen, Zustimmung AO	Vorlesen, Zustimmung AO	(unv.) Fleißarbeit einbeziehen, das ist auch genau das. (.)
152	Vorlesen, Zustimmung AO, Ablehnung AO, Beschreibung, Alternative	Vorlesen, Ablehnung AO, Beschreibung, Bewertung Alternative	Es ist gut, dass du die Mädchen im Unterricht für ihr Handeln im Unterricht lobst, um ihr Selbstkonzept- ja, aber er lobt sie ja eigentlich für Handeln, was nicht lobenswert- wofür (unv.) gar nicht loben wollen, die Mädchen, sondern sie sollen ja selber in der Physik etwas rechnen zum Beispiel oder Durchführen auch, nicht nur schreiben.
153	Vorlesen, Zustimmung AO + Begründung, Alternative	Vorlesen, Zustimmung AO + Begründung, Alternative	Es ist nicht gut, dass du das Lob für die Handschrift von Mädchen und- mit einer inhaltlichen Kritik kombinierst. Das ist genau das Problem. Das Mädchen hat (jetzt?) das Gefühl, dass sie Physik gar nicht kann, sondern nur schön schreiben kann, da sollte er ein bisschen mehr auf die Mädchen eingehen im Unterricht.
154	Vorlesen, Alternative (2x)	Vorlesen, Alternative, Alternative	Nicht gut, dass Schülerin schon vor dem Notieren des zweiten Messwertes erneut auf die Einheit hinweist, ich finde, er sollte auch den Fehler von dem Mädchen erlauben und Schülerfehler möchte er haben, damit die Schüler- innen das Mädchen halt darauf hinweisen können.
155	Zustimmung AO, Alternative	Zustimmung AO, Alternative	Die Einheit finde ich auch nicht gut, er sollte da mehr die Schüler das machen lassen selbstständig ohne da alles vorzugeben.

156	Alternative, (Ablehnung AO) + Begründung	Zustimmung AO, Ablehnung AO + Begründung	Die schöne Handschrift kann er loben, aber es ist gerade eher in der Physik und auch in dem Bild von den Mädchen- was Mädchen von sich selber in der Physik haben nicht anschlussfähig. #00:41:11-9#
157	Vorlesen	Vorlesen	B5: Welche der unten stehenden Hinweise zum Umgang mit den Schüler- innen möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben?
158	Bewertung	Bewertung	Dein Umgang mit den Mädchen ist und Jungs ist gar nicht in Ordnung, das geht nicht.
159	Vorlesen, Beschreibung, Alternative	Vorlesen, Beschreibung, Zustimmung AO, Alternative	Du solltest eine Geschichte darum, das hat er ja gemacht eigentlich, aber an sich sollte die Geschichte vielleicht noch weniger technisch sein und bisschen Mädchen motiverender.
160	Vorlesen, Zustimmung AO + Begründung, Alternative	Vorlesen, Zustimmung AO + Begründung, Alternative	Es wäre besser, die Schüler- innen zu fragen, wer freiwillig schreiben bzw. die Werte ablesen will, genau, damit er das Geschlecht eben nicht einbezieht. Vielleicht auch sogar einen Jungen extra nehmen.
161	Vorlesen, Alternative, Ablehnung AO + Begründung, (Zustimmung AO) + Begründung	Vorlesen, Ablehnung AO + Begründung (2x), Alternative	Es wäre besser, die Aufgabe durch zu tauschen, damit alle Schüler- innen eine Aufgabe wahrnehmen. Ich glaube auch, dass er dadurch mehr die Kontrolle behalten (soll?), das eine würde zu lange dauern, deswegen würde ich das eher nicht sagen, sondern die Schüler- innen sollten das Experiment lieber durchführen gemeinsam und gemeinsam drüber sprechen.
162	Vorlesen, Zustimmung AO + Begründung	Vorlesen, Zustimmung AO + Begründung	(Du solltest?) die Mädchen nicht für ihre Handschrift, sondern explizitgenau, um den Mädchen ein gutes Selbstwergefühl zu geben, weil das die Leistung steigert und das Interesse an der Physik.
163	Vorlesen, Zustimmung AO	Vorlesen, Zustimmung AO	Geschlechterspzifische Zuschreibungen verzichten, da du so ein unangemessenes Verhalten zwischen Geschlechtern in der Physik. Genau, das sollte er auf jeden Fall vermeiden.

164	Vorlesen, Bewertung	Vorlesen, Bewertung	Du solltest grundlegend an deiner Einstellung gegenüber Mädchen und Jungen arbeiten, das stereotypische Verhalten ist sehr schlecht. #00:42:22-1#
165	Vorlesen	Vorlesen	B5: Welches Feedback möchten Sie Robert zu seinem Umgang mit Messunsicherheiten geben?
166	Vorlesen, Beschreibung, Zustimmung AO, Alternative	Vorlesen, Beschreibung, Beschreibung, Bewertung, Alternative	Äh es ist nicht gut, dass du sagst, du könntest exakt oder genau messen, da du so die Schülervorstellung beförderst, man könnte Größen ohne Messunsicherheiten bestimmen. Ich finde im Experiment hat er gerade immer exakt und genau gesagt und die Berechnung der- mit den Schülern auch gar nicht besprochen und auch die Masse nur von einem Gleiter besprochen, das heißt er ist auf Messunsicherheiten nicht eingegangen und auch nur einen der beiden Wagen wiegen, ist auch nicht genau gut da sollte er mit den Schülern vielleicht selber besprechen, ob er jetzt nur einen oder beide wiegen muss.
167	Vorlesen, Zustimmung AO	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du sagst, die Wagen wären gleich schwer, da das da- da das auch wenn sie baugleich sind nicht stimmen muss. Ist genau richtig.
168	Vorlesen, Alternative (2x), Bewertung	Vorlesen, Alternative, Zustimmung AO, Alternative	Es ist nicht gut, dass du das Gewicht mit null Komma ein Kilogramm angibst, da du hier nicht berücksichtigst, welches die signifikanten Nachkommastellen sind. Ich denke, da sollte er dann auf jeden Fall so über das Runden reden, ansonsten finde ich das eigentlich okay, wenn er das so angibt beim Schulunterricht, aber er sollte er halt behandeln und da mit mit den Schülern drüber sprechen vorher oder ein Rundungszeichen vielleicht da hin schreiben.

169	Vorlesen, Zustimmung AO, Alternative	Vorlesen, Zustimmung AO, Alternative	Dass du darauf verzichtest, auf Messunsicherheiten (unv.), um die Komplexität- das finde ich auch okay, aber er sollte- soll trotzdem drauf eingehen, nur sie vielleicht nicht behandeln am Ende.
170	Vorlesen, Zustimmung AO	Vorlesen, Zustimmung AO	Es ist gut, dass du das Gewicht der Wagen mit null Komma ein Kilogramm, um Fehler in der Berechnung des Impuls aufgrund- das finde ich insgesamt richtig,
171	Sonstiges, Bewertung, Alternative	Sonstiges, Alternative	jetzt hätte ich gerne eine Option, dass das Runden auf die null Komma eins Kilogramm sehr gut ist, äh dass er das auf jeden Fall als Rundungseinheit angeben sollte. #00:43:51-6#
172	Vorlesen	Vorlesen	B5: Welche der unten stehenden Hinweise zu seinem Umgang mit Messunsicherheiten möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben?
173	Vorlesen, Bewertung + Begründung	Vorlesen, Bewertung + Begründung	Es ist gut, wie du mit Me- Es ist mit Messunsicherheiten katastrophal der Umgang, er hat sie nicht richtig berücksichtigt.
174	Vorlesen, Zustimmung AO	Vorlesen, Zustimmung AO	Auch Messunsicherheiten einbeziehen, genau das hier ist es.
175	Ablehnung AO, Alternative	Ablehnung AO, Alternative	Alle Nachkommastellen angeben finde ich nicht, er sollte aber über die Nachkommastellen reden.
176	Vorlesen, Zustimmung AO + Begründung	Vorlesen, Zustimmung AO + Begründung	Beide Wagen zu wiegen, das finde ich auch korrekt, damit er ein anschlussfähiges Konzept vom wissenschaftlichen Arbeiten vermittelt.
177	Vorlesen, Zustimmung AO + Begründung, Sonstiges*	Vorlesen, Zustimmung AO + Begründung, Sonstiges	Beide Wagen zu wiegen, da die Wagen- [Anm: B5 wählt AO5] genau, das müsste- das wird den Schülern nochmal deutlich, es geht aber beides eigentlich in ein über (unv.), da würde glaube ich ein Ankreuzaspekt reichen.
178	Vorlesen, Zustimmung AO, Alternative	Vorlesen, Zustimmung AO, Alternative	Da du im Vergleich der Massen auf Messunsicherheiten beziehungs- eingehen kannst. Das finde ich genau sinnvoll, beide Wagen deswegen zu wiegen und auch dann zum Beispiel über Baugleichheit könnte man noch ein bisschen reden mit den Schülern.

179	Vorlesen, Überlegung zum Unterricht	Vorlesen, Zustimmung AO+Begründung	Beide Wagen zu wiegen, damit die Schüler sich selbst überzeugen können. Ich denke mal, dass die Schüler da selber auch drauf ansprechen sollten, warum das jetzt wirklich so sein muss, wenn es baugleich ist. #00:44:52-5#
180	Vorlesen	Vorlesen	B5: Welches Feedback möchten Sie Robert zur Aktivierung der Schülerinnen geben?
181	Beschreibung	Beschreibung, Sonstiges	Die Schüler sind eher nicht aktiv, also es ist nur eine Schülerin aktiv, die müssen nicht mitdenken, das würde ich ihm gerne sagen.
182	Vorlesen, Zustimmung AO, Ablehnung AO + Begründung, Beschreibung	Vorlesen, Zustimmung AO + Begründung, Ablehnung AO + Begründung	Nur oberflächlich involvierst, das ist genau richtig, die schreiben nur. Es bezieht sie zwar ein, aber nur das eine Mädchen vorne, deswegen kreuzwürde ich das nicht ankreuzen.
183	Vorlesen, Ablehnung AO + Begründung	Vorlesen, Ablehnung AO, Alternative	Es ist gut, dass du das Experiment selbst durchführst, ich finde nicht, dass das gut ist, dass er das selbst durchführt. Gerade die Unsich- Unsicherheiten der Schüler, die dabei rauskommen können vielleicht auch für den Unterricht anregend sein.
184	Vorlesen, Ablehnung AO + Begründung, Alternative	Vorlesen, Ablehnung AO + Begründung, Alternative	Ohne Arbeitsauftrag, ja wenn sie einen Arbeitsauftrag hätten, würden sie ja nicht mehr zusehen, deswegen finde ich den Arbeitsauftrag schwer zu geben, ich finde, die Schüler sollten das Experiment eher durchführen. (.)
185	Vorlesen, Zustimmung AO	Vorlesen, Zustimmung AO, Sonstiges	Ist nicht gut, dass du das Experiment- da ist es, lehrerzentriert, das habe ich gerade im Ansatz gesehen.

186	Vorlesen, Zustimmung AO, Alternative (2x), Alternative (2x) + Begründung (2x), Beschreibung (2x), Testinstrument	Vorlesen, Zustimmung AO, Alternative (2x), Beschreibung, Alternative+Begründung, Alternative+Begründung, sonstiges	Es ist nicht gut, dass du die Masse des Wagens und insbesondere die Geschwindigkeit selbst bestimmst, das finde ich genau richtig, (also?) die Schüler sollten die Geschwindigkeit bestimmen und auch wie man das bestimmt sollte er mit den Schülern besprechen, er hat es einfach vorgerechnet. Die Schüler von hinten können es nicht erkennen, wenn dann müssen die (zusammenkommen?), um die Schüler aktiv zu halten, finde ich, dann sollten sie auch nach vorne kommen und eine notiert vorne, das könnte man vielleicht noch als Option für ihn geben, damit man ihn (hier?) verbessern kann. #00:46:07-0#
187	Vorlesen	Vorlesen	B5: Welche der unten stehenden Hinweise zur Aktivierung der Schüler- innen möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben?
188	Vorlesen, Zustimmung AO, Alternative	Vorlesen, Zustimmung AO, Alternative	Es wäre besser, die Schüler- innen zu fragen, wer freiwillig schreiben will beziehungsweise ablesen möchte, das finde ich genau richtig, dabei sollte er aber auch ach- darauf achten, dass vielleicht ein Junge schreibt.
189	Vorlesen, Beschreibung	Vorlesen, Beschreibung	Es wäre besser, die Schüler- innen zu fragen, ob sie den Ablauf verstanden oder noch Schwierigkeiten haben. Seine Rück- seine Rechnung ist ganz alleine, deswegen merkt er gar nicht, ob die Schüler das verstehen, was er macht, oder nicht.
190	Vorlesen, Überlegung zum Unterricht	Vorlesen, Überlegung zum Unterricht	Du hast die Schüler- innen ausreichend aktiviert. Ich glaube, dass die Schüler eher (unv.) oder gar nicht mehr aufgepasst haben, aber es hat man gerade leider auch nicht gesehen (unv.), was mit den Schülern hinten war.
191	Vorlesen	Vorlesen	Es wäre besser, Hypothesen der Schüler- innen abzufragen, um sie zur Beobachtung des Experiments zu motivieren.

192	Vorlesen, Zustimmung AO + Begründung	Vorlesen, Zustimmung AO + Begründung	Es wäre besser, den Schüler- innen Fragen zum Vorgehen während des Experiments oder zur Berechnung der Geschwindigkeit zu stellen, um sie zum Mitdenken anzuregen. Das finde ich genau, dann sind- hat man auch die Schüler unter Kontrolle und hat ein besseres Classroom-Management.
193	Vorlesen, Alternative	Vorlesen, Zustimmung AO, Alternative	Es wäre besser, die Schüler- innen stärker in die Durchführung des Experiments einzubeziehen, zum Beispiel um Wagen anstoßen, Werte ablesen, ich finde auch, dass man das Experiment (zwar?) vorne durchführen kann zentral, aber die Schüler das durchführen sollten und der Lehrer eher Anweisungen oder Hilfestellungen- (unv.) Hilfestellungen geben sollte.
194	Vorlesen, Zustimmung AO, Alternative	Vorlesen, Zustimmung AO, Alternative	Es wäre besser, die Schüler- innen zum Experiment nach vorn kommen zu lassen, damit sie alles gut erkennen können. Auf jeden Fall, das auf jeden Fall und auch selber durchführen können. #00:47:28-5#
195	Vorlesen, Sonstiges	Vorlesen, Sonstiges	B5: Nachdem wir alle Messwerte aufgenommen hatten, habe ich die Beobachtungen der Schüler- innen abgefragt. Um die Allgemeingültigkeit der Impulserhaltung nachzuweisen, habe ich sie dann in einem kurzen Exkurs theoretisch hergeleitet. Im Anschluss haben wir die Impulserhaltung am Experiment überprüft. Da bin ich jetzt gespannt, wie er es theoretisch hergeleitet hat. #00:47:49-9#
196	-	-	[Video Einleitung Reflexion 06] // #00:48:12-6#
197	Überlegung zum Unterricht	Überlegung zum Unterricht	B5: (unv.) liegt ja gerade an deiner Handlung mit der Messunsicherheit, nicht einbezogen. #00:48:10-1#
198	-	-	[Video Szene Reflexion 06] // [Weiterhhin Video Szene Reflexion 06] // [Weiterhhin Video Szene Reflexion 06] // [Weiterhhin Video Szene Reflexion 06] // [Weiterhhin Video Szene Reflexion 06] #00:50:39-9#

199	Überlegung zum Unterricht	Überlegung zum Unterricht	B5: Warum lobt er jetzt wieder die Jungs direkt, das ist ja gerade (unv.). #00:48:55-6#
200	Beschreibung (2x), Bewertung, Überlegung zum Unterricht	Überlegung zum Unterricht, Alternative	B5: Notiert er jetzt, kann ich leider nicht ganz erkennen, ich glaube den Impuls, (aber?) warum fehlen jetzt die Einheiten? (5s) Und (hier?) jetzt so viel Nachkommastellen und eben gar nicht. Ich glaube, da sollte er ein bisschen mehr auf die signifikanten Stellen achten. #00:49:31-1#
201	Überlegung zum Unterricht	Überlegung zum Unterricht	B5: (unv.) Problem würde ich sagen Messunsicherheiten. #00:50:13-0#
202	Beschreibung, Bewertung, Überlegung zum Unterricht	Bewertung, Überlegung zum Unterricht	B5: (Jetzt?) erzählt er über Rundungen und vorher sagt er einfach null Komma eins Kilogramm, warum das jetzt? (.) Ein (unv.) Konzept hinter seiner Einführung der Messungenauigkeiten ist da nicht hinter. #00:50:30-5#
203	Vorlesen	Vorlesen	B5: Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung, Lob und/oder Kritik, erhalten, um etwas für sich beziehungsweise seinen Unterricht zu lernen?
204	Vorlesen, Ablehnung AO	Vorlesen, Zustimmung AO + Bewertung	Der Umgang mit den Schüler- innen ist sch- ist wieder nicht gut.
205	Bewertung, Überlegung zum Unterricht	Vorlesen, Zustimmung AO, Überlegung zum Unterricht	Und die Messunsicherheiten mit dem Einwand von Marie ist genau das Problem, was vorher schon war, (also?) eigentlich ist das Ganze (unv.) Anschlussfähigkeit von (vorher?) wieder das gleiche. #00:51:01-4#
206	Vorlesen	Vorlesen	B5: Robert hat zwar schon einige Rückmeldungen erhalten, würde aber trotzdem gern zu allen drei aufgeführten Aspekten noch eine Rückmeldung erhalten. #00:51:09-3#
207	Vorlesen	Vorlesen	B5: Welches Feedback möchten Sie Robert zu Umgang mit den Schülerinnen geben?
208	Vorlesen, Zustimmung AO	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du Mädels sagst, das finde ich auch.

209	Vorlesen, Beschreibung	Vorlesen, Beschreibung	Es ist nicht- nicht gut, dass du die Mädchen und Jungen so unterschiedlich behandelst, er sagt von Beginn an, dass nur Jungen die richtigen Ergebnisse haben.
210	Vorlesen, Zustimmung AO + Begründung, Beschreibung	Vorlesen, Zustimmung AO + Begründung, Alternative	Es ist nicht gut, dass du die Mädchen und Jungen in ihrem Können so unterschiedlich einschätzt, das finde ich genau, er fördert das eigentlich, also genau das Gegenteil, was man machen sollte.
211	Vorlesen, Zustimmung AO, Beschreibung	Vorlesen, Zustimmung AO, Beschreibung	Du thematisierst die Einheit des Impulses nicht. Das ist nicht gut, genau, eben sollten die Einheiten notiert werden und jetzt, wenn er sie notiert, muss das nicht sein.
212	Vorlesen, Bewertung (2x), Ablehnung AO + Begründung, Testinstrument	Vorlesen, Ablehnung AO+Begründung, Bewertung, Sonstiges	Es ist gut, dass du die Mädchen dazu aufforderst, ihre Tabelle zu korrigieren, damit du sicher sein kannst, dass sie die gleichen Werte, wie die Jungen haben. Das Korrigieren finde ich gut, wie er sie auffordert, finde ich nicht gut und wie er es bespricht, deswegen kreuze ich das jetzt mal nicht an, weil ich da eher halb ge- zugestimmt bin, das würde ich gern noch ein bisschen differenzieren. #00:52:01-2#
213	Vorlesen	Vorlesen	B5: Welche der unten stehenden Hinweise zu seinem Umgang mit den Schüler- innen möchten Sie Robert für sei- eine erneute Durchführung der Unterrichtsstunde geben?
214	Vorlesen, Ablehnung AO	Vorlesen, Ablehnung AO	Dein Umgang mit den Schüler- innen ist- der ist nicht in Ordnung.
215	Vorlesen, Zustimmung AO	Vorlesen, Zustimmung AO	Du solltest die Mädchen stärker einbeziehen, das finde ich.
216	Vorlesen	Vorlesen	Er sollte geschlechtsspre- spezifische Zuschreibungen verwenden äh vermeiden, so.
217	Vorlesen, Überlegung zum Unterricht, Zustimmung AO, Sonstiges	Vorlesen, Überlegung zum Unterricht, sonstiges	Jedem Schüler- jedem nur einen berechneten Wert erfragen, (unv.) das ist je nachdem wie viel Zeit er hat und wie einfach auch die Aufgabe war und vielleicht wenn sie ein bisschen schwerer war oder in Teilen schwerer war finde ich, sollte er das machen, das könnte man noch mehr differenzieren.

218	Vorlesen, Zustimmung AO + Begründung	Vorlesen, Beschreibung, Zustimmung AO	Du solltest verstärkt schwächere Schüler- innen einbinden. Er hat seine Musterschüler wahrscheinlich in der ersten Reihe drangenommen, deswegen finde ich das auf jeden Fall.
219	Vorlesen, Zustimmung AO, Beschreibung, Sonstiges	Vorlesen, Zustimmung AO, Beschreibung, Überlegung zum Unterricht	Unabhängig vom Geschlecht dazu, falsche Ergebnisse der Tabel- genau, er hat jetzt nur die Mädchen korrigiert, aber da sieht man vielleicht auch nicht alles durch den kleinen Ausschnitt.
220	Vorlesen, Zustimmung AO	Vorlesen, Zustimmung AO	Woher die geschlechtsspezifischen Zuschreibungen stammen, die du vertrittst, um sie zu überwinden, auf jeden Fall. #00:52:57-6#
221	Vorlesen, Beschreibung, Überlegung zum Unterricht, Bewertung, Alternative	Vorlesen, Überlegungen zum Unterricht, Bewertung, Alternative, Bewertung	B5: Welches Feedback möchten Sie Robert zu seinem Umgang mit Maries Einwand beziehungsweise mit Messunsicherheiten geben? Maries Einwand ging ja glaube ich über die Messunsicherheiten und das ist glaube ich eher ein Problem von der Anschlussfähigkeit seiner Unterrichts was vorher schon nicht gut war. Ich glaube, er sollte da mehr eingehen, das ist ihm auf jeden Fall nicht gut gelungen.
222	Vorlesen	Vorlesen	Es ist dir nicht gut gelungen da den Fokus von der experimentellen Bestätigung der Impulserhaltung hin zur Diskussion von leichten (.) verschiebst. (.)
223	Vorlesen, Alternative + Begründung	Vorlesen, Ablehnung AO + Begründung	Da (würd?)- da du den Fokus von der experimentellen Bestätigung- (na?), ich finde gerade, dass er den Fokus da verschieben sollte, weil ansonsten versteht man auch die Bestätigung nicht.
224	Vorlesen, Zustimmung AO, Alternative	Vorlesen, Zustimmung AO, Alternative	Es ist nicht gut, dass du nur so knapp- ich finde genau richtig, er sollte darauf eingehen und vorher schon über die Unsicherheiten reden.
225	Vorlesen, Alternative	Vorlesen, Zustimmung AO, Alternative	(Es nur auf den?) knapp, so unklar, Rundungsfehler gemacht haben, ich finde auch, dass er vorher schon, genau auf die- vorher schon auf die Rundungen eingehen sollte.

226	Vorlesen, Zustimmung AO + Begründung, Alternative	Vorlesen, Zustimmung AO + Begründung, Alternative	Dass du die dass du die Ungleichheit der Messwerte mit Rundungsfehlern begründest, es könnten Rundungsfehler sein bei den Schülern, es könnten aber auch eben Mess- Messunsicherheiten sein, die sollte er erwähnen, deswegen würde ich das ankreuzen.
227	Vorlesen, Zustimmung AO, Alternative + Begründung, Bewertung + Begründung	Vorlesen, Zustimmung AO, Alternative, Überlegung zum Unterricht	Es ist nicht gut, dass du die Ungleichheit der Messwerte mit Rundungsfehlern begründest, da du so den Schüler die Schuld dafür gibst, dass die Ergebnisse nicht exakt gleich sind. Das finde ich genau richtig, er sollte eher fördern, dass die Ergebnisse nicht exakt gleich sein sollten, weil das ja schon ziemlich komisch wäre, (wenn es?) zu genau, aber an sich würde ich hier erstmal gerne sagen, dass es an- von ihm nicht anschlussfähig war, weil er vorher schon nicht über die Rundung einge- erwähnt hat, bei den Massen zum Beispiel. #00:54:26-9#
228	Vorlesen	Vorlesen	B5: Welche der unten stehenden Hinweise zu seinem Umgang mit Maries Einwand beziehungsweise mit Messunsicherheiten möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben?
229	Bewertung, Ablehnung AO	Vorlesen, Ablehnung AO	Deine Reaktion auf den Einwand von Marie ist nicht angemessen, aber da steht angemessen, deswegen kreuze ich das mal nicht an.
230	Vorlesen, Ablehnung AO, Alternative	Vorlesen, Ablehnung AO, Alternative	Du solltest den Schüler- innen angeben, auf wie viele Nachkommastellen sie runden sollten. Ich glaube, er sollte es eher nicht angeben, sondern mit den Schülern besprechen, was da sinnvoll ist.
231	Vorlesen, Ablehnung AO, Alternative	Vorlesen, Ablehnung AO, Alternative	Es wäre besser, den Schüler*innen am Anfang mitzuteilen, dass Messunsicherheiten und Rundungsfehler nicht berücksichtigt werden, (.) das finde ich auch gerade falsch, darum sie sollten berücksichtigt werden und sollten aber diskutiert werden.

232	Vorlesen, Zustimmung AO, Ablehnung AO + Begründung	Vorlesen, Ablehnung AO + Begründung	Es wäre besser, ausführlich- es wäre besser, ausführlich zu begründen, wo die Rundungsfehler entstehen, ja, aber er sollte vor allem sich nicht auf die Rundungsfehler versteifen, deswegen würde ich das auch nicht angeben.
233	Vorlesen, Ablehnung AO, Alternative (2x) + Begründung (2x)	Vorlesen, Zustimmung AO+Begründung, Alternative	Du solltest bereits bei der Vorstellung des Experiments über die vorhandenen Reibung und Energieumwandlung beim Aufprall sprechen, damit den Schü- damit das in dieser Phase klar ist, das glaube ich eben nicht, ich glaube, er sollte das weglassen, weil das vermeid- vernachlässigbar trotzdem immernoch ist und eher diese Reibung und die Energieumwandlung nach hinten erstmal (wegkehren?) vielleicht nachher noch kurz diskutieren, das es auch (unv.) daher kommen könnte, aber sich auf die Messunsicherheiten trotzdem fixieren, weil daher die meisten Unsicherheiten kommen.
234	Vorlesen, Zustimmung AO + Begründung	Vorlesen, Zustimmung AO + Begründung	Dass bei der Durchführung des Versuches Messunsicherheiten berücksichtigt werden müssen, genau, und (damit?) kommt es zu den Abweichungen, ansonsten können die Schüler es auch gar nicht verstehen, wie man das jetzt bestätigt hat.
235	Vorlesen, Überlegung zum Unterricht, Ablehnung AO + Begründung	Vorlesen, Ablehnung AO + Begründung	Du solltest das Thema Messunsicherheiten für dich selbst noch einmal (.) aufbereiten, damit du von solchen Fragen nicht vom Thema abgebracht wirst. Da bin ich mir gar nicht sicher, ob er das vielleicht gar nicht- ob er das kann, also ich glaube schon, dass er das kann, nur gar nicht machen wollte, aber deswegen sollte- würde ich jetzt nur das ankreuzen. #00:56:09-8#
236	Vorlesen	Vorlesen	B5: Welches Feedback möchten Sie Robert zum- zu seinem Umgang mit der Einheit des Impulses geben?

237	Beschreibung, Alternative, Alternative + Begründung, Zustimmung AO + Begründung	Überlegungen zum Unterricht, Alternative+Begründung, Überlegung zum Unterricht, Alternative	Die Einheit wurde ja nicht eingeführt, deswegen sollte er auf jeden Fall darüber sprechen. Weil die Einheit ein wichtiger- genau. (Somit?) kann man auch die Formel nachher verstehen über die Einheit oder (einmal?) gucken, ob sie richtig gerechnet haben, das sollte er mit den Schülern vielleicht auch noch besprechen.
238	Vorlesen, Zustimmung AO, Beschreibung, Sonstiges	Vorlesen, Zustimmung AO+Begründung, Beschreibung	Es ist nicht gut, dass du mit- da dies zu inkosis- genau, er se- er kritisiert seine Schülerinnen, aber er- er selber macht es genauso, er ist somit dann kein Vorbild (unv.) vielleicht sollte man ihm noch sagen, dass er keine Vorbildsfunktion so ist. (..)
239	Vorlesen, Ablehnung AO, Alternative + Begründung	Vorlesen, Ablehnung AO, Alternative	Es ist nicht gut, dass du die Schülerinnen nicht dazu aufforderst, die Einheiten anzugeben, da so unklar ist, ob die Impulse wirklich gleich groß sind. Das finde ich jetzt gerade nicht unbedingt, aber er sollte auf jeden Fall die Einheit des Impuls besprechen, die Schüler sollten sollten dann zumindest auch erkennen, dass überall die gleiche Einheit verwendet wird. #00:57:00-5#
240	Vorlesen	Vorlesen	B5: Welche der unten stehenden Hinweise zu seinem Umgang mit der Einheit des Impulses möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben?
241	Vorlesen, Ablehnung AO	Vorlesen, Ablehnung AO	Es ist in Ordnung, dass du die Einheit des Impulses zunächst vernachlässigst, das finde ich nicht, dass du das erneut so machen kannst.
242	Vorlesen, (Zustimmung AO) + Begründung	Vorlesen, Überlegung zum Unterricht	Du solltest die Einheit des Impulses angeben, die Einheit ist ein wichtiger Bestandteil und die Einheit kann auch (meine?) Rechnungen immer wieder korrigieren und auch äh Vorzei- äh Einheitenfehler vermeiden.

243	Vorlesen, Zustimmung AO + Begründung, Beschreibung, Überlegung zum Unterricht	Vorlesen, Zustimmung AO + Begründung, Überlegung zum Unterricht	Du solltest darauf achten, dass die Schüler beim Vorlesen der Werte auch die Einheit angeben, um in deinem- genau, er muss ein Vorbild bleiben, eben kritisiert er es, diesmal jetzt nicht, wieso.
244	Vorlesen, Ablehnung AO	Vorlesen, Ablehnung AO + Begründung	Du solltest die Einheit des Impulses oben über die Spalte notieren, um die Schei- Schreibarbeit in jeder Zeile zu sparen. Das könnte man machen, das ist je nachdem, wie weit die Schüler sind, wenn die Einheit des Impulses aber gerade neu eingeführt ist, würde ich das eher nicht machen.
245	Vorlesen, Zustimmung AO + Begründung	Vorlesen, Überlegung zum Unterricht	Du solltest die Schüler- innen darauf- dazu auffordern, die Einheit des Impulses zu nennen, damit die Schüler sich an die Einheit- die Einheiten gehören (auch?) immer zum Wert, ansonsten bewegt er sich jetzt beim Impuls und der Berechnung eher zur Mathematik anstatt zur Physik. #00:58:03-1#
246	-	-	[Video Einleitung Reflexion 07, Video Szene Reflexion 07] // [Weiterhin Video Szene Reflexion 07] // [Weiterhin Video Szene Reflexion 07] // [Weiterhin Video Szene Reflexion 07] // [Weiterhin Video Szene Reflexion 07] #00:59:55-6#
247	Überlegung zum Unterricht	Überlegung zum Unterricht	B5: Es ist schon sehr mathematisch, ich hoffe, das stellt er jetzt klar. #00:58:54-4#
248	Alternative	Alternative	B5: Ich glaube, er sollte die Impulserhaltung eher verbal formulieren und nicht als Formel unbedingt. (...)
249	Überlegung zum Unterricht, Alternative	Überlegung zum Unterricht, Alternative	Warum ist der jetzt besser? Ich glaube, er sollte das eher mit den Schülern gemeinsam erarbeiten #00:59:13-3#
250	Alternative + Begründung	Alternative + Begründung	B5: Ich glaube, er sollte die Gleichungen tauschen, weil (halt?) aus der zweiten Gleichung die erste erst folgt oder ne Äquivalenz ist, dann wäre es deutlicher auch nochmal. #00:59:43-2#
251	Bewertung	Überlegung zum Unterricht	B5: Wieder der Fehler wie eben, nur abgeschrieben. #00:59:51-8#

252	Vorlesen	Vorlesen	B5: Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung, Lob und/oder Kritik, erhalten, um etwas für sich beziehungsweise seinen seinen Unterricht zu lernen? (.)
253	Vorlesen	Vorlesen	Interaktion mit den Schülern
254	Vorlesen	Vorlesen	oder Formulierung,
255	Zustimmung AO (2x)	Zustimmung AO (2x)	ich würde beides nehmen.
256	Bewertung	Bewertung	Die Formulierung ist an sich okay äh die Formulierung ist eigentlich okay von ihm,
257	Zustimmung AO	Zustimmung AO	die Interaktion mit den Schülern würde ich eher Feedback zu geben. #01:00:17-7#
258	Vorlesen	Vorlesen	B5: Robert hat zwar schon einige Rückmeldungen erhalten, würde aber trotzdem gern zu den beiden aufgeführten Aspekten noch eine Rückmeldung erhalten. #01:00:24-1#
259	Vorlesen	Vorlesen	B5: Welches Feedback möchten Sie Robert zu seiner Formulierung des Merksatzes zur Impulserhaltung geben?
260	Vorlesen	Vorlesen	Die Formulierung des Merksatzes ist dir gut gelungen, da alle wichtigen Informationen korrekt enthalten sind.
261	Vorlesen, Ablehnung AO	Vorlesen, Ablehnung AO	Es ist gut, dass du im Merksatz auf das au- durchgeführte Experiment verweist, finde ich eher nicht.
262	Vorlesen, Ablehnung AO	Vorlesen, Ablehnung AO	Die Formulierung des Merksatzes ist dir nicht gut gelungen, da sie für zu anspruchsvoll ist, finde ich auch nicht gut.
263	Vorlesen, Ablehnung AO + Begründung, Bewertung	Vorlesen, Zustimmung AO+Begründung, Zustimmung AO	Deinen vorgefertigten Merksatz n-, ne das finde ich gerade, er wertschätzt hier die Schüler nicht und nimmt einfach seinen eigenen, aber seine Formulierung ist (ihm?) trotzdem gut gelungen. #01:00:56-5#
264	Vorlesen	Vorlesen	B5: Welche der unten stehenden Hinweise zu seiner Formulierung des Merksatzes zur Impulserhaltung möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben?

265	Vorlesen, Zustimmung AO	Vorlesen, Zustimmung AO	(Ne?) ist angemessen, er kann sie erneut verwenden, das schon.
266	Vorlesen, Bewertung	Vorlesen, Ablehnung AO	Es wäre besser, Beispiele zu trennen, ich finde, Beispiele sind da ganz gut.
267	Vorlesen, Ablehnung AO + Begründung	Vorlesen, Überlegung zum Unterricht	Zusätzlich eine kurze Version, die kurze Version wäre ja die Formel, das finde ich nicht.
268	Vorlesen, Ablehnung AO	Vorlesen, Ablehnung AO	Es wäre besser, den Merksatz nur in Bezug auf einen Stoßprozess, finde ich auch nicht.
269	Alternative + Begründung	Vorlesen, Überlegungen zum Unterricht	Die Einheit des Impulses sollte im Merksatz vorher (unv.), (weil?/wenn?) man mit Einheiten rechnet, wie man auf die Einheit kommt.
270	Vorlesen, Zustimmung AO	Vorlesen	Du solltest den Merksatz schülernäher gestalten, (.) ja,
271	Alternative, Zustimmung AO	Alternative, Zustimmung AO	ich finde, er sollte den vor allem mit den Schülern (.) als Merksatz (unv.), das ist das letzte, das würde ich dann auf jeden Fall nehmen, dass die Schüler den Merksatz mit erstellen, damit sie ihn auch verstehen. #01:01:39-9#
272	Vorlesen	Vorlesen	B5: Welches Feedback möchten Sie Robert zu seiner Interaktion mit den Schüler- innen geben?
273	Vorlesen, Zustimmung AO + Begründung, Alternative	Vorlesen, Zustimmung AO + Begründung, Alternative	Es ist gut, dass die Schüler- innen ihre Merksätze vorlesen, da du so erkennst, das ist richtig, da sieht man die Schülervorstellungen, das finde ich genau, dass sie das vorlesen sollten, aber er sollte (damit?) noch weiter eingehen.
274	Vorlesen, Zustimmung AO, Beschreibung, Sonstiges	Vorlesen, Zustimmung AO, Überlegung zum Unterricht, Sonstiges	Es ist gut, dass die Schüler einen Merksatz formulieren sollen, das ist auch gut, aber er benutzt den Merksatz nicht, ich hoffe, das kommt gleich noch, dass man das kritisieren kann.
275	Vorlesen, Beschreibung	Vorlesen, Beschreibung, Überlegung zum Unterricht	Es ist nicht gut, dass du den Schüler keine Rückmeldung auf die Merksätze gibst, (.) ja ein bisschen Rückmeldung gibt er ja schon, er sagt ja, die sind gut, aber seiner ist noch besser, er wertschätzt sie nicht.

276	Vorlesen, Zustimmung AO	Vorlesen, Zustimmung AO	Dass du deinen vorgefertigten Merksatz nutzt und die Schü- Merksätze nicht einbeziehst, das finde ich genau, das was ich mir grad gewünscht hatte.
277	Vorlesen	Vorlesen	Es ist nicht gut, dass du deine Formulierung als besser bezeichnest, da du die Formulierungen der Schüler so abbewertest.
278	Vorlesen, Zustimmung AO, Ablehnung AO + Begründung	Vorlesen, Ablehnung AO + Begründung	Es ist nicht gut, dass du nur zwei Schüler- innen ihre Merksätze vorstellen, ja, da muss man auch irgendwann auf die Zeit gucken, von daher finde ich das jetzt schon okay.
279	Vorlesen, Alternative + Begründung, Zustimmung AO + Begründung	Vorlesen, Zustimmung AO + Begründung, Alternative	Es ist nicht gut, dass nur zwei Jungen ihre Merksätz- ich finde auch, da sollte er eher auf Meldungen achten, ob es jetzt zwei Jungen oder zwei Mädchen sind, das ist leider je nachdem, wie sie sich melden, er hat sie aber drangenommen, deswegen würde ich auswählen. #01:02:44-0#
280	Vorlesen	Vorlesen	B5: Welche der unten stehenden Hinweise zu seiner Interaktion mit den Schüler- innen möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben? (.)
281	Vorlesen, Zustimmung AO, Alternative, Beschreibung	Vorlesen, Zustimmung AO, Vorlesen, Beschreibung, Alternative	Der Merksatz sollte eher mit den Schülern erklärt werden und besprochen werden, das (ist?) schon mal da, gut. Er sollte ihn gemeinsam erstellen, die Schüler haben ja einen erstellt schon, er könnte sich darauf mehr beziehen auf jeden Fall. (.)
282	Ablehnung AO + Begründung, Alternative	Ablehnung AO + Begründung, Alternative	Mehr vorstellen finde ich nicht unbedingt nötig, eher mit Schülern in (einer?) Diskussion darüber, nur ihren Satz vorlesen, dann hast du tausend Sätze und kannst daraus aber auch nichts mehr machen. (.)
283	Vorlesen, Zustimmung AO, Sonstiges	Vorlesen, Zustimmung AO, Sonstiges	Du solltest die Schülerinnen sich melden lassen, das finde ich genau, das war ja das, was ich gerade noch erwähnt hatte auch schon.
284	Vorlesen, Ablehnung AO	Vorlesen, Ablehnung AO	Der Umgang mit den Schüler- innen war angemessen, so kannst du dich wieder verhalten. Auf keinen Fall.

285	Vorlesen, Zustimmung AO, Alternative (2x)	Vorlesen, Zustimmung AO, Alternative	Du solltest deinen Umgang mit den Schüler- innen überdenken. Dass-dass du sie und ihre Arbeit nicht lobst, das finde ich auch, er sollte sie mehr loben und auch mehr mit den Schülern arbeiten
286	Zustimmung AO	Zustimmung AO	und die Rollenklischees. #01:03:38-7#
287	Vorlesen	Vorlesen	B5: Hier geht es um den Unterricht insgesamt.
288	Vorlesen	Vorlesen	Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung, Lob und/oder Kritik, erhalten, um etwas für sich beziehungsweise seinen Unterricht zu lernen?
289	Vorlesen	Vorlesen	Unterrichtsziele,
290	Vorlesen	Vorlesen	Ablauf der Unterrichtsstunde,
291	Vorlesen	Vorlesen	Auswahl Schrägstrich Umfang der Stundeninhalte,
292	Vorlesen	Vorlesen	Medieneinsatz,
293	Vorlesen, Zustimmung AO	Vorlesen, Zustimmung AO	Auftreten der Lehrkraft würde ich gerne,
294	Vorlesen, Zustimmung AO	Vorlesen, Zustimmung AO	Umgang mit den Schüler- innen würde ich auch gerne noch was erwähnen zu. #01:04:02-6#
295	Vorlesen, Sonstiges	Vorlesen, Sonstiges	B5: Robert hat schon einige Rückmeldungen erhalten. Besonders wichtig ist ihm jetzt noch eine Rückmeldung zu seinem Umgang mit den Schülerinnen und der fachlichen Korrektheit. Schade, da hätte ich gerne trotzdem zu anderen noch ankreuzen können. #01:04:14-4#
296	Vorlesen	Vorlesen	B5: Welches Feedback möchten Sie Robert zu seiner Interaktion mit den Schüler- innen geben?
297	Vorlesen, Zustimmung AO + Begründung	Vorlesen, Zustimmung AO, Beschreibung, Alternative	Es ist nicht gut, dass du die Beiträge der Schüler- innen kaum berücksichtigst, das finde ich genau, die Merksätze sind von ihm erstellt, er könnte es gerade ändern.
298	Bewertung, Alternative	Vorlesen, Bewertung, Alternative	Lehrerzentriert ist eigentlich ganz gut, aber dann sollte er trotzdem als lehrerzentriert vorne stehen, aber mit den Schülern zusammenarbeiten.
299	Bewertung	Ablehnung AO	Unterschiedliche Behandlung ist nicht gut. (.)

300	Vorlesen, Zustimmung AO, Bewertung, Beschreibung, Testinstrument	Vorlesen, Überlegungen zum Unterricht, Bewertung, Zustimmung AO, Überlegungen zum Unterricht	Es ist gut, dass du die Sicherungsphasen in deinen Unterricht eingebaut hast, (eine?) Sicherungsphase ist gut, aber wie er sie macht ist nicht gut, trotzdem kreuze ich das mal an, dass die Phase vorhanden ist. Da hätte ich jetzt (dann?) gern, dass trotzdem etwas fehlt in dieser Phase, dass die Schüler da mehr mitarbeiten.
301	Vorlesen, Alternative	Vorlesen, Alternative	Es ist gut, dass du verschiedene Sozialformen in den Unterricht einbeziehst, ich denke, er sollte da noch mehr Gruppenarbeit auch noch erwähnen beim nächsten.
302	Vorlesen, Zustimmung AO + Begründung	Vorlesen, Zustimmung AO + Begründung	Es ist nicht gut, dass du im Unterricht nur von wenige Schüler- innen Beiträge anhörst, das finde ich gerade auch, dadurch könnte er auch viel mehr von den Schülervorstellungen vo- erfahren und was die Schüler verstanden haben (und?) was noch nicht.
303	Vorlesen, Bewertung	Vorlesen, Bewertung	Mit einer Frage 'habt ihr noch Fragen', das ist gerade die schlechteste Frage, die man stellen kann.
304	Vorlesen, Ablehnung AO, Alternative + Begründung	Vorlesen, Überlegungen zum Unterricht, Zustimmung AO	Du thematisierst die auftretenden Schülervorstellungen nicht. Das ist nicht gut, da diese so verfestigen. Verfestigen vielleicht nicht unbedingt, aber thematisieren könnte (ihm?) ja helfen, um die auch vor allem ja klar zu stellen. #01:05:28-4#
305	Vorlesen	Vorlesen	B5: Welche der unten stehenden Hinweise zu seiner Interaktion mit den Schüler- innen möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben?
306	Vorlesen, Ablehnung AO	Vorlesen, Ablehnung AO	Der Umgang mit den Schüler- innen ist angemessen, das finde ich nicht.
307	Zustimmung AO	Zustimmung AO	Grundsätzlich stärker einbeziehen auf jeden Fall.
308	Zustimmung AO, Alternative	Zustimmung AO + Begründung	Ausführliches Feedback auch, dann kannst du die Schülerbeiträge auch noch diskutieren, vielleicht mit den Schülern zusammen. (.)

309	Zustimmung AO, Überlegung zum Unterricht, Beschreibung	Zustimmung AO, Überlegung zum Unterricht, Beschreibung	Schülervorstellungen ansehen ist auch immer okay, aber ich glaube, dass er da schon halbwegs drauf vorbereitet war, er hat sie aber ja nicht abgefragt die Schülervorstellungen. (.)
310	Vorlesen, Zustimmung AO + Begründung	Vorlesen, Zustimmung AO + Begründung	Du solltest dir grundsätzlich Gedanken darüber machen, wie, zum Beispiel mit welchen Aufgabenstellungen du das Verständnis der Schülerinnen abprüfen kannst, ja, dann muss er auch nicht 'habt ihr noch Fragen' nutzen, weil das das könnte er da durch Aufgabenstellungen ersetzen.
311	Vorlesen, Zustimmung AO	Vorlesen, Zustimmung AO	Deinen Umgang mit Schü- grundsätzlich überdenken, das finde ich auch
312	Alternative	Zustimmung AO	und Rollenklischees sollte er vermeiden. #01:06:17-1#
313	Vorlesen	Vorlesen	B5: Welches Feedback möchten Sie Robert zur fachlichen Korrektheit der Stunde geben?
314	Bewertung	Zustimmung AO	Ich finde schon, dass er überwiegend fachlich korrekt war. (.)
315	Vorlesen, Ablehnung AO	Vorlesen, Ablehnung AO	Du hast fachliche Fehler, das fand ich jetzt nicht so unbedingt.
316	Vorlesen, Zustimmung AO	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du die Abweichungen, das finde ich auf jeden Fall,
317	Vorlesen, Zustimmung AO + Begründung	Vorlesen, Zustimmung AO + Begründung	das Rundungsfehler, das ist fachlich falsch, das stimmt, das könnten welche sein, aber wahrscheinlicher sind eher Messunsicherheiten.
318	Vorlesen, Zustimmung AO + Begründung	Vorlesen, Zustimmung AO + Begründung	Dass du die Einheit des Impulses nicht besprichst, das finde ich auch, Einheitenrechnung ist eine wichtige Sache und auch für das weitere physikalische Arbeiten der Schüler wichtig, um zum Beispiel Rechnungen zu konzi- zu kontrollieren. (.)
319	Zustimmung AO	Zustimmung AO	Die Schülervorstellungen nicht, das finde ich auch.
320	Vorlesen, Zustimmung AO, Ablehnung AO + Begründung	Vorlesen, Ablehnung AO + Begründung	Es ist gut, dass du (unv.) auf den Einheit, ja, aber dafür (dann?) halt später nicht mehr, deswegen lobe ich dich dafür nicht.

321	Vorlesen, Zustimmung AO + Begründung	Vorlesen, Zustimmung AO, Überlegung zum Unterricht	Es ist nicht gut, dass du bei der Durchführung des Demonstrationsexperimentes sagst, du würdest exakt messen, da du den Schülerinnen das Bild vermittelst, du könntest Messunsicherheiten vermeiden. Auf jeden Fall, deswegen Messwegenumgang ist eh relativ kritisch gewesen. #01:07:21-3#
322	Vorlesen	Vorlesen	B5: Welche der unten stehenden Hinweise zur fachlichen Korrektheit der Stunde möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben?
323	Vorlesen, Zustimmung AO, Beschreibung	Vorlesen, Zustimmung AO + Begründung	Du solltest auch bei der Besprechung der Impulse auf die Einheiten achten, ich glaube, das sollte er auf jeden Fall, eben kritisiert er und jetzt nicht.
324	Vorlesen, Zustimmung AO + Begründung	Vorlesen, Zustimmung AO + Begründung	Die im Unterricht auftreten, aufzugreifen, um den Schülern fachlich korrekt- das auf jeden Fall, an Schülervorstellungen sollte angeknüpft werden oder sie sollten um- umstrukturiert werden.
325	Vorlesen, Ablehnung AO + Begründung, Beschreibung	Vorlesen, Ablehnung AO + Begründung	Du solltest dir die Schülervorstellungen zum Thema vor dem Unterricht angucken, ich glaube schon, dass er damit sich vorher damit beschäftigt hat, das würde ich jetzt nicht unbedingt ankreuzen, dass er sich damit nicht beschäftigt hat, er hat sie nur nicht aufgegriffen.
326	Vorlesen, Zustimmung AO	Vorlesen, Zustimmung AO	Es wäre besser, auch auf Messunsicherheiten- auf jeden Fall. (.)
327	Vorlesen, Überlegung zum Unterricht, Bewertung	Vorlesen, Überlegung zum Unterricht	Du solltest darauf achten, dass deine Merksätze fehlerfrei formuliert sind, ich weiß (nicht mehr?), ob da ein Fehler drin war, sie waren vielleicht eher nicht schülernah formuliert.
328	Vorlesen, Ablehnung AO	Vorlesen, Ablehnung AO	Du solltest stärker auf deine Wortwahl achten, um nicht ausversehen falsche Vorstellungen zu befördern. (.) Das fand ich jetzt auch nicht ganz so schlimm. #01:08:17-3#
329	-	-	[Video Prompt Schulnote] #01:08:23-0#

330	Bewertung (2x) + Begründung (2x), Beschreibung (2x)	Überlegung zum Unterricht, Bewertung	B5: Eine Schulnote, eine Eins fällt schon mal raus, eine Zwei glaube ich auch durch den Umgang mit den Schülern, er bezieht sich nicht wirklich auf die f- Schülervorstellungen und die Messunsicherheiten, aber an an sich hat er (unv.) eine fachliche Korrektheit, deswegen würde ich zwischen einer Drei und einer Vier tendieren und würde ihm durch den Umgang mit den Schülerinnen und auch der Selbsterstellung von Merksätzen eine Vier geben- #01:08:46-1#

O.6 Interview 6: LD6

Segment	Kodierung	Transkript
0	-	[Einverständniserklärung, Demographie] #00:06:29-1#
1	-	Okay, das ist dann die Seite, wo es inhaltlich los geht. Ähm (.) genau ich sage gleich noch einmal kurz, was quasi jetzt passiert und dann bin ich wieder ruhig. Ähm genau, es ist so, dass während der Testdurchführung gleich dir immer wieder Videos begegnen und zu denen werden dann im Anschluss Fragen gestellt und wie eben schon, bitte ich dich, alle Gedanken während der Bearbeitung dieser Fragen beziehungsweise Aufgabenstellung immer laut auszusprechen ähm und dann auch wenn du gerade eine Aufgabe bearbeitest oder einen Text liest, das mit vorzulesen. Ähm dabei geht es nicht darum, dass du versuchst, jemandem, also mir irgendwie was zu erklären, sondern eher diese Vorstellung, äh du wärst allein in einem Raum und würdest zu dir selber sprechen. Äh wenn du, genau, aufhörst, (.) laut zu sprechen, äh dann würde ich mich nach einer kurzen Zeit einmal einschalten mit so einem kurzen- einer kurzen Erinnerung und ansonsten würde ich mich jetzt stumm schalten und bin dann, ja, quasi nicht mehr da. #00:07:33-8#
2	Sonstiges	LD6: Mhm[Ja]. Ähm bei den Videos muss ich aber nichts sagen, (.) nur Texte und so vorlesen. #00:07:40-0#
3	-	I: Genau, also du du darfst, wenn dir gerade was einfällt, natürlich immer was sagen, aber ähm es geht hauptsächlich- genau, also wichtig ist (es?) auf jeden Fall bei den bei diesen Fragen, die im Anschluss kommen. #00:07:50-4#
4	Sonstiges, Vorlesen	LD6: Okay. (12s) Also ich fang jetzt auch hier an, vorzulesen. (..) Okay. Stellen Sie sich folgendes vor: Sie absolvieren zurzeit ein Praktikum an der Schule. Ihr Mitpraktikant Robert hat eine Doppelstunde seines Physikunterrichts aufgezeichnet und bittet Sie, ihm dazu eine Rückmeldung zu geben. Er zeigt Ihnen Ausschnitte seines Unterrichts und stellt Ihnen dazu Fragen. Um diese Fragen zu beantworten können Sie Robert unter anderem darum bitten, Ihnen noch einmal das Unterrichtsvideo zu zeigen. Aber beachten Sie bitte die Zeit, denn Robert hat nur 70 Minuten Zeit. Deswegen sollten die Videos maximal einmal wiederholt werden. (.) Schauen Sie sich also nur in dringenden Fällen die Videos ein weiteres Mal an. Robert ist mit seinem Unterricht nicht zufrieden und bittet Sie um ein Feedback. #00:08:46-4#
5	-	[Video Lerngruppe] #00:09:17-9#

6	eigene Erfahrung, Sonstiges	LD6: Ja, da erinnere ich mich gerade an mein Praktikum in der Gesamtschule, wo das auch so ähnlich war, da hatten wir auch sehr differenzierte Schülerinnen und Schüler, (.) die (.) ja vom Leistungsstand her ziemlich unterschiedlich waren, aber an sich auch ziemlich ruhig. (.) Und irgendwie ist es auch, glaube ich, normal, dass es immer mehr Mädchen als Jungen in der Schule gibt. #00:09:41-2#
7	-	[Video Vorwissen] #00:12:03-0#
8	Überlegung zum Unterricht	LD6: Woran ich gerade denke, ist, warum das zweite Newtonsche Axiom mit der Beschleunigung anfängt und nicht die sonst übliche Formel F gleich m mal a (.) benutzt wird als Merksatz. #00:12:18-6#
9	-	[Video Unterrichtsziele] #00:13:15-9#
10	Beschreibung (5x), Überlegung zum Unterricht (2x)	LD6: Achso, das ist ja die Unterrichts- der Unterrichtsplan. (.) Ähm es gibt verschiedene Phasen, in der Unterrichts- im Unterrichtseinstieg stellt der Lehrer das Thema und den Ablauf der Stunde vor und zwar starten sie mit einem Quiz zu Newtonschen Axiomen, (.) dabei erklärt der Lehrer zunächst den Ablauf der Quizrunde und die Schülerinnen und Schüler beantworten die Quizfragen in Partnerarbeit und (das?) alles geschieht auf dem Beamer. Und anschließend moderiert der Lehrer die Besprechung der Quizfragen (.) und die Schülerinnen und Schüler antworten die nochmal beantworten die nochmal (.) oder beziehungsweise geben wahrscheinlich die Antworten vor und das passiert hier im Klassengespräch (äh?) im Plenumsgespräch. Nach dem Unterrichtseinstieg gibt es die Vertiefungsphase, dabei zeigt der Lehrer den Schülerinnen und Schülern einen Comic mit dem Titel Junge und Mädchen sprechen über das dritte Newtonsche Axiom. (.) Ähm dabei diskutieren die Schülerinnen und Schüler die Situation, die im Comic gezeigt (wird) und erklären mit Hilfe des dritten Axioms, welche Aussage richtig und welche (.) welche Aussage richtig ist und welche nicht, auch in dem Klassengespräch und der Comic wird auch (unv.) Beamer gezeigt. (...) Ähm (..) dann gibt es die Sicherungs- und Erarbeitungsphase. In der Erar- zum Impuls, der soll ja in der Stunde eingeführt werden. (.) die der Lehrer führt in die Thematik Impuls mit der Fragestellung ein, warum konnte die Titanic nicht bremsen (.) (im?) Lehrer-Schüler-Gespräch auch auf dem Beamer. Ähm und dann gibt es ein Experiment in der Erarbeitungsphase dazu, ähm wo die Schülerinnen und Schüler das Experiment anscheinend selbst durchführen sollen. (.) Ähm und zwar ja mit zwei Tennisbällen, ähm wobei einer mit Sand gefüllt ist und der andere nicht. (.)

		Dann in der Sicherungsphase sichert der Lehrer die Erkenntnisse aus dem Experiment mit einem Merksatz und die Schülerinnen und Schüler schreiben diesen Merksatz zur Sicherung ab auch vom Beamer, ziemlich viel Beamer. Ähm dann gibt es nach der Sicherungsphase wieder eine Erarbeitungsphase, diesmal nachdem der Impuls eingeführt wurde zur Impulserhaltung, (.) ähm dabei wir wieder ein Experiment durchgeführt und zwar (ein?) Luftkissenfahrbahn. (.) Ähm da präsentiert der Lehrer den Versuchsaufbau und erläutert die Versuchsdurchführung, (.) ähm ja anhand eines Plakates (.) und einer Messwerttabelle, oder beziehungsweise wahrscheinlich ist die Messwerttabelle auf dem Plakat zu sehen. (.) Dann führt der Lehrer dieses Mal den Versuch durch, also nicht die Schülerinnen und Schüler, die assistieren nur bei einigen Aufgaben, weil die Luftkissenfahrbahn anscheinend ein Lehrerexperiment ist. (.) Ähm (.) genau dann werden die ähm Beobachtungen (..) gesammelt und dabei leitet der Lehrer das Gespräch und die Schülerinnen und Schüler sollen ihre Beobachtungen, die sie während des Experiments gemacht haben, nennen. Das passiert auch in einem Plenumsgespräch. (.) Ähm genau also nach dem Versuch soll dann es ein bisschen mathematischer werden, indem der Lehrer die Impulserhaltung aus den Newtonschen Axiomen herleitetm, auch wieder am Beamer als Lehrervortrag. (.) Ähm (.) die Aufgabe ist danach, (.) dass die Schüler- also genau jetzt haben ja die Schülerinnen und Schüler die Impulserhaltung kennengelernt (.) und jetzt sollen die Schülerinnen und Schüler die Impulserhaltung mit denn Messwerten aus dem Experiment überprüfen, also (ob?) auf der Luftkissenfahrbahn auch die Impulserhaltung gilt (.) in Partnerarbeit. Dann wird diese Aufgabe besprochen, wobei der Lehrer das Gespräch leitet und die Schülerinnen und Schüler ihre Ergebnisse jeweils vorlesen, ähm auch im Plenumsgespräch (..) und auch am Plakat, wie genau (.) weiß ich gerade nicht. Vielleicht könnten die ja ihre ihre Messwerte dann einfach in die Tabelle auch noch eintragen, könnte ich könnte ich mir vorstellen. (.) Und dann abschließend gibt es noch eine Sicherungsphase, wo der Lehrer die Erkenntnisse aus dem Experiment wieder in einem Merksatz sichert und die Schülerinnen und Schüler diesen wieder ab- schreiben und auch wieder (im?) Beamer. (.) Ziemlich viel Beamer und gar keine Tafel. #00:17:52-6#

11	Vorlesen, Überlegung zum Unterricht	LD6: Mhm (.) Ich habe die Stunde mit der Vorstellung des Ablaufs begonnen. Als nächstes habe ich ein Quiz zur Wiederholung der Newtonschen Axiome angeleitet, welches die Schülerinnen in Partnerarbeit durchgeführt haben. Die Durchführung des Quiz selbst habe ich zwar nicht gefilmt, dafür aber die Besprechung des Quiz mit den Schülerinnen, die ich dir jetzt zeige. (..) Okay mhm (.) in der St- Einheit befinden wir uns gerade. #00:18:24-0#
12	-	[Video Einleitung Reflexion 01] #00:18:49-8#
13	Bewertung	LD6: Ja, das äh (.) ist immer ein großes Problem. (Ich meine?) Merksätze sind ja gut, aber wenn die Schülerinnen und Schüler es nicht verstehen, dann bringen die auch nicht so viel. #00:19:02-6#
14	-	[Video Szene Reflexion 01] #00:21:07-3#
15	eigene Erfahrung	LD6: Kam mir gerade irgendwie ein bisschen wie in der Uni vor mit dem Beamer und dem Vortrag, das habe ich in der Schule so noch nie gesehen. #00:21:17-2#
16	Vorlesen	LD6: Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung, Lob und/oder Kritik, erhalten, um etwas für sich beziehungsweise- beziehungsweise seinen Unterricht zu lernen? (...)
17	Sonstiges	Achso ich soll jetzt die Punkte auswählen, die Lob und oder Kritik- okay.
18	Vorlesen	Ähm bitte wählen Sie die zutreffen Antworten aus.
19	Vorlesen	Formulierung, Wortlaut des Merksatzes zum ersten Axiom,
20	Vorlesen	Formulierung, Wortlaut des Merksatzes zum zweiten Axiom,
21	Vorlesen	Formulierung, Wortlaut des Merksatzes zum dritten Axiom,
22	Vorlesen	Ergebnissicherung des Quiz,
23	Vorlesen	Gesprächsführung.
24	Vorlesen	Hier können Sie- Hier können Sie zu bestimmten Fragen Antworten von Robert bekommen.
25	Vorlesen	Kann ich die Informationen zur Lerngruppe noch einmal bekommen? (.)
26	Vorlesen	Kannst du mir noch einmal sagen, welches Vorwissen die Schülerinnen und Schüler haben?
27	Vorlesen	Kannst du mir die Unterrichtsziele noch einmal nennen?
28	Vorlesen	Kannst du mir noch einmal das Stundenraster zeigen?
29	Vorlesen	Kann ich den Unterrichtsausschnitt noch einmal sehen?
30	Vorlesen	Kannst du mir die PowerPoint-Folien nochmal zeigen? (..)

31	Verweis auf frühere Stelle im Test, Bewertung + Begründung, eigene Erfahrung	Ähm ja wie ich gerade schon gesagt habe, auf jeden Fall Kritik für die Formulierung des Merksatzes zum zweiten Axiom, weil irgendwie (.) keine Ahnung (..) a gleich F durch m ist irgendwie eine komische (.) Form, das äh aufzuschreiben, ich kenne es (nur?) aus der Schule irgendwie immer mit F gleich m mal a. (.)
32	Bewertung, Überlegung zum Unterricht (2x)	Ähm (..) Formulierung zum ersten Axiom fand ich eigentlich ganz gut, (wie?) war das nochmal? Trägheitsgesetz, (.) ja ist ja generell so ein bisschen schwierig, wenn man (nur?) die Gesetze auswendig kann, aber ich glaube, das finde ich eigentlich ganz gut, es war halt okay, würde ich jetzt auch nicht übertrieben loben. (.)
33	Bewertung	Ähm (.) drittes Axiom mhm fand ich auch okay. (..)
34	eigene Erfahrung, Bewertung, Bewertung + Begründung, Alternative (2x)	Die Ergebnissicherung des Quiz würde ich (...), glaube ich, kritisieren, einfach weil (.) ich w- weiß nicht, irgendwie finde ich das so super komisch mit diesem Beamer im Unterricht, das kenne ich irgendwie nur aus der Universität und irgendwie im Unterricht so einen vorbereiteten Power-Point-Vortrag finde ich irgendwie ein bisschen komisch, da würde ich, glaube ich, lieber die Tafel benutzen, oder den Overhead-Projektor oder halt die Schülerinnen und Schüler das direkt halt irgendwie aufschreiben lassen, (.) (irgendwie?) das fand ich irgendwie ein bisschen komisch. (.)
35	Sonstiges	Ähm (...) mhm (..) gerade denke ich gar nichts, ich weiß nicht, was will ich denn loben?
36	Bewertung + Begründung, Beschriebung, Sonstiges	Ähm (4s) Gesprächsführung, ich weiß nicht, irgendwie war das jetzt nicht so, wo ich mir denke, okay, die Schülerinnen und Schüler waren super motiviert, den Unterricht zu verfolgen, ich weiß nicht, ob ich irgendwas überhaupt loben würde. (5s) Vielleicht (5s) ich hänge so ein bisschen an der Gesprächsführung, aber irgendwie war die jetzt halt auch nicht irgendwie besonders gut, (halt?) einfach Schülerinnen und Schüler drangenommen. Ich glaube, ich würde erstmal (.) vielleicht ein bisschen hart, einfach nichts loben. #00:24:38-9#
37	Vorlesen, Sonstiges	LD6: Robert hat schon einige Rückmeldungen erhalten. Besonders wichtig ist ihm jetzt noch eine Rückmeldung zu seiner Formulierung des Merksatzes zum dritten Newtonschen Axiom. (4s) Okay, (dazu?) muss ich mir den Merksatz gleich nochmal anhören. #00:24:54-7#
38	Vorlesen (3x)	LD6: Mhm welches Feedback möchten Sie Robert zu seiner Formulierung des Merksatzes zum dritten Axiom geben? Wählen Sie die Antwortoptionen aus, die Ihrer Einschätzung am nächsten kommen. Bitte wählen Sie die zutreffenden Antworten aus.
39	Vorlesen	Deine Formulierung des dritten Axioms ist gut, da alles wichtige enthalten ist.
40	Vorlesen, Ablehnung AO	Deine Formulierung des 3. Axioms ist fachlich falsch. (...) Mhm ich glaube nicht.

41	Vorlesen	Die Formulierung des 3. Axioms ist für die Schülerinnen nicht angemessen, da sie zu komplex ist.
42	Vorlesen	Die Formulierung des dritten Axioms ist nicht gut, da die Vermischung von Beispiel und Merksatz die Schülerinnen verwirren kann. (.)
43	Vorlesen, Zustimmung AO, Sonstiges	Es ist nicht gut, dass nur die Kurzformel Kraft gleich Gegenkraft, aber nicht actio gleich reactio im Merksatz vorkommt. Diese Abkürzung ist nämlich etwas, was die Schülerinnen vielleicht kennen. Ja, das stimmt, das (unv.) gleich mal auswählen. (.)
44	Vorlesen	Es ist gut, dass du im Merksatz deutlich machst, dass die Kräfte an verschiedenen Körpern angreifen- Es ist gut, dass du im Merksatz deutlich machst, dass die Kräfte an verschiedenen Körpern angreifen, da du so einer Verwechslung mit dem Kräftegleichgewicht vorbeugst. (.)
45	Vorlesen (2x), Zustimmung AO, Sonstiges, Bewertung, Alternative	Es ist nicht gut, dass die Kurzformel Kraft gleich Gegenkraft im Merksatz vorkommt, da Schülerinnen diese oft mit dem Kräftegleichgewicht assoziieren. (.) Ja, das stimmt, (...) aber ich glaube, (irgendwie?) (unv.) noch einmal kurz anhören, was er nochmal genau gesagt hat. Mhm (7s) (oh?) kann man das überhaupt? (5s) War im Unterrichtsausschnitt die (.) wo ich mir die dritte das dritte Axiom nochmal angucken kann oder (wo?) kann ich mir das dritte Axiom, wie er das formuliert hatte, nochmal angucken? (4s) Mhm (6s) (unv.) hier (LD6 öffnet Hilfe 'Kannst du mir die PowerPoint-Folien nochmal zeigen?'), drittes New- drittes Newtonsches Axiom, Wechselwirkungsprinzip. Wenn ein Körper A eine Kraft F A B, zum Beispiel F A B gleich zehn Newton, auch Körper B ausübt, dann übt Körper B anschließend auch immer eine Gegenkraft F B A auf den Körper A aus. Die Kräfte greifen an verschiedenen Körpern an, sind genau gleich groß, aber einander entgegengerichtet, das heißt Kraft gleich Gegenkraft. (..) Okay (.) mhm (..) Kann ich (das?) zumachen? (8s) Ähm eigentlich fand ich die Formulierung echt gar nicht so schlecht, okay actio gleich reactio würde ich auf jeden Fall mit reinnehmen.
46	Bewertung + Begründung	und mit der Kraft gleich Gegenkraft ist ja wirklich die Gefahr groß, dass sie das mit- es ist generell die Krafähm die (..) die Gefahr ist auf jeden Fall gegeben, dass sie das eh mit dem Kraftbegriff assoziieren, aber ich glaube, das ist nicht so eine gute Formulierung. (..)
47	Vorlesen, (Ablehnung AO) + Begründung	Deine Formulierung des dritten Axioms ist gut, da alles wichtige enthalten ist. (..) Ja, actio gleich reactio fehlt.
48	Vorlesen, Ablehnung AO + Begründung	(unv.) ist fachlich falsch, das (würde?) ich nicht sagen, ist eigentlich richtig, zumindest so, wie ich das sehe. (.)

49	Vorlesen, Ablehnung AO	Die Formulierung des dritten (Merksatzes?) ist für Schülerinnen nicht angemessen, da sie zu komplex ist. (.) Nö, zu komplex ist die eigentlich auch nicht.
50	Vorlesen, Ablehnung AO + Begründung	Die Formulierung des dritten Axioms ist nicht gut, da die Vermischung von Beispiel und Merksatz die Schülerinnen verwirren kann. Eigentlich (finde?) ich das gar nicht so schlecht mit dem Beispiel, weil man da direkt sieht auch das mit dem gegen- (..) gegengesetzten Vorzeichens (unv.) gut.
51	Vorlesen, Zustimmung AO, Bewertung, Alternative (2x)	Es ist gut, dass du im Merksatz deutlich machst, dass die Kräfte an verschiedenen Körpern angreifen, da du so einer Verwechslung mit dem Kräftegleichgewicht vorbeugst. (..) Ja, stimmt, (..) mhm (.) würde ich, glaube ich, auch, ja, finde ich auch gut. Eigentlich ist (das?) gut, aber (man?) muss ja (unv.) (wirklich?) Kraft gleich Gegenkraft sollte man vielleicht vermeiden und dann lieber actio gleich reactio nehmen. #00:29:06-6#
52	Vorlesen (3x)	LD6: Welche der unten stehenden Hinweise zu seiner Formulierung des Merksatzes zum dritten Axiom möchten Sie Robert für eine erneute Durchführung seines Unterrichts- der Unterrichtsstunde geben? Wählen Sie die Antwortoption aus, die Ihrer Einschätzung am nächsten kommen. (..) Bitte wählen Sie die zutreffenden Antworten aus.
53	Vorlesen	Deine Formulierung des (unv.) ist gut so, du kannst sie wieder so verwenden.
54	Vorlesen	Es wäre besser, zuerst das Axiom ohne Beispiel zu notieren und anschließend ein Beispiel, um den Merksatz zu vereinfachen. (.)
55	Vorlesen	Es wäre besser, das Beispiel im Merksatz ausführlicher Darzustellen, damit die Schülerinnen die Anwendung vom dritten Axiom verstehen können.
56	Vorlesen, Zustimmung AO, Überlegung zum Unterricht, Sonstiges	Es wäre sinnvoll, das Axiom um eine Skizze zu ergänzen, damit es anschaulicher wird. (.) Das habe ich mir generell gedacht, hat er überhaupt irgendwie schonmal (..) irgendwie Beispiele gemacht auch mit dem Trägheitsgesetz, (.) das ist irgendwie, (unv.) dass die Schülerinnen und Schüler das (unv.) vorstellen können (unv.) (zu?) theoretisch (wird?). (.) Ähm (..) Skizze oder (unv.) Versuch (unv.).
57	Vorlesen, Zustimmung AO	Es wäre besser, im Merksatz zu verdeutlichen, dass Kraft und Gegenkraft gleichzeitig wirken. (.) Deine Formulierung anschließend übt Körper B eine Gegenkraft auf Körper A aus ist fachlich falsch. (.) Ja, das stimmt. (.)
58	Vorlesen	Du solltest im Merksatz über den Kräften Vektorpfeile notieren, damit die Wirkrichtung der Kräfte deutlich wird.

59	Vorlesen, Überlegung zum Unterricht, Ablehnung AO, Verweis auf frühere Stelle im Test, Sonstiges, Bewertung	Du solltest im Merksatz verdeutlichen, dass Kraft und Gegenkraft in genau entgegengesetzte Richtungen wirken, (.) damit die Schülerinnen die Bedeutung des Vorzeichens im Beispiel verstehen. Hat er das nicht gemacht? Mhm also wieder so verwenden würde ich auf jeden Fall nicht sagen, (unv.) habe ich ja gerade schon was. Aber jetzt, wo hier einige Verbesserungsvorschläge aufgelistet werden, (.) merke ich, dass es es vielleicht doch noch nicht ganz so gut ist. (..)
60	Vorlesen, Ablehnung AO, Bewertung, Alternative	Zuerst das Axiom und anschließend ein Beispiel, um den Merksatz zu vereinfachen, muss ja (jetzt?) nicht unbedingt sein, klar kann man das auch so machen, aber ich finde es jetzt auch nicht schlimm, wenn (er?) direkt hier drin ist. (.)
61	Vorlesen, Zustimmung AO	Es wäre besser, das Beispiel im Merksatz ausführlicher darzustellen, damit (unv.) die Schülerinnen die Anwendung vom dritten Axiom verstehen können. (.) (Ja?).
62	Vorlesen, Zustimmung AO, Alternative + Begründung	Es wäre besser, (unv.) das Axiom eine Skizze zu ergänzen, damit es anschaulicher wird, (..) ja, Skizze ist (da wahrscheinlich?) besser als gar nichts. (.) Vielleicht kann man dann ja auch- ich hab das früher in der Schule mit dem (.) mit einem Ball und einer Wand gemacht, dass halt nicht nur der Ball die Kraft auf die Wand ausübt, sondern auch die Wand auf den Ball, könnte man jetzt auch (irgendwie da vorstellen) irendjemand, ja. (..) Das sieht man ja nicht, aber, weiß ich nicht, Experimente wirken auch ganz gut.
63	Vorlesen, Zustimmung AO, Überlegung zum Unterricht	Es wäre besser, im Merksatz zu verdeutlichen, dass Kraft und Gegenkraft gleichzeitig wirken, Formulieung (unv.) das auf jeden Fall, ist mir gar nicht aufgefallen. (.)
64	Vorlesen, Zustimmung AO + Begründung	Du solltest im Merksatz über den Kräften Vektorpfeile notieren, damit die Wirkrichtung der Kräfte deutlich wird. Mit (Merk-?) Vektorpfeilen ist auf jeden Fall fachlich richtiger, (also?) stimmt.
65	Vorlesen, Überlegung zum Unterricht	Du solltest im Merksatz verdeutlichen, dass Kraft und Gegenkraft genau in genau entgegengesetzte Richtungen wirken, damit die Schülerinnen die Bedeutung des Vorzeichens im Beispiel verstehen. (.) Hat er das nicht gemacht? (4s, LD6 öffnet Hilfe 'Kannst du mir die PowerPoint-Folien nochmal zeigen?')

66	Vorlesen, Beschreibung, Sonstiges, Zustimmung AO + Begründung, Alternative	(unv.) ausübt, dann übt Körper B anschließend auch immer eine Gegenkraft auf Körper A (4s, LD6 flüstert/bewegt die Lippen, liest vmtl. den Merksatz zuende) einander entgegengerichtet, hat er doch geschrieben. (...) Mhm (4s) ja jetzt noch mit dem Beispiel, ob man das in dem Axiom machen kann oder lieber (..) danach, mhm ich glaube, danach ist vielleicht doch besser, dann wird (es?) auf jeden Fall nicht vermischt und dann vielleicht auch noch ausführlicher, (.) kann man es gleich komplett (unv.). #00:32:54-9#
67	Überlegung zum Unterricht	LD6: Ah neue Phase. #00:33:02-2#
68	-	[Video Einleitung Reflexion 02] #00:33:27-0#
69	Überlegung zum Unterricht	LD6: Ah jetzt verstehe ich das auch, er zeigt einen Comic und dann (.) sprechen (.) seine Schülerinnen und Schüler drüber. Ah, okay. (.) (Ja?). #00:33:39-2#
70	-	[Video Szene Reflexion 02] #00:35:22-5#
71	Sonstiges	LD6: Gut, (unv.) ich bin mir auch gerade nichr sicher. #00:35:28-8#
72	Vorlesen (2x)	LD6: Ähm zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung, Lob und/oder Kritik, erhalten, um etwas für sich beziehungsweise seinen Unterricht zu lernen? (.) Bitte wählen Sie die zutreffenden Antworten aus.
73	Vorlesen	Gesprächsführung,
74	Vorlesen	Einsatz des Comics,
75	Vorlesen	Schülerfehlvorstellungen. (..)
76	Vorlesen	(unv.) jeden Fall (unv.) Lob Lob oder Rückmeldung äh Lob und oder Kritik. (..)
77	Überlegung zum Unterricht, Beschreibung, Alternative (2x), Bewertung	Mhm ich weiß irgendwie nicht, was man bei der Gesprächsführung loben soll, es ist halt einfach ein normales Gespräch, (.) irgendwie weiß ich nicht, man könnte halt mit den Schülerfehlvorstellungen vielleicht noch ein bisschen (..) irgendwie einfach direkt entgegenwirken oder halt eventuell sogar gar nicht erst so hochschaukeln lassen, sondern einfach direkt(..) Also Lob auf jeden Fall nicht, eher Kritik. (.) darauf eingehen.
78	Bewertung + Begründung (2x)	Einsatz des Comics, ja eigentlich finde ich das gar nicht so schlecht, Einsatz des Comics war (eigentlich?) gut, (..) weil die können jetzt ja eigentlich alles, (..) beziehungsweise haben alles kennengelernt und ja (darum?) sollen die einfach jetzt (halt?) ihr Wissen ein bisschen vertiefen, eigentlich nicht so schlecht. #00:36:42-2#
79	Vorlesen	LD6: Robert hat schon einige Rückmeldungen erhalten. Besonders wichtig ist ihm jetzt noch eine Rückmeldung zu den aufgetretenen Schülervorstellungen im Unterrichtsausschnitt. #00:36:52-0#

80	Vorlesen (3x)	LD6: Welche der folgenden Schülerfehlvorstellungen sind während des Unterrichtsausschnitts sicht- Unterrichtsausschnitts sichtbar geworden? Wählen Sie die entsprechenden Antwortoptionen aus. Bitte wählen Sie die zutreffenden Antworten aus.
81	Vorlesen, Ablehnung AO	Es sind keine Schül-vo- Schülerfehlvorstellungen sichtbar geworden. Das ist schon falsch.
82	Vorlesen, Ablehnung AO	Schülerinnen denken, dass Kräfte nur dann auftreten, wenn sich zwei Körper berühren. (..) Ne, würde ich auch nicht sagen.
83	Vorlesen	Die Schülerinnen schreiben dem Apfel eine Kraft als Eigenschaft zu. (.)
84	Vorlesen, Zustimmung AO	Schülerinnen denken, dass nur aktive, das heißt sich bewegende, Körper eine Kraft ausüben können, passive leisten nur Widerstand. Ja, das auf jeden Fall. (.)
85	Vorlesen	Die Schülerinnen denken, dass einer Kraft immer eine Beschleunigung zugrunde liegen muss und schließen daher eine Kraftwirkung bei ruhenden Gegenständen aus. (..)
86	Vorlesen, Zustimmung AO	Die Schülerinnen verwechseln Kraft und Gegenkraft mit den wirkenden Kräften, die einen Körper im Kräftegleichgewicht halten. Mhm[Ja] ja. (..)
87	Vorlesen (2x), Sonstiges*, Zustimmung AO	Schülerinnen sehen nicht, dass der Apfel über den Baum mit der Erde verbunden ist- Schülerinnen sehen nicht, dass der Apfel über den Baum mit der Erde verbunden ist und der Baum durch das Festhalten an der Erde eine Gegenkraft zum am Baum ziehenden Apfel ausübt. Was? Schülerinnen sehen nicht, dass der Apfel über den Baum mit der Erde verbunden ist und der Baum durch das Festhalten (..) durch das Festhalten an der Erde (.) eine Gegenkraft zum am Baum ziehenden Apfel ausübt. (..) Mhm[Ja] ja, stimmt, ich glaube, das äh ja, ich glaube auch. (.)
88	Ablehnung AO	Also keine Schülervorstellung ist falsch,
89	Vorlesen, Ablehnung AO	denken, dass Kräfte nur dann auftreten, wenn sich zwei Körper berühren, würde ich jetzt auch nicht sagen. (.)
90	Vorlesen, Ablehnung AO + Begründung	Schülerinnen schreiben dem Apfel eine Kraft als Eigenschaft zu. (.) Eigentlich sagt ja niemand von denen, der Apfel hat Kraft, sondern die gehen ja schon eher von der Gewichtskraft aus, ich glaube, das (ist?) trifft auch nicht zu. (.)
91	Vorlesen, Sonstiges, Sonstiges	Schülerinnen denken, dass einer Kraft immer eine Beschleunigung zugrunde liegen muss und schließen daher eine Kraftwirkung bei ruhenden Gegenständen aus. (...) (Ich?) glaube, das ist eher durch die Schülervorstellung davor (.) schon abgedeckt. (..) Wobei ich (mir?) aber jetzt (nicht?) ganz sicher bin. #00:39:08-7#

92	Vorlesen (3x)	LD6: Welche der unten stehenden Hinweise zum Umgang mit Schülervo- Schülerfehlvorstellungen möchte- möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben? Wählen Sie die Antwortoptionen aus, die Ihrer Einschätzung am nächsten kommen. Bitte wählen Sie die zutreffenden Antworten aus. (5s)
93	Vorlesen, Ablehnung AO + Begründung	Dein Umgang mit den Fehlvo- Schülerfehlvorstellungen ist angemessen, das kannst du wieder so machen. (.) Ja eigentlich geht er ja gar nicht drauf ein, das ist schonmal blöd.
94	Vorlesen	Du solltest den Prozess der Kraftwirkung thematisieren, um die Schülerfehlvorstellungen aus dem Weg zu räumen. (..)
95	Vorlesen	Prozess der Kraftwirkung.
96	Vorlesen	Du solltest den Unterschied zwischen Kraft und Energie thematisieren, um zu verdeutlichen, warum die Schülerfehlvorstellungen falsch sind. (.)
97	Vorlesen	Du solltest erklären, verdeutlichen, dass Kräfte auch dann wirken können, wenn man nichts aktiv tut.
98	Vorlesen, Zustimmung AO	Du solltest erklären, verdeutlichen, dass Kräfte auch dann wirken können, wenn man nichts aktiv tut. Dazu kannst du das Rollbrettexperiment verwenden: Zwei Schülerinnen stehen auf je einem Rollbrett und halten je ein Ende einer Schnur. Egal ob beide Schülerinnen ziehen oder nur einer, treffen sich die Rollbretter in der (Mitte?). Das finde ich sehr gut (.) (auch ein?) Experiment. (.)
99	Vorlesen, Sonstiges, Zustimmung AO	Du solltest mit den Schülerinnen die verwendeten und das (unv.) Du solltest mit den Schülerinnen die verwendeten Begriffe und den Unterschied zwischen dem Kräftegleichgewicht und dem dritten Newtonschen Axiom klären, (.) (unv.) ist das doch nicht, damit ihnen klar wird, warum ihre Vorstellungen falsch sind. (Ja?), die Thematisierung zwischen dem Kräftegleichgewicht und dem dritten Newtonschen Axiom ist auf jeden Fall muss man auf jeden Fall machen. (.)
100	Vorlesen, Ablehnung AO + Begründung	Du solltest die Schülerinnen das Wechselwirkungsprinzip ausführlich mündlich oder schriftlich formulieren lassen, damit du siehst, was die Schülerinnen verstanden haben beziehungsweise wo Fehlvorstellungen bestehen. (..) Ja eigentlich kann man ja die Fehlvorstellungen schon an den mündlichen Aussagen (unv.) da (gut tun?) (.) (Erkenntnisse?) würde ich, glaube ich, nicht sagen. #00:41:03-3#

101	Vorlesen, Sonstiges (2x)	LD6: Bei der Besprechung des Comics hatte ich das Gefühl, dass die Schülerinnen einige Schülervorstellungen haben. Deswegen habe ich die Schülerinnen gebeten, Stellung zu einer Seite des Comics zu beziehen, also Argumente entweder für die Seite des Mädchens oder des Jungens zu sammeln. (Okay?) Diese Argumente haben wir dann besprochen und Anna hat fachlich korrekt erklärt, warum die Position des Mädchens die richtige ist. (5s) Anschließend hatten die Schülerinnen die Möglichkeit, offen gebliebene Fragen zum Comic zu stellen. (.) Da die Schüler*innen hier keine Fragen hatten, hat- habe ich den Unterricht fortgeführt und zum Thema Impuls übergeleitet. (..) Okay #00:41:45-8#
102	-	[Video Einleitung Reflexion 03] #00:42:00-7#
103	Beschreibung	LD6: Nur, weil eine Schülerin (unv.) eine Schülerin das richtig hatte, geht er einfach weiter. (Okay?), weiß ja nicht, ob es alle verstanden haben. #00:42:10-6#
104	-	[Video Szene Reflexion 03] #00:43:55-9#
105	Vorlesen (2x)	LD6: Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung, Lob und/oder Kritik, erhalten, um etwas für sich beziehungsweise seinen Unterricht zu lernen? Zutreffende Antworten auswählen.
106	Vorlesen, Bewertung + Begründung, Zustimmung AO	Einsatz des Beispiels Titanic, das finde ich richtig gut, würde ich, glaube ich, auch (mit?) nehmen, weil einfach das super (äh?) die meisten Schülerinnen und Schüler kennen das, würde ich sagen, (.) finde ich richtig gut.
107	Alternative, Bewertung + Begründung	Aber ich hätte es, glaube ich, ein bisschen anders ähm eingebaut mit der Gesprächsführung, die würde ich ein bisschen kritisieren in dem Fall, weil dadurch, dass es eben so ein alltagsnahes Beispiel ist, könnten ja die Schülerinnen und Schüler ja auch einfach erklären, weil der Lewenn der Lehrer einfach gefragt hätte, ja was genau (ist?) denn mit der Titanic passiert, hätten die Schülerinnen und Schüler das wahrscheinlich auch sagen können, hätte man (einfach die?) noch ein bisschen mehr einziehen- einbeziehen können, deshalb (würde?) ich das ein bisschen kritisieren.
108	Vorlesen	Fachliche Korrektheit,
109	Vorlesen	Fachsprache,
110	Vorlesen, Sonstiges, Beschreibung	Übergang der Unterrichtsphasen. (..) Zum Übergang der Unterrichtsphasen kann ich doch eigentlich gar nichts sagen, oder? (...) Ja, Übergang der Unterrichtsphasen, wie ich gerade schon gesagt hat, irgendwie nur weil die (eine?) die (unv.) Newtonsche Axiom richtig erklären konnte, heißt ja nicht, dass alle das verstanden haben, (.) aber ich bin mir gerade nicht sicher.
111	Überlegung zum Unterricht, Beschreibung	Fachsprache ist mir jetzt gar nicht so aufgefallen. (...) Der Lehrer nutzt eigentlich gar keine Fachsprache, (.) (genauso?) (unv.) Schüler Geschwindigkeit, Schwung, (...)

112	Sonstiges*, Beschreibung	fachliche Korrektheit, (5s) (oh?) da bin ich (auch?) irgendwie unsicher. (...) Weiß nicht, ob das jetzt auf den Lehrer bezogen ist, soviel (Fachliches?) hat er jetzt ja noch nicht gesagt, sondern nur eigentlich nur das Beispiel (.) mhm (unv.) (wissen die was?) (unv.). #00:45:53-2#
113	Vorlesen	LD6: Robert hat schon einige Rückmeldungen erhalten. Besonders wichtig ist ihm jetzt noch eine Rückmeldung zur Nutzung der Fachsprache und zur Gestaltung des Phasenwechsels im Unterrichtsausschnitt. #00:46:03-4#
114	Vorlesen	LD6: Welches Feedback möchten Sie Robert zur Nutzung der Fachsprache geben?
115	Vorlesen, Zustimmung AO	Es ist nicht gut, dass die Schülerinnen den Begriff Schwung nutzen, da die Bedeutung des Begriffs noch nicht geklärt ist und so unklar ist, was genau die Schülerinnen beschreiben. Ja, über Schwung bin ich tatsächlich auch ein bisschen gestolpert. (..)
116	Vorlesen, Zustimmung AO	Es ist gut, dass die Schülerinnen zunächst ihre Alltagssprache verwenden können (.) um zu erklären, warum es zur Kollision gekommen ist. So sinkt die Hemmschwelle für die Schülerinnen, ihre Ideen zu äußern. (Ja?) stimmt auch. (.)
117	Vorlesen, Ablehnung AO, Alternative + Begründung	Es ist gut, zunächst mit dem alltagssprachlichen Begriff Schwung einzusteigen, da der Begriff Schwung ein tragfähiges Konzept für den Impulsbegriff aus dem Alltag der Schülerinnen ist. (.) Ne, das würde ich jetzt nicht sagen, Schwung ist (halt?) irgendwie- ich glaube, das ist ja auch irgendwie so eine Fehlvorstellung, (deswegen?) (unv.) Schwung lieber vermeiden. (...)
118	Vorlesen, Zustimmung AO	Ähm die Formulierung stärker bremsen ist nicht gut gewählt, da auch Körper mit großem Impuls, wie zum Beispiel die Titanic, stärker bremsen könnten, um eine Kollision zu vermeiden. (..) Ja.
119	Vorlesen, Zustimmung AO + Begründung, Bewertung + Begründung	Es ist nicht gut, dass die Schülerinnen mit dem Begriff Sch- Begriff Schwung arbeiten. So werden sie s- So werden sie sich diesen Begriff merken und zukünftig nicht mehr mit dem Impulsbegriff arbeiten. (.) Ja, das sehe ich auch ein bisschen als Problem. (..) (Zum Teil?) ist es gut, weil dann sprechen die wahrscheinlich mehr mit einem, aber irgendwie ist Schwung halt so eine Fehlvorstellung, die das alles ziemlich kaputt macht. (..)
120	Vorlesen, Alternative	Ähm es ist nicht gut, dass die Schülerinnen in der Benutzung des Begriffs Schwung bestärkst. So förderst du ihre Fehlvorstellungen, zum Beispiel dass beide Begriffe gleichgesetzt werden können und dass Körper mehr Kraft haben, wenn sie mehr Schwung besitzen. Das (unv.) (achso?). Also das sollten- sollte man zumindest nicht bestärken, wenn die das nutzen ist okay, aber dann sollte man wenigstens dann direkt sagen, vielleicht das (.) man Schwung nicht- (dass?) das kein physikalisches physikalischer Fachausdruck ist.

121	Vorlesen, Ablehnung AO + Begründung	Es ist nicht gut, dass du so stockend redest und viele Füllwörter nutzt. Das lenkt die Schülerinnen vom Inhalt ab. (Oh?) ist mir gar nicht aufgefallen, also kann es nicht so schlimm gewesen sein. (...)
122	Vorlesen	Es ist nicht gut, da die Bedeutug des Begriffs noch nicht geklärt (unklar?) ist, was genau die Schülerinnen beschreiben.
123	Vorlesen, Zustimmung AO, Alternative	Es ist gut, dass die Schülerinnen zunächst ihre Alltagssprache verwenden können um zu erklären, warum es zur Kollision (.) gekommen ist. So sinkt die Hemmschwelle. Ja, finde ich eigentlich schon, (.) man muss dann da nur aufpassen, dass es nicht zu weit zu lange andauert, wie sie das Wort Schwung benutzen.
124	Vorlesen, Ablehnung AO	Es ist gut, zunächst mit dem alltagssprachlichen Begriff Schwung einzusteigen, ein tragfähiges Konzept- (nein?).
125	Vorlesen, Überlegung zum Unterricht	Die Formulierung stärker bremsen ist nicht gut gewählt, da auch Körper mit großem Impuls stärker bremsen könnten, um eine Kollision zu vermeiden. (..) Mhm (.) ja stärker bremsen ist, ich glaube, (4s) (bin?) mit unsicher.
126	Vorlesen, Zustimmung AO	Es ist nicht gut, dass die Schülerinnen mit dem Begriff Schwung arbeiten. (.) werden sie sich diesen Begriff merken und zukünftig- ja, das ist schon.
127	Ablehnung AO	Bei dem stärker bremsen bin ich mir unsicher, ich würde es jetzt aber, glaube ich, eher (.) erstmal weglassen. #00:49:10-6#
128	Vorlesen	LD6: Welche der unten stehenden Hinweise zu der Nutzung von Fachsprache möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben? (.)
129	Vorlesen, Zustimmung AO, Beschreibung	Es wäre besser, die Begriffe Masse und Impuls gegebenenfalls mit der zugehörigen Formel einzuführen, damit die Schülerinnen die korrekte Fachsprache lernen. Stimmt, das war nicht- die Masse ist (hier?) gar nicht vorgekommen. (.)
130	Vorlesen, Alternative + Begründung	Es wäre besser, wenn die Schülerinnen- wenn du die Schülerinnen dazu aufforderst, mehr Fachbegriffe zu nutzen, zum Beispiel den Massebegriff, damit sie sich so dem Impulsbegriff annähern können. (unv.) die lernen ja den Impuls gerade erst kennen, da muss man ja noch nicht mit ganz großen Fachbegriffen um sich schleudern, dann können die ja ruhig erstmal zumindest, aber das sollte dann (.) zumindest im Laufe der Zeit auf jeden Fall deutlich darauf geachtet werden. (.) Wobei jetzt vielleicht nicht bei den ersten Bemerkungen direkt. (..) (Vielleicht?) erst die einfach sagen lassen und dann danach (erst?) auffordern mit dem Massenbegriff arbeiten. (.)

131	Vorlesen, Zustimmung AO + Begründung, Beschreibung (2x), Bewertung, Alternative (2x)	Es wäre besser, den Impulsbegriff für den Begriff Schwung schon früher einzuführen, um zu verdeutlichen, um welches physikalische Konzept es geht. Ja, das würde ich schon sagen. Impuls fiel (jetzt?) ja im Einstieg eigentlich überhaupt nicht. (..) (Für?) die (Besprechung?) wäre das, glaube ich, schon ganz sinnvoll. (..) Mhm eigentlich der Massebegriff der fiel überhaupt nicht, das ist auch schon ein bisschen komisch. (5s) Ich würde das mal anklicken, (aber?) bin so ein bisschen zwiegespalten, ich würde sie nicht direkt dazu auffordern, also erstmal ihre Begriffe da ihre Aussagen sagen lassen, aber dass man dann schon darauf achtet, dass man den Massebegriff auch benutzt. (.)
132	Vorlesen, Zustimmung AO	Du solltest gegen die Schülervorstellung vorgehen, dass Körper Schwung haben, vergleichbar mit der Schülervorstellung Kraft haben, damit diese sich nicht verfestigt. Ja, auf jeden Fall, ist ganz wichtig. (.)
133	Vorlesen, Ablehnung AO	Die Verwendung der Begrifflichkeiten war angemessen, du kannst so- du kannst- das kannst du wieder so durchführen. Mhm würde ich nicht sagen. (.)
134	Vorlesen, Überlegung zum Unterricht	Du solltest dir in der Vorbereitung genauer überlegen, was du den Schülerinnen erzählen möchtest, damit du flüssiger und motivierender vortragen kannst. Ist mir gar nicht aufgefallen, dass er so gestockt ist. (.)
135	Vorlesen, (Ablehnung AO) + Begründung	Wäre besser, die Schüler*innen zur Verwendung der Fachsprache aufzufordern, um zu vermeiden, dass sie die Konzepte Schwung und Impuls weiterhin verwechseln. (Ich mein?), es ist halt mal gerade Einstieg, (6s)
136	Vorlesen, Alternative	Verwendung der Fachsprache aufzufordern, um (zu vermeiden?) (5s) Fachbegriffe, ich glaube, das reicht, wenn man Fachbegriffe benutzt und nicht (unv.). (.) (Also?) #00:51:46-4#
137	Vorlesen, Sonstiges*	LD6: Welches Feedback möchtest zu Robert zur Gestaltung des Phasenübergangs geben? Ich weiß gar nicht genau, welcher Phasenübergang gemeint ist. (..)
138	Vorlesen	Mhm der Phasenübergang ist dir nicht gut gelungen, da er zu abrupt war.
139	Vorlesen	Der Phasenübergang ist dir nicht gut gelungen, da er zu kurz war.
140	Vorlesen, Zustimmung AO + Begründung, Alternative	Phasenübergang ist die nicht gut gelungen, da die Inhalte zum W- zum Wechselwirkungsprinzip nicht ausreichend verstanden und gesichert wurden. Ja, das finde ich auf jeden Fall. Nur, weil die eine Schülerin das gut erklärt hat, heißt ja nicht, dass der Rest auch ansatzweise gut erstanden hat. (.) Hätte man da noch eine Übung (oder so?) dazwischen packen können. (unv.) würde ich auch sagen, dass (unv.) war.

141	Vorlesen, Ablehnung AO	Es ist unproblematisch, dass der Phasenübergang abrupt war, da du mit dem Impulsbegriff wirklich ein komplett neues Thema beginnst. (...) Mhm (..) weiß ich nicht, (das?) ist, glaube ich, keine Begründung.
142	Vorlesen, Zustimmung AO, Bewertung, Alternative, Überlegung zum Unterricht	Es ist nicht gut, dass du so einen abrupten Übergang zwischen den Themen gemacht hast, da der rote Faden der Stunde so für die Schülerinnen nicht erkennbar ist. (..) Ja, das stimmt schon, wobei man das natürlich vorher einmal ankündigt, wenn man- davor haben die ja die Newtonschen Axiome kennengelernt und dann (..) wird er ja wahrscheinlich schon irgendwie angekündigt haben, okay, wir sichern das nächste Woche nochmal, beginnen wir dann aber auch in Kürze mit dem Impuls, ich glaube, dann ist das schon okay. (.)
143	Vorlesen, Zustimmung AO	Ähm es ist nicht gut, dass du sagst, das Thema ist komplett neu, da den Schülerinnen so das bisher Gelernte nicht mehr relevant erscheint. (...) Ja, eigentlich schon richtig, stimmt schon, würde ich ihm auf jeden Fall mit auf den Weg geben. (..)
144	Vorlesen, Professionalisierung	Es ist nicht gut, dass du sagst, das Thema ist komplett neu, da es einen fachlichen Bezug zwischen den Newtonschen Axiomen und dem Impulsbegriff gibt. (..) Gibt es das? (.) Mhm (bin?) ich mir (grad?) selbst nicht sicher. (..) Manchmal denke ich, (.) ist Physik das richtige Fach, bin mir selbst nicht richtig sicher. #00:53:50-5#
145	Vorlesen	LD6: Ähm (.) welche der unten stehenden Hinweise zur Gestaltung des Phasenübergangs möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben? (...)
146	Vorlesen, Ablehnung AO	Den Phasenübergang hast du gut gestaltet, das kannst du wieder so machen. Ne, würde ich nicht sagen.
147	Vorlesen, Zustimmung AO + Begründung	Es wäre besser, den Zusammenhang zwischen den Newtonschen Axiomen und dem Impulsbegriff zu verdeutlichen, da die Schülerinnen den Zusammenhang nicht unbedingt selbst erkennen können. Ja, ist mir (ja?) selbst aufgefallen.
148	Vorlesen, Zustimmung AO + Begründung, Alternative + Begründung	Es wäre besser, den Einstieg in den Impulsbegriff auf die nächste Stunde zu verschieben, um den Schülerinnen zunächst noch mehr Möglichkeiten zu geben, das Wechselwirkungsprinzip zu verstehen und anzuwenden. Auf jeden Fall auch, weil nur dieses eine- dieser eine (Ausschnitt?) (unv.) gibt es ja auch noch Aufgaben dazu und so weiter, deshalb würde ich das auf jeden Fall machen und tendenziell ist es ja auch eigentlich ganz gut, mhm mit einem neu- mit einem neuen Thema, der Impulseinführung in der nächsten Stunde anzufangen, obwohl das halt nicht nur an- nicht an dem (roten?) Faden liegt, aber (.) (hier?) hätte man auch so-

149	Vorlesen, Überlegung zum Unterricht	Es wäre besser, vor dem Einstieg in das neue Thema auch die Tafel zu wischen, damit die Schülerinnen sich voll auf das neue Thema konzentrieren können und nicht noch über das alte Thema nachdenken. (..) Hat er überhaupt die Tafel benutzt? Er hat doch den Beamer eigentlich an. (...) Ist mir nicht aufgefallen.
150	Vorlesen, Zustimmung AO	Es wäre besser, einen glatteren Übergang zu gestalten, zum Beispiel durch eine kurze Verknüpfung, wie Wir finden jetzt etwas neues mit dem gelernten Wissen heraus, das ist gut, damit der rote Faden der Stunde für die Schülerinnen nachvollziehbar ist. Ja, das ist gut. (.)
151	Vorlesen, Ablehnung AO + Begründung, Alternative	Es wäre besser, den Übergang für- durch das Zeigen eines Experiments zu gestalten, um den Bruch zwischen den Themen für die Schülerinnen weniger hart zu gestalten. Ne, eigentlich nicht, der Übergang muss ja nicht unbedingt mit einem Experiment machen sein, es kann ja auch mit so einem Alltagsbeispiel ist eigentlich ganz okay. #00:55:37-8#
152	Überlegung zum Unterricht	LD6: Mhm (.) (ah?) eine neue Phase. Experiment Tischtennisbälle (..) zum Impuls. #00:55:51-2#
153	-	[Video Einleitung Reflexion 04, Video Szene Reflexion 04] // [Weiterhin Video Szene Reflexion 04] // [Weiterhin Video Szene Reflexion 04] // [Weiterhin Video Szene Reflexion 04] #00:58:13-8#
154	Alternative	LD6: Stimmt, ist echt am Tafelbild so ein Baum. Ja, hätte man machen können, die Tafel wischen, aber ist jetzt auch nicht so wichtig, (das ist?) okay. #00:56:22-5#
155	Sonstiges	LD6: Oh, gleiche Geschwindigkeit, das kann ja (unv.) klappen. #00:56:50-4#
156	Sonstiges	LD6: (unv.) jetzt schnell. #00:57:15-3#
157	Sonstiges	LD6: Mhm (.) (unv.) lieber an die Tafel (unv.). (..)
158	Vorlesen	Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung, (.) Lob oder Kritik, erhalten?
159	Vorlesen	Eingesetzte Beispiele,
160	Vorlesen, Zustimmung AO, Bewertung (2x), Überlegung zum Unterricht, Alternative + Begründung	Fachsprache, ja, Kritik. Warum benutzt er jetzt noch weiter Schwung, frage ich mich. (.) Könnte er doch als Lehrkraft dann einfach auch schon Impuls benutzen, dann merken die Schülerinnen okay, es ist doch nicht (unv.) Schwung, was wir hier machen, sondern wir brauchen etwas anderes. Und hinterher einfach sagen, Schwung ist gleich Impuls, irgendwie komisch. (.)

161	Beschreibung, Bewertung + Begründung, Alternative + Begründung, Sonstiges	Mhm Ergebnissicherung finde ich irgendwie auch jetzt nicht so bombig, ich weiß nicht, ich freunde mich irgendwie nicht mit dem Beamer an. Irgendwie das ist so vorgefertigt und dann müssen die das abschreiben, das für mich wie, keine Ahnung, irgendwie besser, wenn du einfach an der Tafel, dann schreibst du und die Schülerinnen schreiben direkt mit, dann ist das irgendwie so, okay dann macht (ihr?) noch was, probiert das auch mal ab und dann kann man (sich?) selbst erarbeiten, (anstatt?) dass da schon vorgefertigt steht, ich weiß nicht.
162	Vorlesen, Ablehnung AO, Beschreibung (2x), Bewertung (2x) + Begründung (2x), Alternative (2x), Alternative + Begründung	Mhm eingesetzte Beispiele, ja, das Beispiel ist ja eigentlich nur die Tennisbälle und einmal mit einer höheren Geschwindigkeit und einmal mit einer kleineren (..) mhm (...) (das?) eingesetzte Experiment fand ich eigentlich gar nicht so schlecht, weil man (.) an den Tennisbällen irgendwie den Impuls schon gut verdeutlichen kann (.) und eigentlich auch die eingesetzten Beispiele, eigentlich auch, (.) weil einmal langsam werfen, schnell werfen, dann merkt man dass (unv.) irgendwie von der Geschwindigkeit abhängt. Masse, da ist er jetzt vielleicht nicht ganz so viel drauf eingegangen, hätte ich vielleicht noch ein bisschen, okay ich nehm das wieder rau- ja, ich kann es auch kritisieren. Ähm noch ein bisschen (.) das Experiment vielleicht einfach vertiefen, dass man da wirklich (dann?) kann man das den Impuls eigentlich richtig gut dran erklären, aber halt dann wirklich mit der auf die Masse eingehen, schnell, langsam werfen, (.) das fände ich, glaube ich, besser. #01:00:13-3#
163	Vorlesen	LD6: Robert hat schon einige Rückmeldungen erhalten. (.) Besonders wichtig ist ihm jetzt noch eine Rückmeldung zur Wahl des Experiments und zur Ergebnissicherung. #01:00:22-5#
164	Vorlesen	LD6: Zur Wahl des Experiments geben.
165	Vorlesen, Zustimmung AO, (Ablehnung AO) + Begründung	Das Experiment ist gut gewählt, da die beiden Schüler, die das Experiment durchführen, den Impuls am eigenen Körper erfahren können. (.) Ja, aber dann halt auch nur die beiden.
166	Vorlesen, Alternative	Das Experiment ist gut gewählt, da es den Einfluss von Masse und Geschwindigkeit auf den Impuls gut darstellt. (unv.) (einfällt?), dann kann ich theoretisch auch einfach so viele Tennisbälle mitbringen, dass die ganze Klasse das in Zweierteams machen kann, (dann wäre?) das viel besser. Und dann leitet der Lehrer an, (.) okay, jetzt werfen alle den leichteren Ball und gucken was passiert, das (ist?) ja besser, als nach vorne. (..)
167	Vorlesen, Zustimmung AO	Das Experiment ist gut gewählt, da es den Einfluss von Masse und Geschwindigkeit auf den Impuls gut darstellt. Das finde ich zum Teil.

168	Vorlesen, (Ablehnung AO) + Begründung, Alternative, Überlegung zum Unterricht	Es ist nicht gut, da über nur über die Masse und die Geschwindigkeit gesprochen wurde, da unklar bleibt, ob der Impuls noch von anderen Größen, wie zum Beispiel der Form abhängen kann. (.) Mhm das merkt man ja eigentlich hinterher, wenn man die Formel dann (.) anschmeißt. Eventuell könnte man ja auch noch andere Gegenstände (..) mitbringen. Denken die wirklich, dass es von der Form abhängt? Weiß nicht.
169	Vorlesen, Ablehnung AO	Das Experiment ist nicht gut gewählt, da es den Zusammenhang beziehungsweise die Abgrenzung des Impulses zur kinetischen Energie nicht veranschaulicht. (..) Mhm die Gefahr sehe ich jetzt eigentlich nicht.
170	Vorlesen, Zustimmung AO, Alternative	Es ist gut, dass du die beiden Schüler in das Experiment einbezogen hast. So motivierst du und aktivierst du sie. Ja, finde ich an sich gut, (.) aber man könnte auch die ganze Klasse einbeziehen, das ist auch nicht schwer. (5s)
171	Zustimmung AO	Hier (LD6 zeigt auf AO 'Es ist gut, dass du nur zwei Schüler das Experiment durchführen lässt und nicht alle Schülerinnen, da du so Zeit sparst und das Experiment so einfach ist, dass es auch von den sitzenden Schülerinnen gut nachvollzogen werden kann.'), das ist ja genau das, was ich eigentlich wollte. (unv.)
172	Vorlesen, Sonstiges	Es ist gut, dass du vorgibst, wer nach vorne kommen soll. So vermeidest zu lange Organisations-/Wartezeiten nach der Frage 'wer will denn'. Ja (ähm?) alle, (.) (unv.) kennt man das nicht.
173	Sonstiges	Oh, nur eine Rückmeldung. (6s)
174	Zustimmung AO, Alternative	(Ja?) (LD6 zeigt auf Antwortoption 'Das Experiment ist gut gewählt, da die beiden Schüler, die das Experiment durchführen, den Impuls am eigenen Körper erfahren können.') ist ja eigentlich auch richtig, aber ich finde es halt besser, wenn alle (unv.). #01:02:31-1#
175	Vorlesen	LD6: Welche der unten stehenden Hinweise zur Wahl des Experiments möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben? (.)
176	Vorlesen, Zustimmung AO	Experiment ist gut gewählt, du kannst es erneut einsetzen. Ja, finde ich.
177	Sonstiges, Vorlesen, Zustimmung AO + Begründung	Halt nur mit der mit dem Ball. (..) (Ah?) da ist es doch (LD6 wählt AO 'Es wäre besser, wenn du das Experiment von allen Schüler*innen durchführen ließest, da es den Schüler*innen gut tut, sich zwischendurch zu bewegen.'). Dann nehme ich lieber das. Es wäre besser, wenn du das Experiment von allen Schülerinnen durchführen ließest, da es den Schülerinnen gut tut, sich zwischendurch zu bewegen. Ja, das und natürlich, dass sie es alle erfahren.

178	Vorlesen, Alternative + Begründung	Du solltest darauf achten, dass die Schüler vorne so stehen, dass alle sitzenden Schülerinnen erkennen können, wie leicht, schwer es ist, den Ball zu fangen. (.) Es ist eh schwer, von außen zu erkennen, wie schwer es ist, den Ball zu fangen, (da wäre?) es halt besser wenn sie es selbst machen.
179	Vorlesen, Ablehnung AO	Es wäre besser, statt des Experiments ein Video zu zeigen, in welchem gezeigt wird, welchen Einfluss verschiedene Massen und Geschwindigkeiten auf den Impuls haben, da das die Schülerinnen eher anspricht, als ein Experiment. Ne, (auf jeden Fall?) nicht, würde ich sagen.
180	Vorlesen, Zustimmung AO, Alternative	Es wäre besser, ein anderes Experiment durchzuführen, das für alle Schülerinnen anschaulich ist, zum Beispiel eine Person, die auf einem Skateboard steht und einen Ball wegwirft, oder das Werfen beider Bälle in den Sand zur Betrachtung der Krater. (.) Ja, aber wenn man es mit allen Schülern macht, ist es ja irgendwie auch nicht schlecht.
181	Vorlesen, Zustimmung AO, Bewertung, Alternative + Begründung	Es wäre besser, ein Stoßexperiment durchzuführen, da du so gut auf die Impulserhaltung überleiten kannst. (...) Ja, aber eigentlich finde ich das Experiment gar nicht so schlecht, man könnte das Experiment ja auch noch durchführen zusätzlich, um auf die Impulserhaltung zu überleiten.
182	Vorlesen, Überlegung zum Unterricht, Zustimmung AO + Begründung	Es wäre besser, ein Experiment zu wählen, zu dem eine Messreihe aufgenommen werden kann, damit die Formel aus diesen objektiven Messungen abgeleitet werden kann. Was für ein Experiment gibt es denn da? (.) (wo das möglich ist?). (..) Aber den Impuls kann man ja gar nicht messen. Ne, ich glaube (..) ich klicke das (LD6 wählt AO 'Es wäre besser, ein Stoßexperiment durchzuführen, da du so gut auf die Impulserhaltung überleiten kannst.') einfach nochmal an, weil es schon gut ist auch noch, aber halt nicht besser (unv.). #01:04:25-8#
183	Vorlesen, Überlegung zum Unterricht	LD6: Welches Feedback möchten Sie Robert zur Ergebnissicherung geben? Was war nochmal die Ergebnissicherung? Gab es überhaupt eine? Ach ja, stimmt, den Merksatz.
184	Vorlesen, Alternative + Begründung	Es ist nicht gut, dass du den Schülerinnen einen fertigen Merksatz vorgibst, da das reine Lesen und Abschreiben nicht verständnisfördernd ist. Auf jeden Fall, es ist viel besser, den mit denen zusammen zu- abzuleiten. (.) (unv.) dass sie da einfach (unv.) könnte, wenn sie das Experiment gesehen haben. (5s)
185	Vorlesen, Zustimmung AO + Begründung	Es ist gut, dass die Schülerinnen den Merksatz ins Heft übertragen, damit alle Schülerinnen den Merksatz als Sicherung notiert haben. Schon gut, die sollten den schon irgendwo stehen haben.

186	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du die Ergebnisse des Experiments nicht sicherst, so ist unklar, was die Schülerinnen aus dem Experiment mitnehmen sollen. Auch wahr. (..)
187	Vorlesen, Zustimmung AO + Begründung, Alternative	Der Merksatz ist gut gelungen, da die Je-umso-Sätze die Zusammenhänge von Masse, Geschwindigkeit und Impuls nachvollziehbar darstellen. (.) Mhm[Ja] ja, aber da kann man- also an sich schon gut, die Je-umso-Sätze sind wirklich gut, aber vielleicht mit dem (bisher?) Gelernten, dass man sie mit den Schülerinnen und Schülern selbst erarbeiten kann.
188	Vorlesen, Ablehnung AO	Es ist nicht gut, dass du den Impuls als skalare Größe einführst. So grenzt du den Impuls nicht ausreichend von der kinetischen Energie ab. Das wurde jetzt nicht irgendwie (unv.) grad noch gar nicht so.
189	Vorlesen, Zustimmung AO	Es ist gut, dass du den Begriff der Energie nicht erwähnst. So vermeidest du Verwirrung bei den Schülerinnen, die die verschiedenen Fachbegriffe auf Grundlage ihrer Erfahrung nicht trennen können. (..) Ja, Energie (unv.). Ist, glaube ich, eigentlich erstmal ganz gut, das ohne die Energie einzuführen einzuführen. (Ich meine?) ist auch eine Strategie, ja. (.)
190	Vorlesen, Bewertung	In deiner Sicherung fehlt eine Abgrenzung des Impulsbegriffs von der kinetischen Energie und der Kraft, da die Begriffe den Schülerinnen zwar bekannt sind, aber Unterschiede für sie nicht ersichtlich sind. (.) Mhm (..) ich glaube, erstmal nur für die Einführung des Impulses, würde ich sagen, ist der eigentlich schon ganz gut, der Merksatz. #01:06:29-8#
191	Vorlesen	LD6: Welche der unten stehenden Hinweise zur Ergebnissicherung möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben?
192	Vorlesen, Ablehnung AO	Die Ergebnisse des Experiments hast du gut gesichert, das kannst du wieder so machen. Ne.
193	Vorlesen, Zustimmung AO	Es wäre besser, die Beobachtungen des Experiments schriftlich an der Tafel zu sammeln, damit du im Verlauf der Stunde darauf zurückgreifen kannst. Ja, auf jeden Fall ist es wichtig, die zu sammeln.
194	Vorlesen, Zustimmung AO	Es wäre besser, die Schülerinnen in die Sicherung mit einzubeziehen, zum Beispiel einen Merksatz formulieren- Ja, auf jeden Fall.
195	Vorlesen, Zustimmung AO	Es wäre besser, zusätzlich zum Merksatz noch ein Beispiel zu zeigen, um auch die Formel anschaulich darzustellen. (.) Mhm[Ja].
196	Vorlesen, Alternative, Zustimmung AO	Es wäre besser, in der Sicherung auch auf das Experiment einzugehen, (.) (unv.) zum Merksatz auch auf das Experiment eingehst oder Skizzen ergänzt, um die Inhalte stärker zu vernetzen. (6s) Mhm weiß ich nicht, ich würde das mhm (4s) Skizzen (4s) (na?) hier könnte man einen Tennisball malen und dann hat (der?) ja die Masse m, Geschwindigkeit v, ja, vielleicht schon.

197	Vorlesen, Sonstiges	Es wäre besser, den Impuls als vektorielle Größe einzuführen, um ihn von der kinetischen Energie abzugrenzen. (..) Ich weiß es nicht.
198	Vorlesen, Ablehnung AO + Begründung	Es wäre besser, auch auf die kinetische Energie der Bälle einzugehen, um an das Vorwissen der Schülerinnen anzuknüpfen. (..) Ich glaube, man sieht (schon hier?), eigentlich kann man das ja auch so ganz gut verstehen, (deswegen würde ich, glaube ich,?) die kinetische Energie erstmal raus lassen.
199	Zustimmung AO	Aber ich würde das hier noch (LD6 wählt AO 'Es wäre besser, zusätzlich zum Merksatz noch ein Beispiel zu zeigen, um auch die Formel anschaulich darzustellen.') Beispiel (unv.) ja, ist gut.
200	Alternative, Überlegung zum Unterricht	Mit der (Größe?) bin ich mir unsicher. (..) Vielleicht könnte man die auch so einführen, um es an sich zu verstehen und dann später noch die vektorielle Schreibweise. #01:08:18-3#
201	Beschreibung, Vorlesen, Bewertung, Sonstiges	LD6: Luftkissenfahrbahn. (...) Nachdem wir das Experiment mit den Tennisbällen abgeschlossen und die Schülerinnen den Merksatz zum Impuls abgeschrieben haben, (...) habe ich zu dem habe ich zu einem Demonstrationsexperiment übergeleitet, an dem ich die Impulserhaltung veranschaulichen möchte. Ist das eigenltich alles eine Stunde? Bisschen viel. (6s) An dem ich die Impulserhaltung veranschaulichen möchte. Mit dem Demons- Demonstrationsexperiment möchte ich einen Auffahrunfall simulieren. Den Schülerinnen habe ich dann erklärt, wie der Autounfall im Experiment nachgestellt wird und wie die Messwertaufnahme funktioniert. Den experimentellen Aufbau und eine Messwerttabelle habe ich vor der Stunde vorbereitet, hier habe ich dir zwei Fotos vom Aufbau und von der Tabelle mitgebracht. (10s) Als nächstes haben wir das Experiment durchgeführt und die Messwerte aufgenommen. Okay. #01:09:32-5#
202	-	[Vieo Einleitung Reflexion 05] #01:09:41-3#
203	Überlegung zum Unterricht	LD6: Mhm ja, (ist?) beim Experiment ja eigentlich immer so, ne. #01:09:46-1#
204	-	[Video Szene Reflexion 05] // [Weiterhin Video Szene Reflexion 05] // [Weiterhin Video Szene Reflexion 05] #01:12:14-0#
205	Überlegung zum Unterricht (2x)	LD6: Und wie (hat er die jetzt?) ausgerechnet? #01:11:41-5#
		LD6: Warum rechnet er das (eigentlich selber?) #01:11:52-2#
206	Sonstiges	LD6: Ja, gut. (.) #01:12:16-5#
207	Vorlesen	LD6: Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung erhalten, um etwas für sich seinen Unterricht zu lernen?

208	Vorlesen, Zustimmung AO, Überlegung zum Unterricht	Umgang mit den Schülerinnen, ja auf jeden Fall. (..) Wann haben wir das noch besprochen? Das ging überhaupt nicht, einfach (.) Schülerinnen auf die Handschrift zu äh (.) minimieren und Schüler können natürlich viel besser messen, auf jeden Fall. (..)
209	Vorlesen, Bewertung	Ähm (.) Notation der Messwerte war ja (in?) der Tabelle eigentlich ganz gut.
210	Vorlesen, Zustimmung AO, Überlegung zum Unterricht, Alternative, Beschreibung	Durchführung, Auswertung des Experiments durch die Lehrkraft, ja, warum hat (.) er das ausgemessen und nicht der Rest, er hat einfach nur (unv.) Geschwindigkeit jetzt genau bestimmt hat, was er da berechnet hat.
211	Vorlesen, Beschreibung	Messunsicherheiten mhm hat er ja eigentlich gar nicht thematisiert. (.)
212	Vorlesen, Bewertung (2x), Bewertung + Begründung	Wahl des Experiments, experimenteller Aufbau, (ne?), eigentlich (was das alles?) gut. Das Experiment ist eigentlich ziemlich gut geeignet (.) und auch ganz gut, dass er schon (unv.) vorbereitet hat, weil das ist ja ziemlich (.) ähm (4s) aufwendig.
213	Vorlesen, Bewertung, Beschreibung, Alternative	Äh Aktivierung der Schülerinnen, ja, eigentlich ganz gut, dass er da Schülerinnen und Schüler nach vorne holt, über die Art und Weise muss man sich nochmal Gedanken machen. #01:13:35-8#
214	Vorlesen	LD6: Robert hat schon einige Rückmeldungen erhalten. Besonders wichtig ist ihm jetzt noch eine Rückmeldung zum Umgang mit den Schülerinnen, dazu, wie er während des Experiments auf Messunsicherheiten eingegangen ist und zur Aktivierung der Schülerinnen. #01:13:47-6#
215	Vorlesen	LD6: Feedback zu seinem Umgang mit den Schülerinnen. (.)
216	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du sagst, dass Mädchen eine schöne Schrift haben und Jungen besser Werte ablesen können, da du die Mädchen so auf eine schöne Handschrift reduzierst. (Ja?).
217	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du die Mädchen nur für Fleißarbeiten im Unterricht einbeziehst, dass du so Stereotype förderst. Ja.
218	Vorlesen, Zustimmung AO + Begründung	Es ist gut, dass du die Mädchen für ihr Handeln im Unterricht lobst, um ihr Selbstkonzept im Physikunterricht zu stärken. Ja, das wollen die aber nicht hören, dass sie eine schöne Handschrift haben, das ist ja kein Selbstkonzept für Physikunterricht, für äh kein physikalisches Handeln.
219	Vorlesen	Es ist nicht gut, dass du das Lob für die Handschrift von Mädchen mit einer inhaltlichen Kritik, vergessene Einheiten, kombinierst, da die Ursache hier nicht das Geschlecht ist. (..)

220	Vorlesen, Zustimmung AO + Begründung	Es ist nicht gut, dass du das Lob für die Handschrift von Mädchen (unv., LD6 liest murmelnd vor) da die Ursache hier nicht das Geschlecht ist. (...) Ja, eh bisschen überflüssig, die Handschrift der Mädchen zu loben. (.)
221	Vorlesen, Zustimmung AO, Alternative	Es ist nicht gut, dass du die Schülerin schon vor dem Notieren des zweiten Messwertes erneut auf die Einheit hinweist, da du so das Bild festigst, Mädchen wären in Physik nicht kompetent. (..) Ja, ich hätte es wahrscheinlich auch dann eher nochmal danach gesagt. (.)
222	Vorlesen, Zustimmung AO, Alternative	Es ist nicht gut, dass du der Schülerin an der Tafel direkt sagst, welche Einheit fehlt, da du ihr so unterstellst, sie wüsste das nicht. Ja, hätte man theoretisch erstmal nachfragen können, was für eine Einheit das ist.
223	Vorlesen, Ablehnung AO, Alternative + Begründung	Es ist in Ordnung dass du sagst, dass Paula aufgrund ihrer schönen Handschrift die Werte notieren soll, da es wichtig ist, dass alle die Werte an der Tafel gut lesen können. Nein, das finde ich nicht, (..) kann auch einem Jungen sagen, dass der (.) einfach ordentlich schreiben soll, dass (das?) auch Jungen können. #01:15:22-9#
224	Vorlesen	LD6: Ähm welche der unten stehenden Hinweise zum Umgang mit den Schülerinnen möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben?
225	Vorlesen, Ablehnung AO	Dein Umgang mit Mädchen und Jungen ist in Ordnung so, da Paula wirklich eine gut lesbare Handschrift hat. M-hm[Nein].
226	Vorlesen, Ablehnung AO, Alternative + Begründung	Du solltest eine Geschichte um das Experiment herum konstruieren und zum Beispiel in Form eines Videos zeigen, um die Mädchen zu motivieren. (...) (Das?) weiß ich jetzt nicht, ich glaube, eher nicht. (...) Ähm aber was mir dann (bei?) bei Video auf- äh einfällt, es wäre sinnvoll vielleicht, irgendwie eine Video oder so ein so eine Kamera vielleicht auf die Waage zu (unv.), dass die auch wirklich sehen können, das es null Komma eins (.) Kilogramm ist und dass sie das (dem Lehrer?) nicht glauben müssen. (.)
227	Vorlesen, Zustimmung AO	Es wäre besser, die Schülerinnen zu fragen, wer freiwillig schreiben beziehungsweise die Werte ablesen möchte, um den Schülerinnen nicht nach Geschlecht eine Aufgabe zuzuweisen. (.) Ja.
228	Vorlesen, (Ablehnung AO) + Begründung	Es wäre besser, die Aufgaben Aufschreiben, Werte Ablesen durch zu tauschen, damit alle Schülerinnen eine Aufgabe wahrnehmen. Ja, man muss (ja?) jetzt auch nicht übertreiben. (.)
229	Vorlesen, Zustimmung AO	Du solltest darauf achten, die Mädchen nicht für ihre Handschrift, sondern explizit für ihre physikalische Kompetenz zu loben, da du sonst das Bild vermittelst, Physik wäre eine Jungsfach. Auf jeden Fall, ist ganz wichtig.

230	Vorlesen, Zustimmung AO	Du solltest auf geschlechtsspezifische- geschlechtsspezifische Zuschreibungen verzichten, da du so ein unangemessenes Verhältnis zwischen den Geschlechtern im Physikunterricht erzeugst. Mhm[Ja].
231	Vorlesen, Ablehnung AO + Begründung, Zustimmung AO	Du solltest grundlegend an deiner Einstellung gegenüber Mädchen und Jungen arbeiten, um dein stereotypisches Verhalten zu ändern. (...) Schwierig, das als Feedback zu geben, glaube ich, weil, ich glaub, weiß nicht, ob (er?) sich da verändert, aber ich würde es ihm zumindest (einmal) sagen. #01:17:05-7#
232	Vorlesen, Überlegung zum Unterricht	LD6: Welches Feedback möchtest du Robert zu seinem Umgang mit Messunsicherheiten geben? Ich kann mich gar nicht daran erinnern, dass er überhaupt Messunsicherheiten thematisiert hat. (4s)
233	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du sagst, du könntest exakt oder genau messen, da du so die Schülervorstellung beförderst, man könnte Größen ohne Messunsicherheiten bestimmen. (..) Mhm, (ja?).
234	Vorlesen, Alternative + Begründung	Du gehst nicht auf die Messunsicherheiten ein. Das ist nicht gut, da diese eine wichtige Rolle in der Physik spielen. Man muss sie zumindestens, finde ich, eigentlich einmal nennen zumindest, dass man halt hier keine genau exakte Messung vorliegen hat, (.) weil einfach so außer Acht lassen ist vielleicht (unv.) auch nicht so gut. (.)
235	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du nur einen der beiden Wäge-Wagen wiegst, da die Schülerinnen so nicht sicher sein können, dass die Wagen wirklich gleich schwer sind. Das habe ich mich auch schon gefragt. Der andere ist einfach genauso schwer, das ist, glaube ich, nicht so gut.
236	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du sagst, die Wagen wären gleich schwer, da das auch wenn sie baugleich sind nicht stimmen muss. Ja.
237	Vorlesen, Alternative	Es ist nicht gut, dass du das Gewicht mit null Komma einem Kilogramm angibst, da du hier nicht berücksichtigst, (..) welches die signifikan- signifikanten Nachkommastellen-. Ja, okay zu genau würde ich jetzt die Messunsicherheiten im Physikunterricht auch nicht (.) thematisieren.
238	Vorlesen, Ablehnung AO, Alternative	Es ist gut, dass du darauf verzichtest, auch auf Messunsicherheiten einzugehen, (.) um die Komplexität des Experiments überschaubar zu halten. Ja, ähm ne, ich glaube eigentlich nicht, weil ich würde das (unv.) (.) erst Experiment auswerten und dann hinterher nochmal Fragen Messunsicherheiten.
239	Zustimmung AO	Und dann würde ich vielleicht diesen Stichpunkt hier oben mit exakt und genau (.) noch ankreuzen.

240	Vorlesen, Sonstiges	Es ist gut, dass du das Gewicht des Wagens mit null Komma eins Kilogramm angibst, um Fehler in der Berechnung des Impulses aufgrund vieler Nachkommastellen zu vermeiden. (.) (Ist hier?) jetzt angekreuzt, (wäre?) auch jetzt kein Problem. #01:19:02-6#
241	Vorlesen	LD6: Welche der unten stehenden Hinweise zu seinem Umgang mit Messunsicherheiten möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben? (.)
242	Vorlesen, Ablehnung AO	Es ist gut, wie du mit Messunsicherheiten umgegangen bist. Das kannst du erneut so machen. (..) Ne.
243	Vorlesen, Bewertung	Es wäre besser, alle Nachkommastellen anzugeben, die die Laborwaage dir anzeigt, da du den Fokus auf genaue Messungen legst und die Laborwaage genauer, als null Komma eins Kilogramm wiegen kann. Das finde ich zu ausführlich für die Schülerinnen.
244	Vorlesen, Zustimmung AO	Es wäre besser, anstelle der Aussage, du könntest exakt messen, auch Messunsicherheiten einzubeziehen, da Messungen nie wirklich exakt sind. Ja, das finde ich richtig.
245	Vorlesen, Zustimmung AO	Es wäre besser, beide Wagen zu wi- beide Wagen zu wiegen, da du den Schülerinnen ein korrektes Bild von genauem wissenschaftlichen Arbeiten vermittelst. Mhm[Ja].
246	Vorlesen	Es wäre besser, beide Wagen zu wiegen, da die Wagen auch wenn sie baugleich sind nicht gleich schwer sein müssen. (.)
247	Vorlesen, Zustimmung AO, Alternative	Es wäre besser, beide Wagen zu wiegen, da du im Vergleich der Massen auf die Messunsicherheiten beziehungsweise Ungenauigkeiten eingehen kannst. Ja, ich würde sie wahrscheinlich einfach nur einmal nennen, was für Messunsicherheiten wir haben könnten.
248	Vorlesen, Zustimmung AO + Begründung	Es wäre besser, beide Wagen zu wiegen, damit die Schülerinnen sich selbst überzeugen können, dass die Wagen gleich schwer sind. Ja, sonst müssen sie das nur glauben. #01:20:12-4#
249	Vorlesen	LD6: Welches Feedback möchten Sie Robert zur Aktivierung der Schülerinnen geben? (..)
250	Vorlesen, Zustimmung AO + Begründung, Bewertung	Es ist nicht gut, dass du die Schülerinnen nur oberflächlich involvierst, da die Schülerinnen trotz dieser Aktivierung nicht mitdenken müssen. Ja, das stimmt schon, weil zum Beispiel rechnet er die Geschwindigkeit aus, nicht so gut.
251	Vorlesen, Sonstiges	Es ist gut, dass du Schülerinnen in das Experiment einbeziehst, um sie im Unterricht zu aktiveren. (.) Mhm.
252	Vorlesen, Alternative	Es ist gut, dass du das Experiment selbst durchführst, da du in der Durchführung im Vergleich zu den Schülerinnen dabei sehr sicher bist. (Na?) den Wagen anstoßen und den Schalter umlegen kann eigentlich auch ein Schüler machen.

253	Vorlesen, Zustimmung AO	Es ist nicht gut, dass die Schülerinnen dir ohne Arbeitsauftrag zusehen, da die Schülerinnen so keinen Anreiz haben, aufmerksam zu sein. (..) Mhm ja, (das) ist schon richtig.
254	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du die Masse des Wagens und insbesondere die Geschwindigkeit selbst bestimmst, da du den Schülerinnen vorher schon erklärt hast, wie sie das selbst machen könnten. (..) Auf jeden Fall, mhm[ja].
255	Vorlesen, Zustimmung AO, Beschreibung, Alternative	Es ist nicht gut, dass du das Experiment so lehrerzentriert durchführst, da dies für die Schülerinnen langweilig ist. Auch richtig. Also er bezieht die ja ein, aber irgendwie ja nur zum Aufschreiben, er könnte (die?) ja auch anstoßen lassen, an machen, Geschwindigkeit ausrechen und so weiter.
256	Vorlesen, Zustimmung AO, Alternative, Alternative + Begründung	Es ist nicht gut, dass Paula von der Tafel aus und die Schülerinnen aus den hinteren Reihen den experimentellen Aufbau nicht gut erkennen können, da sie- da es für sie (.) schwer ist, dir zu folgen. Mhm[Ja] ich habe mir auch schon überlegt, wenn er da vorne die (.) Wagen wiegt könnte er vielleicht irgendwie so eine Kamera vielleicht irgendwie aufstellen, wo man das dann noch besser sehen kann oder sie halt nach vorne holen, wäre, glaube ich, schon nicht so schlecht. #01:21:49-8#
257	Vorlesen	LD6: Welche der unten stehenden Hinweise zur Aktivierung der Schülerinnen möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben?
258	Vorlesen, Überlgung zum Unterricht	Es wäre besser, die Schülerinnen zu fragen, wer freiwillig schreiben beziehungsweise die Werte ablesen möchte, um den Schülerinnen nicht nach Geschlecht eine Aufgabe zuzuweisen. Es wäre besser- ja, freiwillig (unv.) gut, ansonsten- ja freiwillig ist ja immer so macht (niemand?).
259	Vorlesen, Zustimmung AO	Es wäre besser, die Schülerinnen zu fragen, ob sie den Ablauf verstanden oder noch Schwierigkeiten haben, um den Schülerinnen die Möglichkeit für Rückfragen zu geben. (..) Mhm[Ja].
260	Vorlesen, Ablehnung AO	Du hast die Schülerinnen ausreichend aktiviert. Das kannst du wieder so machen. Ne.
261	Vorlesen, Zustimmung AO	Es wäre besser, Hypothesen der Schülerinnen abzufragen, um sie zur Beobachtung des Experiments zu motivieren. Ja, das ist gut.
262	Vorlesen, Zustimmung AO, Alternative	Es wäre besser, den Schü- Schülerinnen Fragen zum Vorgehen während des Experiments oder zur Berechnung der Geschwindigkeit zu stellen. Ja, auf jeden Fall, nicht alles selbst machen.
263	Vorlesen, Zustimmung AO	Es wäre besser, die Schülerinnen stärker in die Durchführung des Experiments einzubeziehen. Ja. (.)
264	Vorlesen, Zustimmung AO, Alternative	Es wäre besser, die Schülerinnen zum Experiment nach vorne kommen zu lassen, damit sie alles gut erkennen können. Ja, oder halt irgendwie Kamera.

265	Ablehnung AO, Alternative + Begründung	Dieses freiwillig Schreiben, ich weiß- (da meldet?) sich so gefühlt niemand, dann kann er auch einfach jemanden nach vorne mit holen ohne die Begründung, du bist ein Mädchen, hast eine schöne Handschrift und du bist ein Junge, du kannst das gut. (4s) (Kreuz?) raus. #01:23:07-6#
266	Vorlesen, Überlegung zum Unterricht	LD6: Nachdem wir alle Messwerte aufgenommen hatten, habe ich die Beobachtungen der Schülerinnen abgefragt. Um die Allgemeingültigkeit der Impulserhaltung nachzuweisen, habe ich sie dann in einem kurzen Exkurs theoretisch hergeleitet. (..) Im Anschluss haben wir die Impulserhaltung am Experiment überprüft. Wird denn die Impulserhaltung auch theoretisch hergeleitet in der Schule? (Weiß?) ich gar nicht (mehr?). #01:23:33-5#
267	-	[Video Einleitung Reflexion 06, Video Szene Reflexion 06] #01:26:29-0#
268	Sonstiges	LD6: Was ich mich jetzt frage ist, ob der Unterricht wirklich so stattgefunden hat oder ob er einfach nur so (.) exemplarisch ausgewählt wurde, weil das ist- schockiert mich gerade ein bisschen. (..)
269	Vorlesen	Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung erhalten, um etwas für sich beziehungsweise seinen Unterricht zu lernen? (.)
270	Vorlesen, Bewertung + Begründung (2x)	Umgang mit den Schülerinnen kritiseren, einfach, weil es einfach überhaupt nicht geht, (.) da (unv.) kann ja erklären, warum die Schülerinnen nichts mehr mit Physik am Hut haben möchten, wenn die so im Unterricht behandelt werden. (.)
271	Vorlesen, Zustimmung AO, Alternative + Begründung	Messunsicherheiten, Umgang mit der- mit dem Einwand von Marie. Ja, Messunsicherheiten müssen da wenn sowas aufkommt, ist doch der perfekte Zeitpunkt, um Messunsicherheiten zu thematisieren, weil es gibt immer Messunsicherheiten, insbesondere, keine Ahnung, sei es die Reibung auf der Luftkissenfahrbahn, wo man diese Unstimmigkeiten kriegt, gut erklären kann.
272	Vorlesen, Überlegung zum Unterricht, Beschreibung	Einheit des Impuls (.) habe ich gar nicht mitbekommen, dass er da irgendwie drüber gesprochen hat. (..) Ich glaube, er hat darüber auch gar nicht gesprochen, oder ich habe es mal wieder nicht (.) mitbekommen. #01:27:33-9#
273	Vorlesen	LD6: Robert hat zwar schon einige Rückmeldungen erhalten, würde aber trotzdem gern zu allen drei aufgeführten Aspekten noch eine Rückmeldung erhalten. Okay, (kriegst du?). #01:27:44-9#
274	Vorlesen	LD6: Welches Feedback möchten Sie Robert zu Umgang mit den Schülerinnen geben?
275	Vorlesen, Überlegung zum Unterricht	Es ist nicht gut, dass du Mädels sagst, das ist ein sexistischer Ausdruck. (.) Oh mein Gott, das ist mir noch nicht mal so aufgefallen.

276	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du die Mädchen und Jungen so unterschiedlich behandelst. Ja, (.) das ist mir schon eher aufgefallen.
277	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du die Mädchen und Jungen in ihrem Können so unterschiedlich einschätzt, da du so die Vorstellung förderst, Physik wäre ein Jungsfach. Auf jeden Fall, das geht gar nicht.
278	Vorlesen, Zustimmung AO	Du thematisierst die Einheit des Impulses nicht. Das ist nicht gut, da dies inkonsistent zur früheren Szene ist, in der du bei Paula auf die Notation der Einheit der Geschwindigkeit bestanden hast. Mhm[Ja].
279	Vorlesen, Zustimmung AO, Beschreibung, Alternative + Begründung	Es ist gut, dass du die Mädchen dazu aufforderst, ihre Tabelle zu korrigieren, damit du sicher sein kannst, dass sie die gleichen Werte, wie die Jungen notiert haben. Ja, an sich ist es schon gut, dass sie die- ihre Tabelle korrigieren, aber ich würde jetzt auch nicht sagen, dass alle Jungs hundert prozentig die Tabelle richtig haben, das heißt er könnte auch einfach sagen, alle sollen ihre Tabelle nochmal überarbeiten und für die, die es (jetzt?) schon gemacht haben, die können das machen, aber ich klicke es jetzt mal trotzdem an. #01:28:48-1#
280	Vorlesen	LD6: Welche der unten stehenden Hinweise zu seinem Umgang mit den Schülerinnen möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben?
281	Vorlesen, Ablehnung AO	Dein Umgang mit den Schüler*innen ist in Ord- Nein.
282	Vorlesen, Zustimmung AO + Begründung	Du solltest die Mädchen stärker in den Unterricht einbeziehen, (.) berechnete Werte vorlesen lesen lassen. Ja, die können das doch genauso gut.
283	Vorlesen, Zustimmung AO	Du solltest auf geschlechterspezifische Zuschreibungen verzichten. Ja.
284	Vorlesen, Ablehnung AO + Begründung	Du solltest von jedem Schülerin je nur einen berechneten Wert erfragen, da du so die berechneten Werte von mehreren Schülerinnen kontrollieren kannst. (.) Finde ich jetzt gar nicht mal so wichtig, dann ist es nämlich irgendwie auch ziemlich unordent- also dauert viel länger und ja.
285	Vorlesen, Zustimmung AO, Alternative	Du solltest verstärkt schwächere Schülerinnen in den Unterricht einbinden, da wichtig ist, dass sie die Inhalte verstehen. Ja, wenn er doch meint, dass die Mädchen das nicht können, kann er das doch einfach mal vorlesen lassen. (.)
286	Vorlesen, Zustimmung AO	Du solltest die Schülerinnen unabhängig vom Geschlecht dazu auffordern, falsche Ergebnisse in der Tabelle zu korrigieren. Ja, ja.
287	Vorlesen, Zustimmung AO	Du solltest hinterfragen, woher die geschlechterspezifischen Zuschreibungen stammen, die du vertrittst, um sie zu überwinden. Ja, ist wahrscheinlich schwierig, aber ich würde es auf jeden Fall ansprechen. #01:29:55-8#

288	Vorlesen	LD6: Welches Feedback möchten Sie Robert zu seinem Umgang mit Maries Einwand beziehungsweise mit Messunsicherheiten geben?
289	Vorlesen, Ablehnung AO	Die Reaktion auf Maries Einwand ist dir gut gelungen. Nein, kann schonmal (unv.).
290	Vorlesen, Ablehnung AO	Die Reaktion auf Maries Einwand ist dir gut gelungen. Nein.
291	Vorlesen	Die Reaktion auf Maries Einwand ist dir nicht gut gelungen, da du den Fokus von der experimentellen Bestätigung der Impulserhaltung hin zur Diskussion von leichten Ungenauigkeiten verschiebst.
292	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du nur so knapp auf den Einwand von Marie eingehst, da die Begründung nicht ausreicht, um den Schülerinnen zu verdeutlichen, dass die Unterschiede aufgrund von Messunsicherheiten zustande kommen. Ja.
293	Vorlesen, (Ablehnung AO) + Begründung	Es ist nicht gut, dass du nur so knapp auf den Einwand von Marie eingehst, da so unklar bleibt, ob beziehungsweise wo die Schülerinnen Rundungsfehler gemacht haben. Ja, Rundungsfehler schön und gut, aber das sind Messunsicherheiten.
294	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du die Ungleichheit der Messwerte mit Rundungsfehlern begründest. Ja. (.)
295	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du die Ungleichheit der Messwerte mit Rundungsfehlern begründest, da du so den Schülerinnen die Schuld daf- dafür gibst, dass die Ergebnisse nicht exakt gleich sind. (.) Ja, (einfach?) es ist (unv.) auch fachlich falsch. #01:31:01-6#
296	Vorlesen	LD6: Welche der unten stehenden Hinweise zu seinem Umgang mit Maries Einwand beziehungsweise mit Messunsicherheiten möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben?
297	Ablehnung AO	Ne, ist auf jeden Fall schonmal nicht angemessen.
298	Vorlesen	Du solltest den Schülerinnen angeben, auf wie viele Nachkommastellen sie runden sollen. (.)
299	Vorlesen, Ablehnung AO	Es wäre besser, den Schülerinnen am Anfang mitzuteilen, dass Messunsicherheiten und Rundungsfehler nicht berücksichtigt werden. Nein.
300	Vorlesen, Ablehnung AO	Es wäre besser, ausführlicher zu begründen, wo die Rundungsfehler entstehen. Nein.
301	Vorlesen, Zustimmung AO	Du solltest bereits bei der Vorstellung des Experiments über die vorhandene Reibung und Energieumwandlung beim Aufprall sprechen, damit den Schülerinnen in dieser Phase klar ist, dass die Ergebnisse aufgrund dieser Einflüsse nicht exakt gleich sein können. Ja.
302	Vorlesen, Zustimmung AO	Es wäre besser, zu thematisieren, dass es bei der Durchführung des Versuches Messunsicherheiten berücksichtigt werden müssen, damit die Schülerinnen nachvollziehen können, wieso es zu Abweichungen kommt. Ja.

303	Vorlesen, Überlegung zum Unterricht, Zustimmung AO	Du solltest das Thema Messunsicherheiten für dich selbst noch einmal aufbereiten, damit du von solchen Fragen nicht vom Thema abgebracht wirst. (.) Ich weiß ja nicht, ob die (Messunsicherheiten?) nicht thematisieren wollte, aber (vielleicht?) wenn er die nicht verstanden hat, sollte er es auf jeden Fall nochmal machen. #01:32:07-7#
304	Vorlesen	LD6: Welches Feedback möchten Sie Robert zu seinem Umgang mit der Einheit des Impulses geben?
305	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du nicht über die Einheit des Impulses sprichst, da die Einheit ein wichtiger Bestandteil ist. (.) Ja.
306	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du nicht über die Einheit des Impulses sprichst, da dies inkonsistent zur früheren Szene ist. Ja. (.)
307	Vorlesen, Ablehnung AO + Begründung	Es ist nicht äh Es ist gut, dass du die Einheit des Impulses nicht so stark in den Fokus gerückt hast. Ne, das ist (wichtig?).
308	Vorlesen, Ablehnung AO	Es ist gut, dass du den Fokus auf die berechneten Werte und nicht auf die Einheit gelegt hast. Nein.
309	Vorlesen, Ablehnung AO + Begründung	Es ist nicht gut, dass du die Schülerinnen nicht dazu aufforderst, die Einheiten anzugeben, da so unklar ist, ob die Impulse wirklich gleich groß sind. (.) Ne, die (Begründung?) würde ich jetzt auch mal nicht gelten lassen. #01:32:48-1#
310	Vorlesen	LD6: Welche der unten stehenden Hinweise zu seinem Umgang mit der Einheit des Impulses möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben?
311	Vorlesen, Ablehnung O	Es ist in Ordnung. (Nein?) (unv.) die Impuls Einheiten müssen (äh?) ja.
312	Vorlesen, Zustimmung AO	Du solltest die Einheiten des Du solltest die Einheit des Impulses angeben, da die Einheit ein wichtiger Bestandteil der Angabe ist, (..) ja.
313	Vorlesen, Zustimmung AO	Du solltest darauf achten, dass die Schüler beim Vorlesen der Werte auch die Einheit angeben, um in deinem eigenen Verhalten konsistent zu bleiben. Mhm[Ja].
314	Vorlesen, (Zustimmung AO) + Begründung, Alternative	Du solltest die Einheit des Impulses oben über die Spalte der Tabelle notieren, um die Schreibarbeit in jeder Zeile zu sparen. Irgendwie muss ja auf jeden Fall (eine?) erwähnt werden. Sei es in der Spalte, äh in der o- über der Spalte.
315	Vorlesen, Zustimmung AO	Du solltest die Schülerinnen dazu auffordern, die Einheit des Impulses zu nennen, damit die Schülerinnen sich an die Einheit des Impulses gewöhnen. Ja. (.)
316	Sonstiges	(Vier Stück?) (Anm: Vermutlich ist die Anzahl der angekreuzten AOs gemeint). #01:33:49-7#
317	Überlegung zum Unterricht	LD6: (Dann?) letzte Phase, (Sicherungsphase?). Jetzt kommt bestimmt wieder // (unv.). #01:33:58-2#
318	-	[Video Einleitung Reflexion 07] #01:34:10-2#

319	Überlegung zum Unterricht	LD6: Bestimmt wieder vorgefertigt auf dem Beamer. #01:34:14-0#
320	-	[Video Szene Reflexion 07] #01:35:51-0#
321	Sonstiges	LD6: Diese Merksätze von ihm. (.)
322	Vorlesen	Zu welchen der folgenden Aspekte sollen sollte Robert eine Rückmeldung erhalten, um etwas für sich beziehungsweise seinen Unterricht zu lernen?
323	Vorlesen, Bewertung (3x), Alternative	Formulierung des Merksatzes zur Impulserhaltung. An sich ist er fachlich ja richtig, aber warum warum macht er das nicht mit den Schülerinnen und Schülern zusammen? (..) Also Lob der äh Merksatz ist gut, aber Kritik die Interaktion mit den Schülerinnen, warum macht er das nicht mit denen zusammen. #01:36:15-8#
324	Vorlesen	LD6: Robert hat zwar schon einige Rückmeldungen erhalten, würde aber trotzdem gern zu den beiden aufgeführten Aspekten noch eine Rückmeldung erhalten. #01:36:23-3#
325	Vorlesen	LD6: Welches Feedback möchten Sie Robert zu seiner Formulierung des Merksatzes zum Impuls zur Impulserhaltung geben? (.)
326	Vorlesen, Zustimmung AO	Die Formulierung des Merksatzes ist dir gut gelungen, da alle wichtigen Informationen korrekt enthalten sind. (.) Ja, an sich schon, stimmt ja.
327	Vorlesen	Es ist gut, dass du im Merksatz auf das durchgeführte Experiment verweist, da die Schülerinnen den Merksatz direkt mit der Realität verbinden können.
328	Vorlesen, Ablehnung AO	Die Formulierung des Merksatzes ist dir nicht gut gelungen, da sie für die Schülerinnen zu anspruchsvoll ist. Mhm würde ich nicht sagen.
329	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du deinen vorgefertigten Merksatz nutzt und die Merksätze der Schülerinnen nicht einbeziehst. Ja, auf jeden Fall. #01:36:57-2#
330	Vorlesen	LD6: Welche der unten stehenden Hinweise (.) zu seiner Formulierung des Merksatzes zur Impulserhaltung möchten Sie Robert für eine erneute Durchführung in der Unterrichtsstunde geben?
331	Vorlesen, Bewertung, Alternative	Die Formulierung des Merksatzes ist angemessen, du kannst sie erneut verwenden. (.) (Ja?) die Formulierung ist (an sich?) gut, aber die kann man ja auch erarbeiten mit den Schülerinnen und Schülern (unv.) so ein Ding mit dem Merksatz.
332	Vorlesen, Ablehnung AO	Es wäre besser, Beispiel und Merksatz zu trennen. (..) Ist wie am Anfang, finde ich jetzt nicht so wichtig.
333	Vorlesen, Zustimmung AO + Begründung	Es wäre besser, zusätzlich zum langen Merksatz eine kurze Version zu formulieren, da sich die Schülerinnen eine kurze Version besser merken können. Ja, die Schüler merken sich wahrscheinlich eh nur die Formel hinten.

334	Vorlesen	Es wäre besser, den Merksatz nur in Bezug auf einen Stoßprozess zu formulieren, damit der Merksatz näher am besprochenen Experiment bleibt.
335	Vorlesen, Zustimmung AO	Wäre besser, den Merksatz um die Einheit des Impulses zu ergänzen, da die Einheit stets- Ja.
336	Vorlesen, Zustimmung AO, Sonstiges	Du solltest den Merksatz schulnäher gestalten, zum Beispiel indem du eine Skizze ergänzt- oh ja, Skizze, das wäre natürlich auch gut. (..) Dann mache ich das oben mal raus (LD6 entfernt das Kreuz bei AO 'Die Formulierung des Merksatzes ist angemessen, du kannst sie erneut verwenden.').
337	Vorlesen, Alternative, Sonstiges	Es wäre besser, eine Formulierung der Schülerinnen als Merksatz zu verwenden, da diese Formulierungen für sie besser verständlich sind. Es muss ja noch nicht mal eine Formulierung der Schülerinnen direkt sein, aber halt auf deren Grundlage einfach noch ein bisschen was ergänzen an der Tafel. (...) Deshalb (.) ich lasse es mal drin. #01:38:19-9#
338	Vorlesen	LD6: Welches Feedback möchten Sie Robert zu seiner Interaktion mit den Schülerinnen geben?
339	Vorlesen, Alternative	Es ist gut, dass die Schüler*innen ihre Merksätze vorlesen, da du so erkennen kannst, ob sie die Impulserhaltung umfassend verstanden haben. Und sie auch benutzen.
340	Vorlesen, Zustimmung AO	Es ist gut, dass die Schülerinnen einen Merksatz formulieren sollen, um ein tiefergehendes Verständnis zu fördern. (...) Ja. (.)
341	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du den Schülerinnen keine Rückmeldung auf die Merksätze gibst, da die Schülerinnen so nicht einordnen können, ob ihre Merksätze angemessen formuliert sind. Ja, das ist so.
342	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du deinen vorgefertigten Merksatz nutzt und die Merksätze der Schülerinnen nicht einbeziehst. Ja.
343	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du deine Formulierung als besser be- Ja mhm[ja]. (.)
344	Vorlesen, Alternative + Begründung	Es ist nicht gut, dass du nur zwei Schülerinnen ihre Merksätze vorgestellt haben, da du so nicht weißt, welche Merksätze die restlichen Schülerinnen formuliert haben. Das könnte man ja umgehen, indem er einfach Gruppenarbeit macht und dann die (Formulierung?) zu viert, keine Ahnung, dann kann man alle vorlesen. (..) Ja.
345	Vorlesen, Zustimmung AO, Überlegung zum Unterricht	Es ist nicht gut, dass nur zwei Jungs ihre Merksätze vorgestellt haben. Ja, das (ist?) habe ich mich auch gefragt, warum nimmt er immer nur Jungen dran. #01:39:27-4#
346	Vorlesen	LD6: Welche der unten stehenden Hinweise zu seiner Interaktion mit den Schülerinnen möchten Sie Robert für eine erneute Durchführung der Unterrichtsstunde geben?

347	Vorlesen, Sonstiges	Du solltest den Merksatz ausführlicher erklären beziehungsweise mit den Schülerinnen besprechen, da einige Aspekte des Merksatzes nicht selbsterklärend sind. Wenn man das erarbeiten (unv.), (unv.) gibt es das gar nicht.
348	Vorlesen, Zustimmung AO	Du solltest gemeinsam mit den Schülerinnen einen Merksatz erstellen, um sie stärker in den Unterricht einzubeziehen. Ja. (.)
349	Vorlesen, Zustimmung AO, Alternative + Begründung	Es wäre besser, insgesamt mehr Schülerinnen ihre Merksätze vorstellen zu lassen, da du aus den Antworten Rückschlüsse auf die Vorstellungen der Schüler*innen ziehen kannst. Ja, oder halt in Gruppen, das man halt die von (jeder?) Gruppe (...) kennt.
350	Vorlesen, Zustimmung AO, Beschreibung, Alternative	Du solltest die Schülerinnen sich melden lassen, wenn sie ihren Merksatz vorstellen wollen. (.) Ja, zum Beispiel. Oder halt auch direkt Mädchen dran nehmen. Generell melden sich da richtig wenige, er nimmt einfach immer dran.
351	Vorlesen, Ablehnung AO	Der Umgang mit den Schülerinnen war angemessen, so kannst du dich wieder verhalten. Nein.
352	Vorlesen, Zustimmung AO	Du solltest deinen Umgang mit den Schülerinnen überdenken. Dass du sie und ihre Arbeit nicht lobst, kann dem Verhältnis der Schülerinnen zum Physikunterricht nachhaltig schaden. (.) Mhm[Ja].
353	Vorlesen, Zustimmung AO	Du solltest darauf aufpassen, deine Rollenklischees äh keine Rollenklischees in den Unterricht einzubringen, um die Schülerinnen nicht in diese Muster zu drängen. Ja. #01:40:46-0#
354	Vorlesen	LD6: Hier geht es um den Unterricht insgesamt. (.)
355	Vorlesen	Zu welchen (.) der folgenden Aspekte sollte Robert eine Rückmeldung erhalten?
356	Bewertung (2x), Beschreibung (2x)	(unv.) mhm Unterrichtsziele, die sind schon ziemlich anspruchsvoll, (unv.) ziemlich viel für eine Unterrichtsstunde hat er ja (..) Wechselwirkung- äh ne Newtonsche Axiome gemacht, dann Impuls und dann noch In- Impulserhaltung plus verschiedenste Experimente, das ist schon ziemlich viel. (..)
357	Alternative	Ja, Auswahl, Umfang der Stundeninhalte (halte ich?) (unv.) viel (unv.) hätte er (lieber?) mit den Schülerinnen und Schülern das irgendwie zusammen erarbeitet und einfach weniger gemacht. (.)
358	Vorlesen, Bewertung	Medieneinsatz, ich finde diese PowerPoint, dass er da diese fe- vorgefertigen Sachen hat, einfach echt nicht gut. (.)
359	Vorlesen, Bewertung	Umgang mit den Schülerinnen und Schülern (.) schrecklich, wirklich schrecklich. (.)
360	Vorlesen, Bewertung	Fachliche Korrektheit war größtenteils okay, aber mit dem Schwung war jetzt, dass er das dann auch benutzt hat, fand ich jetzt auch nicht so gut.

361	Vorlesen, Bewertung	Auftreten der Lehrkraft würde ich erstmal (.) war okay. (.)
362	Vorlesen, Bewertung, Zustimmung AO + Begründung	Ablauf der Unterrichtsstunde war auch okay. (...) Ja, wobei, (das?) ja, die Interakation ist ja hier schon mit drin.
363	Vorlesen, Zustimmung AO	Unterrichtsziele (.) ja, könnte man eigentlich auch (gut?) erkennen. #01:42:09-6#
364	Vorlesen, Sonstiges	LD6: Besonders wichtig ist ihm jetzt noch eine Rückmeldung zu seinem Umgang mit den Schülerinnen und der fachlichen Korrektheit. Ja. #01:42:17-9#
365	Vorlesen	LD6: Feedback zu der Interaktion.
366	Vorlesen, Zustimmung AO, Beschreibung	Es ist nicht gut, dass du die Beiträge von Schüler*innen kaum berücksichtigst. (4s) Ja, wenn sie mal (ihren?) Merksatz vorgelesen haben und so weiter, ja, generell schon immer ziemlich viel vorgefertigt.
367	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du so lehrerzentriert unterrichtest. Ja, (.) mir auch aufgefallen.
368	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du die Jungen und Mädchen unter- Ja mhm[ja].
369	Sonstiges	Ich habe (ja?) alles angeklickt.
370	Vorlesen, Zustimmung AO, Überlegung zum Unterricht, Bewertung, Alternative, Sonstiges	Es ist gut, dass du Sicherungsphasen in deinen Unterricht eingebaut hast, damit du sichergehst. Ja, Sicherungsphasen sind an sich gut, (.) aber er muss sie halt auch (4s) halt nicht nur abschreiben, sondern er sollte (unv.) mit (denen?) richtig erarbeiten. Ich glaube, ich lasse es dann einfach mal raus. Also an sich ist eine Sicherung gut, dass er das eingebaut hat, aber halt anders.
371	Vorlesen, Überlegung zum Unterricht, Zustimmung AO, Bewertung + Begründung	Es ist gut, dass du verschiedene Sozialformen im Unterricht einplanst, Einzel-, Partnerarbeit. Hatte er Einzelarbeit? (.) (unv.) (.) ja stimmt, ach doch, bei der Formulierung der Merksätze wahrscheinlich. Ja, das ist doch auf jeden Fall gut, ein bisschen Abwechslung tut gut.
372	Vorlesen, Zustimmung AO, Alternative + Begründung	Es ist nicht gut, dass du im Unterricht nur von wenige Schülerinnen Beiträge anhörst. Mhm[Ja] ja, wobei, er kann ja auch nicht so viele Beitrage, (denn?) wenn dann müsste man irgendwie Partnerarbeit oder Gruppenarbeit machen, dann kann man das von vielen- allen kann man ja eh (nicht?) immer nicht so gerecht werden. (.)
373	Vorlesen, Zustimmung AO + Begründung	Deine Verständnisversicherung mit einer Frage, wie 'Habt ihr noch Fragen?' ist nicht zielführend, da Schülerinnen auf diese Frage eher nicht antworten. Ja, das stimmt schon, das ist so, (..) dann bekommt er zumindest (unv.) Antworten.
374	Vorlesen, Zustimmung AO	Du thematisierst die auftretenden Schülervorstellungen nicht. Das ist nicht gut, da sich diese so verfestigen. Ja. #01:44:01-5#
375	Vorlesen	LD6: Welche der unten stehenden Hinweise zu seiner Interaktion mit den Schülerinnen möchten Sie Robert für eine erneute Durchführung geben?

376	Ablehnung AO	Ne, angemessen ist (sie?) auf jeden Fall nicht.
377	Vorlesen, Zustimmung AO	Du solltest die Schülerinnen grundsätzlich stärker in den Unterricht einbeziehen. (.) Ja, (finde?) ich schon.
378	Vorlesen, Zustimmung AO, Alternative	Du solltest den Schülerinnen ausführlicher Feedback zu ihren Beiträgen geben. (.) Ja, nicht immer nur ja, okay, ja.
379	Vorlesen, Zustimmung AO, Alternative	Du solltest dir die Schülervorstellungen zum Thema vor dem Unterricht ansehen. Ja, vor allem solltest du sel- (die?) selbst nicht benutzen.
380	Vorlesen, Überlegung zum Unterricht	Du solltest dir grundsätzlich Gedanken darüber machen, wie das Verständnis der Schülerinnen abprüfen kannst, um sicherzugehen, dass die Schülerinnen die Ziele der Stunde erreichen. (.) (unv.) das bin ich mir gerade gar nicht mal so sicher, weil (die?) ja vor allem bei der Erarbeitung waren. (.)
381	Vorlesen, Bewertung	Du solltest keinen- deinen- Und das Quiz hat ja am Anfang war ja auch nicht so schlimm.
382	Vorlesen, Zustimmung AO	Du solltest deinen Umgang mit den Schülerinnen grundsätzlich überdenken. Ja. (.)
383	Vorlesen, Zustimmung AO	Es ist wichtig- Ja, Rollenklischees (unv.). #01:45:00-4#
384	Vorlesen (2x)	LD6: Deinen Un- Achso welches Feedback möchten Sie Robert zur fachlichen Korrektheit der Stunde geben?
385	Vorlesen	Deinen Unterricht hast du insgesamt gut durchgeführt, da er überwiegend fachlich korrekt war.
386	Vorlesen, Bewertung	Du hast fachliche Fehler in der Unterrichtsstunde gemacht. Das ist nicht gut, da du den Schülerinnen so falsche Inhalte beibringst. (..) (unv.) vielleicht bis auf diesen Schwung eigentlich auch gar nicht richtig erkannt.
387	Vorlesen, Sonstiges	Es ist nicht gut, dass du die Ab- achso doch, Messunsicherheiten.
388	Vorlesen + Zustimmung AO	Es ist nicht gut, dass du die Abweichungen beim experimentellen Nachweis der Impuls- Ja. Das ist fachlich falsch, (ja?).
389	Vorlesen, Zustimmung AO	Es ist nicht gut, dass du die Einheit des Impulses nicht besprichst. Auch richtig.
390	Vorlesen, Zustimmung AO, Sonstiges	Du thematisierst die auftretenden Schülervorstellungen- vorstellungen nicht. Das ist nicht gut, da sich diese so verfestigen. (.) Auch richtig, ups, dann kann ich doch mehr sagen, (die?) fachlich falsch (sind?).
391	Vorlesen, Zustimmung AO, Alternative	Es ist gut, dass du so großen Wert auf die Einheit der Geschwindigkeit in der in der Tabelle legst, da die Einheit zentraler Bestandteil einer Größen-angabe ist. Ja, aber dann sollte er auch schon konsequent bleiben.
392	Vorlesen, Zustimmung AO, Verweis auf frühere Stelle im Test	Es ist nicht gut, dass du bei der Durchführung des Demonstrationsexperimentes sagst, du würdest exakt messen. Ja. Habe ich doch gerade schonmal angekreuzt. (...)

393	Sonstiges* (2x)	Das (LD6 zeigt auf AO 'Du hast fachliche Fehler in der Unterrichtsstunde gemacht. Das ist nicht gut, da du den Schüler*innen so falsche Inhalte beibringst.') lasse ich jetzt mal unangekreuzt, weil so so (wichtig?) das (unv.) hier (Anm: in den anderen AOn) alles schon drin ist, (dann lasse?) ich das mal unangekreuzt. #01:46:13-3#
394	Vorlesen	LD6: Welche der unten stehenden Hinweise zur fachlichen Korrektheit der Stunde möchtest du Robert für (seine?) Unterrichtsstunde geben?
395	Vorlesen, Zustimmung AO	Du solltest auch bei der Besprechung der Impulse auf die Einheiten achten. Ja.
396	Vorlesen, Zustimmung AO	Es wäre besser, die Schülervorstellungen, die im Unterricht auftreten, aufzugreifen. Mhm[Ja]. (.)
397	Vorlesen, Zustimmung AO	Du solltest dir die Schülervorstellungen zum Thema vor dem Unterricht angucken. Ja.
398	Vorlesen, Zustimmung AO	Es wäre besser, auch auf Messunsicherheiten beim Fahrbahn-Experiment einzugehen, damit die Schülerinnen die Impulserhaltung trotz der leicht unterschiedlichen Werte akzeptieren. Auf jeden Fall, ist mega wichtig.
399	Vorlesen, Überlegung zum Unterricht	Du solltest darauf achten, dass deine Merksätze fehlerfrei formuliert sind. (unv.) (waren?) die doch eigentlich, (..) ich habe zumindest jetzt keine größeren Fehler (beziehungsweise?) gar keine Fehler erkannt.
400	Vorlesen, Zustimmung AO + Begründung	Du solltest stärker auf deine Wortwahl achten, um nicht ausversehen falsche Vorstel- Ja, Schwung, dass er Schwung benutzt hat, das war nicht so gut. #01:47:04-1#
401	-	[Video Prompt Schulnote] #01:47:11-8#
402	Beschreibung, Bewertung (7x), Alternative.	LD6: Keine gute, (.) (mir?) überwiegt halt irgendwie total, dass er so mit den Schülern und Schüler- Schülerinnen und Schülern umgegangen ist. (.) Das ist ja dieses diese Klischees, also (unv.) weiß ich nicht, das wirkt sich bei mir irgendwie super negativ aus. (.) Dann (..) die ja Fehlvorstellung Schwung hat er, dann (kamen?) die Messunsicherheiten nicht, also irgendwie weiß ich nicht, also mega gut fand ich das auf jeden Fall nicht, das war einfach viel zu viel für die Unterrichtsstunde. (.) Dann lieber halt weniger und (.) mehr mit den Schülerinnen und Schülern erarbeiten, ich glaube, ich würde ihm (.) eine Vier (geben?) (.) eine Drei? (.) Was war denn gut? (..) An sich die Experimente, die (Auswahl?) war schon okay. (...) Aber auch, wie er das in den Merk- ich glaube, ich würde ihm einfach eine Vier geben. #01:48:08-1#

O.7 Interview 7: LD7

Segment	Kodierung	Transkript
0	-	[Einverständniserklärung, Demographie] #00:05:15-9#
1	-	I: Okay, das ist die Seite, wo es dann gleich inhaltlich los geht. Ähm genau, ich würde gern einfach nochmal diesen ja quasi Arbeitsauftrag, also diese Sprechaufforderung nochmal vorlesen und dann würde ich mich wieder stumm schalten (.) // zum- #00:05:34-4#
2	Sonstiges	LD7: Okay (.) äh ich habe auch nochmal eine kurze Frage #00:05:35-8#
3	-	I: Ja. #00:05:36-4#
4	Sonstiges	LD7: Die Aufgaben muss ich jetzt aber nicht laut vorlesen, die also das zählt dann nicht zum lauten Denken, ne? #00:05:41-4#
5	-	I: Ja, jein tatsächlich, also- #00:05:43-1#
6	Sonstiges	LD7: Okay. #00:05:43-7#
7	-	I: Ähm genau das ist tatsächlich hilfreich, wenn du- also diese diese längere Situationsbeschreibung hier, die ist nicht super entscheidend für mich, aber später kommen so kürzere Aufgaben mit verschiedenen Antwortoptionen und wenn du die so ein bisschen mit vorliest, hilft mir das halt total, mitzuverfolgen, wo du bist, also mir ist da jetzt nicht wichtig, dass jedes Wort richtig ausgesprochen vorgelesen wird, sondern (.) so ein bisschen mitsprechen, mitlesen, wo du bist quasi. #00:06:10-5#
8	Sonstiges	LD7: Okay mhm[ja]. #00:06:12-2#
9	-	I: Genau, aber das äh genau sage ich jetzt quasi auch nochmal. #00:06:15-0#
10	Sonstiges	LD7: Okay. #00:06:15-8#
11	-	I: Aber gut, dass du fragst. Ähm (.) genau, also es ist so, dass während der Sonstiges gleich immer wieder Unterrichtsszenen zu sehen sind und zu denen werden dann im Anschluss fragen gestellt und meine Bitte wäre dann, dass du da alle Gedanken während der Bearbeitung dieser Fragen beziehungsweise Aufgabenstellung immer laut aussprichst und wenn du gerade einen Text oder so eine Aufgabenstellung liest, da auch mit vorliest und es geht, wie vorhin schon gesagt, es geht überhaupt nicht darum, mir zu erklären, was du denkst, sondern #00:06:46-7#
12	Sonstiges	LD7: Mhm[Ja]. #00:06:47-1#
13	-	I: die Vorstellung ist eher, du wärst alleine in einem Raum, würdest zu dir selber sprechen ähm // genau. #00:06:52-8#
14	Sonstiges	LD7: Okay. #00:06:52-6#
15	-	I: Und wie eben auch schon, wenn du aufhörst zu sprechen, dann würde ich mich melden, würde irgendwie was sagen oder // nachfragen und ansonsten schalte ich mich stumm. #00:07:04-4#
		LD7: Mhm[Ja]. #00:07:01-6#

16	Sonstiges (2x)	LD7: Alles klar. #00:07:06-2#
17	-	I: Gibt es noch Fragen? #00:07:09-3#
18	Sonstiges	LD7: Äh, also mitschreiben muss ich auch immern- och nichts, ich muss einfach nur hier ankreuzen, okay. #00:07:14-8#
19	-	I: Ne, mitschreiben musst du // gar nicht, ne. Alles klar. #00:07:18-1#
20	Sonstiges (2x)	LD7: Alles klar. #00:07:16-5#
		LD7: Okay, mhm[ja]. (18s) mhm (LD7 liest Situationsbeschreibung). #00:07:40-7#
21	-	[Video Beschreibung Lerngruppe, Video Vorwissen, Video Unterrichtsziele, Video Stundenraster] #00:12:10-8#
22	Vorlesen, Sonstiges	LD7: (7s) Schülerinnen und Schüler beantworten Quizfragen, moderiert Besprechung der Quizfragen, Vertiefung mit einem Comic, (8s) Sicherung und Erarbeitung zum Thema Impuls, (5s) dann Experiment (.) und in der Erarbeitung kommt auch wieder ein Experiment, (12s) und nachher werden die Experimente ausgewertet, (..) in einem Plenumsgespräch und mit Partnerarbeit (..) und zum Schluss machen sie einen Merksatz mhm. (5s) #00:13:24-4#
23	Vorlesen, Sonstiges	LD7: Habe die Stunde mit der Vorstellung des Ablaufs begonnen, als nächstes habe ich ein Quiz zur Wiederholung angeleitet, welches die Schüler- innen in Partnerarbeit durchgeführt haben. Die Durchführung des Quiz selbst habe ich zwar nicht gefilmt, dafür aber die Besprechung des Quiz mit den Schüler- innen, die ich dir jetzt zeige. (unv.) (5s) mhm. #00:13:a52-1#
24	-	[Video Einleitung Reflexion 01, Video Szene Reflexion 01] #00:16:28-0#
25	Vorlesen	LD7: Zu welchen Aspekten sollte Robert Rückmeldung erhalten, um etwas für sich zu lernen?
26	Ablehnung AO	Mhm (.) also das zum ersten Axiom ist mir jetzt nichts groß aufgefallen,
27	Zustimmung AO	aber die Formulierung vom zweiten Axiom müsste er irgendwie Rückmeldung zu erhalten (.)
28	Zustimmung AO + Begründung	und ich finde bei der Gesprächsführung auch, war immer sehr kurz die Gesprächsanteile der Schülerinnen und Schüler. (6s). #00:17:06-3#
29	Vorlesen	LD7: (6s) Rückmeldung zur Formulierung zum dritten Axiom. #00:17:16-1#
30	Vorlesen	LD7: Mhm (...) mhm (6s) nicht angemessen, (..)
31	Vorlesen	Vermischung von Beipiel und Merksatz, (..)
32	Beschreibung	im Video selber hat er den Merksatz ja gar nicht nochmal vorgelesen. (...)
33	Vorlesen	Es ist nicht gut, dass nur die Kurzformel actio- (5s)

34	Zustimmung AO + Begründung	Also ich finde auf jeden Fall auch gut, dass die Kräfte an verschiedenen Körpern angreifen, (.) da haben, glaube ich, viele Schülerinnen und Schüler Schwierigkeiten mit. (..)
35	Ablehnung AO + Begründung	Mhm (4s) und das mit dem Beispiel hier drin, finde ich eigentlich gut, weil im Merksatz selber hatte er (.) ein Beispiel genannt, was den Schülerinnen und Schülern hilft, das irgendwie einzuordnen. (.)
36	Überlegung zum Unterricht	Mhm (5s) das kann ich jetzt irgendwie hier nicht einschätzen, ob die Schülerinnen und Schüler Kraft gleich Gegenkraft dann irgendwie zu oft mit einem Kräftegleichgewicht assoziieren. #00:18:37-1#
37	Vorlesen	LD7: Mhm (.) welche Hinweise zur Formulierung des Merksatzes zum dritten Axiom möchten Sie Robert für die erneute Durchführung des Unterrichts geben? (5s)
38	Zustimmung AO, Alternative	Äh erst das Axiom, dann ein Beispiel, das kann auch gut sein, aber vielleicht ist es auch mit dem Beispiel direkt ein bisschen anschaulicher, (.) aber dann sollte das Beispiel, finde ich, schon ausführlicher sein, (.) damit man das besser verstehen kann. (.)
39	Zustimmung AO	Und eine Skizze im Axiom finde ich auch sinnvoll. (3s)
40	Alternative + Begründung	Aber über diese Formulierung mit anschließend übt der Körper B die Gegenkraft auf, die ist mir auch aufgefallen, das muss er, glaube ich, auch besser machen, damit man sieht, dass die Kräfte gleichzeitig wirken. (4s)
41	Ablehnung AO + Begründung, Zustimmung AO	Mhm das mit den Vektorpfeilen über den Kräften, (..) finde ich ein bisschen schwierig, damit die Wirkrichtung deutlich wird, weil wenn man die über die Buchstaben schreibt, sind die Vektorpfeile eh in die gleiche Richtung, (...) aber die Verdeutlichung von den entgegengesetzten Richtungen finde ich auch sinnvoll. #00:19:55-3#
42	-	[Video Einleitung Reflexion 02] #00:20:30-6#
43	Überlegung zum Unterricht	LD7: Okay, also jetzt geht es um den (.) Comic zum dritten Axiom. #00:20:35-1#
44	-	[Video Szene Reflexion 02] #00:22:23-9#
45	Zustimmung AO + Begründung	LD7: (6s) Mhm also Robert müsste auf jeden Fall (.) Rückmeldung zu den Schülerfehlvorstellungen erhalten, weil er gar nicht darauf eingeht, wenn irgendwie was gesagt wird mit es hat Kraft. (..)
46	Bewertung	Gesprächsführung fand ich eigentlich in Ordnung. #00:22:46-7#
47	-	LD7: (5s, liest 'Robert hat schon einige Rückmeldungen erhalten...') #00:22:57-0#
48	Vorlesen	LD7: Welche der folgenden Schülerfehlvorstellungen sind während des Unterrichtsausschnitts sichtbar geworden?
49	Ablehnung AO	Also die erste Antwort stimmt ja schon nicht. (.)
50	Vorlesen, Zustimmung AO	Mhm (.) die Schülerinnen und Schüler (unv.) schreiben dem Apfel Kraft als Eigenschaft zu, das auf jeden Fall. (.)

51	Vorlesen	Und sie denken auch, dass nur aktive, also sich bewegende Körper Kraft ausüben können, das (hat?) man daran gesehen, dass sie gesagt haben, der Baum bewegt sich nicht, also übt er keine Kraft aus. (5s)
52	Überlegung zum Unterricht, Beschreibung	Mhm das mit der Beschleunigung bin ich mir gerade nicht so sicher, weil irgendwie das Wort Beschleunigung gar nicht gefallen ist. (.)
53	Zustimmung AO + Begründung	Ähm (..) aber die Schülerinnen und Schüler verwechseln Kraft und Gegenkraft mit den wirkenden Kräften im Kräftegleichgewicht, (..) weil sie sich da nicht sicher waren, also sie meinten, dass der Baum die Erdanziehungskraft ausgleichen muss. (5s)
54	Vorlesen, Zustimmung AO	Und insgesamt sehen die Schülerinnen und Schüler nicht, dass der Baum mit der Erde verbunden ist und der Baum durch das Festhalten an der Erde eine Gegenkraft zum am Baum ziehenden Apfel ausübt, das finde ich auch. #00:24:17-1#
55	Vorlesen	LD7: (Die?) unten stehenden Hinweise zum Umgang mit Schülerfehlvorstellungen möchten Sie Robert für eine erneute Durchführunge geben? (.)
56	Sonstiges*	Äh (6s) der Prozess der Kraftwirkung, ich bin mir jetzt nicht ganz sicher, was damit gemeint ist. (..)
57	Zustimmung AO + Begründung	Aber der Unterschied zwischen Kraft und Energie sollte thematisiert werden, damit die Schülerinnen und Schüler nicht sagen, dass der Apfel (hier?) eine Kraft hat. (..)
58	Vorlesen, Zustimmung AO + Begründung	Erklären oder verdeutlichen, dass Kräfte auch dann wirken können, wenn man nichts aktiv tut, dazu das Rollbrettexperiment, (..) ja das finde ich sinnvoll, damit sie auch merken, dass nicht nur aktive Körper Kräfte ausüben. (.)
59	Vorlesen, Zustimmung AO + Begründung	Verwendeten Begriffe und den Unterschied zwischen dem Kräftegleichgewicht und Axiom erklären, (..) das finde ich auch gut, damit die Schülerinnen und Schüler Kräftegleichgewicht und das dritte Axiom nicht mehr verwechseln. (4s)
60	Zustimmung AO + Begründung (2x)	Und vielleicht ist es auch sinnvoll, wenn die Schülerinnen und Schüler das selber einmal formulieren, damit man sehen kann, was sie verstanden haben und was nicht, das finde ich auch gut, dann hat man es auch nochmal von allen Schülerinnen und Schülern und nicht nur von denen, die sich im Unterricht melden. #00:25:38-9#
61	Vorlesen, Beschreibung, Sonstiges	LD7: (4s) Schülervorstellungen an der Stelle, (..) also die sollen sich entscheiden für eine Seite (.) mhm[ja]. (4s) (unv.) das wieder die Unterrichtsphase (...) mhm. #00:26:04-8#
62	-	[Video Einleitung Reflexion 03, Video Szene Reflexion 03] #00:28:16-5#
63	Vorlesen	LD7: Hier ist wieder die Rückmeldung, was Robert dann (.) die Robert brauchen kann, um etwas für seinen Unterricht zu lernen. (.)

64	Bewertung, Alternative	Bei der Notation der Messwerte hat er aber darauf geachtet, dass die Einheiten mit notiert werden, das fand ich gut. (...) ich weiß nicht, ob man da irgendwie nochmal drüber sprechen müsste. (.)
65	Zustimmung AO, Überlegung zum Unterricht, Bewertung	Mhm vielleicht müsste er auch über die Fachsprache, also ich weiß nicht, ob das dann nachher vielleicht noch kommt, aber wenn er erstmal das immer bei Schwung belässt, obwohl er den Impuls einführen will, finde ich ein bisschen schwierig.
66	Vorlesen, Beschreibung, Bewertung	Und die Übergänge von den Unterrichtsphasen, wenn er erst das mit der Kraft erklärt und dann irgendwie so einen Cut macht, wo sogar noch der Apfelbaum (da?) an der Tafel war und dann einfach einen Cut zum Impuls macht, (ist?) irgendwie (.) kommt mir auch ein bisschen komisch vor. #00:29:08-9#
67	-	LD7: (6s, LD7 liest 'Robert hat schon einige Rückmeldungen erhalten...') #00:29:17-1#
68	Vorlesen	LD7: Feedback für Robert zur Nutzung der Fachsprache. (...)
69	Vorlesen, Ablehnung AO, Alternative	Es ist nicht gut, dass die Schüler- innen den Begriff Schwung nutzen, da die Bedeutung des Begriffs noch nicht geklärt ist und so unklar ist, was genau die Schülerinnen beschreiben. Ich finde, wenn das ganz am Anfang passiert, ist es ja an sich nicht so schlimm, aber irgendwie sollte es dann (.) umgewandelt werden,
70	Vorlesen, Zustimmung AO	also hier, es ist gut, dass die Schüler- innen zunächst ihre Alltagssprache verwenden können, (..) das finde ich sinnvoll, dann könnten die hier steht dann können die ihre Ideen vielleicht besser äußern. (...)
71	Ablehnung AO + Begründung	Aber ob der Begriff Schwung so wirklich tragfähiges Konzept für den Impulsbegriff ist, weiß ich nicht, (.) also mit so Alltagsbegriffen ist, glaube ich, immer ein bisschen schwierig. (...)
72	Zustimmung AO + Begründung	Und das mit dem stärker bremsen sehe ich auch ein bisschen so, theoretisch könnte, also je nachdem, wie gut die Bremseinrichtungen sind, könnte ja die Titanic bestimmt auch stärker bremsen. (.)
73	Vorlesen, Zustimmung AO + Begründung, Alternative, Sonstiges	Mhm (5s) aber hier, es ist nicht gut, dass du die Schülerinnen in der Benutzung des Begriffs Schwung bestärkst, das finde ich sinnvoll, (...) damit man nicht irgendwie dahin abrutscht, dass dann Schwung irgendwie gleich Kraft ist oder sonst irgendwas, also wenn sie das am Anfang machen, finde ich das nicht schlimm, aber es sollte schnell in die fachlich korrekte (.), also in den Fachbegriff umgewandelt werden. (4s)
74	Überlegung zum Unterricht, (Zustimmung AO) + Begründung	Das mit dem stockend Reden und vielen Füllwörtern ist mir jetzt irgendwie nicht ganz so aufgefallen, also wenn das der Fall ist, dann kann das die Schüler- innen schon vom Inhalt ablenken. #00:31:03-3#

75	Vorlesen	LD4: Und noch zur Nutzung von Fachsprache für die erneute Durchführung des Unterrichts, (5s)
76	Vorlesen, Zustimmung AO	hier, finde ich auch, es wäre besser, die Begriffe Masse und Impuls (.) gegebenenfalls mit den Formeln einzuführen, damit die Schülerinnen und Schüler die korrekte Fachsprache lernen, das finde ich sehr wichtig. (5s)
77	Zustimmung AO + Begründung	Das mit der Aufforderung zu mehr Fachbegriffen finde ich auch gut, damit sie dann sich die Fachsprache auch mehr angewöhnen. (..)
78	Zustimmung AO + Begründung, Überlegung zum Unterricht, Alternative	Und der Impulsbegriff sollte auch eher eingeführt werden, (.) damit die erstens wissen, worum es geht und auch schneller die physikalisch fachlich richtigen Begriffe benutzen. (7s) Wenn die Schülerinnen und Schüler Schwung für Impuls sehen, dann ist die Vorstellung von Schwung haben ja eigentlich nicht so falsch (...) (und?) wenn die den Schwung natürlich mit Kraft gleichsetzen irgendwie schon, also muss man da auf jeden Fall aufpassen. (..)
79	Zustimmung AO	Mhm (5s) aber hier die Verwendung der Fachsprache finde ich auf jeden Fall gut (..) und sinnvoll. #00:32:24-7#
80	Vorlesen	LD4: Und jetzt kommt das Feedback zur Gestaltung des Phasenübergangs. (...)
81	Vorlesen, Zustimmung AO	Mhm (...) zu abrupt irgendwie ich fand ihn schon sehr abrupt (5s)
82	Ablehnung AO + Begründung	und hier finde ich es auch schwierig zu sagen, ob die Inhalt dann wirklich ausreichend gesichert wurden, wenn einmal über diesen Comic diskutiert wurde und dann gesagt wurde, das ist jetzt die richtige Position. (...)
83	Vorlesen, Ablehnung AO + Begründung	Und unproblematisch, weil der Impulsbegriff ein komplett neues Thema ist, stimmt ja irgendwie auch nicht, Kraft und Impuls hängen ja schon stark zusammen. (..)
84	Zustimmung AO	Und ich könnte mir auch vorstellen, dass die Schülerinnen und Schüler den roten Faden irgendwie ein bisschen verlieren in der Stunde. (...)
85	Vorlesen, Zustimmung AO + Begründung (2x)	Und da, hier, das finde ich auch, wenn man sagt komplett neu, (scheinen?) den Schülerinnen das bisher gelernte nicht mehr so relevant, das könnte ich mir auch vorstellen, gerade weil halt eben Kraft und Impuls auf den ersten Blick vielleicht erstmal unterschiedlich wirken für die Schüler- innen (.) und wenn er das dann auch noch so betont, könnte ich mir vorstellen, dass sie den Kraftbegriff, den sie vielleicht eh nicht gut verstanden haben, (.) wieder vergessen. (5s)
86	Zustimmung AO + Begründung	Genau und es gibt ja auch wirklich einen fachlichen Bezug zwischen Axiom und Impulsbegriff. #00:33:51-4#
87	Vorlesen	LD4: Hinweise zur Gestaltung des Phasenübergangs für die erneute Durchführung. (.)

88	Vorlesen, Zustimmung AO + Begründung, Beschreibung	Mhm (4s) es wäre besser, den Zusammenhang zwischen Newtonschen Axiomen und Impulsbegriff zu verdeutlichen, weil die Schüler- innen das nicht selbst erkennen können, das sehe ich auch so, jetzt hat er gesagt, das es irgendwie ein komplett neues Thema ist und dann kommen die Schüler- innen eh nie drauf, dass es da einen Zusammenhang gibt, vermutlich. (.)
89	Zustimmung AO + Begründung	Mhm (..) vielleicht wäre es auch wirklich sinnvoll, den Impulsbegriff wann anders zu machen, damit sie erstmal wirklich das Wechselwirkungsprinzip verstehen und es auch wirklich verinnerlicht haben. (...)
90	Zustimmung AO, Alternative	Und das mit der Tafel ist mir vorhin auch schon aufgefallen, dass die hätte gewischt werden sollen, genau. (.)
91	Vorlesen, Zustimmung AO + Begründung	Einen glatteren Übergang zu gestalten, (4s) das finde ich auch (.) sinnvoll, (.) ähm damit man die Inhalte verknüpft und dann auch wieder ein roter Faden erkennbar ist. (..)
92	Ablehnung AO + Begründung	Aber ein Experiment muss eigentlich nicht zwingend sein, es muss ja nicht jede Einheit mit einem Experiment begonnen werden. #00:35:08-2#
93	-	[Video Einleitung Reflexion 04] #00:35:25-2#
94	Überleung zum Unterricht	LD7: Also sind wir jetzt in der Phase, wo ein Experiment kommt (..) und nachher noch eine Sicherung kommt. Experiment mit Tennisbällen, okay. #00:35:36-6#
95	-	[Video Szene Reflexion 04] #00:37:41-7#
96	Zustimmung AO + Begründung	LD7: Also ich finde, Robert sollte wieder (.) Rückmeldung zur Fachsprache erhalten, weil er (.) ganz oft noch von Schwung geredet hat und das auch (.) bei den Schüler- innen noch bestätigt hat. (.)
97	Bewertung + Begründung	Das Experiment fand ich irgendwie auch nicht so (.) super gewählt, weil es dann irgendwie was ganz anderes war, als mit der Titanic, das hat dann auch sehr voneinander abgewichen. (.)
98	Beschreibung, Überlegung zum Unterricht	Und die Ergebnissicherung war auch ein bisschen mhm das war einfach nur schreib das mal ab, also glaube ich nicht (.) irgendwie wo (sie?) Schülerinnen und Schüler selber aktiv drüber nachdenken könnten. (..) Ja. #00:38:26-8#
99	-	LD7: (8s, LD7 liest 'Robert hat schon einige Rückmeldungen erhalten...') #00:38:36-6#
100	Vorlesen	LD7: Feedback zur Wahl des Experiments. (.)
101	Zustimmung AO, Bewertung	Also an sich fand ich es gut, dass es irgendwas war, was man am eigenen Körper erfahren kann, (.) aber so richtig gut umgesetzt war es, glaube ich, nicht. (.)
102	Beschreibung	Mhm (8s) mhm genau, hier über- es wurde nur über Masse und Geschwindigkeit gesprochen und man geht einfach davon aus, dass das die einzigen Größen sind, von denen der Impuls abhängen kann. (..)

103	Beschreibung	Mhm (..) es gibt keine Abgrenzung zur kinetischen Energie, das wurde auch gar nicht erwähnt, obwohl die Schüler- innen das ja bestimmt schon kennen, also es wurde ja gesagt, dass sie das Vorwissen eigentlich haben. (...)
104	Beschreibung, Alternative	Es wurden irgendwie nur zwei Schüler mit einbezogen, also klar, das hat die wahrscheinlich motiviert und aktiviert, aber der Rest wurde nicht mit einbezogen, wenn man so was simples macht, wie mit den Tennisbällen, (.) könnten das eigentlich auch mehr machen. (5s)
105	Zustimmung AO, Überlegung zum Unterricht	Das mit der Vorgabe kann auch sinnvoll sein, (.) dass man eben keine langen Wartezeiten hat (..) und aber der Zeitspareffekt ist ja nicht immer (.) das Ziel von Unterricht. (..)
106	Zustimmung AO + Begründung	Mhm (.) also das (Anm: AO 'Das Experiment ist gut gewählt, weil die beiden Schüler, die das Experiment durchführen, den Impuls am eigenen Körper erfahren können) klicke ich jetzt doch an, weil man das ja auch irgendwie am eigenen Körper erfahren kann,
107	Zustimmung AO + Begründung	aber ich finde auch, (.) dass das wirf den Ball mit der gleichen Geschwindigkeit, das (.) ist ja auch irgendwie nur so ein relatives Ding, das kann man ja gar nicht richtig beurteilen. #00:40:14-8#
108	Vorlesen	LD7: Hinweise zur Wahl des Experiments für die erneute Durchführung. (.)
109	Ablehnung AO + Begründung	Mhm (.) statt des Experiments ein Video zeigen, (.) das würde ich vielleicht nicht unbedingt machen, also ist so ein Realexperiment vielleicht doch schöner. (..)
110	Vorlesen	Mhm (...) ein anderes Experiment mit einem Skateboard einen Ball weg werfen, (4s)
111	Zustimmung AO + Begründung	(so?) ein Stoßexperiment ist ja schon irgendwie so das (.) klassische zur Impulserhaltung, da würde ich mir das vielleicht auch eher angucken (...)
112	Zustimmung AO + Begründung	und irgendwas mit einer Messreihe finde ich auch sinnvoller, dann können die Schülerinnen und Schüler nämlich auch selber auf die Formel kommen und die wird denen nicht einfach da hingeklatscht. (4s)
113	Zustimmung AO + Begründung	Wenn das von allen durchgeführt wird, fände ich das auch besser, also jetzt nicht unbedingt, weil sie sich bewegen, aber weil dann auch alle irgendwie mal was praktisches machen können. (4s)
114	Zustimmung AO, Überlegung zum Unterricht	Und es sollten auf jeden Fall auch alle gut sehen können, wie leicht oder schwer es ist, den Ball zu fangen, da bin ich mir jetzt auch nicht sicher, ob er dadrauf geachtet hat im Unterricht. #00:41:37-6#
115	Vorlesen	LD7: Feedback zur Ergebnissicherung. (...)
116	Vorlesen, Zustimmung AO	Das also hier, es ist nicht gut, dass der Merksatz fertig vorgegeben wird, weil das reine Lesen und Abschreiben nicht verständnisfördernd ist, das sehe ich auch so. (..)

117	Zustimmung AO + Begründung, Beschreibung (2x)	Also an sich gut, dass sie den ins Heft übertragen, damit man das als Sicherung einfach selber nochmal notiert hat, (..) aber die Ergebnissicherung vom Experiment selber hat ja gar nicht wirklich stattgefunden, (.) dann können die auch gar nicht wissen, was sie wirklich mitnehmen sollen aus dem Experiment. (...)
118	Vorlesen, Zustimmung AO, Bewertung	Mhm, Merksatz gelungen, da die je-umso-Sätze die Zusammenhänge von Masse, Geschwindigkeit und Impuls nachvollziehbar darstellen. (...) An sich ja, aber die Art, wie der Merksatz gewonnen wurde, ist nicht gut. (...)
119	Zustimmung AO + Begründung	Und dass der Impuls auch nur skalar ist, also alles direkt ohne Vektorpfeile auch nicht gut, (.) weil man dann auch eben gar nicht die Richtungsänderung mit einbeziehen kann. (..)
120	Ablehnung AO + Begründung	Mhm (..) aber ich hätte den Begriff wahrscheinlich doch eher von der Energie abgegrenzt, (.) weil die kinetische Energie den Schüler- innen ja schon bekannt war und die Formel vom Impuls vielleicht doch sehr ähnlich dazu aussieht (5s) und die dann das wieder leicht verwechseln könnten. #00:43:09-7#
121	Vorlesen	LD7: Hinweise zur Ergebnissicherung für eine erneute Durchführung. (..)
122	Zustimmung AO + Begründung (2x)	Die Beobachtung vom Experiment würde ich auf jeden Fall auch schriftlich an der Tafel sammeln, damit man das sich später nochmal angucken kann und damit man auch irgendwie was Konkreteres dazu hat. (..)
123	Zustimmung AO + Begründung	Und ich würde die Schüler- innen auch selber in die Sicherung einbeziehen, damit sie nicht einfach nur das abschreiben müssen. (..)
124	Zustimmung AO + Begründung, Überlegung zum Unterricht	Ein Beispiel zur Formel ist eigentlich auch immer gut, wenn es passend gewählt ist, damit die Schüler- innen auch wirklich mehr was mit der Formel anfangen können. (5s)
125	Zustimmung AO + Begründung	In der Sicherung sollte auch nochmal auf das Experiment eingegangen werden, damit man den Bezug zum Anfang auch nochmal hat. (..)
126	Zustimmung AO + begründung	Und ich fänd es auch besser, wenn der Impuls als vektorielle Größe eingeführt wird, um ihn von der kinetischen Energie abzugrenzen
127	Zustimmung AO	und wenn man auf die kinetische Energie eingeht, kann man ja auch an äh an das Vorwissen der Schüler an- anknüpfen. #00:44:21-4#

128	Vorlesen, Beschreibung (2x), Überlegung zum Unterricht (2x)	LD7: Demonstrationsexperiment zur Impulserhaltung, (..) (beim?) Auffahrunfall, (...) mit Messwertaufnahme und das wurde vor der Stunde vorbereitet, (.) also so eine Luftkissenbahn, wo zwei Wagen drauf sind (...) und eine Tabelle mit den Massen und Geschwindigkeiten vor und nach dem Stoß, (..) okay, wie genau die Messwerte aufgenommen werden, kann man hier jetzt nicht sehen, aber das wird ja bestimmt mit irgendeinem (.) Computeprogramm. (4s) Und dann ist das diese Phase, wo noch ein Experiment durchgeführt wird mit der Luftkissen- äh Fahrbahn. (..) Mhm und die Schüler- innen können bei einigen Aufgaben assistieren. #00:45:25-4#
129	-	[Video Einleitung Reflexion 05] #00:45:35-3#
130	Sonstiges	LD7: Mhm[Ja]. #00:45:38-3#
131	-	[Video Szene Reflexion 05] #00:48:14-5#
132	Vorlesen	LD7: Was Robert- welche- Zu welchen Aspekten er Rückmeldung erhalten sollte,
133	Zustimmung AO + Begründung, Bewertung	auf jeden Fall zum Umgang mit den Schüler- innen, weil er die irgendwie so komisch getrennt, die Mädchen haben eine schöne Handschrift, dann dürfen die das aufschreiben und Jungs können besser ablesen, dann dürfen die das machen, (.) das fand ich nicht gut.
134	Beschreibung, Bewertung	Bei der Notation der Messwerte hat er aber darauf geachtet, dass die Einheiten mit notiert werden, das fand ich gut. (...)
135	Ablehnung AO	Die Auswertung des Experiments, dazu kann man jetzt irgendwie nichts sagen
136	Beschreibung	und Messunsicherheiten hat er zum Beispiel auch nicht thematisiert. (.)
137	Beschreibung, Bewertung	Den Aufbau finde ich in Ordnung, das ist ja so die typische Luftkissenfahrbahn. (.)
138	Bewertung + Begründung	Irgendwie wie die Schüler- innen aktiviert wurden, finde ich ein bisschen schwierig, weil er das so geschlechtergetrennt hat und mit so (.) Stereotypen, das fand ich nicht gut.
139	Alternative (2x)	Ähm die Durchführung vom Experiment, da hätte er irgendwie noch ein bisschen mehr erklären können, so (.) warum er den wie genau anstößt oder warum jetzt der andere in Ruhe ist (.) und auch welche weiteren Messwerte da noch aufgenommen werden, hätte irgendwie noch ein bisschen mehr erklärt werden müssen. #00:49:20-0#
140	-	LD7: (8s, LD7 liest 'Robert hat schon einige Rückmeldungen erhalten...') #00:49:29-2#
141	Vorlesen	LD7: Feedback zum Umgang mit den Schüler- innen. (.)
142	Vorlesen, Zustimmung AO	(Und?) der erste Punkt, nicht gut, dass Mädchen eine schöne Schrift haben und Jungs besser Werte ablesen können, da die Mädchen so auf eine schöne Handschrift reduziert werden, das finde ich auch, das ist mir sofort aufgefallen. (...)

143	Vorlesen, Zustimmung AO	Und die Mädchen werden irgendwie nur für die Fleißarbeiten, also für das Aufschreiben mit einbezogen, das (fördert?) Stereotypen, finde ich auch. (..)
144	Vorlesen, Ablehnung AO + Begründung	Aber hier, die Mädchen für ihr Handeln im Unterricht loben, um das Selbstkonzept zu stärken, die haben ja gar nicht richtig gehandelt, also das sehe ich jetzt irgendwie nicht so. (.)
145	Vorlesen, Zustimmung AO, Ablehnung AO + Begründung, Alternative	Mhm (...) Lob für die Handschrift vom Mädchen mit der inhaltlichen Kritik kombinieren, (.) das finde ich auch schwierig, also da muss er auch irgendwie nochmal (.) arbeiten. (...) An sich finde ich es nicht schlimm, dass er (.) äh an die Einheit erinnert, weil man das ja (.) (es?) ist, glaube ich, eher häufiger, dass die Einheiten generell vergessen werden und das nicht nur bei Mädchen. (5s)
146	Zustimmung AO, Überlegung zum Unterricht	Ja, hier, sie wäre bestimmt auch selbst drauf gekommen, dass sie Kilogramm hätte anschreiben müssen. (...)
147	Überlegung zum Unterricht, Beschreibung, Ablehnung AO + Begründung	Natürlich ist es wichtig, dass die Werte an der Tafel gelesen werden können, aber wenn die Jungs sich anstrengen, oder Jungs können auch eine schöne Handschrift haben, also irgendwie das ist schon wieder so mit Stereotypen, das ist eigentlich nicht nötig von ihm. #00:51:00-8#
148	Vorlesen	LD7: Hinweise zum Umgang mit den Schüler- innen für eine erneute Durchführung. (..)
149	Ablehnung AO, Alternative, Überlegung zum Unterricht	Mhm (..) (ja?) man muss sich- weiß ich nicht, ob man zwingend eine Geschichte um das Experiment herum konstruieren muss, um die Mädchen zu motivieren, also vielleicht könnte man (.) generell sagen, dass man so einen Autounfall simulieren will, ich bin mir jetzt nicht sicher, wie das im Unterricht dann konkret umgesetzt wurde. (..)
150	Zustimmung AO + Begründung	Aber hier, wer freiwillig schreiben will, das finde ich auf jeden Fall besser, (.) damit man halt eben nicht diese (.) Stereotypen bedient. (...)
151	Ablehnung AO + Begründung, Bewertung	Mhm es müssen, finde ich, nicht alle Schüler- innen eine Aufgabe wahrnehmen, also klar ist es gut, wenn alle aktiviert werden, aber wenn das eh ein Demonstrationsexperiment ist, (könnte?) es vielleicht auch eher unruhig werden, wenn man (.) das immer durch tauscht und immer irgendwer anders was schreibt, (..) dann könnte es auch einfach zu unruhig werden. (4s)
152	Zustimmung AO + Begründung	Das sollte er auf jeden Fall machen, also Mädchen auch für physikalische Kompetenzen zu loben, damit man gegen diese Stereotypen vorgehen kann.
153	Zustimmung AO	Und generell auf geschlechtsspezifische Zuschreibungen verzichten, (.) das ist auch wichtig. (..)
154	Zustimmung AO + Begründung	Und generell sollte er da vielleicht ein bisschen an seiner Einstellung arbeiten, weil das jetzt auch nichts unterschwelliges war. #00:52:30-1#

155	Vorlesen	LD7: Ähm Feedback zu seinem Umgang mit Messunsicherheiten. (4s)
156	Überleung zum Unterricht, Alternative + Begründung	(Bin mir?) gerade gar nicht mehr sicher, ob er das gesagt hat, dass man nur exakt, also dass er exakt oder genau messen kann, (.) vielleicht sollte er dann auch wirklich lieber möglichst genau sagen, weil man ja wirklich nicht (.) ohne Messunsicherheiten messen kann. (.)
157	Alternative	Er geht gar nicht auf die Messunsicherheiten ein, also natürlich muss er da jetzt keine große Fehlerrechnung machen, aber (.) irgendwie wenigstens erwähnen könnte er sie schon. (..)
158	Zustimmung AO + Begründung	Ich hätte auch lieber beide Wagen gewogen, damit man sieht, dass die gleich schwer sind, (.) mhm (..) das muss ja nicht zwingend- also es können ja sonst auch immer Abweichungen geben. (.)
159	Überlegung zum Unterricht	Mhm (6s) mit den Nachkommastellen bin ich mir jetzt irgendwie nicht so sicher, äh kommt halt irgendwie drauf an, wie genau die Waage gemessen hat. (4s)
160	Ablehnung AO + Begründung	Also ich glaube nicht, dass Messunsicherheiten die Komplexität des Experiments so stark (.) erhöhen würden, dass man da nicht mehr mitkommt, wenn man das irgendwie einmal sagt, dass man das (.) nicht zu hundert Prozent genau messen kann, überfordert das die Schülerinnen, glaube ich, noch nicht. (4s)
161	Ablehnung AO + Begründung	Aber irgendwie Fehler aufgrund vieler Nachkommastellen, die Schüler- innen haben ja vermutlich eh einen Taschenrechner (..) (und?) deswegen würde ich die Fehlerquellen da ein bisschen geringer einschätzen. (...)
162	Ablehnung AO + Begründung	Und irgendwie diese null Komma ein Kilogramm, die sind dann auch so ein (.) das leitet schnell zu so einem Idealwert irgendwie hin, (.) (genau?). #00:54:19-3#
163	Vorlesen	LD7: Und (jetzt?) Hinweise zum Umgang mit den Messunsicherheiten für eine erneute Durchführung.
164	Zustimmung AO	Mhm (.) also ich finde es eigentlich auch besser, wenn alle Nachkommastellen angegeben werden, (.) und (so?) diese Waage bestimmt auch genauer ist, als die null Komma ein Kilogramm. (...)
165	Zustimmung AO + Begründung	Die Messunsicherheiten sollten hier mit einbezogen werden, da die Messungen nicht wirklich exakt sind
166	Zustimmung AO + Begründung, eigene Erfahrung	und es wäre auch besser, beide Wagen zu messen, damit die Schüler- innen ein korrektes Bild von genauem wissenschaftlichen Arbeiten erhalten, genau. Wenn ich mir vorstelle, ich hätte jetzt so ein Experiment in der Uni gemacht, dann hätte ich ja auch beide Wagen messen müssen. (...)
167	Zustimmung AO	Genau, auch wenn sie baugleich sind, es kann ja trotzdem irgendwelche kleinen Unterschiede noch geben. (..)
168	Zustimmung AO	Und dann dann kann man hier auch direkt schon erste Messunsicherheiten thematisieren. (..)

169	Zustimmung AO + Begründung, Beschreibung	Und vielleicht ist (das?) auch für die Überzeugung der Schüler- innen gut, dass die gleich schwer sind, dann (.) geht man mit dem Experiment auch nicht zu schnell um, das hatte er, glaube ich, selbst schon gesagt, dass er das Gefühl hatte, dass nicht alle mitgekommen sind, dann kann man das auch ein bisschen entzerren. #00:55:29-3#
170	Vorlesen	LD7: Feedback zur Aktivierung der Schüler- innen. (..)
171	Vorlesen, Zustimmung AO, Beschreibung, Überlegung zum Unterricht	Mhm die Schüler- innen werden nur oberflächlich involviert und können trotz der Aktivierung nicht mitdenken. (.) Ja, sie müssen ja irgendwie nur aufschreiben und dann kann (ich?) schlecht einschätzen, was die nachher im Experiment noch machen müssen, wahrscheinlich nur (einen?) Wagen anstoßen oder so und ich glaube nicht, dass da so viel drüber gesprochen wird dann,
172	Zustimmung AO + Begründung, Alternative + Begründung	aber an sich ist es gut, dass er generell welche ins Experiment einbeziehen einbezieht, um sie im Unterricht zu aktivieren. Auch, wenn es ein Demonstrationsexperiment ist, gibt es ja meistens immer oder bei vielen Experimenten irgendwelche kleinen Tätigkeiten, die die durchführen können, (.) damit sie nicht ganz nur zugucken müssen. (...)
173	Sonstiges, Beschreibung	Mhm ja dass er das Experiment selbst durchführt, äh quasi kann ich jetzt irgendwie nichts zu sagen, weil er (hat?) ja nachher noch einen Schüler nach vorne holt für weitere Messungen. (..)
174	Zustimmung AO	Mhm aber ich- hier, das mit dem Arbeitsauftrag, dass sie keinen haben, das könne wirklich sein, dass sie dann nicht aufmerksam genug sind. (..)
175	Beschreibung, Überlegung zum Unterricht	Aber ich- mir ist auch aufgefallen, die Geschwindigkeiten zu bestimmen, da hat er gar nicht gesagt, wie das geht, ähm (..) und irgendwie nimmt er den Schülern da, glaube ich, was vorweg. (..)
176	Vorlesen, Beschreibung, Alternative + Begründung, Sonstiges	Mhm (..) es ist nicht gut, dass du das Experiment lehrerzentriert durchführst, da den Schülern (jetzt?) langweilig ist, irgendwie auch so eine Auslegungssache, also es werden ja welche mit involviert und ich denke, wenn man selbst wenn man ein Experiment sehr lehrerzentriert durchführt, kann man die Schüler- innen immernoch motivieren, da aufmerksam zuzugucken mit einem geeigneten Arbeitsauftrag. (5s)
177	Zustimmung AO + Begründung, Beschreibung	Mhm und um den experimentellen Aufbau näher zu betrachten, könnten vielleicht wirklich alle Schüler- innen einmal nach vorne kommen, um sich das näher anzugucken, aber andererseits ist die Klasse auch nicht so groß, aber ich nehme es jetzt trotzdem mal mit rein. #00:57:48-1#
178	Vorlesen	LD7: (unv.) Hinweise zur Aktivierung für eine erneute Durchführung. (...)

179	Zustimmung AO	Hier finde ich es auch wieder, es wäre besser, wenn die Schüler- innen sich freiwillig melden für das Notieren der Werte, (..)
180	Zustimmung AO + Begründung	mhm (..) und hier, es wäre auch besser, zu Fragen, ob es irgendwie Schwierigkeiten oder Unklarheiten gibt zum Ablauf, damit die Schüler- innen auch wirklich alle mitkommen. (.)
181	Ablehnung AO	Ähm (ah?) das kann ich jetzt nicht so sagen, dass er sie ausreichend aktiviert hat.
182	Zustimmung AO + Begründung	Ähm (.) aber ich würde vielleicht auch schon vorher Hypothesen abfragen, damit sie motivierter sind, sich das anzugucken. (.)
183	Vorlesen, Zustimmung AO + Begründung, Bewertung	Äh es wäre besser, den Schüler- innen Fragen zum Vorgehen während des Experiments oder zur Berechnung der Geschwindigkeit zu stellen, um sie zum Mitdenken anzuregen, das finde ich auch, weil er hat ja einfach nur gesagt, ich rechne jetzt die Geschwindigkeit aus und hat das dann gemacht und den Schüler- innen den Rücken komplett zugedreht, das war unterrichtsmäßig jetzt, glaube ich, nicht so sinnvoll. (...)
184	Vorlesen, Zustimmung AO	Mhm die Schüler- innen stärker in die Durchführung des Experiments einzubeziehen, (.) Werte ablesen können die da auf jeden Fall machen, denke ich äh und die Geschwindigkeit bestimmen eigentlich auch, dann kann man die Schüler- innen aktiver einbinden. (7s)
185	Zustimmung AO + Begründung	Und eigentlich finde ich auch, dass man die Schüler- innen alle noch vorne kommen lassen müsste, damit sie die Luftkissenbahn besser sehen. Je nachdem, wie groß so eine Klasse dann wird, (.) äh kann ich mir schon vorstellen, dass man die Luftkissenbahn von hinten nicht mehr so gut sieht. #00:59:27-0#
186	Vorlesen, Überlegung zum Unterricht	LD7: Nach der Aufnahme des Messwerte kommt die Beobachtung, das hat er von den Schüler- innen abgefragt, Allgemeingültigkeit der Impulserhaltung, (.) in einem kurzen Exkurs theoretisch hergeleitet, das hat er dann auch wieder gemacht (.) und die Impulserhaltung nochmal am Experiment überprüft. Das ist dann die Phase hier. #00:59:56-2#
187	-	[Video Einleitung Reflexion 06] #01:00:22-0#
188	Beschreibung	LD7: (Na?) da kommt es ja irgendwie wieder mit den Messunsicherheiten. #01:00:27-6#
189	-	[Video Szene Reflexion 06] #01:02:58-0#
190	Vorlesen	LD7: Mhm (die?) Aspekte zum zur Rückmeldung mit den also wo er Lob oder Kritik für erhalten sollte.
191	Zustimmung AO + Begründung	Auf jeden Fall wieder Umgang mit den Schüler- innen, weil er wieder so geschlechtergetrennt damit umgegangen ist. (.)

192	Beschreibung, Bewertung	Äh irgendwie er ist auch gar nicht richtig auf Messunsicherheiten selbst eingegangen, sondern hat das mehr als äh (..) als Rundungsfehler abgetan, was ja auch irgendwie nicht so ganz richtig ist.
193	Überlegung zum Unterricht	Aber zur Einheit des Impulses bin ich mir gerade gar nicht so sicher. #01:03:35-2#
194	-	LD7: (6s, LD7 liest 'Robert hat schon einige Rückmeldungen erhalten...') (Äh?). #01:03:41-4#
195	Vorlesen	LD7: Dann das Feedback zum Umgang mit den Schülerinnen. (..)
196	Ablehnung AO	Also mit dem Begriff Mädels sehe ich an sich jetzt an sich nicht das Problem,
197	Zustimmung AO + Begründung (2x)	aber er behandelt die sehr unterschiedlich (.) und das ist auch irgendwie wenig wertschätzend gegenüber den Mädchen, vor allem wenn er, glaube ich, ganz am Anfang gesagt hat, dass es mehr Mädchen sind und wenn er denen dann gar keine physikalischen Kompetenzen zutraut, ist das schon irgendwie sehr komisch. (.)
198	Zustimmung AO	Und er schätzt die Mädchen und Jungen auch in ihrem Können sehr unterschiedlich ein, (.) was dann ja auch wieder die Vorstellung fördert, dass Physik so ein Jungsfach ist. (..)
199	Vorlesen, Zustimmung AO	Mhm und er thematisiert die Einheit des Impulses auch gar nicht, wenn ich das (.) richtig gesehen habe, das ist nicht gut, da dies zu inkonsistent- da dies inkonsistent zur früheren Szene ist,
200	Vorlesen, Zustimmung AO, Beschreibung (2x), Alternative	genau, bei Paula hat er ja auf die Einheit der Geschwindigkeit bestanden und von der Masse. (4s) Aber hier mit der dass er nur die Mädchen auffordert, das könnten ja auch andere sein, die äh (.) das nicht korrigiert haben in ihren Tabellen, also wenn sollte er auch alle auffordern und nicht nur die Mädchen. #01:05:00-7#
201	Vorlesen	LD7: Welche Hinweise zum Umgang mit den Schüler- innen für die erneute Durchführung der Unterrichtsstunde.
202	Ablehnung AO	Ähm also ich fand den Umgang nicht in Ordnung, das sollte er nicht so beibehalten.
203	Zustimmung AO + Begründung	Er sollte die Mädchen auch stärker im Unterricht einbeziehen und die zum Beispiel auch ihre Werte vorlesen lassen, das (fand?) ich nämlich auch nicht richtig, dass er da nur die Jungs drangenommen hat.
204	Zustimmung AO	Er sollte auch definitiv auf geschlechtsspezifische Zuschreibungen verzichten. (.)
205	Zustimmung AO + Begründung, Alternative	Ähm (..) ich finde, er sollte auch mehrere Schüler- innen befragen nach den Werten, damit er die nicht nur von einem hört, also müssten jetzt nicht zwingend alle sein, (.) aber schon mehrere. (.) Mhm ich kreuze es trotzdem mal an.

206	Vorlesen, Zustimmung AO + Begründung (2x)	Ähm verstärkt schwächere Schüler- innen in den Unterricht einbinden, damit auch sie die Inhalte verstehen, das finde ich auch wichtig, also dass man (.) wenigstens irgendwie sicher gehen kann, dass die auch soweit mitkommen, oder (man?) dann besser herausfinden kann, an welchen Stellen es genau hapert. (5s)
207	Zustimmung AO + Begründung	Und (wenn?) soll auch alle also auch andere Schülerinnen irgendwie aufgefordert werden, die falschen Ergebnisse zu korrigieren, damit das alle nachher richtig haben. (.)
208	Zustimmung AO + Begründung	Und dann sollte er sich auch irgendwie mal selbst hinterfragen, woher seine Zuschreibungen immer kommt, zu den Geschlechtern, weil er die ja schon sehr stark vertritt. #01:06:28-4#
209	Vorlesen	LD7: Mhm Feedback für den Umgang mit Maries Einwand und mit den Messunsicherheiten. (..)
210	Ablehnung AO + Begründung	Mhm (.) also ich fand es nicht gut gelungen, also es geht ja auch im Unterricht nicht immer nur um Effizienz, also dass man irgendwie möglichst zeitsparend zu einer gewissen Erkenntnis kommt. (.)
211	Beschreibung	Mhm die Rundungsfehler sind zwar schon spielen da auch irgendwie mit rein, aber sind ja nicht die hauptsächliche Ursache dafür. (.)
212	Vorlesen, Ablehnung AO + Begründung	Mhm (...) mhm nicht gut gelungen, da du den Fokus von der experimentellen Bestätigung der Impulserhaltung hin zur Diskussion von Messungenauigkeiten verschiebst. Das finde ich irgendwie nicht, weil er hat ja immernoch den Fokus (darauf?), dass die Impulserhaltung unbedingt gelten muss (.)
213	Beschreibung	und ich finde, er geht generell nur sehr knapp darauf ein, (weil?) er gar nicht wirklich Messunsicherheiten thematisiert, sondern nur Rundungsfehler
214	Beschreibung, Zustimmung AO	und er sagt ja auch gar nicht, wo genau die Rundungsfehler gemacht wurden, das ist ja dieser Punkt hier.
215	Vorlesen, Zustimmung AO + Begründung	Ähm (...) genau, es ist nicht gut, dass du die Ungleichheit der Messwerte mit Rundungsfehlern begründest, da die Begründung falsch ist, das liegt ja eigentlich an Messunsicherheiten von Geschwindigkeit und zum Beispiel von der Masse. (..)
216	Zustimmung AO	Mhm (.) genau und (das?) könnten die Schülerinnen und Schüler ja auch denken, (.) dass sie selbst irgendwo die Rundungsfehler gemacht haben, (.) das ist ja auch nicht richtig. #01:07:57-0#
217	Vorlesen	LD7: Hinweise zum Umgang mit Maries Einwand und Messunsicherheiten für eine erneute Durchführung der Stunde. Äh die Reaktion war nicht angemessen, (..)
218	Zustimmung AO + Begründung	er sollte den Schüler- innen generell irgendwie angeben, auf wie viele Nachkommastellen sie runden sollen, damit man damit angemessener umgehen kann. (.)

219	Zustimmung AO	Ähm und er sollte ihnen auch am Anfang mitteilen, dass Messunsicherheiten und Rundungsfehler nicht berücksichtigt werden, (.) oder also wenigstens irgendwie einmal darüber sprechen und dann nachher da drauf eingehen, dass das im Rahmen von den Messunsicherheiten übereinstimmt. (..)
220	Ablehnung AO + Begründung	Mhm vielleicht muss er gar nicht richtig begründen, wo die Rundungsfehler liegen, (.) aber (unv.) man das, glaube ich, schlecht sagen kann, war die Hauptursache für die Abweichung waren ja auch die Messunsicherheiten. (.)
221	Vorlesen, Alternative	Mhm (5s) (aber?) hier, das mit der Reibung und Energieumwandlung beim Aufprall, damit den Schüler- innen klar ist, dass die Ergebnisse auch (unterhalb?) dass sie nicht exakt gleich sein können, muss nicht zwingend bei der Vorstellung des Experiments, (aber?) irgendwann sollte es schon erwähnt werden. (..)
222	Vorlesen, Zustimmung AO + Begründung	Und bei der Durchführung des Versuchs müssen Messunsicherheiten berücksichtigt werden, damit die Schülerinnen nachvollziehen können, wieso es zu den Abweichungen kommt, das sollte er auf jeden Fall nächstes Mal mit berücksichtigen, weil die Messunsicherheiten, die hat er wirklich gar nicht thematisiert (.)
223	Zustimmung AO + Begründung	ähm (..) und wenn er auch sagt, dass ihn solche Sachen aus dem Konzept gebracht haben, dann sollte er irgendwie auch nochmal selber über die Messunsicherheiten nachdenken. #01:09:43-9#
224	Vorlesen	LD7: Feedback zum Umgang mit der Einheit des Impulses. (..)
225	Zustimmung AO + Begründung	Ähm es ist nicht gut, dass er irgendwie gar keine Einheit thematisiert, die Einheit gehört ja immer zu einer äh Größe mit dazu und zu einem Formelzeichen auch, wenn man das aufschreibt. (.)
226	Vorlesen	Ähm (.) es ist inkonsistent zur früheren Szene wo er richtig darauf bestanden hat, dass die Einheit aufgeschrieben wird. (...)
227	Ablehnung AO + Begründung, Bewertung	Und irgendwo ist es vielleicht auch halbwegs sinnvoll, dass man das nicht zu stark in den Fokus rückt, aber er hat es halt gar nicht in den Fokus gerückt, deswegen finde ich den Punkt nicht gut.
228	Vorlesen, Ablehnung AO + Begründung	Äh es ist gut, dass du den Fokus auf die berechneten Werte und nicht auf die Einheit legst, (.) mhm (na?) es sind ja nicht nur die Zahlen relevant, es ist ja auch immer die Einheit relevant, in der die Werte angegeben werden, achso also dann das nicht. (.)
229	Vorlesen	Äh und es ist nicht gut, dass die Schüler-innen gar nicht aufgefordert werden, eine Einheit anzugeben, da man so ja eigentlich gar nicht weiß, ob die Impulse wirklich gleich groß sind. #01:10:52-0#
230	Vorlesen	LD7: Äh Hinweise zum Umgang mit der Einheit für eine erneute Durchführung. (.)

231	Ablehnung AO, Alternative + Begründugn	Äh es ist nicht in Ordnung, dass er die Einheit zunächst vernachlässigt hat, er sollte die Einheit des Impulses angeben, da die Einheit ein wichtiger Bestandteil der Angabe ist.
232	Vorlesen, Zustimmung AO	Ähm er sollte darauf achten, dass die Schüler beim Vorlesen auch die Einheit angeben, um im eigenen Verhalten konsistent zu bleiben, das sehe ich auch so. (.)
233	Zustimmung AO, Bewertung + Begründung	Die Einheit sollte irgendwo auf jeden Fall notiert werden, am besten finde ich es da auch über der Tabelle, also in der Spalte, da kann man sich Schreibarbeit sparen, aber sie ist trotzdem thematisiert worden. (.)
234	Vorlesen	Und die Schüler- innen sollten die Einheit des Impulses auch nennen, damit sie sich daran gewöhnen. #01:11:42-4#
235	-	[Video Einleitung Reflexion 07] #01:12:01-6#
236	Überlegung zum Unterricht	LD7: Also dann sind wir in der Unterrichtsphase, wo er noch einen Merksatz formuliert, das ist die letzte Sicherungsphase. #01:12:09-0#
237	-	[Video Szene Reflexion 07] #01:13:51-2#
238	Zustimmung AO + Begründung, Alternative	LD7: Mhm also er sollte irgendwie auf jeden Fall Rückmeldung zur Formulierung des Merksatzes erhalten, weil die dann ja doch stark abgewichen ist von der Formulierung, die die Schüler vorgeschlagen hatten, da hätte er irgendwie nochmal drauf eingehen können, warum man bestimmte Aspekte jetzt noch berücksichtigen muss.
239	Alternative	Und auch die Interaktion mit den Schüler- innen hätte er irgendwie besser gestalten können, also zum Beispiel auch, dass er von den Schüler- innen welche fragt, was die für einen Merksatz formuliert haben und irgendwie generell hätte er mehr auf die Schüler eingehen können. #01:14:25-9#
240	-	LD7: (4s, LD7 liest 'Robert hat schon einige Rückmeldungen erhalten...') #01:14:30-5#
241	Vorlesen	LD7: Robert äh Feedback zur Formulierung des Merksatzes.
242	Bewertung	Also ich fand die Formulierung an sich gut gelungen. (.)
243	Ablehnung AO + Begründung	Mhm (...) aber er hat hier ja auch gar nicht im Merksatz auf das durchgeführte Experiment verwiesen, also kann ich den Punkt jetzt auch gar nicht (.) ankreuzen.
244	Zustimmung AO + Begründung (2x)	Äh ich könnte mir schon vorstellen, dass die Formulierung irgendwie zu anspruchsvoll ist, weil die ja eben stark abweicht von den Sachen, die die Schüler- innen gesagt haben. (.) Äh und da (er?) auch in einem geschlossenen System und sowas, ich weiß nicht, ob die (.) das direkt so mit dem Versuch verbunden haben. (..)

245	Beschreibung	Und in dem vorgefertigten Merksatz werden die Merksätze der Schüler- innen gar nicht mit einbezogen, (.) also er sagt irgendwie, ist okay, was ihr gemacht habt, aber eigentlich muss es so sein, also er geht halt eben gar nicht richtig darauf ein. #01:15:28-5#
246	Vorlesen	LD7: Hinweise zur Formulierung des Merksatzes für eine erneute Durchführung. (.)
247	Vorlesen, Ablehnung AO + Begründung	Ähm (..) also irgendwie Beispiel und Merksatz zu trennen, er hatte ja gar kein Beispiel da drin, also (.) muss man das jetzt auch irgendwie nicht trennen, äh (.)
248	Vorlesen, Alternative	ein- (zusätzlich?) zum langen Merksatz eine kurze Version zu formulieren, da die sich die Schüler- innen besser merken könne, also im Grunde formuliert man ja auch eine kurze Version, wenn man sagt also wenn man sich die Formel anguckt, oder wenn man sagt, es gilt Impulserhaltung.
249	Vorlesen, Zustimmung AO, (Ablehnung AO) + Begründung	Ähm den Merksatz generell nur in Bezug auf einen Stoßprozess zu formulieren, damit der Merksatz näher beim besprochenen Experiment bleibt, äh generell sollte der schon am Experiment sein, also vor allem jetzt wenn der eingeführt wird, (.) insgesamt braucht man aber auch irgendwie so einen allgemeingültigen Merksatz, finde ich. (.)
250	Vorlesen, Alternative	Mhm (..) den Merksatz um die Einheit des Impulses zu ergänzen, da die Einheit stets zur Angabe einer physikalischen Größe gehört. Die Einheit zum Impuls müsste dann ja eher in die Definition vom Impuls und nicht in den Merksatz zur Impulserhaltung,
251	Vorlesen, Zustimmung AO	aber der Merksatz sollte auf jeden Fall schülernäher gestaltet werden, also zum Beispiel mit einer Skizze oder mit einer sprachlich vereinfachten Formulierung, das (ist?) wäre auf jeden Fall besser. (.)
252	Zustimmung AO, Alternative + Begründung	Ähm und vielleicht auch irgendwie eine Schüler- innenformulierung verwenden oder die aufgreifen und dann weiter ausbauen, damit das dann auch für die Schüler- innen besser verständlich ist. #01:17:05-4#
253	Vorlesen	LD7: Feedback zur Interaktion mit den Schüler- innen.
254	Zustimmung AO + Begründung	Mhm an sich finde ich es gut, dass die die Merksätze vorlesen, dann kann man auch sehen, ob sie das verstanden haben oder nicht. (..)
255	Zustimmung AO, Alternative	Und es ist auch gut, dass sie das selber formulieren sollen, aber insgesamt hätte er da mehr drauf eingehen müssen. (.)
256	Zustimmung AO, Beschreibung (2x)	Genau, er hatte (denen?) irgendwie nicht so eine richtig ausführliche Rückmeldung gegeben zu den Merksätzen, er hat dann einfach seinen vorgefertigten da vorgestellt. (.)
257	Beschreibung	Ähm genau er bezieht die Merksätze der Schüler- innen gar nicht mit ein. (..)

258	Beschreibung	Er bezeichnet seinen Merksatz dann irgendwie als besser, also auf einer gewissen Ebene stimmt das vielleicht, aber die Formulierungen der Schüler- innen werden dann schon eher abgewertet. (...)
259	Zustimmung AO	Und ich hätte bei den kurzen Merksätzen, die die Schülerinnen hatten vielleicht auch mehr vorlesen lassen. (.)
260	Zustimmung AO + Begründung, Alternative	Und ich hätte auch nicht nur Merksätze von den Jungen vorlesen lassen, weil irgendwie er fokussiert sich, wenn es um die fachlichen oder um die physikalischen Kompetenzen geht immer sehr stark auf die Jungs, da hätte ich auch mal ein Mädchen dran genommen. #01:18:18-0#
261	Vorlesen	LD7: Hinweise zur Interaktion mit den Schüler- innen für eine erneute Durchführung. (..)
262	Zustimmung AO + Begründung	Er sollte die Merksätze auf jeden Fall ausführlicher erklären, gerade seinen vorgefertigten, der war schon ähm für die Schüler- innen bestimmt sprachlich sehr anspruchsvoll. (.)
263	Zustimmung AO + Begründung, Beschreibung	Ähm (.) er sollte den Merksatz vielleicht auch gemeinsam mit den Schüler- innen erstellen, vor allem, wenn er eh schon verlangt, dass sie einen eigenen aufschreiben, damit er dann nicht den einfach nur mit seinem eigenen überdecken kann, da also da sollte er das gemeinsam erarbeiten. (..)
264	Zustimmung AO + Begründung	Und vielleicht, wenn er eh die Merksätze formulieren lässt, auch mehr Schüler- innen die Merksätze vorstellen zu lassen, dann hat man auch ein breiteres Bild, wie die Schüler- innen das verstanden haben. (..)
265	Zustimmung AO, Alternative	Ähm (..) vielleicht sollte er die Schüler- innen sich auch selbst melden lassen, dann kann er auch die Mädchen eher dran nehmen, (..) aber a- vielleicht ist es dann auch so, dass sich dann gar keiner meldet, aber dann kann man immernoch von sich aus wen dran nehmen.
266	Bewertung	Ähm den Umgang an sich fand ich jetzt nicht so ganz angemessen,
267	Beschreibung, Zustimmung AO	ähm (.) er lobt die Schüler- innen nicht in ihrer Arbeit, (.) dass kann ja ist dann ja wahrscheinlich nicht gut für den Physikunterricht, oder wie die sich im Physikunterricht verhalten und fühlen.
268	Vorlesen, Zustimmung AO	Und er sollte auch hier wieder aufpassen, dass er keine Rollenklischees im Unterricht einbringt. (..) Ja. #01:19:48-8#
269	Vorlesen	LD7: Jetzt zum Unterricht insgesamt.
270	Vorlesen	Zu welchen der folgenden Aspekte sollte Robert eine Rückmeldung bekommen. (.)
271	Sonstiges	Äh (..) (ah?) ich weiß nicht,
272	Bewertung	insgesamt ging es ja auch so um den Ablauf der Unterrichtsstunde, wo er diesen harten Cut irgendwie einmal drin hatte zwischen (.) Kraft und Impuls, das war ja nicht ganz so gut.

273	Bewertung, Alternative	Ich fand irgendwie auch den Umfang der Stundeninhalte ein bisschen viel, dass man die Axiome, ja okay die wurden dann wahrscheinlich nochmal wiederholt, aber dann hätte ich mich lieber da drum ein bisschen ausführlicher gekümmert und dann den Impuls wann anders eingeführt. (.)
274	Beschreibung, Überlegung zum Unterricht (2x)	Den Medieneinsatz fand ich an sich interessant, also er hat halt viel mit dem Beamer gearbeitet, (.) wahrscheinlich damit er nicht so viel an die Tafel schreiben muss, warum er die Messwerttabelle (auf?) einem Plakat und nicht an der Tafel gemacht hat, habe ich irgendwie auch noch nicht so ganz verstanden. (.)
275	Zustimmung AO	Ähm auf jeden Fall braucht er Rückmeldung zum Umgang mit den Schüler- innen,
276	Bewertung	die fachliche Korrektheit, da ist mir jetzt an sich aber nichts aufgefallen, (.) das war gut soweit,
277	Alternative, Beschreibung	aber beim Auftreten ist mir auch aufgefallen, dass er irgendwie (.) äh weniger Füllwörter benutzen könnte, also er sagt halt sehr oft 'äh' und braucht dann irgendwie manchmal ein bisschen länger, um seine (.) Sätze deutlich zu formulieren. #01:21:13-6#
278	Vorlesen	LD7: (4s, LD7 liest 'Robert hat schon einige Rückmeldungen erhalten...') (Zur?) Rückmeldung zum Umgang mit den Schüler- innen und zur fachlichen Korrektheit der ganzen Stunde. #01:21:25-2#
279	Vorlesen	LD7: Erstmal zur Interaktion mit den Schüler- innen. (.)
280	Zustimmung AO, Beschreibung	Ich- genau ist nicht gut, dass er die Beitrage von Schülerinnen kaum berücksichtigt, (.) manchmal geht er ja wirklich gar nicht drauf ein, um das mal gehört zu haben und präsentiert dann die (.) (ich sage mal?) richtig Lösung. (.)
281	Überlegung zum Unterricht, Alternative	Mhm (.) an sich ist es ja nicht schlimm, wenn man eher lehrerzentriert unterrichtet, (..) aber er müsste halt die Schüler- innen irgendwie mehr aktivieren dabei. (.)
282	Zustimmung AO + Begründung	Aber er behandelt Jungen und Mädchen stark unterschiedlich, was eben nicht gut für die Motivation ist und auch für das Selbstkonzept von den Schüler- innen. (..)
283	Bewertung	Die Sicherungsphasen, dass er welche einbaut ist an sich sehr gut. (..)
284	Vorlesen, Zustimmung AO, Beschreibung	Ähm (...) er verwendet verschiedene Sozialformen im Unterricht, Einzelarbeit, Partnerarbeit, ähm ja das ist auch gut, Partnerarbeit ist mir jetzt nicht ganz so viel aufgefallen, glaube ich nur bei der Formulierung der Merksätze.
285	Beschreibung	Aber er hat sich gefühlt immer nur Beiträge von denselben Schüler- innen angehört, (.) und den Rest hat er irgendwie ein bisschen ignoriert. (..)
286	Vorlesen, Zustimmung AO + Begründung	Und sowas mit habt ihr noch Fragen ist nicht zielführend, genau, da antworten dann halt eben viele gar nicht drauf, weil sie dann auch gar nicht wissen, was sie fragen sollen.

287	Zustimmung AO	Ähm und über auftretende Schülervorstellungen (.) sagt er halt irgendwie auch gar nichts, (.) das ist eben auch nicht gut. #01:23:03-8#
288	Vorlesen	LD7: Hinweise zur Interaktion mit den Schüler- innen für eine erneute Durchführung.
289	Ablehnung AO	Ähm der Umgang war (aber?) nicht angemessen,
290	Zustimmung AO + Begründung	er sollte sie grundsätzlich stärker irgendwie mit einbeziehen, zum Beispiel in die Diskussion oder Durchführung von Experimenten. Und besonders in die Formulierung der Merksätze, dann kann er die Impulserhaltung nämlich auch gemeinsam erarbeiten und (.) es ist irgendwie nicht so ein- eine Präsentation von den Inhalten, was die Schüler- innen dann irgendwie schlucken müssen. (.)
291	Zustimmung AO + Begründung	Er sollte den Schüler- innen ein besseres Feedback geben, also ein ausführlicheres, damit sie auch wissen, ob das jetzt richtig ist, was sie sagen. (..)
292	Vorlesen, Zustimmung AO + Begründung	Ähm Schülervorstellung zum Thema vor dem Unterricht ansehen, damit er da gem- angemessen drauf reagieren kann, finde ich auch, er hat ja an mehreren Stellen gesagt, dass (ihn?) die Schülerbeiträge irgendwie aus dem Konzept gebracht haben. (..)
293	Zustimmung AO + Begründung, Alternative	Ähm, vielleicht sollte er sich auch Gedanken dadrüber machen, wie man das Verständnis von Schüler- innen abprüfen kann, damit man auch sicher sein kann, (.) dass sie das verstanden haben, was er im Unterricht gemacht hat und nicht einfach (.) nachfragt mit habt ihr noch Fragen, ja. (...)
294	Zustimmung AO + Begründung	Äh und generell sollte er irgendwie seinen Umgang nochmal überdenken, damit er die Schüler- innen mehr lobt
295	Zustimmung AO	und (.) vor allem sollte er auf die Rollenklischees aufpassen, (.) dass er die nicht mit in den Unterricht bringt. #01:24:36-3#
296	Vorlesen	LD7: Feedback zur fachlichen Korrektheit. (.)
297	Bewertung	Also ich fand (ihn?) gut, mir ist da eigentlich nichts (.) aufgefallen, ähm das einzige, was mir, ja okay doch, das mit dem Rundungsfehler und der (.) äh und den Messunsicherheiten, das ist eben fachlich falsch.
298	Vorlesen, Zustimmung AO, Überlegung zum Unterricht	Okay, die Einheit des Impulses, okay wenn man das so sieht, dann ist es schon äh irgendwie was fachlich falsches da drin. (.) Ähm (..) aber ob das (dann?) so ein gravierender fachlicher Fehler ist, bin ich mir jetzt auch nicht so sicher.
299	Zustimmung AO	Äh er sollte die Schülervorstellungen, (..) die auftreten irgendwie thematisieren, irgendwie besser damit umgehen. (.)
300	Zustimmung AO + Begründung	Mhm (..) wenn er so einen großen Wert auf die Geschwindigkeit, also auf die Einheit der Geschwindigkeit in der Tabelle legt, sollte er (jetzt?) das beim Impuls auch machen. (.)

301	Zustimmung AO	(Ja und?) wenn er sagt, er würde exakt messen, ist das eben auch nicht richtig,
302	Vorlesen, Bewertung, Zustimmung AO	also da sind dann doch einige fachliche Fehler, die er da irgendwie gemacht hat. (...) Genau, dann kann ich das ja auch ankreuzen, er hat fachliche Fehler gemacht, das ist nicht gut, da man den Schüler- innen so falsche Inhalte beibringt. #01:25:53-8#
303	Vorlesen	LD7: Hinweise zur fachlichen Korrektheit für die erneute Durchführung.
304	Zustimmung AO + Begründung	Er sollte auch bei der Besprechung der Impulse auf die Einheiten achten, dann kann er nämlich auch einheitlich vorgehen im Vergleich zu anderen Sachen, wo er die Einheit explizit gefordert hat nochmal.
305	Zustimmung AO + Begründung (2x)	Ähm (..) er sollte auch sowas aufgreifen, wie dass ein Apfel Kraft hat, damit äh sich das bei den Schüler- innen nicht verfestigt und die die fachlich korrekte Sichtweise lernen. (.)
306	Zustimmung AO + Begründung	Dafür muss er sich die Schülervorstellungen auch vorher einmal angucken, damit er (.) weiß, was auf ihn da zu kommen könnte und wie er damit umgehen könnte. (.)
307	Zustimmung AO + Begründung	Und er sollte auch auf die Messunsicherheiten bei der Luftkissenbahn eingehen, damit die Schüler- innen die Impulserhaltung trotz der leicht unterschiedlichen Werte akzeptieren. (..)
308	Zustimmung AO, Bewertung	Ähm (.) genau, Merksätze sollten fehlerfrei formuliert sein, (aber?) seine Merksätze waren das, soweit ich das beurteil- äh beurteilen kann. (..) Aber generell ist natürlich wichtig, darauf zu achten. (..)
309	Zustimmung AO + Begründung, Beschreibung, Alternative	Und auf Wortwahl sollte auch immer geachtet werden, damit man nicht ausversehen irgendwas fördert, zum Beispiel als der Impuls eingeführt wurde, hat er ja selber noch oft von Schwung geredet, dass man da das schon ein bisschen klarer abgrenzt. #01:27:12-5#
310	-	[Video Prompt Schulnote] #01:27:23-6#
311	Sonstiges, Bewertung (2x), Bewertung (2x) + Begründung (2x)	LD7: Okay mhm ich finde es irgendwie ein bisschen schwierig zu beurteilen, also es gab schon einige fachliche Fehler auf die man irgendwie besser eingehen könnte und (.) der Umgang mit den Schüler- innen, den fand ich an vielen Stellen auch nicht so gut. (.) Äh weiß nicht, ob es jetzt ein bisschen zu hart ist, aber ich glaube, (.) besonders gut war es dann halt eben nicht, wenn man die fachlichen Fehler berücksichtigt und den Umgang, also da gibt es schon noch einige Sachen, wo er dran arbeiten müsste, deswegen würde ich, glaube ich, tatsächlich zur Vier tendieren. #01:27:55-6#

P Transkripte der Interviews zur Prüfung der Anforderung an Teil 4 des INA (Evaluation der Rückmeldung)

Die Interviews zur Evaluation des Assessment-Feedbacks wurden vollständig transkribiert. Dabei wurde analog zur Transkription der Think-Aloud-Interviews vorgegangen (s. Kapitel G.1).[208]

Nachfolgend sind die vollständigen Transkripte angehängt. Grau hinterlegt sind jeweils die Beiträge der Interviewleitung sowie organisatorische Passagen.

P.1 Interviews zur Pilotierung (Kohorte A)

P.1.1 Interview 1: A1

1 [...](Hier: Einleitung, Einverständniserklärung)
2 I: Genau, dann würde ich einmal zum Einstieg einmal fragen, ob Sie sich noch an die Rückmeldung erinnern können, die Sie zu Ihrer Bearbeitung bekommen haben, oder ob // das schon zu lang- #00:02:00-6#
3 A1: Ja. // #00:01:58-1#
4 I: Ja, okay, und können Sie sich noch dran erinnern, welche Schlüsse Sie daraus gezogen haben, was so die Schlussfolgerungen waren? #00:02:05-9#

[208] Leichte Anpassungen ergeben sich für den Transkriptionsbeginn und das Transkriptionsende (die Transkription beginnt in der Kohorte A mit dem inhaltlichen Start des Interviews (bei der Transkription ausgelassen wurden die Begrüßung und Einverständniserklärung zur Aufzeichnung der Interviews); für die Kohorten B und C wurden die Aufnahmen vollständig transkribiert) sowie die Wahl der Sprechercodes (die Teilnehmenden werden entsprechend der jeweiligen Kohorte mit „A1“ bis „A6“, „B1“ bis „B6“ bzw. „C1“ bis „C5“ benannt; die Interviewleitung wird mit „I“ abgekürzt).

5 A1: Also für mich war dann sozusagen die Schlussforderung- folgerung, ach guck mal, bist ja gar nicht so falsch in dem Job, weil ich noch nicht in der Schule war ähm wegen Corona hat sich das mit dem Eignungs-Orientierungspraktikum für mein Studium etwas nach hinten verschoben, das mach ich erst noch und dann war es halt schön, ähm da die Rückmeldung einmal zu bekommen, weil es ja auch sehr praxisorientiert war, fand ich in dem Reflexionstest und ich hab den halt auch mit voller Mühe und Gedankeneinsatz sozusagen bearbeitet und dementsprechend war das für mich eigentlich ne sehr schöne Rückmeldung. #00:02:40-3#
6 I: Okay, danke. [...](Ab hier: Öffnen und Lesen der Beispielrückmeldung) #00:04:00-7#
7 A1: Dopplung äh können können, das könnte etwas missverständlich sein, ich mein, ich verstehe es aber bei Reflexionskompetenz was ist das, im Rahmen einer Reflexion können unterschiedliche Aussagen getätigt werden, beobachtete Aspekte können beschrieben werden im ersten Satz mit im Rahmen einer Reflexion können folgende Aussagen getätigt werden da ist diese handelnde Person die diese Aussagen tätigt das sind ja diejenigen die das bewerten während dann bei der Beschreibung von beobachtete Aspekte können beschrieben werden und so das beschreibt ja die Fähigkeit von demjenigen der den Reflexionstest gemacht hat. #00:04:37-8#
8 I: Mhm[Ja]. #00:04:38-1#
9 A1: Und diese Unterscheidung, die kommt irgendwie nicht so gut raus, weil es immer in diesem können, können, können bleiben, also vielleicht muss man da noch einmal ne Person nennen oder den Satz mit im Rahmen einer Reflexion können unterschiedliche Aussagen getätigt werden verändern #00:04:53-1#
10 I: Ja, danke. #00:04:55-2#
11 A1: Mhm. [...](Ab hier: Weiterlesen) #00:05:59-2#
12 A1: Vielleicht kleiner- kleine Formsache, ich nehme an, dass all diese drei ähm Kompetenzen beziehungsweise alle diese drei Punkte Bewertung, Alternativen und Schülervorstellungen erkennen, dass die sich alle um zentrale Punkte oder zentrale Aspekte in der Physikdidaktik drehen, vielleicht ist es deshalb nicht so super, ausgerechnet einen da besonders hervorzuheben mit die Diagnose ist ein zentraler Aspekt der Physikdidaktik, (.) oder? (.) Vermute ich jetzt, // dass die eigenltich gleichwertig sind. #00:06:30-8#
13 I: Ja, das (unv.) // Ja, ja, klar. Also grundsätzlich sind das alles (Sachen?), genau, ich nehm das aber auch- also selbst wenn das nicht so wäre, würde ich das auf jeden Fall jetzt äh mitnehmen, also deswegen einfach gern alles sagen, wenn sie Gedanken (haben?). #00:06:44-5#
14 A1: [...](Ab hier: Weiterlesen) #00:07:15-1#
15 A1: Vielleicht nochmal bei der Beschreibung von den Antennen und den schwarzen Punkten, also die schwarzen Punkte sind ja jeweils am Ende von ähm den Antennen dieses Boxplots, also wenn wir nochmal auf das Diagramm gucken, (.) genau ähm und dann sprechen Sie einmal von etwaigen schwarzen Punkten, sehe ich jetzt keine hier drin, ich vermute, dass das das oberste und das unterste Ende jeweils von diesem Boxplot ist, also von der Linie. #00:07:43-1#
16 I: Mhm. #00:07:44-6#
17 A1: Stimmt das, oder wie ist das gemeint? #00:07:47-1#
18 I: Ähm tatsächlich sind hier einfach keine Punkte drin, also etwaige Punkte meint, falls da Punkte sind, aber // es ist ja trotzdem gut zu wissen auf jeden Fall. #00:07:56-0#
19 A1: Mhm[Ja] // (...) wie sehen die dann aus, würden die auch auf (den?) Linien liegen oder im Boxplot oder nebendran? #00:08:01-9#
20 I: Genau, also die würden quasi auf dieser Mittellinie, auf der Linie, wo auch die Antennen sind, liegen und dann sind das in der Größe und in der Dicke wie die roten, nur, dass sie schwarz sind und irgendwie hier oben (zeigt im PDF) oder hier unten (zeigt im PDF) dann liegen würden. #00:08:14-4#
21 A1: Und das bedeutet dann, dass die nicht in den Boxplot mit eingegangen sind, weil sie zu äh stark vom Durchschnitt abgewichen sind, oder wie? #00:08:20-6#

22 I: Ne, die sind schon mit eingegangen also weil der Boxplot ja quasi ein bisschen- also der- der Boxplot- wenn ich den Boxplot als gesamte Betrachte, sind die ja mit eingegangen wenn ich jetzt nur die Box betrachte, sind da eben nur die mittleren fünfzig Prozent eingegangen (.) ja #00:08:38-5#
23 A1: Okay also ich kann mir jetzt noch nicht vorstellen, wann dann dieser Fall zustande kommen sollte, dass da noch ein Punkt ist (.) über der Linie oder unter der Linie #00:08:52-8#
24 I: Ähm (..) okay, ich schlage vor, für den Moment ignorieren wir diese möglichen schwarzen Punkte // und ich nehm auf jeden Fall mit, dass ich entweder diesen Satz rausnehme, wenn da keine Punkte sind oder das nochmal anders erklär. #00:09:06-4#
25 A1: Mhm[Ja]. // (7) Vielleicht ja, #00:09:08-4#
26 I: Ja, okay. #00:09:09-1#
27 A1: (Gute?) Idee. #00:09:10-1#
28 I: Gut. #00:09:12-0#
29 A1: Genau weiter. [...](Ab hier: Weiterlesen) #00:09:18-0#
30 I: Gut, dann erstmal danke dafür, ich hab jetzt noch so ne handvoll Fragen, die ich da einfach nochmal anschließen würde, es sei denn, es gibt grad von Ihrer Seite noch was. #00:09:27-5#
31 A1: Äh vielleicht noch ne Empfehlung, ähm es ist jetzt vielleicht dann nur für mich so gewesen, meinen Namen löschen Sie ja, ne? #00:09:35-5#
32 I: Ja, #00:09:36-4#
33 A1: Ja, dann kann ich das sagen, also ich hatte ähm das höchste Testergebnis, also ich hatte überall die roten Punkte am oberen Ende, und im ersten Moment, als ich drauf geguckt hab, hab ich das halt nicht verstanden, Moment, wo bin ich denn jetzt hier eingetragen, weil es halt gerade ganz oben war, deshalb wär vielleicht irgendwie die Möglichkeit, diesen roten Punkt noch n bisschen dicker oder so zu machen oder vom Format her noch mehr davon abzuheben, damit er nicht mehr aussieht, wie das Ende von der Skala. #00:10:01-3#
34 I: Mhm[Ja] (.) okay, danke. (4) Gut, dann äh würde ich einmal mit den Fragen anschließen. #00:10:12-5#
35 A1: Mhm[Ja] #00:10:13-3#
36 I: Die erste wäre, ob sie zuvor schonmal mit- also mit solchen Boxplots, also mit dieser Art von Diagrammen gearbeitet haben, und falls ja, in welchem Kontext in welchem Rahmen? #00:10:25-2#
37 A1: Ja und zwar habe ich vor meinem Studium eine Ausbildung gemacht (zum Industriekaufmann/zur Industriekauffrau) und hab da häufig mit ähm (.) Boxplots zu tun gehabt, wenn wir ähm Benchmarks gemacht haben, also Verkaufspreise verglichen oder durchschnittliche Mengen oder sowas, um halt die ähm Verteilung auf Lager oder auf die Märkte oder so darzustellen. #00:10:49-1#
38 I: Alles klar, äh aber im Studium dann noch nicht // wahrscheinlich. #00:10:53-3#
39 A1: Im // Studium noch nicht, nein. #00:10:54-2#
40 I: Okay und ähm, wenn Sie sich jetzt die Ergebnisse von dieser Rückmeldung, also nicht Ihrer eigenen sondern die, die jetzt hier liegt, ähm ja- wenn Sie sich diese Rückmeldung angucken, welche Schlüsse würden Sie daraus ziehen? #00:11:08-2#

41 A1: (6) Also ich würde insgesamt daraus ziehen, dass ich eindeutig noch viel zu lernen habe, weil ich unter der fünfzig Prozent-Marke liege insgesamt, aber trotzdem nicht bei den schlechtesten aus meiner Vergleichsgruppe oder halt aus meinem Studienjahrgang, nenn ichs jetzt mal, wir sind ja in Physik dadurch, dass es so strukturiert ist, wenn wir überall durchgekommen sind, so ziemlich in ähnlichen Semestern eigentlich, ähm dass ich da im Vergleich zur Gruppe eigentlich sogar leicht über dem Durchschnitt bin, aber für mich selbst, wenn ich halt ein gutes Ergebnis erzielen möchte, (.) noch nicht mal fünfzig Prozent erreicht hab, also eigentlich weiß, okay da kann ich noch was machen und bei der Verteilung unten (Ergebnisse nach Kategorien aufgeteilt) würde ich mir dann eben genau angucken, okay, ich war am besten in der Lage, die Bewertungen (.) durchzuführen und das äh Verhalten der Lehrperson zu bewerten, beim Alternativen aufzeigen war ich ungefähr so gut wie meine Mitstudenten, was ja auch Sinn ergeben würde, weil wir ja in ähnlichen Semestern sind, haben wir wahrscheinlich schon ähnliches gelernt oder nicht gelernt, also da würde ich mir wahrscheinlich weniger Sorgen drum machen, bei den Schülervorstellungen da würde ich n bisschen äh große Augen kriegen und würd mir denken, oh, da bin ich ganz am unteren Rand hier von dem Boxplot, da muss ich aber auf jeden Fall noch was machen, was kann ich denn da vielleicht mal nachschlagen, nachlesen, ausprobieren? #00:12:41-2#

42 I: Okay, danke (.) ähm dann daran anschließend, (.) auch ruhig auf ganze Rückmeldung bezogen und auch die Diagramme, also, ne, die Rückmeldung als Ganze, äh welche Informationen in dieser Rückmeldung für Sie besonders hilfreich sind oder besonders interessant oder sonst besonders rausstechen? #00:13:02-9#

43 A1: (7) Das ist schwierig. (...) Ich glaube, was mir auf jeden Fall heraussticht, sind eben die Dinge, die in den Diagrammen visualisiert sind, also immer der Vergleich von mir zu dem Mittelstrich sozusagen, das ist etwas, was raus sticht, und ähm (..) ja ansonsten eben (..) die einzelnen Kategorien, die Sie nennen, da wirkt jetzt keine irgendwie wichtiger oder weniger wichtig, ich finde das ist eigentlich ziemlich ausgewogen, (.) deshalb kann ich nicht sagen, dass mir da inhaltlich irgendwas besonders ins Auge sticht aber für mich wahrscheinlich dann jeweils der Abstand zwischen dem roten Punkt und der Mittellinie #00:13:56-3#

44 I: Mhm[Ja] und ähm, wenn ich jetzt nochmal- also ich cluster das jetzt mal so ganz grob irgendwie in Bereiche, also wir haben ja hier oben irgendwie so ne Beschreibung von was haben wir eigentlich gemessen, die Rückmeldung und die verschiedenen Bereiche, da haben Sie jetzt, wenn ich das richtig verstanden hab, grad gesagt die sind irgendwie alle gleich wichtig und unten noch die Hinweise, also das sind ja grob gesagt so drei Bereiche, ähm (.) waren die für Sie dann auch ungefähr gleich wichtig (.) zum Verständnis oder zur Einordnung? #00:14:26-6#

45 A1: (..) Also zum ähm Verständnis ist erstmal der erste Absatz, denk ich, ziemlich wichtig, damit man halt weiß, was überhaupt betrachtet wurde, die Inhalte so richtig hat dann der zweite Abschnitt und der dritte hilft dann eben bei der Interpretation der Diagramme. Also ich würde nicht sagen, dass der dritte Abschnitt fürs Verständnis weniger wichtig ist, aber er transportiert halt weniger Inhalt, sondern hat mehr so Hilfestellungen. #00:14:52-4#

46 I: Okay, (.) ähm und dann quasi andersrum ins andere Extrem, sind Informationen überflüssig und wenn ja, welche? #00:15:02-1#

47 A1: Die schwarzen Punkte. #00:15:09-2#

48 I: Das ist eingespeichert. #00:15:10-3#

49 A1: Ne, ansonsten eigentlich nicht. Also ich find das gut, dass Sie äh am Anfang nochmal nennen, was alles beobachtet wurde, das sind ja noch mehr Aspekte, als die die jetzt äh da stehen, aber dass Sie dann sagen, äh in dem Reflexiontest war jetzt der Fokus da und da rauf, da bekommt man halt so n bisschen in- n Einblick darein, wie denn diese empirische Forschung da funktioniert. #00:15:29-8#

50 I: Mhm[Ja] (.) okay und ähm wenn Sie die Rückmeldung jetzt gelesen haben, welche Fragen bleiben dann für Sie vielleicht noch offen? #00:15:39-1#

51 A1: (6) Es ist vielleicht, was was jetzt auch son bisschen mit dem Test zu tun hat, nämlich bei dem Punkt Alternativen, (...) wenn ich mich recht entsinne, hatte man die Möglichkeit, noch zusätzliche Alternativen reinzuschreiben, oder war das bei dem anderen Test? #00:16:00-5#
52 I: Das war der andere. #00:16:02-3#
53 A1: War der andere // genau. #00:16:03-8#
54 I: Das war (Erklären?) // ja. #00:16:04-9#
55 A1: Dann ähm das sind ja alles nur Bewertungen für das, was ich jeweils auswählen konnte und nicht eine Bewertung für das, was ich vielleicht selbst in der Situation gemacht hätte, also es ist ja nochmal ne andere Anforderung, ob man jetzt entscheiden muss und hat vier Möglichkeiten gegeben, oder ob man alleine da steht und es fällt einem vielleicht nur eine ein, (.) also es wär für mich jetzt interessant, nachdem ich jetzt gesehen hab, dass ich das mit dem Auswählen vielleicht schon ganz gut kann, wie es denn dann wäre, wenn ich das selbst reproduzieren und anwenden muss. Also vielleicht wär dann so ein aufbauender Test irgendwie in die Richtung ähm, was würde die Person dann selbst sagen, schreiben, vorschlagen oder so interessant. #00:16:48-6#
56 I: Ja okay (.) aber fürs- quasi für die- für die Einordnung der Rückmeldung ist dann weniger konkret was, sondern das ist dann mehr das große Ganze versteh ich richtig? #00:16:59-7#
57 A1: Genau richtig. #00:17:01-6#
58 I: Okay (..) ähm (.) und dann wär noch die Frage, welche weiteren Informationen Sie vielleicht gerne gehabt hätten, gibts da was? #00:17:11-2#
59 A1: (7) Ne #00:17:18-0#
60 I: (..) Okay, ähm (.) ich frag jetzt trotzdem nochmal so ne Liste ab an an quasi Dingen, über die wir nachgedacht haben, die man vielleicht noch ergänzen könnte, und ähm (.) ich würd sagen, ich geh die einfach einmal kurz durch, es sind so- son paar Stichworte, und dann (.) sagen Sie einfach jeweils dazu, bringt mir gar nichts bringt mir was, weil, falls Ihnen da grad was einfällt, okay? #00:17:41-6#
61 A1: Okay #00:17:42-6#
62 I: Okay, das erste wär, ähm (.) Information über das Ziel vom Test oder Information über das Projekt in dem Rahmen in dem dieser Test entstanden ist. (...) Ich scroll dazu vielleicht noch // einmal hoch- #00:17:59-4#
63 A1: Mhm // (.) (bringt?) denjenigen was, die sich insgesamt äh für Didaktik begeistern und interessieren. Mir persönlich hätte es was gebracht, weil, wie gesagt, ich hab ja auch schon gesagt, dass ich das eigentlich schön finde, dass Sie nochmal erzählen, was Sie untersucht haben, ich denke aber dass es eventuell zu weit ausufern könnte oder zu kurz kommen könnte, weil, wenn man es nur einmal kurz mit rein schreibt, dann hat man so einmal kurz den Merker gesetzt, aber (.) man hat trotzdem noch keinen Eindruck vom ganzen Projekt und wenn mans im ganzen beschreibt, hat man noch ne Zusatzseite, glaube ich. Ich denke, es ist vielleicht besser, das deshalb rauszulassen, weil man dem nicht ganz gerecht werden kann. #00:18:34-8#
64 I: Mhm[Ja] okay, äh das nächste wäre, inwieweit man vielleicht explizite Informationen darüber ergänzt, wie denn dieser Zielzustand der Reflexionsfähigkeit aussieht, also was ist- was ist da das Optimum, was sollte man erreichen? #00:18:49-2#
65 A1: (11) Schwierig. (..) Könnte man machen, (..) ja, wir sind ja alle erwachsen, könnte man machen, ja . #00:19:06-8#
66 I: Äh und wieso schwierig? #00:19:09-1#

67 A1: Äh ich hab halt erst überlegt ob das vielleicht ähm n bisschen ähm entmutigend wirkt, wenn man eben selbst noch ähm sich dann in diesem Diagramm als recht niedrig eingestuft sieht und dann noch zusätzlich das Bild liest, was irgendwie ganz viel zusammenfasst, was man eben ähm selber noch nicht bei sich erkannt hat, oder wo man ähm sich noch total unsicher ist und so. Nicht, dass dann so ein bisschen Überforderung geschieht, aber ich denke dadurch, dass wir alle Studierende sind und deshalb hab ich gesagt, wir sind ja alle erwachsen, haben wir das schon besser gelernt, diese Selbsteinschätzung zu machen, als jetzt zum Beispiel Schulkinder, also in der Schule würde ich das zum Beispiel nicht machen. #00:19:55-2#
68 I: (..) Okay, das nächste wäre eine feinere Aufschlüsselung der ähm Bewertung, also der Punkte, dass man zum Beispiel sagen könnte- ich könnte zum Beispiel so Plots oder Diagramme für jede einzelne Aufgabe oder sowas da rein machen, (.) das wäre ja theoretisch denkbar #00:20:18-2#
69 A1: (5) Nein, finde ich nicht sinnvoll? #00:20:25-5#
70 I: Warum? #00:20:27-4#
71 A1: Weil ich denke, dass wenn man diesen Test gemacht hat, und da liegt ja dann wieder noch n bisschen Zeit zwischen, dass man nicht die Erinnerung an die einzelnen Fragen und so weiter hat und es eigentlich besser ist, eine ähm zusammengefasste Rückmeldung zu bekommen, weil die einem in dem Moment mehr bringt und man vielleicht auch nicht die Muße hat, sich da jede einzelne Aufgabe nochmal anzugucken oder sich gar nicht richtig erinnert. #00:20:48-4#
72 I: Alles klar, dann die Option, ähm andere Diagramme entweder statt der Boxplots oder zusätzlich zu den Boxplots, also man könnte ja zum Beispiel auch sich Säulendiagramme oder so vorstellen oder noch andere Punkte oder was auch immer. #00:21:04-5#
73 A1: (5) Also für ne Darstellung äh (.) Durchschnittsleistung und so finde ich das eigentlich wirklich gut gewählt das Boxplot, (..) gerade weil es halt auch schon in ähnlichen Kontexten irgendwie begegnet ist. Bei Säulen wüsste ich jetzt gar nicht, wie man das dann umsetzen sollte (.) so richtig // (.) ja #00:21:30-6#
74 I: Ja also // ich sag mal so ganz platt, da gäbs verschiedene Möglichkeiten, man könnte einfach zum Beispiel die eigenen Prozentscores da in Säulen einfach angeben, oder so (.) aber (.) // genau (.) da also (unv.) #00:21:42-3#
75 A1: Achso (.) (das man?) (unv.) // dazu schreibt, das kann man machen, s tut keinem weh. #00:21:46-1#
76 I: Klingt aber auch nicht nach nem Vorteil, okay. #00:21:49-3#
77 A1: Ne es ist- es ist so ein nice to have würde ich sagen. #00:21:52-9#
78 I: Okay, dann ähm ne Beispielaufgabe, also entweder aus dem Test oder eine, die ganz ähnlich ist zu den- zu den Testaufgaben, dass man die noch mit angibt. #00:22:06-2#
79 A1: Und dann sozusagen daran zeigt, wie bewertet wurde? #00:22:10-1#
80 I: Genau, also man könnte ne Beispielaufgabe zeigen, über die man nachdenken kann, oder auch Beispielaufgabe mit Musterlösung und Begründung oder sowas oder auch nur die Musterlösung ohne Begründung, je nachdem. #00:22:22-2#
81 A1: Das könnte noch interessant sein, ja, (.) also mit Musterlösung und Begründung, weil man dann eben sehr gut nachvollziehen kann, wie das Ganze funktioniert, und wie es zu dem Wert geführt hat, den man erreicht hat. #00:22:38-4#
82 I: Okay, ähm zwei noch und dann haben wirs auch. Das erste sind Literaturhinweise oder so Hinweise auf physikdidaktische Grundlagentexte, um das besser einzuordnen quasi mit Literatur. #00:22:55-6#
83 A1: (6) Die könnte ich höchstens sehen, dass mans halt für die wissenschaftliche Richtigkeit macht, (.) weil man ja immer nichts behaupten darf, ohne es zu zitieren. (.) Ansonsten seh ich dadurch für uns Studenten keinen großen Vorteil drin. #00:23:13-4#
84 I: (Ja?) #00:23:15-3#

85 A1: Ich mein, es wird nur ganz vereinzelte geben, die dann ausgerechnet in diesen Reflexionstext nochmal rein gucken, was da unten in der Fußnote stand, wenn sie nach Literatur für ne Hausarbeit suchen. #00:23:23-2#
86 I: Okay ähm und dann die Option, dass man Material reingibt oder vielleicht auch Hinweis auf n Text oder, oder- also verschiedenste Materialien oder Hinweise, wie man die Reflexionfähigkeit schulen kann. #00:23:39-1#
87 A1: (7) Hm (..) ja (...) aber ich glaube, dass man damit nicht dann (.) das dann als gelöst betrachten kann, ne, also ich denke dass diese Ergebnisse vom Reflexionstest dann vielleicht doch Indikationen geben können, wie vielleicht die äh Lehre noch angepasst werden kann, um das besser zu schulen. Ich glaube, da ist nicht nur der Student selbst in der Verantwortung sich da irgendwie (.) ranzuarbeiten (..) ich meine, es ist ja dann auch immer schwierig, solche Reflexionsdinge nur dann aus der eigenen Brille zu lesen, (.) deshalb denk ich dass man das auf jeden Fall machen könnte, um halt so die Möglichkeit zu geben, sich da was anzugucken, wenn man das möchte, aber dass man- das dann nicht damit als abgeschlossen betrachten kann. #00:24:39-7#
88 I: Oka,y dann sind wir mit diesem Fragenblock durch, es sei denn, Sie möchten noch was ergänzen, und dann gibts gleich noch zwei ganz kurze und dann würden wir auch ganz durch. #00:24:51-3#
89 A1: Ja, bitte weiter. #00:24:53-4#
90 I: Alles klar äh und zwar gehts jetzt noch einmal darum wie Sie an Ihre Rückmeldung gekommen sind und zwar gabs-, ähm wenn Sie sich erinnern, gabs da diese Website, da haben Sie Ihren Code nochmal eingegeben und konnten das dann runterladen und da wäre meine erste Frage, obs dabei irgendwelche Schwierigkeiten gab, wie Sie die Rückmeldung erhalten haben und, falls ja, welche das waren? #00:25:15-1#
91 A1: Also es gab keine Schwierigkeiten, ähm der Herr Doktor N hat das bei uns im Seminar erwähnt und hatte den Link in den Chat geteilt und ausgehend von dem bin ich auf die Seite gekommen und hab mir das runtergeladen. #00:25:31-1#
92 I: Okay. Und ähm dann noch die Frage, in welcher Form Sie sich das wünschen, diese Rückmeldung zu bekommen. Also ob sie da- also ob das so rund ist, oder ob das besser wär, die (.) direkt per Mail zu erhalten oder direkt online abrufen zu können oder sowas also was da quasi die Wunschvorstellung wäre. #00:25:52-0#
93 A1: (6) Wenn man sie per Mail schickt, geht man natürlich sicher, dass (.) ähm viele sie auch sehen und wahrnehmen und lesen, während, wenn man die halt zum runterladen anbietet, dann vielleicht Leute doch wieder nicht dazu kommen und es wieder vergessen und so und dann vielleicht ähm Bedarfe, die sie eigentlich haben, da noch weiter zu arbeiten, dann untergehen, weil sie halt (die?) ähm Rückmeldung nicht sehen. Ja, ich glaube, es wäre gut, die per Mail zu schicken. #00:26:24-9#
94 I: Okay äh wie wäre das, ähm also Notiz am Rande, das können wir gerade technisch noch nicht, aber wir reden ja gerade über Wunschvorstellungen, wenn man diese Auswertung quasi auf der letzten Testseite bekommt, also ich glaub, Sie haben auch den Erklärtest gemacht, da war das ja so. #00:26:41-8#
95 A1: // Mhm[Ja]. #00:26:42-9#
96 I: Da gabs // dann ganz am Ende n Ergebnis, wenn man das so umsetzen könnte, wär das besser? #00:26:46-3#
97 A1: Das könnte man auch auf jeden Fall machen, weil da hat man direkt ne Rückmeldung dazu, es ist noch alles ganz frisch, was man grade gemacht hat, und man hat dann sozusagen ähm die Sache direkt abgeschlossen, also s wär eigentlich ne runde Sache, ja. #00:27:02-2#
98 [...](Ab hier: Abschluss)

P.1.2 Interview 2: A2

1 [...][Hier: Einleitung, Einverständniserklärung etc.]
2 I: Dann würde ich einmal zum Einstieg fragen, ob du dich noch dran erinnern kannst äh an deine Rückmeldung, die du bekommen hast zum Test? #00:01:58-5#
3 A2: Äh, ich habe gerade sicherheitshalber nochmal reingeguckt, um mich zu erinnern. #00:02:02-5#
4 I: Alles klar. Und weißt du noch, welche, äh, welche Schlüsse du daraus gezogen hast oder was für Konsequenzen oder so? #00:02:10-8#
5 A2: Ähm, also im Wesentlichen hab ich nichts überraschendes da gesehen, sondern mich eher in dem bestätigt gesehen, was ich sowieso für mein weiteres Studium so im Kopf hatte und das zeigt auch quasi auch einfach nur welche Teile im Studium vielleicht noch weniger im Vordergrund standen. Das war so n bisschen mein erster Eindruck dabei. Und auch son bisschen, was man sowieso halt vielleicht schon aus persönlichem Interesse kannte. Also zum Beispiel in einem Bereich, den ich sowieso, auch dadurch, dass ich Nachhilfe privat gebe, sehr, sehr schon für mich dachte, dass ich ihn gut kann, in dem wars dann auch eben entsprechend. Und andere Bereiche, die da vielleicht auch eher noch so sind, die muss man noch vielleicht mal sich im Studium erarbeiten explizit, (weil?) man sich da privat nicht mit beschäftigen würde, das war dann auch eher der Bereich, wo man dann vielleicht etwas noch Verbesserung sich hätte- Worte ... wo man sich hätte verbessern können. #00:03:03-8#
6 I: Alles klar. Ähm, gut, danke. [...][Ab hier: Öffnen und Lesen der Beispielrückmeldung] #00:05:26-1#
7 I: Gut. (.) Ähm (.) Genau. Gibts da von deiner Seite direkt irgendwie Anmerkungen, Fragen oder so? #00:05:34-5#
8 A2: Ähm, wenn du nochmal auf die Boxplots, oder wie auch immer du die Dinger nennst, scrollst. #00:05:38-6#
9 I: Ja. #00:05:40-0#
10 A2: Ähm (4) Direkt an der Stelle würde ich zum Beispiel interessieren: 'Ergebnis ist Prozent' Was für Prozent? Prozent von was? Prozent von 'das was ihr als richtig erachtet'? (.) [I nickt] Okay. (..) // (unv.) Das steht da nirgendwo. #00:06:00-7#
11 I: Was (unv.) // (.) Was wär die Alternative? Also was könntest du dir noch vortellen? #00:06:04-5#
12 A2: Ähm gerade weil ihr auch in den Boxplots quasi drin steckt: Vergleich zu anderen Studierenden. (.) Aber dafür gibts ja den Boxplot. #00:06:15-5#
13 I: Mhm[Ja]. #00:06:16-0#
14 A2: Also sonst müsste der ja irgendwie von Hundert bis Null gehen, stimmt schon. #00:06:20-0#
15 I: Ah, okay. Ja, aber ich- ich versteh den Gedanken, nehm ich auf jeden Fall mit. (.) Okay. (..) Dann äh wär meine erste Frage, ob du schonmal mit solchen Boxplots gearbeitet hast und wenn ja, in welchem Rahmen? Also // jetzt vor dieser Rückmeldung. #00:06:39-1#
16 A2: Äh (5) ich glaube, gesehen habe ich die nur in der Schule in welcher Klasse auch immer man das macht. (.) Ansonsten tatsächlich noch gar nicht, aber ähm (.) nachdem ich dann deine kurze Erläuterung dazu gesehen hatte, wusste ich auch wieder, wie sie funktionieren, deswegen, die fand ich sehr verständlich. #00:06:57-7#
17 I: Alles klar, danke. Dann wär das nächste: Wenn du dir diese Rückmeldung jetzt quasi ja so angeschaut hast, welche Schlüsse würdest du aus dieser Rückmeldung ziehen? #00:07:09-2#

18 A2: Hab ich tatsächlich im Vorhinein schon kurz drüber nachgedacht und äh abgesehen davon, dass ich mich vielleicht, je nachdem wie sehr ich (jetzt?) über einzelne Punkte im Vorhinein reflektiert war ähm vielleicht dadrin bestätigt gesehen hab oder vielleicht festgestellt habe: 'Okay, da gibts einen Bereich in dem ich vielleicht schlechter bin, als gedacht', habe ich für mich (..) noch irgendwie nicht so den Punkt gesehen, wo ich jetzt aus diesem Bereich 'okay, da hätte ich noch Verbesserungspotential' konkrete Handlungen ableiten könnte. (..) Das heißt, zum Beispiel hätte mir jetzt einfach an ein, zwei Beispielen wahrscheinlich ganz gut getan, einfach nochmal eine der Situationen aus dem Test zu sehen und zu sagen: 'Okay, das wären Situationen, die man vielleicht aus didaktischer Sicht jetzt auch oder halt nach den Kriterien, die man da anlegen kann, so und so aus den und den Gründen lösen könnte, sollte. Du hast das gemacht, das hat vielleicht die Vor- und Nachteile' und das man vielleicht da an der Stelle n bisschen tiefer reingeht, weil so war das jetzt so: 'Joa, ich hab son Test gemacht, ich hab jetzt son groben Überblick', aber jetzt wirklich konkret da Handlungsschritte draus abzuleiten, find ich sehr, sehr schwierig, gerade wenn man jetzt noch von also ich jetzt, Stand viertes Semester, von sämtlichen ähm (...) Unterrichten selber völlig weit entfernt ist (..) und einfach nur jetzt fürs Studium sich zu sagen, ich möchte mich jetzt auf diesen Bereich speziell mehr fixieren, ist irgendwie find ich schwierig, weil ja doch in Verstaltungen sehr- sehr, sehr eng vorgegeben ist, was gemacht wird. #00:08:45-5#
19 I: Okay. Einmal, dass ich das richtig verstanden hab: Also, den Schluss, den man ziehen kann, ist irgendwie 'An dieser oder jener Stelle ist Entwicklungspotential, aber wie das geht, das fehlt.' Hab ich das richtig verstanden? #00:08:59-3#
20 A2: Im wesentlichen ja. #00:09:00-7#
21 I: Ja, okay. Ich hörs mir auch nochmal an, also so ist nicht aber- genau. Okay- #00:09:06-1#
22 A2: Lange Sätze ohne Punkt und Komma. #00:09:08-2#
23 I: Ja, alles gut. Die dürfen. Gerade hier. Ähm dann wär die nächste Frage, welche Informationen aus dieser Rückmeldung jetzt die besonders hilfreichen oder besonders interessanten sind. #00:09:20-3#
24 A2: (9) Also (..) wirklich hilfreich und interessant ähm hab ich die diesen Reflexionstest eben als Reflexionsanlass empfunden. Also überhaupt nur aufgrund des Testes mal anzufangen, zu überlegen, (.) wie gut bin ich überhaupt in einzelnen Punkten, wo hab ich denn noch Verbesserungspotential und das war die (.) eine Sache, wo ich den Test als wirklich hilfreich empfunden hab. Äh jetzt zum Beispiel grade bezüglich der Bods- äh Boxplots, ich mein, klar ist es interessant, ne- n Vergleich zu anderen Studierenden zu haben, aber der sagt einem ja letzten Endes auch nicht, ob man jetzt wirklich gut oder schlecht ist, sondern eigenltich nur, ob man schlechter als andere ist. (.) Das heißt, selbst wenn man jetzt hier zum Beipsiel- in der Beispielrückmeldung ist ja im Punkt Bewertungen ein relativ gutes Ergebnis da, dann sind das ja trotzdem nur fünfzig Prozent und ich fand gerade bei diesen Boxplots neigte man dazu, nur aufgrund seiner Position im Boxplot zu ziehen, ob man schon gut oder schlecht in einem Bereich ist. Und gerade wenn man jetzt vielleicht überlegt, als Lehrer sollte das Ziel sein, da überall deutlich über dem, was auch die Boxplots da jetzt zeigen, zu sein. Ich mein klar, noch im Bachelorstudium und so, kann noch gar nicht perfekt sein, aber deswegen- ich habs mehr als Reflexionsanlass empfunden und weniger jetzt so als absolutes 'da bin ich schon top und da so überhaupt nicht', sondern mehr so als 'die Bereiche gibt es und in (den?) hab ich vielleicht noch besonders noch das äh (.) Entwicklungspotential. #00:10:59-3#
25 I: Okay. (.) Ähm (.) und dann ja quasi die Frage in die andere Richtung: Sind Informationen in dieser Rückmeldung überflüssig und falls ja, welche? #00:11:11-1#

26 A2: (9) Ähm (11) Was ich mich son bisschen gefragt habe, ähm die Boxplots sind ja jetzt- da steht einundzwanzig Studierende, fünftes Fachsemester. (..) Das heißt, ihr habt nen sehr, sehr kleinen Datenpool. Was ich persönlich interessanter gefunden hätte, wäre jetzt wirklich mit Studierenden aus demselben Fachsemester verglichen zu werden. (.) Das heißt, indirekt ist für mich jetzt unnötig die Information was können Studierende aus dem achten Fachsemester, das interessiert mich eigentlich nicht. (..) Ich mein, das ist- es ist nett zu wissen, (.) aber irgendwie sorgt es für mich auch dafür, dass es irgendwie so schwammig von 'ja, da sind jetzt irgendwie son Mix aus allen Studierenden drin und der Mix wird sich halt durchschnittlich so auf das fünfte Fachsemester beziehen'. Das heißt, ich werde eigentlich bei dieser Datenlage wahrscheinlich immer nur so ein fünftes Fachsemester im Durchschnitt haben, mit dem ich mich vergleichen kann und interessanter fänd ichs tatsächlich, wenn man konkret sich jetzt auf der Entwicklungsstufe, wo man gerade ist, einordnen kann. #00:12:31-2#
27 I: (...) Ja. Okay. (..) Aber ähm von den- von den, weiß ich nicht, Texten, Beschreibungs- Beschreibungstexten oder diesem einleitenden Texten, da war jetzt nichts bei, wo du beim Lesen gedacht hast: 'Okay, das hätt man jetzt auch echt einfach weglassen können'? #00:12:49-2#
28 A2: Nein, also die Sachen, die da sind, find ich gut und ganz ehrlich, wenn man weiß, wie n boxplot funktioniert, kann man halt den kurzen Absatz überspringen. (.) Also- #00:12:57-0#
29 I: (.) Okay. (.) Ähm dann wir hams schon an son paar Stellen angesprochen, aber ich frags trotzdem nochmal, vielleicht gibts ja noch mehr und zwar: Welche Fragen jetzt noch offen sind für dich? (..) Also in Bezug auf die Rückmeldung und die Interpretation der Rückmeldung und so weiter. #00:13:14-9#
30 A2: Also zum einen natürlich das, was ich gerade meinte, mit konkret Vergleich zu meiner aktuellen Entwicklungsstufe, also wirklich Studierende auf derselben (.) äh Stufe, also vielleicht dann auch gar nicht nach Fachsemester, sondern vielleicht eher nach ordnen, was für Veranstaltungen hat man schon gehört, weil das ist glaub ich der interessante Punkt, gerade, wenn Leute Veranstaltungen durcheinander hören, (.) das man guckt: 'Leute, die diese Veranstaltung schon gehört haben, sind durchschnittlich so gut', das fänd ich vielleicht auch so jetzt als Lerneffekt, wie viel nehm ich eigentlich aus meinem Studium mit, ne ganz interessante Info. (..) Ähm (...) und zum anderen den Punkte, den ich vorher schon angesprochen hatte bezüglich: Wie kann ich jetzt konkret die Bereiche, die ich jetzt hier feststelle als die haben Verbesserungspotential eben zum einen auf den Test beziehen, um die für mich ein bisschen klarer zu haben, welche Punkte ich da eigentlich vielleicht nochmal genauer reingucken sollte, und dass dann eben auch auf andere Situationen besser transferieren zu können. (.) Das sind, denk ich, die beiden Hauptpunkte. (.) Und ansonsten (.) fällt mir noch irgendwas ein? Gute Frage. Mhm (8) glaube tatsächlich für die Rückmeldung nicht, ne. #00:14:41-7#
31 I: Okay. Beim Test sonst noch? Das (.) Oder- also weil du jetzt so gesagt hast für die Rückmeldung nicht? #00:14:48-0#
32 A2: Äh ich- ich hab gerade überlegt, bei dem Test. Ich erinnere mich natürlich jetzt nur noch äh teilweise an den Test, aber der Test war doch immer so aufgebaut: Wir sehen einen Videoausschnitt, dann durften wir uns einen Bereich auswählen, für den wir Rückmeldung geben mussten oder durften und dann durften wir uns eine von vorgegebenen Rückmeldungen aussuchen. #00:15:06-7#
33 I: Ja, also mehrere, aber ja. Also mehrere Rückmeldungen // aber- #00:15:10-3#
34 A2: Ja, genau. // Ähm was mir jetzt zum Beispiel an einer Stelle passiert ist, einfach mal die Frage, ob das immer so ist? An einer Stelle zum Beispiel (.) äh habe ich von dem Test die Nachricht bekommen 'zu diesem Thema wurde schon genug Rückmeldung gegeben, da- das überspringen wir jetzt. #00:15:26-7#
35 I: Mhm[Ja]. #00:15:28-0#
36 A2: Ist das einfach nur, wenn ich immer wieder zu dem selben Punkt Rückmeldung gebe, oder wann kommt diese Nachricht? #00:15:32-9#

37 I: Äh, das sind- das sind festgelegte Aspekte. Aber das heißt nicht, dass das dann irgendwie grundsätzlich gut oder grundsätzlich schlecht ist, sondern das sind einfach Dinge, wo diese Aufschlüsselung nicht effizient ist für n Mehrwert an Informationen. (.) Also das kann passieren, wenn es einen Punkt gibt, der ähm in dieser Situation einfach nicht so wichtig ist, dass man sagt, es ist alles nicht so wichtig, egal, welche Rückmeldung man hier gibt, oder es ist sowieso alles wichtig äh und in beiden Fällen ergibt dann eben so ne Aufschlüsselung nicht mehr unbedingt Sinn, sondern dauert einfach nur n Moment länger und so macht mans n bisschen zeit- äh zeiteffizienter. #00:16:15-8#

38 A2: Weil an der Stelle fänd ichs dann vielleicht interessant- ich mein, ich weiß nicht, inwiefern das für diese Art von Test sinnvoll ist, da hab ich keine Ahnung von, aber ähm wenn man einfach auch an der ein oder anderen Stelle einen relevanten Aspekt vorgibt zu dem Rückmeldung gegeben werden soll. Also gerade ich habe jetzt mit dem Test auch festgestellt, ja, ich such mir halt die Punkte raus, zu denen ich Rückmeldung geben möchte, zu denen mir auch was einfällt, (..) aber das ist ja nicht zwingend was sinnvolles. Also, es wäre vielleicht auch mal sinnvoll, über Dinge (.) zu reflektieren und Rückmeldung zu geben, wo einem nicht sofort einfällt, dies das Ananas, er hat halt die ganze Zeit Mädchen nur als 'du kannst schön schreiben' dargestellt, ne? Also (..) das man gerade an der Stelle vielleicht auch guckt: Okay, dieser Punkt, da ist was Interessantes, was fällt dir denn da auf? Ich weiß nicht, ob das aus eurer Perspektive sinnvoll ist, das zu tun, (aber?) das (.) hätte mir glaub ich in dem Test ganz gut getan, weil ich dann nicht immer nur zu Punkten, die ich sehr offensichtlich fand, Rückmeldung gegeben hätte. #00:17:18-7#

39 I: Ja. Ne, wir haben das auf jeden Fall auch in der Diskussion. Das ist- ist n Punkt auf jeden Fall. (..) Okay. Dann äh hatten wir die Fragen, die offen sind und jetzt auch noch ne Frage, da hast du auch schon einiges zu angesprochen, aber welche weiteren Informationen noch hilfreich gewesen wären, welche du gern noch gehabt hättest in der Rückmeldung? #00:17:41-5#

40 A2: (4) Hm ist das nicht die selbe Frage wie grade? #00:17:48-8#

41 I: Ja, je nachdem, wie man drauf antwortet, bei dir kommt jetzt wahrscheinlich das gleiche, das ist auch okay. Manchmal ist so ne Umformulierung bringt nochmal neue Ideen. #00:18:00-3#

42 A2: Welche Informationen hätte ich gerne gehabt. Ähm (.) Vergleich zu Studierenden aus selber Entwicklungsstufe, eher bezüglich welche- welchen Input hatten sie schon in der Uni. #00:18:15-5#

43 I: Ja. #00:18:16-2#

44 A2: Als bezüglich wie alt sind sie, (.) also welches Fachsemester. (.) Ähm (.) außerdem konkret: Wie kann ich mich denn in dem Bereich verbessern? Und auch: An welchen Punkten im Test kann man den jetzt festmachen 'ich bin auf- ich bin so gut oder so schlecht'. Schlecht und gut natürlich in Anführungsstrichen. (.) Ähm die beiden Punkte. (.) Ähm (6) was ich tatsächlich dann an der Stelle auch noch interessant fände puncto wie kann ich mich denn verbessern (.) also wenn man jetzt konkret versucht, zu den einzelnen Bereichen zu gucken, was sind da die Punkte gewesen, an denen man das festmacht und wie kann man sich da verbessern, auch einfach vielleicht konkret sagen: 'Okay, das sind Sachen, die kommen dann erst in der Veranstaltung, das musst du noch gar nicht können.' oder vielleicht auch zu sagen: 'Okay, das waren Dinge, die waren in der Veranstaltung, die hast du schon gehört, guck da doch nochmal rein.' (.) (Also?) dass man vielleicht konkret diesen Bezug herstellt, wenn man am Anfang gleich abfragt, die und die Veranstaltung gibt es, mach n Haken, welche du schon gehört hast und welche noch nicht, dass man da konkret sagt: 'Okay, der Punkt, die Veranstaltung, guck da nochmal rein, der Punkt kommt noch, in der Veranstaltung, kannst du dich drauf freun, bist du jetzt motiviert für', keine Ahnung. (..) Sowas in die Richtung fände ich vielleicht dann an der Stelle noch ganz interessant. #00:19:39-9#

45 I: Mhm[Ja]. Okay, nehm ich mit. Ähm ich hab jetzt hier tatsächlich auch noch so ne Liste so von, was ist das, sechs, sieben Punkten. Ich würd die einmal durchgehen, also das sind Optioenn, die man noch ergänzen könnte (.) ähm und da würd mich dann interessieren, wie du das einschätzt, also ist das ne sinnvolle Ergänzung oder ist das irgendwie unnötiger mehr Leseaufwand oder so. (.) Ähm das erste wäre mehr Information oder- oder ja klarere Information über das Ziel des Tests oder auch das Ziel des Projekts in dem Rahmen, in dem der Test entstanden ist. (.) Also quasi was so das- (.) was soll das hier, was wir machen. #00:20:19-9#
46 A2: (...) Ähm also an der Stelle muss ich dann gleich noch einfach mal Rückfragen, was denn überhaupt euer Ziel ist, weil bisher ist das einzige Ziel, was son bisschen durchgekommen ist, dass ihr uns einfach ne Rückmeldung geben wollt, auf welchem Stand wir sind. Beziehungsweise Herr N hatte das dann auch so formuliert, äh dass er uns einfach n bisschen besser kennenlernen möchte, welche Leute in seiner Veranstaltung sind. (.) Wenn das denn das Ziel einzig und allein ist, dann ist das gut rübergekommen, wenns noch n anderes gibt, dann müsste man das vielleicht formulieren, je nachdem, was euer Ziel ist. #00:20:53-7#
47 I: Ne, das kommt schon ganz gut hin mit dem Ziel. Also wenn das rübergekommen ist, ist gut. (.) Okay. Ähm das nächste wäre explizite Informationen darüber, was ist denn jetzt der Zielzustand Reflexionsfähigkeit? Ist der Zielzustand überall Hundert Prozent? Oder ist der Zielzu- Woran mach ich den fest? (..) Ob man sowas noch ergänzt. #00:21:16-8#
48 A2: (6) Ähm ja, das fänd ich tatsächlich gar nicht unsinnig zu wissen, was habt ihr jetzt als Ausgangszustand genommen? Also hat Herr N einmal den Test bearbeitet und das sind jetzt Hundert Prozent, das wär jetzt zum Beispiel interessant zu wissen, oder ist das jetzt jemand, der grad seinen Bachelor hat oder jemand, der grad seinen Master hat oder jemand der jetzt schon etwas länger als Lehrer schon arbeitet. Das wäre tatsächlich interessant, weil das ja wahrscheinlich auch nen ziemlichen Unterschied macht, äh aus welchem Blickwinkel man da drauf guckt. (.) Ähm (.) das wäre tatsächlich interessant, weil ich jetzt zum Beispiel glaube nicht, dass ich es für mich als sinnvoll empfinden würde, wenn ich jetzt einen Test bearbeite, der mich im Vergleich zu jemandem setzt, der jetzt schon seit zwanzig Jahren Lehrer ist. Ich glaube, dieser- diese Person würde ganz anders an viele Sachen herangehen als ich und auch teilweise einfach nur aus Gewohnheit, während ich jetzt vielleicht noch versuche, da irgendwelches Wissen aus Veranstaltungen drauf anzuwenden. Ich glaube, das wäre zum Beispiel n Vergleich, der wäre mir dann gar nicht so wichtig. (.) Also zumindest die absolute Zahl, wie gut ich dann abgeschnitten hab, wäre mir nicht so wichtig. (.) Ähm deswegen wäre das in der Tat interessant. Ich bin jetzt auch einfach davon ausgegangen, dass ihr von so einem universitären Standpunkt äh was ihr gelernt haben solltet da ran geht. #00:22:41-4#
49 I: Mhm[Ja]. (.) Okay. Äh das nächste wäre, inwiefern ne feinere Aufschlüsselung der Bewertung interessant wär. (.) Also dass man zum Beispiel- verschiedene Optionen denkbar, ein Beispiel wäre, dass man sagt, man geht die einzelnen ähm Aufgaben durch und und gibt dazu immer irgendwie Prozenzangaben oder Punkte oder sowas. (.) Oder teilt ähm teilt nicht- also wir habens ja hier in Alternativen und Bewertung, sondern versucht das inhaltlich noch weiter aufzusplitten. #00:23:13-9#
50 A2: (..) Da steckt ja im Wesentlichen son bisschen der Punkt drin, der ich jetzt immer schon wieder mal so bisschen umstreift habe, dass zumindestens ich mit diesen Begriffen Alternative, Bewertung, Schülervorstellung, auch mit den Erläuterungen dadrunter zwar als Begriff was anfangen kann, aber die nicht wirklich im Test verorten kann. (.) Und an der Stelle, wenn man wirklich- Anders: Wenn ich da jetzt wirklich für mich konkret Handlungsoptionen aus dieser Reflexion ableiten wollte, dann bräuchte ich das, dass ich konkret jetzt in die Aufgabe reingucken kann und an- sage, an der Stelle kann man sowas oder was anderes halt festmachen. Das bräuchte ich, um Handlungsoptionen ableiten zu können. #00:24:01-8#

51 I: Okay. Ähm da hör ich jetzt raus, korrigier mich bitte, wenn das falsch ist, dass so ne Beispielaufgabe ne sinnvolle Sache wär oder ne Beispielaufgabe oder Musterlösung mit Begründung, sowas da mal reinzu- (.) also ne- in die Rückmeldung zu ergänzen, um das da am Beispiel zu haben. Sehe ich das richtig? #00:24:20-2#
52 A2: (...) Ja, das wäre auf jeden Fall sinnvoll. Ähm (..) ich ganz persönlich würde mir nach so einem Test auch die Zeit nehmen wollen, dann nochmal durch den ganzen Test durchzugehen und an jeder einzelnen Frage zu gucken äh das ist aber natürlich sehr zeitintensiv. Aber eine Musteraufgabe würde ich auf jeden Fall äh in der Lösung als sinvoll erachten. #00:24:46-4#
53 I: Okay. (.) Dann ähm die Frage nach dem Diagrammtyp, also ob diese Boxplots gut sind oder obs schöner wäre andere Diagramme zu haben, entweder stattdessen oder zusätzlich. #00:25:00-1#
54 A2: Was wär denn ne Alternative? #00:25:03-6#
55 I: Man könnte Säulendiagramme machen, oder Punkte oder so, also zum Beispiel ne Säule machen mit 'hier bin ich' und daneben ne Säule mit 'hier ist der Durchschnitt der Studierenden' idealerweise dann in meinem Semester oder bei unserem Datenpool dann die, die wir haben. #00:25:21-9#
56 A2: (..) was ich an den Boxplots eigentlich ganz schön finde ist, dass man ja in den allermeisten Fällen mit seinem roten Punkt innerhalb der Box liegen wird. (.) Und zumindestens für mich heißt das so lange ich innerhalb der Box liege, das, was ich mache ist halbwegs okay, ich muss mich jetzt nicht vollkommen schlecht fühlen, sondern ich kann das Ganze als Motivation nehmen. (.) Ich glaube gerade, wenn man jetzt zu nem Säulendiagramm übergehen würde, dann würde dieser Unterschied von 'die anderen sind da oben und ich bin da' n bisschen stärker noch in den Vordergrund treten, als er das jetzt tut und ich glaube, dass es dann eher das Potential hätte, eben wenn man denn schlechter ist, son bisschen zu frustrieren und so ist es eben so halbwegs 'ja, ich bin halt irgendwo da in dem Durchschnitt, aber da hab ich auch das Potenzial bisschen besser zu werden' (..) ähm das find ich glaub ich ganz sinnvoll ähm (8) Eine Sache, die ich mich jetzt tatsächlich bezüglich der Boxplots gefragt habe, einfach weil ich keine Ahnung von Boxplots hab: Diese ähm Striche, die noch bis ganz oben und unten gehen, (.) heißt das, dass jemand auch bis zu diesem Punkt mal ein Ergebnis geschafft hat? Also dass das das- Okay, das sind die Maximalausreißer, okay. #00:26:39-9#
57 I: Genau #00:26:41-0#
58 A2: Ähm (...) ja, ist auch ne ganz interessante Info, die einfach zu Boxplots dazu gehört, ich glaube, das macht es jetzt nicht besser oder schlechter, aber die bräuchte ich zum Beispiel gar nicht, aber die gehört halt dazu. #00:26:53-4#
59 I: Du meinst diese Antennen jetzt, also // (unv.) Mitte reichen. #00:26:57-7#
60 A2: Ja, genau die. // (..) Ja. #00:26:58-7#
61 I: Okay. #00:26:59-8#
62 A2: Also wenn mans anders darstellen wollte, könnte man halt auch irgendwie son- hier ist son Durchschnittsbereich und du bist halt so im- also (.) es ist halt genau ne Box, die man auch anders anordnen könnte. Also 'hier ist so der Durschnittsbereich [A2 deutet mit Gesten eine horizontale Box an] und du bist irgendwo in diesem Durchschnitt' finde ich eigentlich ne ganz gute Sache sofern man denn im Durschnitt ist. (..) Also wenn man halt außerhalb dieser Durschnittsbox ist, dann sind diese Antennen halt wahrscheinlich praktisch, um zu sehen, ja okay, ich bin doch noch irgendwie bei den anderen mit dabei (.) ähm (..) ja, aber ich finde halt grundsätzlich diese Boxen umgehen halt dieses (..) wenn man jetzt n Säulendiagramm hätte, die andern sind soo hoch und ich bin nur irgendwo da unten. Das wirkt halt hier son bisschen freundlicher. #00:27:43-0#

63 I: Okay. Ähm genau dann hattest du jetzt schon öfter angesprochen, habe ich nämlich auch auf meiner Liste: Material oder Hinweise, um die Reflexionsfähigkeit- also gut, ich hab mir jetzt augeschrieben 'um das zu schulen', du hast gesagt 'wie mach ich denn jetzt weiter?' ähm da nehm ich einfach mit: das ist sinnvoll. Oder hast du da noch ganz konkret irgendwie ne Vorstellung, wie man das am besten macht? #00:28:07-5#
64 A2: Also zumindestens ich in meinem Studium muss ziemlich häufig reflektieren. Sei es jetzt fürs EOP [Eignungs-Orientierungs-Praktikum?] oder jetzt gerade fürs BFP [Berufsfeld-praktikum?] (.) oder eben auch die Abschlussreflexion im physikalischen Grundpraktikum oder oder oder. Also so grundsätzlich diesen Dreischritt oder wie viel Schritte auch immer. Es ist ja auch manchmal ein Fünfschritt, wenn du nach Bräuer vom PLAZ [Institut anonymisieren?] gehst oder was weiß ich was. (.) Ich glaube, den hat man bis zum gewissen Punkt verstanden. Zumindestens an der Stelle im Semest- äh im Studium, wo ich jetzt bin. (..) Das würde ich dann vielleicht auch son bisschen daran anpassen, in welcher Phase des Studiums man gerade ist, weil an ner gewissen Stelle muss man, finde ich, nicht mehr erklären, wie reflektiert man, das kann man irgendwann, das hat man oft genug gemacht. Was man hingegen tatsächlich bräuchte ist eben konkret jetzt: Was mache ich mit meiner Reflexion? Was mache ich jetzt mit dieser Erkenntnis 'okay, in dem Bereich A habe ich halt noch Verbesserungspotential', wie setze ich das um? #00:29:04-3#
65 I: (..) Okay. Und ähm der letzte Punkt dieser Liste: Inwiefern das gut wäre, Literaturhinweise oder Hinweise auf irgendwie physikdidaktische Grundlagentexte oder ähnliches zu geben, zum Beispiel- gut, wenn mans jetzt dafür macht, was du gerade angesprochen hast, ne, um irgendwie, wie setz ich das um? Da könnte das- könnte mans natürlich ansiedeln, man könnte das aber auch zum Beispiel zum Begriff der Reflexion oder zum Konzept noch weiter machen. Ähm, also es an verschieden Stellen denkbar. #00:29:40-3#
66 A2: (4) Grundsätzlich finde ich es sinnvoll, wenn man die Möglichkeit hat, an Stellen, die einen interessieren, nachzulesen. (.) Das heißt, wenn man die Möglichkeit hat, (.) 'wenn es dich interessiert, weitere Informationen, klick hier'. (..) Aber eher so in so nem optionalen Stil, das sind dann Dinge verlinkt, die man anklicken kann, wenns einen interessiert. Ich glaube, es zum festen Bestandteil zu machen, halte ich an den meisten Stellen nicht für sinnvoll. Das einzige, was ich halt ähm durchaus für sinnvoll halten würde ist das, was ich vorhin schonmal so bisschen angerissen hatte, (.) dass man eben, um konkrete Handlungsoptionen haben zu können sagt: 'Okay, den Punkt hattest du schon, oder soll- oder der kommt noch und für den Fall, du hattest den schon: Hier kannst du nochmal nachlesen und für den Fall kommt noch: Wenn du schonmal n bisschen vorweg lesen willst gibts hier was.' (.) Also den Punkt (..) gerade halt in dem Fall, dass man vielleicht an irgendeinem Bereich eher schlecht abgeschnitten hat, würde ich (.) als sinnvoll erachten, (.) ähm allgemein zum Reflexionsbegriff: Ich bräuchte es glaube ich nicht, aber es irgendwo reinzupacken zum 'für die, die es interessiert, hier nochmal ne kurze (.) Literatur', warum nicht. Mich würds nicht stören, wenns da n kleinen Link gäbe in dem Ganzen. #00:31:02-3#
67 I: (...) Alles klar. Ähm dann hab ich noch n kurzen letzten Fragenblock, also sind auch nur zwei Fragen, ähm zum ja dazu, wie du die Rückmeldung bekommen hast. Also es war ja so, dass ihr den Link bekommen habt zu dieser Website, wo ihr euern Code eingeben musstet und die runterladen konntet. (.) Da wär einmal die Frage, obs da irgendwie Schwierigkeiten gab, diese Rückmeldung zu bekommen und wenn ja, welche das waren? #00:31:31-8#
68 A2: (5) Ähm (..) Schwierigkeiten gabs nicht, wir haben ja die äh den Link einfach in einer Präsenz- also Onlinepräsenzveranstaltung von Herrn N bekommen (.) und das war eigneltich ziemlich selbsterklärend. Also (.) das war völlig unproblematisch, da gibts nicht viel zu zu sagen. #00:31:59-3#

69 I: Okay. Und ähm die Anschlussfrage ja quasi in welcher- in welcher Form diese Rückmeldung (.) sinnvoll ist. Also ist- ist der Erhalt wie der war, also über diesen Download äh Download ist das sinnvoll oder wär da n anderes Format wünschenswert? #00:32:17-3#
70 A2: (5) Was mich tatsächlich hier an der Rückmeldung so n bisschen äh gestört ist falsch, aber ich fands n bisschen suboptimal: Wir haben ja zwei Reflexionstests bearbeitet und bei dem anderen zum Beispiel, also diesem äh hier diesem Erklärtest, ich weiß nicht, ob du weißt, was ich meine? #00:32:39-3#
71 I: Ja. #00:32:39-8#
72 A2: Ähm bei dem ham wir direkt im Anschluss ne Rückmeldung bekommen, als man auch noch präsent hatte, (..) äh was man gerade gemacht hat. Und wenn man das jetzt, ich glaube, hier wars jetzt ein, zwei Wochen später, (.) ich musste erstmal reingucken und war so: Joa, ich hab jetzt in den Bereichen so abgeschnitten, aber was jetzt überhaupt in dem Test nochmal bezüglich den Bereichen passiert ist, kann ich mich jetzt da zwei Wochen noch später dran erinnern? Nö. Deswegen wenn das machbar ist, fände ichs gerade, wenn man dann schon nen Datenpool hat, auf den man das beziehen kann, (.) ähm sinnvoll, direkt ein Ergebnis anzugeben (.) und das eben nicht erst ähm, wenn das technisch machbar ist bei euch, eben jetzt ne Zeit später (..) ähm sich angucken zu lassen, weil ich unterstell auch einfach mal, dass je länger das Ganze her ist, desto weniger interessiert man sich für die Rückmeldung. #00:33:33-6#
73 [...][Ab hier: Abschluss]

P.1.3 Interview 3: A3

1 [...][Hier: Einleitung, Einverständniserklärung etc.]
2 I: Genau, ich würd, äh, zum Einstieg einmal fragen wollen, ob Sie sich noch dran erinnern- also an Ihre Rückmeldung zur Bearbeitung des Reflexionstests erinnern können? Und falls ja, welche Schlüsse Sie daraus gezogen haben? #00:01:50-3#
3 A3: Also die hab ich hier vorliegen. Ähm, meine Schlüsse daraus waren eben, dass ich im Vergleich zu den anderen Testergebnissen relativ okay- also zu den anderen Teilnehmern relativ okay abgeschlossen habe. Ähm, auf jeden Fall besser als äh das, ja, das mittlere, also der mittlere Teil. Aber eben, dass noch sehr viel Potential. Zum einen im Bereich der Teilnehmer, die überhaupt teilgenommen haben, aber auch im Bezug auf das ganze Ergeb- äh Gesamtergebnis. #00:02:21-5#
4 I: Okay, danke. [...][Ab hier: Öffnen und Lesen der Beispielrückmeldung] #00:03:20-3#
5 A3: Okay, das ist ja noch allgemeiner Wortlaut [...] Ergebnis liegt knapp über dem Median, scheint der gleiche Datensatz zu sein wie bei der Rückmeldung, die ich erhalten habe. #00:03:38-6#
6 I: Ja, der Datensatz ist der gleiche. #00:03:40-2#
7 A3: Mhm, alles klar, dann weiter. (8) Okay, anscheinend hat die Person unterschiedlich stark in den verschiedenen Bereichen abgeschnitten. Ähm, bei den Alternativen eben sehr im Mittel, also direkt auf dem Median. Bei den Bewertungen tatsächlich auf der oberen Quantilgrenze und bei den Schülervorstellungen auf der unteren. ... Okay ... [...][Ab hier: Weiterlesen] #00:04:41-6#
8 A3: Oh, mit fällt noch auf, dass das korrigiert wurde mit den- bei mir steht noch Eins Komma Null Null oben. Mittlerweise steht da ja hundert Prozent, was ja formal richtig ist. #00:04:50-7#
9 I: Ja, ja, danke für den Hinweis trotzdem, das ist mir erst zu spät aufgefallen. #00:04:55-0#
10 A3: Kein Problem. #00:04:56-8#
11 I: Ähm, genau, dann wäre meine Frage- Achso ich scroll natürlich auch in Zukunft noch hoch und runter, ne, wenn da irgendwie Bedarf besteht. #00:05:04-2#
12 A3: Mhm[Ja]. #00:05:04-5#
13 I: Ähm genau meine erste Frage wär, ob Sie vor dieser Rückmeldung schonmal mit- mit solchen Boxplots gearbeitet haben und wenn ja, in welchem Rahmen? #00:05:13-3#
14 A3: Nein, hab ich nicht. (.) Ich äh (muss-?) hab mich tatsächlich damit beschäftigt, was das überhaupt ist ähm also ich hab mir grad jetzt davor zehn Minuten genommen um da genauer reinzuschauen ähm und wie halt son Boxplot generiert wird. #00:05:28-3#
15 I: Okay. Und äh haben Sie das Gefühl, das hat soweit geklappt oder (.) äh (.) // war das doof? #00:05:39-3#
16 A3: Also ähm (unv.) (.) ja also tatsächlich vielleicht in Bezug auf diesen Reflexionstest ähm ist es tasächlich ein Teil den ich äh eher als negativ angesehen habe ähm dass es nicht sofort klar war, was dieses Ergebnis bedeutet. Ähm zum einen auch, dass das Ergebnis nicht in ähm also nochmal in Worten geschildert wurde, also was quasi ich dort sehe also wo m- wo ich mich quasi einmal zum Verhältnis zur ähm Testgruppe oder zu der ähm Referenzgruppe befinde und im Verhältnis zum Gesamtergebnis. Nachdem ich mich damit dann beschäftigt hab, was n Boxplot ist ähm ist es aber dann klar gewesen. Es ist natürlich dann immer die Frage, was ist hier die Zielgruppe für den Reflexionstest. Also setzen wir voraus, dass das schon bekannt ist? Dafür ist mir aufgefallen, dass dafür, dass man äh Voraussetzt, dass es bekannt ist, wird dann ja doch noch relativ viel zum Diagramm erklärt im wei- im Teil weiter unten. (.) // Das war so mein Eindruck dazu. #00:06:38-4#
17 I: Mhm[Ja]. Und- (..) // Ja. Ähm wenn Sie jetzt die Beispielrückmeldung ansehen, also Sie hatten eben schon bisschen angedeutet, da ist was überm Median, unterm Median, aber welche Schlüsse würden Sie jetzt aus dieser Rückmeldung ziehen? #00:06:50-8#

18 A3: (5) Okay ähm ja ich würd erstmal ähm mit den Schülervorstellungen beginnen und da würde ich schätzen, dass die Testperson die Schülervorstellung ja nicht so gut abbilden kann, das heißt wahrscheinlich selber andere Vorstellungen hat als die Schüler. (Da?) könnt ich jetzt vielleicht ähm überlegen, dass diese Person vielleicht schon länger aus der Schule raus ist ähm diesen Test gemacht hat oder halt sich noch nicht so sehr mit dem Thema Schülervorstellungen überhaupt auseinandergesetzt hat, das heißt, vielleicht davon ausgehend, dass Sachen bekannt sind, die bei Schülern dann ähm ja eher un- unbekannt sind, sag ich mal. Mhm in Bezug auf die Alternativen würde ich schätzen, dass diese Person schon nen ähm recht gutes Fachwissen hat mhm zum einen auch, was den Didaktikbereich angeht, ähm so dass sie die Möglichkeit hat halt Alternativen vorzuschlagen und das das spielt dann wahrscheinlich auch weiter in die- ähm in dieses positive Ergebnis bei der Bewertung mit ein, ähm weil sie ja da durchaus besser abgeschnitten hat als der Großteil im Test (unv.). #00:08:04-3#
19 I: Mhm[Ja]. Und ähm wie ist das, also ich scroll einmal hoch zu dem andern Plot zum- ja quasi zu diesem Gesamtergebnis. #00:08:14-0#
20 A3: Mhm[Ja]. #00:08:15-0#
21 I: Ähm (.) ja genau, wie ordnen Sie das ein und vielleicht auch welches Gewicht würden dem Gesamtergebnis oder den Einzelergebnissen geben. #00:08:26-6#
22 A3: Also ähm so wie das hier dargestellt ist, macht das für mich den Eindruck, als wärn die Schülervorstellungen etwas weniger stark gewichtet, weil wenn alles äh relativ gleich stark gewichtet werden sollte, dann müssten die Ergebnisse ja auf dem Median liegen #00:08:42-0#
23 I: Mhm[Ja]. (.) Okay. Danke. Ähm wenn Sie jetzt die Rückmeldung angucken bzw. angeschaut haben, welche Informationen sind da die, die sich also irgendwie besonders hilfreich oder besonders interessant erachten? #00:08:59-9#
24 A3: (4) Ja im Prinzip das, was der Boxplot liefert, also dass man gleichzeitig nen ähm ja nen Ergebnis hat (so?) wo man sieht, wie man im Vergleich zu anderen Personen steht, die diesen Test gemacht haben, dass man aber auch gleichzeitig sehen kann, wo man im Vergleich zum Maximum so liegt und dass man eben auch sieht, dass das Maximum, was zu erreichen ist also Testergebnis auch nicht äh von ja Personen- oder oder noch von keiner Person der Testgruppe erreicht wurde. Das ist durchaus auch interessant zu sehen. #00:09:34-7#
25 I: Okay und dann andersrum quasi die Frage: Sind Informationen überflüssig und welche wären das? #00:09:44-1#
26 A3: (.) Ähm würd ich tatsächlich nicht sagen, dann würd ich eher sagen, dass noch äh vielleicht Informationen fehlen und zwar das Ganze nochmal verschriftlicht einfach, was ja was dieses Diagramm aussagt und am besten das auch gleich in dem ersten Absatz, damit man sein Ergebnis irgendwie einordnen kann (.) bevor man überhaupt diesen Plot gesehen hatte. #00:10:08-5#
27 I: Der- ähm nur, dass ich das richtig versteh: Der erste Absatz ist immer der Absatz vor dem Diagramm // oder? #00:10:15-1#
28 A3: Ne, // ganz oben, ganz oben also sobald man das aufmacht das Erste, was man liest ähm (.) also das erste, was mich interessiert hatte, als ich diesen äh Test also diesen Rückmeldung zum Test bekommen hab war natürlich ähm wie hab ich abgeschnitten im Vergleich zum dem Möglichen, im Vergleich zu den anderen und das könnte man ja einfach in nem Satz oder zwei Sätzen dort mit aufführen. (unv.) man halt einmal beschreibt, wie es im Vergleich zu der Testgruppe ist und einmal im Vergleich zum Gesamtergebnis. #00:10:40-9#
29 I: Also quasi ganz oben n Satz drüber 'Sie haben insgesamt (.) überdurchschnittlich, unterdurchschnittlich, wie auch immer abgeschnitten.' Versteh ich das richtig? #00:10:50-1#
30 A3: Ähm vom vom Prinzip her // also äh genau. #00:10:53-3#
31 I: Ja, nicht die // Formulierung ja. #00:10:54-6#
32 A3: Ja. #00:10:55-5#

33 I: Genau. Okay. Ähm gibts denn- also da kommt schon- eine Sache ist schon bisschen durchgekommen, aber gibts noch weitere (.) Fragen, die für Sie offen bleiben, wenn Sie die Reflexion so- äh die äh Rückmeldung so lesen? #00:11:10-4#
34 A3: Ähm also so für mich ist jetzt- na doch (unv.)[A3 murmelt, sucht Textstellt] ähm tatsächlich genau die Bewertungskriterien, ähm was das dann konkret bedeutet also ähm für mich fehlte halt so der direkte Bezug zu den Antworten, die ich dann in diesem Test gegeben hab. Irgendwie konnte ich mir das so einigermaßen vorstellen, aber ähm (.) so ne direkte Verknüpfung fehlte dann tatsächlich irgendwie. #00:11:43-1#
35 I: Mh wie könnte die aussehen? #00:11:45-8#
36 A3: (.) Das äh war ich mir jetzt auch nicht sicher, wie man die darstellen könnte, ohne alle Testergebnisse ähm da hab ich leider keine Antwort drauf gefunden, ähm aber die Frage, also die man sich ja stellt: Man beantwortet dort ja Fragen oder äh klickt eben mögliche Antworten an und ähm die Frage geht eher dahin, was sind eigentlich Hundert Prozent oder was wäre Null Prozent (also?) was war da richtig, was war da falsch oder sowas. (.) Mhm und alternativ was mit noch einfällt: Ähm vielleicht ähm so als Ergebnis äh ja ne Handlungsempfehlung für einen selber also ähm im Verhältnis- also wenn man jetzt das Ergebnis betrachtet, ähm das ist jetzt zum Beispiel bei diesem Test- oder bei diesem Beispielergebnis hier wäre, dass man der- äh also das dem ja (.) Teilnehmer empfohlen wird, dass er sich nochmal mit Schülervorstellungen auseinandersetzt. #00:12:42-4#
37 I: Okay. #00:12:45-0#
38 A3: Und äh dann vielleicht auch konkretisiert ähm zum Beispiel nen entsprechendes Buch oder sowas äh empfohlen oder ja entsprechende Links oder ähnliches direkt. Das wär natürlich noch ne Möglichkeit, damit man das Ganze dann auch praktisch nutzen kann, wenn das denn angedacht ist. #00:13:03-6#
39 I: Also Literaturhinweise quasi dann zum Weiterlesen. #00:13:07-4#
40 A3: Genau, genau. Einfach womit man sich nochmal- also je nachdem, was es jetzt ist oder ähm alternativ halt auch Möglichkeiten oder Tipps, wie man das Ganze verbessern könnte. Es kann ja mal n- also bei Schülervorstellungen ist es jetzt vielleicht nen äh nen Literaturhinweis, bei Alternativen oder so, da ist es vielleicht eher so, dass man nen Denkanstoß ge- äh geben muss, wie man wie man da ran geht, also vielleicht einfach mal, selbst wenn man ne gute Lösung gefunden hat, nochmal über ne zweite Nachdenken (und sowas?). #00:13:37-4#
41 I: Mhm[Ja]. (..) Okay, jetzt sind wir äh tatsächlich auch schon mittendrin in meiner nächsten Frage. Ich stell sie trotzdem nochmal als Frage, falls dann noch was einfällt. Und zwar wärs gewesen: Welche weiteren Informationen hätten Sie gerne? #00:13:52-3#
42 A3: (..) Okay, das ähm war ja dann glaub ich in der letzten Frage // etwas beantwortet. #00:13:58-8#
43 I: Ja, genau. // Ähm ich hab da noch so ne kleine Liste an Dingen, die quasi denkbar wären. #00:14:05-0#
44 A3: Mhm[Ja]. #00:14:05-8#
45 I: Ähm ich würd die vielleicht nochmal durchgehen äh und dann würd ich drum bitten, dass sie einfach vielleicht (.) äh einmal dazu sagen: Halt ich für sinnvoll, halt ich überhaupt nicht für sinnvoll und // genau. #00:14:18-0#
46 A3: Mhm[Ja]. // #00:14:17-3#
47 I: Ähm das erste wäre obs ja von Interesse wäre, weitere Informationen oder konkretere Informationen über das Ziel des Testes und vielleicht auch das Ziel des Projekts in dem dieser Test entstanden ist, zu bekommen. #00:14:31-7#
48 A3: (4) Mhm (..) also ich glaube auf jeden Fall ähm was das Ziel- was das Ziel des Testes ist- äh Ziel im Sinne von Ziel für einen selber oder das Projekt des Tests jetzt oder? #00:14:50-8#
49 I: Also quasi (.) ähm wenn ich diesen Test bearbeite, // was krieg ich dann raus? #00:14:57-5#

50 A3: Mhm[Ja]. //#00:14:56-5#
51 I: Was ist das Ziel, was das Ergebnis vom Test mir geben kann sozusagen. #00:15:01-6#
52 A3: (.) Ja, das wär auf jeden Fall interessant, denk ich. #00:15:05-0#
53 I: Okay. (..) Äh das nächste wär, das war auch schonmal n bisschen angerissen, ähm inwiefern explizite Informationen darüber interessant wären, was denn jetzt dieser Zielzustand Reflekti- äh Reflexionsfähigkeit ist. Also sind es immer die Hundert Prozent oder ist der Zielzustand- macht der sich vielleicht auch an anderen Kriterien fest oder (..) sowas in die Richtung. #00:15:29-8#
54 A3: (..) Auf jeden Fall, ja. Das wär interessant. #00:15:35-2#
55 I: Okay. Einmal nur als Nachfrage: Hab ich das richtig verstanden, dass Sie als Zielzustand die Hundert Prozent interpretiert haben? #00:15:43-0#
56 A3: Ähm tatsächlich hab ich als Zielzustand die Hundert Prozent interpretiert ähm hab aber auch gleichzeitig gesehen, dass es noch nicht in der Testgruppe erreicht wurde. #00:15:52-2#
57 I: (...) Ähm das nächste wär ne feinere Aufschlüsselung der Bewertung. Also dass man jetzt- Wir haben ja hier zum Beispiel die Aufschlüsselung in diese drei Teilergebnisse, man könnte sich ja aber auch vorstellen, das noch feiner zu gliedern oder sogar nach Aufgaben zu gliedern oder was auch immer. Da sind ja verschiedene Sachen denkbar. #00:16:16-4#
58 A3: (5) Mhm, das würde meiner Ansicht nach nur Sinn machen, wenn das gleichzeitig mit ner Konkretisierung einhergeht, also dass man es wirklich dann auf die Aufgaben bezieht, die dort gestellt wurden im- im Test. Vielleicht ähm wäre es aber auch ähm (.) schon alleine hilfreich, dass man ähm die Aufgaben also im- aus dem Test jetzt ähm irgendwie diesen Schülervorstellungen nochmal zu ordnet oder diesen diesen drei Bewertungskriterien vielmehr, also das äh in den- also das heißt 'hier wurde vor allem auf die Schülervorstellungen Wert gelegt' oder 'diese Aufgabe zielte dahin' oder 'hier ging es um die Alternativen'. #00:16:57-5#
59 I: Also dass man dann alle Aufgaben quasi nochmal sieht und die zugeordnet hat, versteh ich das richtig? #00:17:03-6#
60 A3: Genau in in irgendeiner Form oder zumindest ähm also es muss ja nicht die komplette Aufgabe sein, aber das es irgendnen Wiedererkennungswert gibt. Zum Beispiel gabs ja diese eine Aufgabe, wo es um die Titanic als Beispiel geht oder sowas, dass man halt äh den Aufgaben, sagen wir mal, Namen gibt, die dann im Test auch wiedererkannt werden können, wenn dann nachher nur der Name da steht, dann müsste nicht mehr die ganze Aufgabe dort stehen wahrscheinlich, weil das wird ja dann sehr (unv.) umfangreich. #00:17:30-1#
61 I: Mhm[Ja]. Okay. Und ähm also gerade wenn Sie jetzt gesagt haben, dass Sie die Boxplots vorher noch nicht kannten oder vorher noch nicht damit gearbeitet haben, hätten Sie sich da andere Diagramme oder ne andere Art der Ergebnisdarstellung gewünscht? #00:17:45-9#

62 A3: Ähm jetzt im Nachhinein ist das eigentlich ziemlich gute Art der Ergebnisdarstellung. Am Anfang konnte ich damit eben nichts anfangen mhm es ist aber wahrscheinlich ein Problem, was man ganz leicht beheben könnte, indem man vielleicht ähm das Diagramm selber nochmal zusätzlich beschriftet. Also man könnte ja einfach mal diesen ähm Bereich also vom oberen Teil der Antenne bis zum unteren Teil der Antenne einmal ähm ja, ich würds markieren, entweder mit ner geschweiften Klammer oder halt ähm nach außen rausziehen und kurz dranschreiben, dass das eben äh ja quasi alle Ergebnisse sind. Und dann eben das, was quasi auch unten im Text nochmal dran steht halt einfach nur an die Diagramme beschriften, dass man dann eben äh ja an den Median dran schreibt, dass es der Median ist und äh an dieses äh ja Mittlere, diesen mittleren Teil eben, dass es die mittleren fünfzig Prozent sind und äh ja dann entsprechend könnte man das sehen und müsste ohne oder könnte ohne, dass man weiß, was n Boxplot ist eigentlich schon verstehen, was einem dieses Diagramm aussagt, was dann ähm ja im ersten Draufgucken sehr einfach macht. Das würde meiner Ansicht dann ähm auch im hinteren Teil diese ganzen Erklärungen zu dem Box- Boxplot relativ hinfällig machen, also (die?) könnten dann wahrscheinlich rausgenommen werden ähm unds würde dann reichen eben zu erwähnen, dass es n Boxplot ist, weil dann haben die Personen, die ähm sich näher damit beschäftigen möchten, einfach die Möglichkeit einfach nochmal nach dem Schlagwort Boxplot zu suchen und zu gucken, was das ist, während die Personen, die noch keine Ahnung von nem Boxplot haben äh und das auch nicht möchten, also sich nicht damit näher befassen möchten, die können eben trotzdem alleine mit dem Diagramm etwas anfangen und haben im Prinzip das gleiche Ergebnis. #00:19:30-4#
63 I: Mhm[Ja]. (.) Okay. Ich hab jetzt noch zwei, drei Punkte auf dieser Liste von denen ich aber glaube, wir haben schon drüber gesprochen und zwar einmal, dass man ne Beis- vielleicht ne Beispielaufgabe gibt äh und da ne Musterlösung mit Begründung zu? #00:19:49-5#
64 A3: Mhm[Ja], ja. #00:19:50-4#
65 I: Ist von Interesse. #00:19:52-3#
66 A3: Auf jeden Fall, ja. Haben wir ja- also hatten wir drüber gesprochen, genau. (unv.) würd ich auch sagen, macht Sinn. #00:19:58-0#
67 I: Mhm[Ja]. Okay und das andere waren äh genau entweder Literaturhinweise oder Hinweise auf physikdidaktische Grundlagentexte ähm entweder eben ja zu Handlungsempfehlungen, gegebenenfalls könnte man sich das ja auch denken ähm zum Beispiel am Anfang in der Rückmeldung wo irgendwie dieser Begriff der Reflexionskompetenz beschrieben wird, dass man da nochmal n Verweis zu gibt? #00:20:22-3#
68 A3: Mhm find ich tatsächlich da gar nicht so wichtig beziehungsweise ähm vielleicht auch schon zu viel ähm sondern eher im Zuge der Handlungsempfehlung. #00:20:33-0#
69 I: Okay. Gut. Dann ähm wär ich mit diesem Block soweit durch ähm es sei denn, Sie wollen da grad noch was ergänzen? #00:20:44-7#
70 A3: (unv.), aber wenns dabei um (15) verbessern könnte. [Verbindungsschwierigkeiten] #00:21:03-6#
71 [...][Verbindungsschwierigkeiten, Rückkopplung] #00:21:40-8#
72 I: Ich (hab?) jetzt leider die den letzten Punkt von Ihnen nicht mehr mitbekommen, (.) können Sie das einfach nochmal sagen? #00:21:47-5#
73 A3: Ähm genau also im Prinzip ging es ja um Literaturempfehlungen beziehungsweise Literaturverweise und ähm da bin ich eben der Meinung, dass es nur im Zuge einer Handlungsempfehlung Sinn macht. Also bei dem äh ja dem 'was denn reflektiert wird' [gemeint ist: Beschreibung Reflexionsbegriff/-modell], da würde ichs eben nicht einfügen, sondern später, damit man sich dann konkret damit befassen kann, // (.) wo man sich noch verbessern kann. #00:22:12-3#

74 I: Okay. // (..) Alles klar. Ähm gut dann hätt ich jetzt noch son letzten kleinen Block dazu, wie Sie die Rückmeldung erhalten haben. Also es war ja so, dass- huch- dass die per per Download ähm abrufbar war auf dieser Website, den Link haben Sie von Herrn N erhalten. #00:22:34-1#
75 A3: Mhm[Ja]. #00:22:34-4#
76 I: Da wäre einmal die Frage, obs da beim Erhalten irgendwelche Schwierigkeiten gab und wenn ja, welche? #00:22:40-9#
77 A3: Nein, da gabs soweit keine Schwierigkeiten. Ähm (.) die Frage ist- also vielleicht wäre es auch sinnvoll, die Rückmel- dass man die Rückmeldung direkt ähm nach dem Test schon einmal erhält, also dass man da sein Ergebnis schon einmal sehen kann (unv.). #00:22:57-9#
78 I: Also quasi auf der letzten Testseite? #00:23:01-2#
79 A3: Genau. Also das andere, dass mans noch zusätzlich nochmal abrufen kann ist natürlich gut, wenn man das dann nicht gespeichert hat ähm aber durchaus wenn man den Test gemacht hat, dann möchte man ja auch gern sein Ergebnis sehen und dann kann man meiner Meinung nach auch am meisten damit anfangen, weil man noch die diese Verbindung dazu hat, also noch weiß, was man gerade bearbeitet hat, während äh ja wenn mans dann etwas später bekommt, vielleicht n paar Wochen, dann äh weiß man gar nicht mehr so konkret, was war eigentlich in den Aufgaben drin, sondern hat nur noch so ne grobe Vorstellung und hat vielleicht nur noch n paar Aufgaben, die einem wirklich im Gedächt- Gedächtnis geblieben sind. #00:23:35-6#
80 I: Ja. Okay dann äh soweit erstmal vielen Dank. Ich bin tatsächlich durch mit meinen Fragen. #00:23:45-5#
81 A3: Mhm[Ja]. #00:23:45-9#
82 I: Wenns von Ihrer Seite noch irgendwie Fragen, Anmerkungen gibt, dann gerne jetzt? #00:23:50-4#
83 A3: Okay, also was mir halt aufgefallen ist ähm so der Gesamteindruck war halt, dass es für mich erstmal- das Ganze erstmal bisschen überwältigend war als ähm ich mich noch nicht damit befasst habe. Ähm und das bisschen lange gedauert hat, bis ich überhaupt dahin gekommen bin, was denn jetzt dieses Ergebnis ist. Als Vergleich: Wir haben dann nochmal nen andern Test gemacht äh gleichzeitig und da hatte man halt direkt so ne- son kleines Ergebnis am Anfang und konnte dann von da aus- also da warn dann auch natürlich noch mehrere Seiten dran, ich glaub zwei warens auch ähm konnte man von da aus ausgehend dann das Ganze lesen, aber man hatte schonmal son Bezug ähm was das Ganze in irgendeiner Form bedeutet und das macht es dann einfacher sich vielleicht damit zu befassen. #00:24:33-9#
84 I: Darf ich einmal Fragen: Welcher Test war das? Dieser Erklärtest? #00:24:37-7#
85 A3: Genau, // das war dieser Erklärtest. #00:24:39-6#
86 I: Gut, // dann weiß ich, wie das da aussieht (und?) dann kann ich das zuordnen. #00:24:43-7#
87 [...][Ab hier: Abschluss]

P.1.4 Interview 4: A4

1 [...][Hier: Einleitung, Einverständniserklärung etc.]
2 I: Dann wäre meine erste Frage zum Einstieg einmal: Können Sie sich noch an Ihre Rückmeldung erinnern, die sich nach dem Reflexionstest erhalten haben? Und wenn ja: Welche Schlüsse haben Sie daraus gezogen? #00:02:02-6#
3 A4: Ähm, ich kann mich da noch recht gut dran erinnern. Ähm, ich muss vielleicht sagen, ich bin noch recht früh im Studium, hab da nicht so viel Ahnung, deswegen habe ich mir nicht so viel erhofft. Aber ich war relativ überrascht, dass es doch relativ- also dass ich gut im Mittel war immer und mich nicht so ausgeschlagen bin ... Da war ich auf jeden Fall sehr überrascht und dass es ja eigentlich dann schon ganz gut klappt. #00:02:22-5#
4 I: Okay. [...][Ab hier: Öffnen und Lesen der Beispielrückmeldung] #00:04:00-7#
5 A4: Mhm, ja, okay, das ist ja alles relativ selbsterklärend noch. [Bezieht sich auf Einleitung und Absatz zur Reflexionskompetenz] #00:04:04-5#
6 I: Mhm. [Ab hier: Weiterlesen] #00:04:26-7#
7 A4: Sind denn da jetzt ähm schon all unsere Ergebnisse auch schon mit drin oder jetzt nur aus der Datenbank vorher quasi? #00:04:34-0#
8 I: Ähm, ich weiß jetzt nicht ganz genau, an welchem Stichtag Sie den Test gemacht haben, aber da sind auf jeden Fall Bearbeitungen aus Ihrem Kurs im Datenpool schon mit drin. #00:04:43-5#
9 A4: Okay. #00:04:44-3#
10 I: Ob jetzt- #00:04:44-8#
11 A4: Ja, (weil?) da einundzwanzig Studierende steht, da hätte ich jetzt gedacht, das wäre schon mehr, aber- #00:04:48-3#
12 I: Genau, es sind noch ein paar andere, aber ... ja, also welche von Ihnen sind auch dabei sozusagen. #00:04:54-4#
13 A4: Okay. Äh, gerne weiter. #00:04:57-3#
14 I: Mhm[Ja]. [Ab hier: Weiterlesen] #00:05:23-5#
15 A4: (Na?) ich persönlich kann halt noch nicht so viel mit Schülervorstellungen (.) anfangen, weil ich dazu jetzt noch nichts gehört hab, also (.) Physikdidaktik ist jetzt das erste bei mir in der Richtung. [...][Ab hier: Weiterlesen] #00:05:54-4#
16 I: Ähm genau dann wär meine erste Frage dazu, ob Sie schonmal mit solchen Boxplots gearbeitet haben und wenn ja, in welchem Rahmen? #00:06:03-2#
17 A4: Also gearbeitet noch nie mit ähm in der Schule schonmal besprochen halt, ne. #00:06:08-1#
18 I: Als Schülerin dann, oder? #00:06:10-0#
19 A4: Genau, ja, aber (ich?) noch nie mit gearbeitet. #00:06:12-7#
20 I: Okay. (.) Und äh wenn Sie die Rückmeldung jetzt so lesen beziehungsweise gelesen haben, welche Schlüsse würde Sie da draus ziehen? #00:06:21-9#
21 A4: (..) Mhm (..) ja auf jeden Fall aufm rich- also aufm guten Weg, ne, also es war ja jetzt im Prinzip alles relativ Mitte. Die Schülervorstellung war jetzt nicht ganz so gut, aber dafür warn jetzt, ich glaub, Bewertungen besser. (..) Genau mhm (...) wahrscheinlich könnte- also man könnte (da?) wahrscheinlich draus ziehen, dass man besser mit den Schülervorstellungen dann arbeitet, ne, dass man da sich nochmal n bisschen was raneignet. #00:06:49-9#
22 I: (..) Okay. Und ähm aus dieser Rückmeldung, welche der Informationen sind da besonder hilfreich oder besonders interessant für Sie? #00:07:00-2#
23 A4: (..) Mhm (...) ja ich würd auch wieder hier auf die Schülervorstellungen gehen, also weil der Rest, würde ich ja sagen, ist schon schon (.) zumindest im Mittel ganz gut (.) natürlich immernoch immernoch verbesserungswürdig, aber da sollte man dann vielleicht nochmal dran arbeiten an den Schülervorstellungen. #00:07:21-1#

24 I: Mhm[Ja]. Und wie ist das, wenn Sie die äh Rückmeldung, also dieses, quasi das Dokument, als Ganze und nicht nur die ganz konkrete, ich sag mal, 'Punkte-Rückmeldung' betrachten, äh gibts da Dinge, die besonders interessant sind oder landen wir dann wieder beim Plot? #00:07:37-4#
25 A4: (..) Naja im Prinzip ist es ja ganz spannend, dass das Mittel oder der Median vor allem auch trotzdem relativ weit unten ist, also dass es ja eigentlich im Mittel noch weiter nach oben gehen könnte. #00:07:50-0#
26 I: Mhm[Ja]. (...) Okay und dann anders rum die Frage ähm welche Informationen sind vielleicht auch Überflüssig in dieser Rückmeldung? #00:08:02-5#
27 A4: (5) Hm da würde ich jetzt erstmal (..) nichts zu sagen können. #00:08:11-0#
28 I: (.) Okay. Ähm dann: Gibt es Fragen, die für Sie offen geblieben sind, wenn Sie die Rückmeldung so gelesen haben? Und wenn ja, welche sind das? #00:08:21-9#
29 A4: (.) Naja es wär ja eigentlich ganz cool, wenn man direkt dann Feedback kriegen könnte, wo- also an welchen Stellen (.) es zu dem Ergebnis gekommen ist. Also ich weiß jetzt zum Beispiel bei diesem: Okay, Schülervorstellungen sind nicht so gut. Aber ich würd ganz gerne wissen, an welchen Stellen wurde das so bewertet. Also in diesem (.) Fragebogen quasi. #00:08:41-6#
30 I: Mhm[Ja]. Wie könnte das aussehen? Also dass man die Fragen nochmal auflistet dazu? Oder- #00:08:47-3#
31 A4: Ja, zum Beispiel dass man die Punkte, die- also ich persönlich hab jetzt nicht erkannt, an welcher Stelle des Fragebogens welche Kompetenz da herausgekitzelt wurde. Und dass man da das son bisschen- dass man den Fragebogen vielleicht nochmal komplett einbindet oder nochmal sehen kann und dann pro Frage sieht: Okay, diese Frage war jetzt darauf abgezielt und die Frage war darauf abgezielt. #00:09:07-3#
32 I: (.) Mhm[Ja]. (.) Was würde das bringen? #00:09:12-0#
33 A4: (.) Mhm vielleicht, dass ich einfach direkt sehe: Okay, in dem Moment hab ich dann vielleicht falsch gehandelt oder nicht so gut gehandelt und das wäre dann ne Möglichkeit, anders zu- also besser zu handeln oder so. #00:09:24-3#
34 I: (.) Als- #00:09:25-6#
35 A4: Und das kann man ja dann verallgemeinern auf auf den normalen Unterricht. #00:09:28-4#
36 I: (.) Okay.(.) Ähm genau jetzt sind wir auch schon son bisschen in in meiner nächsten Frage quasi und zwar welche weiteren Informationen Sie vielleicht gerne gehabt hätten? Also gut, diese Einordnung ham wir jetzt schon, gibts vielleicht noch mehr? #00:09:43-7#
37 A4: (...) Mhm (..) ne, eigentlich eigentlich ganz gut so. Äh vielleicht vielleicht noch äh ich hab ja grad gesagt, dass ich jetzt noch ganz früh in dem Studium bin ähm und hier steht ja, dass es im Schnitt beim fünften Semester gemacht wurde, da wär es ja für mich ganz interessant, an welcher Stelle ich stehe. Also (.) ich kann mich ja eigentlich nicht mit Leuten aus dem fünften Semester vergleichen, weil ich ja jetzt persönlich didaktisch und erziehungstechnisch noch nicht so viel weiß (.) und da hätte ich mir- wär vielleicht ganz interessant, wenn der Datenpool n bisschen größer ist, dass man dann noch ne Einordnung hat: Wie siehts aus, wenn ich im ersten Semester stehe und wie siehts aus, wenn ich im fünften Semester stehe. #00:10:22-7#
38 I: (..) Ja. Okay. (.) Ähm ich hab auch noch so ne ja ungefähr ne handvoll Dinge, die irgendwie mögliche Optionen wären, die man noch ergänzen könnte. Ich würd die einmal durchgehen und Sie bitten, dass Sie mir einfach jedes Mal sagen: (.) 'Klingt sinnvoll, weil' oder 'bringt mir überhaupt nichts' genau einfach so als kurzes Feedback. Das erste wäre, dass man (.) äh in die Rückmeldung in diesem Text noch Informationen über das Z- äh das Ziel des Testes einbindet. #00:10:56-5#

39 A4: (.) Mhm naja eigentlich find ich ist das Ziel doch relativ (..) klar formuliert also ich meine, der Test zielt doch darauf ab, zu erkennen, wo ich- wo ich Schwächen hab, wo ich Stärken hab oder so. (...) Wenn das das einzige Ziel ist, hab ich das ja direkt schon verstanden, also da brauch. bräuchte man das nicht unbedingt noch die Information. #00:11:17-8#
40 I: Okay. Wie ist das mit ähm Informationen zum Beispiel über das Projekt in dem Rahmen diser Test entstanden ist? #00:11:25-1#
41 A4: (.) Ja, da hat, glaub ich, Herr N schon bisschen was zu gesagt, also zumindest hab ich da auch zumindest was im Hinterkopf (.) ähm wenn man den Test jetzt natürlich einfach so, ich sag mal random, durchführt, dann wär das vielleicht ganz interessant, ja. #00:11:38-0#
42 I: (..) Okay. Ähm gibts da was Konkretes, was Sie da interessieren würde oder wo Sie sagen würden, das hat Herr N erwähnt, das ist was, was jeder wissen sollte, der diesen Test macht, zum Beispiel? #00:11:49-7#
43 A4: (Ja?) mir war am Anfang als ich den noch durchgeführt hab, war mir nicht ganz klar, dass das zum Beispiel n Projekt ist, was noch im Aufbau ist ähm und dass das auch von hier kommt. Ich war mir nicht ganz sicher, von wo der Test jetzt son kommt. Ähm einfach- äh so das war einfach nur so 'Achso, so ist ist, das ist ja ganz interessant', dass man das ja auch n bisschen besser einordnen kann, wenn das noch im ein- im Aufbau ist, dass es da in Größe ja noch nicht so groß ist. Das ist alles klar geworden, als wir das- die Auswertung gekriegt haben, aber am Anfang war das halt noch nicht so ganz klar. #00:12:17-1#
44 I: Okay. (.) Ähm nächste Option, die man ergänzen könnte, wäre, dass man explizite Informationen darüber gibt, was der 'Zielzustand' Reflexionsfähigkeit ist, also was- wo sollte man am Ende stehen quasi. #00:12:32-0#
45 A4: (.) Ja, das seh ich auch relativ sinnvoll. (..) Mhm einfach, um nochmal n bisschen besser so (.) klar zu werden, was man da gerade macht. (.) Also- #00:12:45-8#
46 I: Wie ähm (.) wie haben Sie denn diese Rückmeldung gelesen, also- (.) was ist für Sie dieser Zielzustand, wenn Sie die Rückmeldung lesen? Können Sie da was (.) rauslesen oder bleibt das komplett offen? #00:12:58-1#
47 A4: (...) Mhm (..) ne, das versteh ich jetzt- also (..) ja, ich weiß ja, dass es quasi die ne Reflexion über meine Reflexion ist also (..) und einzuschätzen, wie ich da stehe halt. #00:13:13-0#
48 I: (.) Okay. Ähm ne andere Option wäre ne feinere Aufschlüsselung dieser Bewertung, also zum Beispiel in den Plots, genau, das genauer aufzuschlüsseln. #00:13:25-6#
49 A4: Mhm vielleicht nicht unbedingt in den Plots, sondern ja eher wieder in die Richtung, was ich eben schon mal angedeutet hab, dass ich jetzt- ja, ich seh jetzt, wo ich stehe, aber ich seh noch nicht, was ich verbessern kann oder wie man da- oder wodran- wodrans halt auch gelegen hat. #00:13:41-4#
50 I: (.) Mhm[Ja]. (.) Also ich hab, vielleicht passt es ganz gut, ich hab hier noch als eine Option, dass man eine Beispielaufgabe angibt oder eine Musterlösung mit Begründung. Ist das sowas in die Richtung? Oder sollns wirklich ganz konkret quasi die- dien- (..) Ich stell die Frage erstmal so: Ist das eine Option? #00:14:03-0#
51 A4: Ja, ist auf jeden Fall auch ne Opotion, ja. #00:14:07-1#
52 I: Okay. // (.) Ja. #00:14:08-1#
53 A4: Aber (.) // ist ja das gleiche (quasi?), wenn ich- wie wenn man ne ne Aufgabe aus dem Fragebogen dann direkt auch bespricht also besprechen würde. #00:14:16-0#
54 I: (..) Okay. (..) Ähm (.) eine Option wäre noch, dass man andere Diagramme (.) verwendet. Also entweder stattdessen oder als Ergänzung? #00:14:28-8#
55 A4: (.) Ja, als Ergänzung würd sichs vielleicht auch anbieten hier wieder das mit dem 'aufgeschlüsselt in die Semester'. Also dass ich da halt dann die Information kriege, wo stehe ich zu meinem Lernstandpunkt (.) ähm ansonsten die Boxplots find ich eigentlich ne ganz gute Methode. #00:14:42-8#

56 I: (..) Okay und die warn auch äh gut nachvollziehbar, obwohl Sie da jetzt noch nicht- #00:14:48-2#
57 A4: Ja, auf jeden Fall. War ja auch unten erklärt. Spätestens damit hätte mans dann ja (.) verstanden. #00:14:54-5#
58 I: Okay und dann äh genau wir hattens jetzt n paar Mal son bisschen angedeutet irgendwie welche Schlüsse ziehe ich jetzt, wie mache ich jetzt weiter mit diesem Ergebnis ähm da wär meine Frage ob ja obs ob Hinweise oder Material dazu, wie ich jetzt meine Reflexionsfähigkeit schulen kann, wie ich da dran arbeiten kann, ob das was ist, was ergänzt werden sollte? #00:15:17-5#
59 A4: Ja, also wenn dann halt explizit dann halt auch das, was das Ergebnis ist. Also (.) in diesem Fall ist es ja jetzt vielleicht nicht so ratsam, sich noch viel mehr über die Bewertung anzugucken, sondern wirklich da, wo die größeren Schwächen sind, (.) dass man da nochmal irgendwie Beispiele zu gibt, oder so, wie man besser reagieren kann. #00:15:33-9#
60 I: Mhm[Ja]. Also quasi angepasst auf das Ergebnis- #00:15:37-1#
61 A4: Ja. #00:15:38-2#
62 I: rausfiltern. Mhm[Ja]. (.) Okay und der letzte Punkt wär noch weitere Literaturhinweise oder so Hinweise auf so physikdidaktische Grundlagentexte ähm genau, die könnt man auch ergänzen. #00:15:50-0#
63 A4: (..) Joa, ich bezweifel zwar, dass da sehr viel (.) also sehr viele das benutzen würden, aber wäre auf jeden Fall sinnvoll für die Leute, die das dann wirklich versuchen. #00:15:58-9#
64 I: Und an welcher Stelle würden Sie die sehn? #00:16:01-5#
65 A4: (..) Ähm vielleicht auch dann, wenn das nochmal kommen würde mit ähm expliziten Hinweisen zur Verbesserung, (.) da dann mit einbauen. #00:16:10-4#
66 I: (..) Okay. (.) Ähm genau dann hätt ich noch eine Frage zu dieser Rückmeldung und zwar ähm (4) ähm sie ham ja- also wir ham jetzt sehr viel über die Plots und und über diese Aspekte hier gesprochen (.) wie ist das mit diesen quasi mit diesen ganzen Infos in Textform, sag ich jetzt mal, irgendwie davor zum Beispiel. Welches Gewicht haben die für Sie oder warn die hilfreich? #00:16:38-6#
67 A4: Ja, ich hab dadurch halt überhaupt erstmal verstanden, was ähm was da bewertet wurde, oder was was damit gemeint ist und da- also unter dem Boxplot stand ja glaub ich dann nur 'bewertet' und dann wusst ich ja nicht, was das bedeutet (..) ähm und von daher war das gut, dass das da so aufgeschlüsselt ist. (.) Vielleicht noch n bisschen (.) expliziter auch hier ähm (.) an welchen Punkten man das festhalten kann. #00:17:02-4#
68 I: Im Test, // also die- #00:17:04-8#
69 A4: (unv.) // (..) Ja oder generell- also (.) ja hier steht was was- was das bedeutet, aber so direkt aufn didaktischen Punkt bringts ja das noch nicht, meiner Meinung nach. #00:17:15-8#
70 I: Also Unterrichtsbeispiele? Oder so beschreibende Situationen? #00:17:21-5#
71 A4: Ja. #00:17:21-7#
72 [Ab hier: Erhalt der Rückmeldung problemlos, Schwierig ist es nur, wenn man den Link zu Test/Rückmeldung verliert, Idee der Verknüpfung mit e-Learning-Profil ist nicht praktikabel; Abschluss]

P.1.5 Interview 5: A5

1 [...][Hier: Einleitung, Einverständniserklärung etc.]
2 I: Dann wäre meine erste Frage zum Einstieg, ob Sie sich noch an Ihre Rückmeldung erinnern können, die Sie bekommen haben und wenn ja, welche Schlüsse sie daraus gezogen haben. #00:01:52-7#
3 A5: Äh, ja, also ich hab mir die Diagramme runtergeladen und ausgedruckt. Das war ja die Rückmeldung, glaube ich, die wir bekommen haben. Und, ähm, also erstmal- das erste was mir aufgefallen ist: Das ich ungefähr immer so in den Bereichen lag, wo auch die anderen einundzwandig Studierenden lagen, also die Ergebnisse, jetzt nicht unbedingt aus dem Rahmen gefallen bin in Anführungsstrichen. Ich hätte aber auch ne Frage zu dem Diagramm? #00:02:16-9#
4 I: Ja. #00:02:17-5#
5 A5: Äh, ich kann nicht so richtig was mit dieser Prozentangabe anfangen. Also man hat ja dieses, äh, diese Box und dann den Punkt mit dem eigenen Ergebnis. Und was genau sagt mir jetzt die Prozentangabe? Wie gut ich sozusagen in der Reflexion war oder ... also ... die konnte ich nicht so richtig einordnen, wenn ich jetzt zum Beispiel bei Gesamt ungefähr ein bisschen unter Null Komma Fünf liege. #00:02:41-3#
6 I: Ja, ähm, da ist tatsächlich ein Fehler drin, weil die- das sind theoretisch Prozentangaben von Null bis Hundert und nicht von Null bis Eins. // Ähm-# #00:02:50-7#
7 A5: Mmh// #00:02:51-3#
8 I: Das ist- Und dann ist es quasi- ja, wie viel, wie viel Prozent der absolut möglichen ... äh Punkte im Test man, oder auch eben der Boxplot, also alle anderen, bekommen haben. #00:03:05-4#
9 A5: Mhm. ... Okay (4) Ähm, ... Jo. ... Äh, ich les mir das grad nochmal kurz durch. #00:03:20-0#
10 I: Ja, alles gut. #00:03:21-1#
11 A5: (5) Äh, dann wär mir aufgefallen, dass ich bei den Schülervorstellungen relativ ... über den anderen Ergebnissen war. (4) Also da ne relativ hohe Prozentangabe hatte und ja sonst eigentlich immer sehr im Durchschnitt lag sozusagen. #00:03:42-8#
12 I: Mhm[Ja]. #00:03:43-5#
13 A5: (10) Äh, ja.# #00:03:54-6#
14 I: Ja, okay. Danke. [...][Ab hier: Öffnen und Lesen der Beispielrückmeldung] #00:06:08-8#
15 I: Ähm dann wär meine erste Frage dazu, ob Sie schonmal mit äh mit solchen Diagrammen, mit Boxplots gearbeitet haben und wenn ja, in welchem Rahmen #00:06:19-3#
16 A5: Ne, noch gar nicht. #00:06:21-1#
17 I: Auch noch gar nicht gesehen vorher? #00:06:23-7#
18 A5: Äh ne, nicht wirklich. #00:06:26-1#
19 I: (...) Okay und wie war das dann? Also ham s- Oder wie wie haben Sie dann ja mit dieser Rückmeldung gearbeitet? #00:06:35-1#
20 A5: Äh also es war halt relativ neu und äh (also?) ich musste mir das dann erstmal genauer durchlesen, um das dann irgendwie verstehen zu können. Sonst gabs halt irgendwie immer mehr Rückmeldung in schriftlicher Form (.) und nicht so mit Diagrammen. #00:06:49-1#
21 I: Mhm[Ja] und haben da die Hinweise aus der Rückmeldung gereicht oder ha- äh ham Sie noch (..) // (was weiß ich?) im Internet geguckt, oder sonstwo? #00:06:58-6#
22 A5: Äh // (...) äh die Hinweise haben eigentlich gereicht. #00:07:01-5#
23 I: Okay. (.) Ähm dann wär die nächste Frage, wenn Sie dich jetzt die Rückmeldung, also die sie sich gerade angeguckt haben, ähm wenn Sie sich die ansehen, welche Schlüsse würden Sie aus dieser Rückmeldung ziehen? #00:07:15-3#
24 A5: (..) Äh kann ich da nochmal die (.) // Diagramme- (.) Ah okay. #00:07:20-5#

25 I: ja (..) ich krieg, ja, ungefähr alle. #00:07:24-4#
26 A5: Äh dass man sich dann wahrscheinlich noch mehr im Bereich Schülervorstellungen nochmal reinarbeiten sollte, weil da ja die Bewertung relativ weit unten war (.) ähm (..) und dann auch nochmal im Bezug auf Alternativen, dass man sich da dann vielleicht nochmal das Vorgehen anschaut, wie man denn gute Alternativen im Unterricht findet zu Situationen (..) und ja, das wäre das spontan. #00:07:51-6#
27 I: Okay und ähm (.) für das Verständnis oder beim Lesen der Rückmeldung, welche Informationen sind da besonders hilfreich oder besonders interessant für Sie? #00:08:03-4#
28 A5: (..) Äh jetzt von der Rückmeldung, die wir bekommen haben, was am hilfreichsten war? (..) (oder?) was- #00:08:11-2#
29 I: Genau also ähm ich zeig Ihnen hier diese Beispielrückmeldung, wenn Sie das jetzt auf Ihre eigene beziehen, ist das für mich auch okay, das müssen Sie nur einfach nicht, um die Anonymität irgendwie zu gewährleisten, (.) deswegen also stell ich die Fragen so, dass sie in jedem Fall an dieser Beispielrückmeldung beantwortet werden könnten, wenn Sie jetzt sagen, zum Beispiel: 'Besonders interessant war für mich, ich weiß es nicht, hier das Diagramm oder am Anfang der Text oder am Ende der Text' // ähm- #00:08:37-0#
30 A5: Also // am hilfreichsten fand ichs nochmal am Anfang, dass dann nochmal erklärt wurde, was jetzt die einzelnen Aspekte nochmal genau bedeuten, also äh von wegen Alternativen, Schülervorstellungen, was man sich (daran?) nochmal vorstellen konnte. #00:08:48-0#
31 I: (...) Also die- sind das diese drei- #00:08:51-8#
32 A5: Äh ja. #00:08:52-7#
33 I: Ja, okay, also unterm zweiten Diagramm. Ja. (..) Okay. Äh und dann anders rum quasi die Frage: Gibts Informationen, die überflüssig sind und wenn ja, welche sind das? #00:09:06-3#
34 A5: Ähm (.) überflüssig hätt- würd ich jetzt nicht bezeichnen, aber vielleicht die Erklärung dann zu den äh (.) zu dem äh Kasten mit dem Median und so, also ich fand, das konnte man sich denken, dass das der Median ist, aber also ist jetzt auch nicht unbedingt blöd, dass es einfach mal steht, aber das fand ich jetzt nicht war die wichtigste Information. #00:09:30-5#
35 I: Okay, meinen Sie da die die ganzen Hinweise oder ganz konkret dies- den den Satz zum Median? #00:09:37-7#
36 A5: Äh ganz konkret den (.) ähm (5) ja (..) ganz konkret den eigentlich. #00:09:50-8#
37 I: Okay (4) und ähm (.) jetzt so im Nachhinein, wenn Sie die Rückmeldung äh wenn Sie die Rückmeldung hier gelesen haben, welche Fragen bleiben dann noch offen für Sie? #00:10:06-4#
38 A5: (..) Mhm (13) mhm (..) ne gute Frage.(...) Äh also vielleicht halt allgemein nochmal, was ich mir dann fürs spätere Lehrerleben daraus ziehn kann, aber das ist dann ja trotzdem immernoch Situationsabhängig. #00:10:39-3#
39 I: (..) Meinen Sie da äh konkrete Beispielsituationen, oder // (...) was könnte da helfen? #00:10:49-3#
40 A5: Äh // (16) äh tschuldigung drei Sekunden. #00:11:06-9#
41 I: Alles gut, kein Stress. #00:11:08-3#
42 A5: (...) Mhm (10) also eher so anhand der Diagramme würden mir jetzt nicht unbedingt nochmal Fragen einfallen. Vielleicht, wenn man irgendwie genaueres Feedback nochmal bekommen hätte, also vielleicht auch in schriftlicher Form, was vielleicht schlecht in Anführungsstrichen an dem Test war, dass man daraus dann nochmal seine Schlüsse ziehen könnte, (.) (nur?) anhand der Diagramme (.) find ichs jetzt schwierig, Fragen zu formulieren. #00:11:42-5#
43 I: (..) Mhm[Ja] das- also nur damit, ich das richtig versteh: Das wär dann zum Beispiel, dass man in Schriftform sowas schreibt, wie ja, weiß ich nicht, zum Beispiel 'bei den Schülervorstellungen ähm liefs nicht so gut, machen Sie doch nochmal xy' oder 'die sind im Unterricht wichtig, weil-' sowas? #00:12:05-9#
44 A5: Äh ja. #00:12:07-4#

45 I: Okay, also quasi ähm (.) ja quasi so so Schlussfolgerungen für die Weiterentwicklung. #00:12:16-1#
46 A5: Ja. #00:12:16-6#
47 I: Ja, okay. (.) Ähm dann wär meine nächste Frage, wir sind auch quasi schon son bisschen drin, ähm obs Dinge gibt, die Sie gerne noch gehabt hätten, also welche weiteren Informationen ja in der Rückmeldung noch ergänzt werden sollten vielleicht. #00:12:35-3#
48 A5: (..) Mhm (ja?) der Test ist jetzt auch schon n paar Wochen her (.) das- #00:12:41-3#
49 I: Ja, klar. #00:12:42-2#
50 A5: (5) Ähm (..) da würd mir jetzt spontan nichts direkt einfalln (15) äh ne, ne leider nicht. #00:13:10-0#
51 I: Okay ähm ich hab zu diesem Punkt noch so ne ungefähr ne handvoll mögliche Dinge, die man ergänzen könnte. Ähm ich würd die gleich einmal durgehen und Sie dann bitten, dass sie einmal danach sagen irgendwie 'halt ich für sinnvoll' oder 'überhaupt nicht sinnvoll, weil' (.) äh dass ich das son bisschen einordnen kann. (.) Das erste wäre ähm ob man weitere Informationen über das Ziel des Tests gibt. #00:13:37-5#
52 A5: (...) Ja, fänd ich auf jeden Fall nicht verkehrt, um das alles dann nochmal n bisschen besser einordnen zu können, also fänd ich hilfreich. #00:13:46-1#
53 I: (.) Mhm und gibts da was, wo Sie irgendwie ganz konkret sagen 'das hätte ich gebraucht' oder ähm oder 'das würd ich mir wünschen' oder mehr so zum ja zum Gesamteindruck zum Einordnen. #00:14:00-1#
54 A5: Mhm ich glaub also zum Gesamteindruck. #00:14:02-6#
55 I: Okay (.) und wie ist es mit Informationen über das Projekt, in wes- also in dem Rahmen dieser Test entstanden ist? #00:14:11-5#
56 A5: (..) Ich glaub, das fänd ich nicht unbedingt hilfreich. Vielleicht noch so als Zusatzinformation für besonders Interessierte. #00:14:20-6#
57 I: (..) Okay, aber nicht- nicht hilfreich zur Einordnung und zum // Verständnis. #00:14:26-1#
58 A5: Ja. // #00:14:26-0#
59 I: Okay. (.) Ähm das nächste wäre, inwiefern explizite Informationen darüber hilfreich wären, was denn eigentlich dieser 'Zielzustand' ist, also (.) wann ist meine Reflexionsfähigkeit denn am Ziel quasi. #00:14:41-2#
60 A5: (...) Mhm (..) ja fänd ich glaub ich auch sinnvoll. #00:14:49-5#
61 I: (...) Ähm (..) wenn Sie die Rückmeldung ähm so angucken oder gelesen haben. #00:14:58-7#
62 A5: Mhm? #00:14:59-6#
63 I: Haben Sie da- Sind Sie da von nem bestimmten Zielzustand ausgegangen? #00:15:03-7#
64 A5: (.) Mhm nicht direkt, weil ich eigentlich in meinem Kopf hatte, dass man nie unbedingt (.) fertig ist, sozusagen, sondern sich halt immernoch (.) immer verbessern kann und irgendwas dran ändern kann an seiner Reflexionskompetenz. #00:15:17-6#
65 I: (.) Okay. (.) Ähm (.) das nächste wäre eine feinere Aufschlüsselung der Bewertung. #00:15:26-1#
66 A5: Mhm ja das fänd ich auf jeden Fall sehr hilfreich, dass es n bisschen differenzierter und mehr vielleicht ist. #00:15:33-7#
67 I: Und ähm in welcher Form würden Sie sich das wünschen? #00:15:38-7#
68 A5: (..) Äh also in Bezug auf die Darstellungsweise? Oder? #00:15:43-8#
69 I: Alles, was Sie wollen. #00:15:45-4#
70 A5: Ähm also mir persönlich äh hätten glaub ich Text- wär ne Textform (..) für mich persönlich wäre eine Textform glaub ich hilfreicher gewesen (.) äh als zum Beispiel die Diagramme aber ich- ja hängt dann halt auch immer vom Studierenden ab. #00:16:02-6#
71 I: (9) Okay und ähm wie wär das- also man könnte sich ja jetzt auch vorstellen zum Beispiel

die Ergebnisse auf- also noch noch feiner aufzugliedern zum Beispiel dass man sagt (..) ja, ne, 'bei Frage eins liefs so gut, bei Frage zwei liefs so und so gut' oder 'bei den Fragen zu Video eins liefs so und so gut', (.) wär das was von Interesse? #00:16:37-4#
72 A5: (..) Mhm (.) fänd ich persönlich auf jeden Fall interessant, das alles so einzeln betrachten zu können und nicht nur jetzt unter einem Punkt, zum Beispiel unter dem Punkt Schülervorstellungen, sondern dass man da nochmal jede einzelne Frage hätte. #00:16:51-9#
73 I: (.) Okay und dann also quasi die die Fragen aufführen und dazu das individuelle Ergebnis rückmelden. #00:16:59-9#
74 A5: Ja. #00:17:01-1#
75 I: (.) Okay (..) ähm ne andere Option wäre (.) man könnte überlegen, also wir haben ja über die Textform schon gesprochen, aber man könnte auch überlegen, entweder statt diesen Boxplots oder zusätzlich dazu noch andere Diagramme zu ergänzen. #00:17:18-5#
76 A5: (8) Ähm (..) ja also ich glaub zusätzlich noch andere Diagramme fänd ich auf jeden Fall noch hilfreich, vor allem, wenn man jetzt vorher halt noch nie mit diesen Boxplots gearbeitet hat, fand ichs halt auf den ersten Blick n bisschen- also musste man sich erstmal reinlesen. Dann (können?) andere Diagramme auf jeden Fall gut helfen. #00:17:41-0#
77 I: Gibts da eins, was Sie bevorzugen würden? #00:17:44-4#
78 A5: (...) Äh eigentlich nicht. #00:17:48-2#
79 I: Einfach irgendwas, was man quasi eher gewohnt ist, // oder? #00:17:52-1#
80 A5: Ja. // #00:17:52-3#
81 I: Ja okay.(.) Ähm (.) noch ne Option wär es, ne Beispielaufgabe zu ergänzen oder auch mehrere, um sich an den Test zu erinnern? #00:18:04-0#
82 A5: (...) Äh ja find ich auf jeden Fall gut jetzt im Hinblick auf die Zeit. Also wenn man die Reflexion vielleicht n bisschen direkter nach dem Test bekommen hätte, was- also (.) kann ich gut nachvollziehen, wenn das- die Auswertung länger dauert, aber dann nochmal so kleine Hilfen zur Erinnerung find ich auf jeden Fall gut. #00:18:23-4#
83 I: Okay und ähm (.) für die Beispielaufgabe könnte man dann auch Musterlösung mit- zum Beispiel eben mit Begründungen, warum bestimmte Aspekte wichtig sind oder nicht, ergänzen. (.) Wie sehen Sie das? #00:18:37-2#
84 A5: (..) Ähm (...) sinnvoll auf jeden Fall, um sich da vielleicht dran zu orientieren. Vielleicht würd ichs nicht unbedingt 'Musterlösung' nennen, weil ja vielleicht nicht allein ein Weg der richtige unbedingt- also nicht nur ein Weg der richtige ist (..) aber fänd ich auf jeden Fall gut. #00:18:57-3#
85 I: Okay (.) ähm noch zwei Optionen, dann sind wir mit diesem Block auch durch. Ähm das eine wäre inwiefern man zum Beispiel Literaturhinweise gibt oder Hinweise auf physikdidaktische Grundlagentexte (.) an verschiedenen Stellen denkbar, also man könnte es ja zum Beispiel machen, um noch zum (.) zum Begriff der Reflexionskompetenz oder eben zu so Begriffen, wie Schülervorstellungen (.) oder auch noch an anderen Stellen denkbar. #00:19:28-9#
86 A5: (..) Mhm also ich kann mir gut vorstellen, dass es jetzt vielleicht für die Studierenden, die jetzt gerade eh in der Veranstaltung sind, nicht unbedingt notwendig ist, weil man ja eh mit den Thematiken konfrontiert wird und das dann vielleicht eh schon einordnen kann ohne noch zusätzliche Literatur. #00:19:44-0#
87 I: (..) Okay. (.) Und dann noch die Frage danach, ob man Material ergänzen sollte oder Hinweise dazu, wie man diese Reflexionsfähigkeit schulen könnte. Also wie kann ich jetzt weiterarbeiten damit. #00:20:00-7#
88 A5: (.) Äh das fänd ich auf jeden Fall sehr hilfreich, dass man dann noch irgendwie weitere Ideen hätte, wie man jetzt nach dem Test dann noch verfahren könnte, um dann fürs spätere Berufsleben gewappnet zu sein sozusagen, dass man dann weiter arbeiten könnte. #00:20:15-5#
89 [...][Ab hier: Keine Probleme beim Erhalt der Rückmeldung, Format angemeseen; Abschluss]

P.1.6 Interview 6: A6

1 [...][Hier: Einleitung, Einverständniserklärung, Lesen der Rückmeldung] #00:07:26-5#
2 I: Ähm dann wär meine erste Frage, ob Sie vor dieser Rückmeldung schonmal mit solchen Diagrammen, also den äh sogenannten Boxplots gearbeitet haben und wenn ja, in welchem Rahmen? #00:07:38-8#
3 A6: Ähm also so noch nicht, also mit dem ähm mit dem Median und so, mit den oberen und unteren Grenzen jeweils äh eher nicht ähm aber son bisschen äh son bisschen ähm (.) ähnelt das ja den den Feh- den graphischen Fehlergrenzen in der- in der Physik, also mit dem ähm (.) also wenn man die Fehler halt graphisch darstellt, dass die ja-, da gibts ja auch immer die oberen und unteren Grenzen, ähm aber jetzt direkt so mit dem- mit dem Median und den Punkten, äh den roten Punkten als Abweichung und schwarzen äh als als weitere Abweichung, äh so noch nicht. #00:08:22-5#
4 I: Okay, danke. Ähm wenn Sie diese Rückmeldung, also die Beispielrückmeldung hier jetzt, gelesen haben, welche Schlüsse ziehn Sie aus dieser Rückmeldung? #00:08:32-6#
5 A6: (5) Mhm also s- sie hilft mir auf jeden Fall, die- die- das Diagramm zu verstehen, auch dass nochmal auf die Schülerkompetenzen beziehungsweise auch die Bewertung, was die jeweiligen ähm die jeweiligen Elemente überhaupt bedeuten, dass das nochmal aufgelistet wird, das hilft auf jeden Fall (.) und ähm ja also besonders halt die Interpretation der Diragramme, also ich denk, wenn jemand das aufn ersten Blick sieht jetzt, das Diagramm, der könnte sich jetzt schon son bisschen was dabei denken, aber was die roten Punkte beziehungsweise die Boxen und der Strich, der äh horizontale Strich in der Mitte direkt bedeuten, ähm wüsste der jetzt vielleicht nicht, aber das hilft auf jeden Fall zur Interpretation, ja. #00:09:23-7#
6 I: Mhm[Ja] und wenn Sie jetzt- also wenn Sie sich mal vorstellen, das wär jetzt Ihr Ergebnis, ich versuch einmal, beide Plots gleichzeitig aufzumachen, zumindest so grob, ähm ich zoom sonst auch wieder ran, wenns- Wenn Sie sich vorstellen, das wär Ihr Ergebnis, wie würden Sie für sich diese Diagramme interpretieren oder welche Schlüsse dadraus ziehen. #00:09:44-0#
7 A6: (6) Also bei dem- bei dem oberen äh Diagramm, wär ich ja dann praktisch dann über dem Median im Ergebnis, dann so bei (..) äh vierzig Prozent rund (.) und (..) ja jeweils dann unten würde ich dann bei den ähm bei den ähm Alternativen äh (da?) wär ich dann genau im Median, ähm bei den Bewertungen (.) ähm überdurchschnittlich äh im Vergleich zu den- zu den anderen Ergebnissen und bei den Schülervorstellungen dann im unteren Teil der anderen praktisch. #00:10:28-9#
8 I: (...) Okay (.) ähm und würden sich für Sie da jetzt Schlussfolgerungen draus ergeben? Also dass Sie sagen 'aufgrund dieser Rückmeldung (..) mach ich dies oder jenes, also nehm ich irgendwelche schließ- ähm zieh ich irgendwelche Konsequenzen? (.) Oder wär das jetzt erstmal die die Information, die für den Moment ausreicht, zu wissen, 'hier bin ich eben auf Höhe des Medians und hier drüber, da drunter'. #00:10:59-3#
9 A6: (.) Ich würde das jetzt so als Momentaufnahme deuten, also dass ich jetzt genau weiß äh wo ich mich auch im Vergleich zu den anderen ähm Studierenden, die den Test durchgeführt haben, ähm in etwa befinde. #00:11:14-9#
10 I: Okay (..) ähm meine nächste Frage wär, obs in dieser Rückmeldung (..) äh welche Informationen da für Sie besonders hilfreich oder besonders interessant sind. #00:11:27-9#
11 A6: (...) Mhm (.) das nochmal ähm ja wie gesagt nochmal auf die Alternativen, Konsequenzen äh beziehungsweise auf die auf die Bewertungen und auf die ähm Schülervorstellungen nochmal ähm allgemein eingegangen wird (.) ähm (.) ja. #00:11:48-2#
12 I: (..) Da ähm hups, da meinen Sie diese- diese Beschreibungen // unter dem zweiten Boxplot, ne? #00:11:54-9#
13 A6: Genau, diese Beschreibungen // da, genau. #00:11:55-9#
14 I: Alles klar. #00:11:56-8#

15 A6: Ja. #00:11:57-8#
16 I: (.) Okay (.) und äh anders rum gefragt: Sind Informationen auch überflüssig? #00:12:05-4#
17 A6: (..) Mhm (.) würd ich jetzt nicht direkt sagen oder könnten Sie vielleicht nochmal ganz kurz hoch nochmal machen bitte? #00:12:15-1#
18 I: Nach ganz oben? #00:12:15-9#
19 A6: Nach ganz oben, ja. #00:12:17-8#
20 I: Ich kann auch mal wieder dichter ran gehen[zoomen]. #00:12:20-8#
21 A6: (10) Also ich find- also ich find jetzt direkt überflüssig jetzt nichts und es hat sich ja auch gut lesen lassen ähm (..) auch oben mit der- mit dem Einstieg nochmal, also ich find das- ich find das schon gut gemacht. #00:12:43-3#
22 I: (.) Alles klar. Ähm (.) und (..) bleiben für Sie Fragen offen, wenn Sie die Rückmeldung gelesen haben, zu denen Sie gern noch ne Antwort hätten oder was Sie sich jetzt fragen? #00:12:56-7#
23 A6: (...) Ähm (6) ja ähm (4) wie (.) ähm (...) ähm ja jetzt halt irgendwie, dass man das- dass man das ähm also wenn ich jetzt ähm bei Schülervorstellungen sehr niedrig abgeschnitten habe beziehungsweise in dem Test darauf äh nicht unbedingt Wert gelegt hab, inwiefern ich das äh (.) sp- äh für den späteren Beruf beziehungsweise auch äh ja denn auch jetzt noch für die Uni und so ähm (..) ja, wie das verbessert werden kann, beziehungsweise ob wir in der Uni da nochmal drüber sprechen (.) ähm (.) ja ob wir die Aspekte nochmal so einzeln behandeln so irgendwie, würde mir vielleicht- würde mich jetzt vielleicht nochmal so interessieren, also inwiefern mich die Aspekte dann begleiten und mich dahingehend dann irgendwie nochmal zu verbessern oder (.) sowas, ja. #00:14:09-5#
24 I: (.) Okay, danke. Ähm (..) dann, das schließt eigentlich auch n bisschen an die Frage an ähm, welche weiteren Informationen hätten Sie vielleicht gerne gehabt? #00:14:20-9#
25 A6: (5) Mhm jetzt in der Rückmeldung? #00:14:27-2#
26 I: (...) Mhm. #00:14:29-5#
27 A6: (...) weitere Informationen[geflüstert] (..) äh das war ja jetzt dann praktisch nur n Ausschnitt dann zu den- äh zu den andern oder wie war das? #00:14:44-1#
28 I: Was heißt zu den andern? #00:14:45-9#
29 A6: Ähm warn da noch- achja ne genau, das warn die Elemente, okay, genau, das warn die drei Elemente. #00:14:51-6#
30 I: Achso. #00:14:52-0#
31 A6: Ja, ähm was würd ich noch- weitere Informationen (.) ähm (4) so an sich passt das ja, also Sie können mir ja jetzt auch keine weitere Information geben nachdem ich das jetzt äh nachdem ich- nachdem ich das jetzt durchgeführt habe, also Sie haben das ja gut ausgewertet dann (.) und auch nochmal die- die äh (...) die Elemente nochmal ähm (..) abgeschrieben ähm (.) ja. #00:15:25-7#
32 I: (..) Okay äh Sie haben eben gefragt, obs um weitere Informationen zur Rückmeldung geht. Hätten Sie bei ner andern Stelle noch Bedarf an Informationen oder war das ne allgemeine Frage? #00:15:39-0#
33 A6: Ne, das war jetzt so allgemein so, ja //(unv.). #00:15:41-9#
34 I: Okay. // (..) Ne, hätt nur sein können, dass Sie an irgendwas anderes grad gedacht haben, // aber-. #00:15:48-5#
35 A6: Ja, ne ne // allgemein, ja. #00:15:50-0#

36 I: Okay. Ähm ich hab hier noch so ne kleine Liste ähm an Punkten, die man möglicherweise in dieser Rückmeldung ergänzen könnte und ich würde die gerne einmal durchgehen und Sie bitten, dass Sie mir Feedback dazu geben, ob Sie das für sinnvoll erachten oder eben auch nicht. (.) Also dass ich- genau. Ähm (.) das erste wäre, dass man Informationen darüber gibt, (.) was das Ziel des Tests ist oder auch, was das Projekt ist, in dessen Rahmen dieser Test entstanden ist. #00:16:23-2#
37 A6: (4) Ja okay ja ähm (.) ja vielleicht doch, ja. #00:16:33-5#
38 I: (..) Und was davon? #00:16:36-4#
39 A6: (..) Äh ja so vielleicht, wie Sie schon gesagt haben, allgemein Informationen zu dem Projekt, äh was vielleicht das Ziel ist, ähm (...) ja, ja. #00:16:52-6#
40 I: (..) Ähm ham Sie konkret was was im Kopf, wie das der Einordnung helfen- ähm ja Einordnung hlefen könnte oder ist das (.) klingt es plausibel quasi? #00:17:04-1#
41 A6: (...) Vielleicht wüsste man dann ähm was man was man als Auszufüllender mit dem Test äh allgemein äh auch bezweckt irgendwie, wie äh (.) und äh wie Sie wie Sie damit forschen irgendwie, wie Sie da- damit umgehen, vielleicht (.) irgendwie sowas, (.) ja. #00:17:28-1#
42 I: (..) Ähm das nächste wäre, inwiefern es von Interesse wäre, dass ähm explizite Informationen darüber ergänzt werden, (.) was denn eigentlich der 'Zielzustand', also die Zielmarke der Reflexionsfähigkeit ist. #00:17:45-0#
43 A6: (...) Zi- Zielmarke. Also dann praktisch (...) dass man dann über- äh dass man dann äh (.) im Median oder über dem Median abschneidet, oder wie meinen Sie jetzt? #00:18:00-4#
44 I: Ja zum Beispiel- also das ähm das wär dann ne Angabe dazu, (.) quasi 'wie sollte ich in diesem Test abschneiden, (.) ähm damit ich, ich sag jetzt mal ganz platt, damit ich gut bin'. #00:18:14-2#
45 A6: Okay ja (.) (ah?) da das ist eher so da (bin?) eher so zweifelhaft würd ich sagen äh (.) äh also auf der einen Seite äh genau, das war nämlich in dem anderen Test [Erklärtest] dann im Nachhinein genannt äh im im auf der einen Seite ähm (.) ja erfährt man halt direkte Rückmeldung, man weiß halt äh wo man steht, wie man den Test ausgefüllt hat, aber auf der anderen Seite kann das halt auch (.) äh Verunsicherung bringen und (.) ähm das man (dann?) so im Nachhinein drüber nachdenkt 'oh (.) äh jetzt bin ich ja hier komplett im unterdurschnittlichen Bereich, obwohl ich mich eigentlich ähm gut eingeschätzt hab beziehungsweise den Test (.) ähm ernsthaft und ähm (.) ja (..) äh ernsthaft und äh mit hohem Interesse durchgeführt habe (.) äh, ja das ist (ah?) es ist son zweifelhaftes Ding, würde ich sagen. (.) Also bringt bringt Vor und Nachteile #00:19:15-2#
46 I: (.) Okay (..) ähm ne andere Option wäre, die Ergebnisse feiner aufzuschlüsseln. #00:19:24-0#
47 A6: (4) Also jetzt bei äh bei Schülervorstellungen da praktisch dann äh beispielsweise dann noch irgendwie nochmal so zwei, drei Diagramme, oder? #00:19:35-4#
48 I: Zum Beispiel (..) also ich- man könnte ähm ja vielleicht mal so zwei, drei verschiedene Sachen sagen und eine Möglichkeit wäre, dass man sagt, ähm ich mach zum Beispiel diese Diagramme nicht für die drei Bereiche Alternativen, Bewertung, Schülervorstellungen, sondern ich könnte die theoretisch auch für jede Aufgabe machen, oder zu jedem Video oder versuchen, das inhaltlich noch anders zu sortieren. Ne andere Option wär auch, dass man versucht, dass äh geht mit dem Datenpool aktuell noch nicht, aber dass man zum Beispiel versucht, (.) die Vergleichsgruppe klarer zu setzten, also wir ham ja jetzt hier einundzwanzig Studierende, die durchschnittlich im fünften Fachsemester sind, (.) könnte man zum Beispiel sagen, man nimmt nur Bachelorstudierende oder nur Masterstudierende oder was auch immer. Also es ist so auf verschiedenen Ebenen sind da Dinge denkbar. #00:20:29-7#

49 A6: (.) Ähm wenn ich ehrlich bin, ich find das- ich find das gut zusammengefasst in einem Diagramm (.) und ähm wenn wenn ich ehrlich bin, wenn Sie das in mehrere Diagramm da auch unterteilen würden, (.) dann würde das ja auch äh erstmal noch mehr Arbeit für Sie dann bedeuten und auch noch ähm die Gefahr, (.) äh dass man sich das erste Diagramm anguckt und dann ins zweite dann übergeht, (..) guckt sichs auch noch son bisschen an und dann das dritte und dann die Konzentration beziehungsweise das Interesse dann mit äh wenn das zehn Diagramme jetzt wärn beispielsweise, dann son bisschen auch abnimmt. Also (.) wenn ich ehrlich bin, ich find das- ich find das so schon ganz gut gemacht. #00:21:12-8#
50 I: (.) Okay ähm dann wäre eine Option, entweder- (..) ich fang anders an. Eine Option wäre, ähm (.) nicht diese Boxplots zu zeigen, sondern ne andere Art von Diagramm. (.) Wär das für Sie von Interesse, gerade wenn Sie jetzt zum Beispiel gesagt haben, Sie haben vorher noch nicht mit Boxplots gearbeitet? #00:21:38-0#
51 A6: (4) Ja ähm (...) ja also wär auch nicht schlecht, aber ich find das jetzt auch, wie gesagt, mit den- mit den Boxplots jetzt auch nicht- (.) jetzt auch nicht irgendwie- also find ich auch gut angebracht, (.) ja, also ich- ja. #00:21:59-1#
52 I: Ähm und wenn- wenn Sie sich jetzt andere Diagramme vorstellen, was wär da n Diagramm, was Sie sich vorstellen könnten? #00:22:06-0#
53 A6: (...) (unv.) gute Frage an sich, ich hab auch grad schon drüber nachgedacht, (.) äh (6) jetzt son (10) ja viele- viele Diagrammtypen machen ja da eigentlich gar keinen Sinn da mit den mit der Abweichung dann zum Median, wenn ich mir jetzt irgendwie son (.) klassisches äh Kreisdiagramm oder jetzt auch äh son normalen Graphen oder sowas anguck, äh (..) macht das ja gar keinen Sinn dann mit der- mit der- mit der Abweichung mit der Darstellung des Median ähm (.) ne ich find das ich find das so schon ich find das so schon sehr gut (.) ja. #00:22:58-1#
54 I: Okay (.) ähm ne weitere Option wäre eine Beispielaufgabe. Also (.) ne Aufgabe aus dem Test beziehungsweise eine eine von der Struktur sehr ähnliche Aufgabe zu ergänzen als Anschauung. #00:23:12-9#
55 A6: (8) Ja also könnte man machen, (.) aber ähm wenn man den Test absolviert hat, (.) ähm ist einem ja eigentlich- sind einem ja eigentlich die ähm die Aufgaben beziehungsweise die ähm (...) ähm ja, was man halt ausfüllen musste in dem Test, äh ja eigentlich bekannt. #00:23:42-5#
56 I: Mhm[Ja]. #00:23:43-5#
57 A6: Ähm also (..) find- also (..) ja wenn ich ehrlich bin müsste meiner Meinung nach nicht unbedingt rein (.) ähm ja. #00:23:54-1#
58 I: Ähm ne ähnliche Option wäre, dass man Musterlösungen mit Begründungen ergänzt (..) zum Beispiel eben für so eine Beispielaufgabe (.) dann. #00:24:07-8#
59 A6: (4) Begründung dann noch, gut, dann ist es natürlich wieder (..) ja, ist dann vielleicht mit- wenn da noch ne Musterlösung mit dabei ist äh ist dann vielleicht auch wieder nicht schlecht auf der einen Seite (.) äh man (.) man wüsste dann vielleicht äh (.) ähm was bei der- was bei der Aufgabe (eher?) im angebrachten Bereich lag, wie man wie man die hätte ausfüllen können (.) ähm (.) aber es kann natürlich auch wieder son bisschen für für Verwirrung dann wieder sorgen, wenn man die grade die Aufgabe dann äh (..) ja komplett falsch angekreuzt hat äh beziehungsweise ausgefüllt hat und man dann denkt, man hat äh man hat alle andern Aufgaben genauso äh (.) äh nicht unbedingt angebracht äh gelöst, wies in der Musterlösung dann von der Aufgabe dann in dem Reflexionsbogen dann da gefordert äh beschrieben ist. #00:25:08-2#
60 I: (..) Okay also auch so ne zweischneidige // Sache. #00:25:13-7#
61 A6: (unv.) // auch son zwe- zwiespältiges Ding, ja. #00:25:15-3#
62 I: Okay. (.) Ähm dann gäbs noch die Option, Literaturhinweise oder so Hinweise auf physikdidaktische Grundlagentexte zu ergänzen. #00:25:25-5#

63 A6: (...) Ja, find ich find ich nicht schlecht ähm (..) grade, wenn man ähm (..) wenn man ähm ja in der Physikdidaktik irgendwie vor ner vor ner wichtigen Prüfung oder ähm ne wichtige Hausarbeit oder so schreiben muss, auch. Wenns- wenns das oder son ähnliches Thema ist oder auch wenns einen weiterhin natürlich noch begleitet (.) ähm würd ich würd ich in der Tat nicht schlecht finden, ja. #00:26:00-6#
64 I: Ähm wozu (.) wozu genau, wären dann solche Hinweise sinnvoll? (.) // Oder an welcher Stelle? #00:26:10-4#
65 A6: Ähm (..) mhm jetzt für ähm für Seminare, die die man gerade hat, jetzt auch zum Beispiel wir in Heterogenität, da müssen wir auch jetzt son Vortrag über Schülervorstellungen und auch über ähm über Formulierungen von Lernzielen und so halten, also ich denk, das wär schon so ähm so ne zusätzliche (.) so ne zusätzliche Hilfe auch, auch wenns ma- wenns einem in der (.) in der unmittelbaren Zukunft, (.) ähm wenn man denkt, dass es dass es einem gar nicht äh (..) ähm (.) ja gar nicht ähm dahin bringt, aber irgendwie trotzdem auch so für später so find ich- find ich das schon nicht schlecht, ja. #00:26:53-3#
66 I: Okay und das wärn dann zum Beispiel Hinweise hier zur zur Erläuterung dieser (.) der einzelnen Begriffe [Erläuterung der Kategorien unter dem 2. Diagramm], hab ich das richtig verstanden? #00:27:01-8#
67 A6: Genau, ja, nochmal irgendwie sowas vertieftes irgendwie. #00:27:04-8#
68 I: Mhm[Ja]. #00:27:05-8#
69 A6: // Ja. #00:27:06-6#
70 I: Und // auch zum zum Reflexionsbegriff dann allgemein oder dann schon konkreter zu den Ergebnissen? #00:27:13-0#
71 A6: (...) Vielleicht auch zu zum Reflexionsbegriff allgemein (..) also wär vielleicht auch gar nicht so (.) schlecht, ja. (..) // Ja. #00:27:24-5#
72 I: (..) Okay. // (.) Äh und eine weitere Option wäre noch, dass man Material ergänzt oder äh also auf Material verlinkt oder Hinweise ergänzt, wie man ja an der an seiner Reflexionsfähigkeit arbeiten kann, wie man die schulen kann. #00:27:41-1#
73 A6: Ja (.) finde- fänd ich fänd ich fänd ich top (.) ähm (.) auch irgendwie ähm (.) ja genauson paar Hinweise irgendwie und dass man das dann nach den Hinweisen, dass man da irgendwie son bisschen (.) bisschen üben kann, bisschen trainieren kann und dann den Test irgendwie in ein oder zwei Wochen oder etwas später, (der?) Zeitpunkt später irgendwie dann nochmal später durchführt, man da dann ne ne ne Verbesserung sieht. #00:28:10-8#
74 I: Mhm[Ja]. #00:28:11-5#
75 A6: Vielleicht, das wär vielleicht ganz gut, ja. #00:28:14-5#
76 I: Hätten Sie da ähm ja spontan irgendwie ne ne Wunschvorstellunge oder ne Idee, (.) wie genau solche Hilfen aufgebaut sein sollten, oder was es für Material sein sollte, was Ihnen da helfen würde? #00:28:27-6#
77 A6: (.) Das ist schwierig (.) äh (5) mhm (5) ja jetzt vielleicht mit dem äh (...) ja ob einen jetzt son (.) ja ob einen jetzt so ne son (.) son Literaturtext jetzt- ich denke, der hilft schon, aber ob der jetzt so krass auch hilft, äh in der Reflexionsfähigkeit, (.) äh (.) ist irgendwie zu bezweifeln, aber vielleicht könnte man irgendwie (.) das irgendwie mit den Dozenten irgendwie absprechen, dass man das irgendwie mal in so ner (.) so ner Seminarstunde irgendwie mal ähm (.) irgendwie mal trainiert oder so, ja. #00:29:17-5#
78 I: Also konkret die die Handlung üben. #00:29:21-1#
79 A6: Genau, ja irgendwie son bisschen auch (.) vielleicht auch praktisch irgendwie son bisschen äh (.) ja, aber es- ja, es ist (ist?) es ist glaub ich ja, gar nicht so einfach. #00:29:34-3#
80 [Ab hier: Keine Probleme beim Erhalt der Rückmeldung, das Format hat Sinn ergeben, die Rückmeldung hat in der Art gut gefallen; Abschluss]

P.2 Interviews zur 1. Evaluation (Kohorte B)

P.2.1 Interview 1: B1

1 I: Okay. Ähm dann nochmal offiziell auch danke, dass du das Interview mit mir führst. Als erstes, ich habe es eben schon angekündigt, muss ich um das Einverständnis für die Verwendung der Daten bitten. Ich lese einmal die Einverständniserklärung vor und bitte dich danach, dass du einfach einmal sagst, dass du damit einverstanden bist. Äh die Einverständniserklärung lautet ich bin damit einverstanden, dass dieses Interview aufgezeichnet wird. Das Interview wird für eine Verschriftlichung, Verwendung oder Weitergabe anonymisiert. Die Einverständniserklärung zur Nutzung der erhobenen Daten kann jederzeit mit Wirkung für die Zukunft widerrufen werden. Bist du damit einverstanden? #00:00:36-9#
2 B1: Ja, das bin ich.
3 I: Super. Okay. Genau. Jetzt, ich hatte es- oder N (Anm: Dozent der Veranstaltung) und ich hatten es eben schon so ein bisschen angekündigt, es geht um den Reflexionstest. Ihr habt den vor ein paar Wochen bearbeitet, ihr habt eine Rückmeldung bekommen und ähm in der letzten Woche auch so ein bisschen mit dem Fördermaterial gearbeitet, falls du da da warst. Und ich würde gerne gleich ein bisschen was darüber erfahren, wie du die Rückmeldung wahrgenommen hast oder was du da draus mitgenommen hast. Es geht mir überhaupt nicht darum, dich irgendwie zu bewerten, sondern darum, das Format und die Gestaltung der Rückmeldung zu verbessern und ich hatte es eben auch schon gesagt, die Aussagen in dem Interview nehmen überhaupt keinen Einfluss auf hier irgendwie Ihre Teilnahme in der Veranstaltung. (.) Gibt es noch Fragen bevor wir anfangen? #00:01:21-5#
4 B1: Nein, überhaupt keine. #00:01:23-5#
5 I: Alles klar. Dann äh geht es zum Einstieg um deine eigene Rückmeldung und zwar einmal ganz platt die Frage kannst du dich noch an die Rückmeldung erinnern? #00:01:32-7#
6 B1: So vage. Es ist so zwei Wochen her, seit ich mir sie angesehen habe. #00:01:36-5#
7 I: Ja. #00:01:37-5#
8 B1: Aber so grobe Züge daran kann ich mich noch erinnern. #00:01:40-0#
9 I: Okay. Weißt du auch noch, welche Schlüsse du daraus gezogen hast? #00:01:43-9#
10 B1: (.) Ähm das weiß ich nicht mehr so ganz vollständig. (.) Auf jeden Fall konnte ich (daraus?) sehen, woran ich- also in welchen Sachen ich mich noch verbessern muss. #00:01:56-0#
11 I: Okay. Ähm (.) genau, ich habe hier, einfach damit wir auch was haben, worüber wir reden können, eine Beispielrückmeldung, also die ist anonym, das ist höchstwahrscheinlich nicht deine. Ähm ich geb dir die sonst einfach mal, der ist vom Aufbau genau gleich, nur die die tatsächlichen Ergebnisse sind natürlich andere. (.) Ähm kannst dir gerne einen Moment Zeit nehmen, da drüber zu gucken (.) und dann hätte ich da ein paar Fragen zu. #00:02:24-9#
12 B1: (39s, liest) Ja, jetzt kann ich mich schon besser erinnern, (.) wie das aussah. Stimmt ja. (...) Also was ich besonders gut fand an der Rückmeldung war auch hier, dass das (also?) die ganze Reflexion aufgeteilt war in Bewertung, Alternativen und Schülervorstellungen, das man sehen konnte, in welcher Kategorie wo- wo man stark ist und wo man eventuell noch was ausbessern muss. Das fand ich (.) das fand ich auch vor allen Dingen durch die Graphen ziemlich anschaulich dargestellt. Und auch wo man sich so insgesamt befindet. #00:03:35-2#
13 I: Okay. #00:03:36-3#
14 B1: Das war sehr ausführlich. (..) Ja. #00:03:41-7#
15 I: Okay. Dann ähm starte ich mal mit meinen Fragen. Das erste wäre, ob du vorher äh mit solchen Boxplots also dieser Diagrammart gearbeitet hast vor der Rückmeldung #00:03:52-9#
16 B1: Äh nein, bisher noch gar nicht. #00:03:56-4#

17 I: Noch gar nicht. Okay. Aber die Ergebnisse fandst du trotzdem- also hattest das Gefühl, // du kannst damit was anfangen #00:04:04-0#
18 B1: Ja, die fand- die fand ich schlüssig. Mein Zweitfach ist ja auch Mathematik, das heißt, damit komme ich ein bisschen besser klar. #00:04:08-7#
19 I: Okay. Ähm und das andere ist, hast du vorher schonmal mit der Effektstärke gearbeitet? #00:04:14-4#
20 B1: Nein, noch nie. #00:04:15-6#
21 I: Nicht. Okay. Ähm gut dann wäre die nächste Frage, wenn du dir diese Rückmeldung jetzt anschaust, diese Beispielrückmeldung, welche Schlüsse würdest du da draus ziehen? #00:04:24-5#
22 B1: (...) Naja (.) anhand der Diagramme kann man ja sehen, äh die Ergebnisse sind etwas besser als der Schnitt, aber wenn man jetzt hier genau rein sieht, (.) kann man feststellen, ja, die bewerten- (in Sachen?) Bewertungen ist man am oberen Ende, in den- in Sachen Alternativen ist man so im Median, (.) aber an den Schülervorstellungen müsste man noch feilen, weil man da ziemlich weit unter dem Schnitt ist. (.) Das sind die Rückschlüsse, die ich zumindest daraus ziehen würde, also an den Schülervorstellungen arbeiten. #00:04:57-8#
23 I: Mhm[Ja]. Äh wenn wir jetzt noch einmal auf den ersten Plot, wo das Gesamtergebnis steht, einmal schauen, da ist auf der rechten Seite die Effektstärke aufgetragen. #00:05:08-2#
24 B1: // Ja. #00:05:09-4#
25 I: Wenn du dir das (.) anschaust, was für einen Schluss würdest du aus dieser Information ziehen oder wie interpretierst du das? #00:05:16-8#
26 B1: (12s) (Ah?) ich weiß jetzt nicht genau, wie ich das interpretieren soll. Da bin ich etwas ratlos. #00:05:33-2#
27 I: Ist auch in Ordnung #00:05:34-9#
28 B1: Ja. #00:05:36-3#
29 I: Okay. Ähm genau dann nochmal zurück zur ganzen Rückmeldung. Welche Informationen aus dieser Rückmeldung sind besonders interessant oder besonders hilfreich für dich? #00:05:48-5#
30 B1: (..) Auch wenn ich mich jetzt wiederhole, also die Graphen sind besonders hilfreich für mich. Auch ja halt eben die Aufteilung. (.) Und auch dass hier dass hier im Test auch noch so Erwartungen gegeben sind, was genau man beachten soll bei Schülervorstellungen oder bei Alternativen und bei Bewertungen. Dass das ausführlich ähm geschildert ist und man sehen kann, was man verbessern muss. Das man vielleicht auf die eigene Idee kommt, was noch optimiert werden muss. #00:06:17-9#
31 I: Mhm[Ja]. Und gibt es andersrum auch Inhalt, die überflüssig sind und welche wären das? #00:06:23-9#
32 B1: (..) Ähm eigentlich (4s) also ich finde da an sich ist nichts überflüssig. (...) Das ist auch ziemlich- (.) ja es ist alles ziemlich kompakt beschrieben, also ich finde die Rückmeldung wirklich gut, ich finde da nichts überflüssig. #00:06:43-9#
33 I: Okay. Und bleiben Fragen für dich offen? Also nachdem du die Rückmeldung anschaust auf die du noch eine Antwort hätte- gerne hättest. #00:06:50-9#
34 B1: (..) Äh eigentlich nein, es bleibt nichts offen. Es ist alles ziemlich gut. #00:06:58-5#
35 I: Okay. Und gibt es weitere Informationen, die du gerne hättest? #00:07:02-2#
36 B1: Mir fallen gerade keine ein. #00:07:04-2#
37 I: Okay. Ähm genau dann war das quasi waren das meine Fragen direkt zur Rückmeldung. Und in der Rückmeldung ist ja so ein Fördermaterial oder so ein Lernmaterial irgendwie verlinkt. Das wurde in der letzten Woche in der letzten Seminarsitzung angeschaut. Jetzt frage ich glaube ich einmal vorweg, ob du da da warst und da auch mit gearbeitet hast. #00:07:24-6#

38 B1: Ich war da (.) und habe da ich habe da auch mit gearbeitet und ich habe mir das dann auch durchgelesen, aber es waren- Es war schon- Es war ziemlich viel, aber es waren- (.) Das Fördermaterial war ja quasi dann (.) sozusagen eine Musterlösung, (.) wo man sehen konnte aha, das hätte man vielleicht (auch?) so machen können. (.) Das fand ich auch ziemlich gut, damit man- Man konnte das selbst bearbeiten und dann konnte man sich angucken, was so die Musterlösung war und konnte das dann abgleichen und das war- Da konnte man dann auch sehen, wo man sich noch bessern musste. #00:08:00-2#
39 I: (..) Okay. Dann ähm (.) wenn wir jetzt einmal quasi zurück gehen in die Arbeit mit dem Material wäre meine Frage, wie du damit gearbeitet hast. Also wie bist du (.) vorgegangen oder hast du dich auf was fokussiert oder also wie war die Arbeit damit? #00:08:18-1#
40 B1: Äh ich fokussiere mich bei Aufgabenstellungen meistens nicht genau darauf und versuche das einfach von vorne bis hinten genau durchzuarbeiten und (.) ich habe mir die einzelnen Videos angeguckt, habe dann erstmal versucht, die Fragestellung selbst zu beantworten. Nachdem ich dann mir gedacht habe, gut, jetzt habe ich die Antwort gefunden, habe ich mir dann die Lösung angeguckt und habe dann das abgeglichen. #00:08:40-3#
41 I: Mhm[Ja]. #00:08:41-2#
42 B1: Und geschaut, wo ich mich noch verbessern konnte. #00:08:43-5#
43 I: Und was nimmst du aus der Arbeit mit dem Material mit? #00:08:47-7#
44 B1: (...) Äh zumindest nehme ich daraus mit, wo ich mich noch verbessern kann (.) und wo ich auch schon gut bin. Also es gibt mir einen guten Überblick, wo wo ich jetzt genau bin und wo ich hin möchte. #00:09:02-3#
45 I: Mhm[Ja]. (.) Ähm das Material jetzt als solches. Bietet das aus deiner Perspektive einen Mehrwert und warum? #00:09:11-1#
46 B1: (...) Es bietet einen Mehrwert, (.) weil man halt auch ähm (unv.) sehen kann, (.) anhand der Musterlösung, wie das wie das genau aufgebaut ist. Also man lernt aus den Musterlösungen auch daraus. Man sieht die Unterschiede zwischen sich und dem ange- den angegebenen Lösungen und man kann auch genau sehen, weil das so detailliert ist, (.) in welchen Unterkategorien man sich noch verbessern muss. (Das ist?) Deshalb ist es ähm sehr sinnvoll und sehr gut. #00:09:43-6#
47 I: Okay. Und welche Inhalt aus diesem Material waren da besonders hilfreich oder besonders interessant? #00:09:50-7#
48 B1: Daran kann ich mich nicht- kann ich mich nicht mehr so genau erinnern. Eigentlich war (es?) war alles hilfreich, sagen wir es mal so (lacht). #00:09:57-3#
49 I: Okay. Und auch wieder analog dann andersrum die Frage gab es da was, was überflüssig war und was war das? #00:10:04-1#
50 B1: Nein, das war alles (.) das war alles detailliert, aber es war auch nicht zu viel. Es war genau richtig. #00:10:11-5#
51 I: Okay. Und gibt es bei dem äh bei diesem Lernmaterial da Fragen, die für dich offen geblieben sind oder jetzt noch offen sind? #00:10:19-6#
52 B1: Nein, da wurde alles geklärt. // Da bin ich zufrieden mit. #00:10:22-8#
53 I: Okay. (.) Ähm genau dann wäre ich soweit durch mit meinen Fragen. Wenn du noch irgendwelche Fragen, Anmerkungen hast zum Test, zur Rückmeldung oder zum Material gerne. #00:10:33-5#
54 B1: (Ähm?) habe ich nicht. #00:10:35-0#
55 I: Hast du nicht. Okay. Dann // mache ich einmal die Aufnahme aus #00:10:38-8#
56 B1: Ja. #00:10:37-1#

P.2.2 Interview 2: B2

1 I: Okay. Dann nochmal offiziell vielen Dank, dass du das Interview mit mir führst. Ähm genau ich brauche zuerst gleich einmal das Einverständnis für die Verwendung der Daten. Dazu lese ich jetzt einmal die Einverständniserklärung vor und bitte danach, dass du einfach mündlich einmal sagst, dass du einverstanden bist, wenn du es bist. Ähm Einverständniserklärung lautet ich bin damit einverstanden, dass dieses Interview aufgezeichnet wird. Dieses Interview wird für eine Verschriftlichung, Verwendung oder Weitergabe anonymisiert. Diese Einverständniserklärung zur Nutzung der erhobenen Daten kann jederzeit mit Wirkung für die Zukunft widerrufen werden. Bist du damit einverstanden? #00:00:35-5#
2 B2: Ich bin damit einverstanden (lacht). #00:00:36-5#
3 I: Super, dann können wir auch anfangen (lacht). Okay. Genau. Du hast vor ein paar Wochen höchstwahrscheinlich diesen Reflexionstest bearbeitet und danach eine Rückmeldung bekommen und ihr habt in der letzten Woche, meine ich, auch mit diesem Onlinematerial gearbeitet. (.) Äh ich würde gerne gleich so ein bisschen mehr darüber erfahren, wie du die Rückmeldung wahrgenommen hast oder wahrnimmst. (.) Ähm mir geht es überhaupt nicht darum, dich zu bewerten oder deine Aussagen, sondern darum, dass das Format und die Gestaltung von der Rückmeldung und diesem Material verbessert wird. Das heißt aber auch natürlich, dass das, was du hier sagst überhaupt keinen Einfluss auf die Veranstaltung hat auf irgendwie Bewertung oder Bestehen oder so. Das hat damit nichts zu tun. (.) Hast du noch Fragen bevor wir anfangen? #00:01:21-0#
4 B2: Nicht soweit, nein. #00:01:22-3#
5 I: Okay. Dann äh gehen wir zum Einstieg einmal vielleicht zu deiner eigenen Rückmeldung und zwar mit der Frage kannst du dich an die Rückmeldung noch dran erinnern und wenn ja, welche Schlüsse hast du daraus gezogen? #00:01:34-1#
6 B2: Ähm das waren ja die drei äh Felder in denen wir die Rückmeldung bekommen hatten. Äh in dem ersten Feld wusste ich- also habe ich halt- weiß ich noch, dass ich ähm knapp über dem Durchschnitt war, aber bei den anderen zwei Bereichen im Durchschnitt. (.) Ähm was ich halt daraus mitgenommen habe war, dass ich dann schon auf mein Gefühl, sage ich mal, schon vertrauen kann ähm und dass ich dann aber leider bei der Auswertung danach und ähm ich weiß jetzt nicht mehr, welches zweite und dritte Feld wie die hießen oder worauf sie sich bezogen hatten ähm aber dass ich da halt noch ein bisschen nachbessern muss vielleicht oder (also?) ich glaube, das lag einfach nur noch daran, dass ich mich halt nicht so doll angestrengt hätte oder habe, wie ich es hätte tun sollen vielleicht ähm das kann halt auch ein Punkt sein. Aber sonst ähm fand ich den Test eigentlich relativ gut und auch gut für die Rückmeldung jetzt für die Veranstaltung noch (mit?). #00:02:29-9#
7 I: (.) Okay. Ähm genau dann habe ich jetzt, damit es so ein bisschen konkreter wird, einmal so eine Beispielrückmeldung mitgebracht. Das sind mit ziemlicher Sicherheit nicht deine Ergebnisse. Es ist anonym, du kannst gerne reingucken, einmal kurz irgendwie genau reinschauen zum Erinnern und danach habe ich ein paar Fragen. #00:02:47-8#
8 B2: (83s, liest) #00:04:11-0#
9 I: Okay. Äh dann wäre meine erste Frage, ob du vor dieser Rückmeldung schonmal mit äh solchen Boxplots gearbeitet hast? #00:04:18-4#
10 B2: Ähm nur in der Schulzeit, aber das auch nur ganz kurz zum Thema Stochastik war das glaube ich da noch. Genau. Also ich habe die- ich kenne die, aber ähm genauso werten oder bewerten konnte ich die tatsächlich nicht, aber ich habe mich halt länger mit der Graphik dann auseinander gesetzt und dann habe- konnte ich das schon dann- #00:04:37-6#
11 I: Mit dieser ersten Graphik meinst du // oder? #00:04:39-6#
12 B2: Genau. #00:04:39-7#
13 I: Mhm[Ja]. #00:04:40-3#

14 B2: (Na?) und die anderen sind ja ähnlich dann. Genau. (.) Also unbekannt waren die dann noch nicht. #00:04:47-1#
15 I: Okay. Und haben dir die Informationen gereicht, die in der Rückmeldung gegeben sind, um das zu verstehen, oder? #00:04:52-3#
16 B2: Ähm ja, das fand ich sogar richtig hilfreich. Ich habe mir natürlich erstmal die- weil es so eine große äh Darstellung erstmal war, habe ich mir erstmal den äh Boxplot angeschaut, aber danach, als ich mir die Rückmeldung nochmal durchgelesen habe, dann war es nachvollziehbar und alles verständlich. Ja. #00:05:06-3#
17 I: Okay. Und dann ist in dem gleichen Plot dieser Begriff der Effektstärke. Hast du damit vorher schon einmal gearbeitet? #00:05:14-0#
18 B2: Äh nein, tatsächlich noch nicht. Nein. #00:05:16-0#
19 I: Okay. Ähm dann wäre die nächste Frage wenn du die diese Rückmeldung jetzt anschaust, welche Schlüsse ziehst du aus der Rückmeldung? #00:05:22-7#
20 B2: Ähm (..) genau, dass da äh das Gesamtergebnis, das wäre über dem Durchschnitt ähm das eben die Person (.) ähm die da äh diesen Test gemacht hat ähm (.) über (.) ja über dem Durchschnitt war oder (.) (das?) Mittlere (unv.) (ja?) also die meisten, sage ich jetzt mal, ähm aus diesem Test, die den Test gemacht haben, ähm waren schon in dem Boxplot natürlich im oberen Bereich, in der oberen Hälfte und da dann nochmal im Vergleich zu den anderen äh Teilnehmern lag sie nochmal über dem Durchschnitt, würde ich jetzt sagen. (.) Äh genau hier sieht man nochmal den Mittelwert und dass sie dann mit ihrem Ergebnis über dem Durchschnitt sa- äh steht. Ähm (..) genau. Über dem Durchschnitt, aber nicht zu hundert Prozent würde ich dann jetzt nochmal sagen, wenn da oben da nochmal Luft nach oben ist, wo natürlich andere Teilnehmer nochmal besser abgeschnitten haben, als sie. (.) Genau. (.) Aber das ist, glaube ich, der größte Mittelwert dann in dem sie (oder?) wo die meisten Teilnehmer dann drin stehen oder drin liegen. (.) Ja. (.) Hätte ich jetzt so (unv.) interpretiert. #00:06:44-1#
21 I: (..) Mhm[Ja]. (.) Ähm (.) wenn du jetzt, ich habe jetzt eben schonmal diese Effektstärke angesprochen, die da rechts ist. Kannst du da auch einen Schluss draus ziehen? Oder wie würdest du die interpretieren? #00:06:58-3#
22 B2: Ähm (...) äh (.) (unv.) (..) mhm (.) ja und bei der Effektstärke sieht man halt, dass ähm null null bei dem Mittelwert der Vergleichsgruppe ist. (.) Mhm (...) da würde ich wahrscheinlich mit der Abweichung mitgehen, also ja mit welcher- oder wie groß die Abweichung zum mitt- äh (.) zum Mittelwert ist von der Gruppe würde ich mal schätzen. (...) Genau. Also Effektstärke habe ich nämlich so noch nie gehört oder habe auch daraus jetzt nicht wirklich schlau werden können, aber ich würde jetzt schätzen, das es (..) ja die Abweichung ist (zum Mittelwert?). #00:07:50-5#
23 I: Okay. Ähm dann wäre meine nächste Frage zu dieser Rückmeldung welche Informationen sind da besonders hilfreich oder besonders interessant? #00:08:00-6#
24 B2: Natürlich der Vergleich zum Mittelwert, also wie man im Vergleich zu den anderen Teilnehmern steht ähm (.) inwieweit man auch vielleicht über dem Durchschnitt steht oder ähm was natürlich noch verbesserungswürdig ist. (.) Also das habe ich so zumindest für mich dann äh festgestellt, als ich die ganzen Boxplots gesehen habe, dass ich entweder im Mittelwert bin, vielleicht darüber, ähm wo ich mich hätte verbessern können oder wo ich hätte aufmerksamer sein sollen, aber so im Allgemeinen glaube ich dann ähm ist der Vergleich zum Mittelwert ganz hilfreich, um zu sehen, ob man ähm (.) da (.) zumindest im Vergleich zu den anderen Teilnehmern noch äh gut dabei ist oder (eben?) ein Einzelfall ist oder auch selber vielleicht festzumachen, ob man hätte sich mehr daran äh anstrengen sollen oder nicht, das wird ja dann in dem Test auch noch gefragt, (.) ob man das so nebenbei gemacht hat oder sich wirklich intensiv damit ausinander gesetzt hat, aber ich würde schätzen, das ist das dann. #00:08:58-4#
25 I: Okay. Und andersrum gefragt gibt es Informationen oder Inhalte dieser Rückmeldung, die überflüssig sind? #00:09:05-7#

26 B2: Mhm (12s) also ich find- also mit diesen äh unteren fünfundzwanzig Prozent äh in der Vergleichsgruppe oder mit den oberen fünfundzwanzig Prozent hätte ich jetzt zum Beispiel nichts damit anfangen können, denn im Allgemeinen weiß ich ja nie, wie viele Leute da jetzt so nochmal mitgemacht haben. (.) Also (sind?)- und naja je nachdem, wie der Boxplot sich halt verschoben hat oder der Mittelwert eben, kann ich ja sehen, im Vergleich zu der Anzahl, die da halt mitgemacht haben ähm inwieweit ich da dann eben stehe, aber so mit den unteren und oberen fünfundzwanzig Prozent hätte ich jetzt zum Beispiel nichts anfangen können oder könnte die dann auch nicht deuten. (.) Ähm ja sonst ähm (..) nein, ich glaube, das wäre es dann eigentlich auch. #00:10:06-5#
27 I: Okay. Und gibt es irgendwie andere oder weitere Informationen, die du gerne gehabt hättest oder Fragen, die offen geblieben sind? #00:10:15-6#
28 B2: (...) Ähm (4s) tatsächlich nicht, weil ähm das Gesamtergebnis stand- wurde ja dann nochmal ähm in in einzelnen ähm Ergebnis immer nochmal dargestellt. Das fand ich sehr hilfreich, um zu sehen, in welchem Bereich das nochmal einzeln gewertet wurde oder wo ich mich in einem einzelnen Bereich nochmal verbessern mustse. (.) Ähm es wurde natürlich auch erklärt also vieles erklärt ähm wie man im Vergleich zu den anderen ist ähm (..) genau also das fand ich halt auch schon sehr hilfreich. Dann ähm (.) sonst vermissen, nein. Also ich fand das auch gut, dass dann am Ende des Tests nochmal gefragt wurde, wie intensiv man sich auf diesen Test dann nochmal eingelassen hat, ähm wie doll man sich damit auseinander gesetzt hat, denn dann wurde ja auch anhand dessen das Ergebnis festgemacht und konnte man dann auch sehen okay, selbst wenn ich da nicht so viel aufgepasst habe (unv.) in dem Bereich, hätte ich mehr aufgepasst oder mitgemacht, dann hätte ich oder wäre ich in dem Bereich gewesen. Das fand ich sehr ähm hilfreich und nochmal motivierend dann im Vergleich zu dem Ergebnis, äh das ich dann (da?) bekommen hatte. Ähm (.) also sich halt mehr darauf einzulassen, intensiver damit zu arbeiten, um dann nochmal ein besseres Ergebnis mitzubekommen und vielleicht noch mehr über dem Durchschnitt zu sein, aber sonst ähm genau ja. #00:11:41-2#
29 I: Okay. Genau dann wären das meine Fragen zu diesem Rück- oder zu dieser Rückmeldung genau gewesen. Dann hätte ich jetzt noch ein paar zu diesem Onlinematerial. #00:11:51-1#
30 B2: Mhm[Ja]. #00:11:51-8#
31 I: Da habt ihr in der letzten Woche hier in der Veranstaltung was gemacht. #00:11:54-2#
32 B2: Genau. // Wir haben- #00:11:54-9#
33 I: Du nickst so, das heißt du hast da auch mit- du warst auch da? #00:11:57-2#
34 B2: Genau, // ich war da. #00:11:57-8#
35 I: Ja, okay. #00:11:58-4#
36 B2: Wir haben die äh oder wir sollten die durchlesen, durchgesprochen hatten wir die tatsächlich nicht. Es gab nur eine Gruppe, die die zusammen bearbeitet hatte und ähm besprochen hätte. Ähm ich für meinen Teil war allein, ich bin da einzeln mit durchgegangen, ähm habe mir die Situation nochmal angeschaut, ähm habe damit gearbeitet, inwieweit sich da alles äh oder wie das halt vorgegeben wurde, ob ich hätte das genauso gemacht oder ob ich das genauso interpretiert hätte. Ähm also so eher eine Selbstreflexion damit durchgeführt. #00:12:31-3#
37 I: Mhm[Ja]. Du hast jetzt auch schon so ein bisschen angeschnitten, aber wie hast du dann mit dem Material gearbeitet? #00:12:36-8#

38 B2: Ähm ich habe mir das Video oder das waren ja diese Situationen ähm die habe ich mir dann nochmal einzeln angeschaut. Ähm habe da versucht, da so viel mitzunehmen, eigentlich aus der Situation, wie ich- oder als ob ich so im Unterricht säße und das bewerten müsste. Ähm habe dann äh genau versucht, das alles mitzunehmen und danach eben bin ich auf die Kommentare oder auf die ähm auf die Plots da unten (da?) dann nochmal gegangen und ähm habe die mir durchgelesen, ob ich das hätte genauso gemacht, ob ich das nachvollziehen konnte, ähm ob ich die Situation genauso eingeschätzt hätte, wie ähm das halt vorgegeben war und ähm wenn ich da (.) einen Unterschied gefunden habe, dann bin ich nochmal auf das Video gegangen, habe es mir nochmal angeschaut und ähm wie gesagt, eine kleine Selbstreflexion dann durchgeführt, ob das alles genauso gewesen ist oder ob ich das genauso eingeschätzt hätte, wie es vorgegeben war. #00:13:31-1#

39 I: Mhm[Ja]. (.) Und aus der Arbeit mit dem Material, was kannst du daraus für dich mitnehmen oder was nimmst du daraus für dich mit? #00:13:38-3#

40 B2: Ähm meinem eigenen Gefühl zu vertrauen. Ähm also dass ich da schon die Schülervorstellungen ein bisschen nachvollziehen kann, ob ich da- also kann auch sein, dass ich da gefühlt selber erst noch aus der Schule raus komme, aber dass ich nachvollziehen kann, wie die Schüler was vors- also wie die Schüler den Lehrer, die Lehrkraft sehen, wie ich das auch hätte anders gemacht teilweise oder es- meiner Meinung nach war das wie- (also?) sind Situationen ein bisschen übertrieben dargestellt gewesen, aber ähm das zeigt ja nochmal, dass es auch eben einer Lehrkraft passieren kann, dass man da irgendwie so ein bisschen rein rutscht indirekt und dass eben darauf noch ein bisschen aufpassen muss und ähm (.) beispielsweise jetzt mit (.) äh (.) äh (.) mit der Schülersituation, wo die Lehrkraft ein bisschen ähm die Schüler so ein bisschen, sage ich mal, dumm dargestellt hat, ähm ich weiß nicht mehr, welche Situation das war, aber ähm es gab da einen und ähm (man?) das halt- das wurde natürlich übertrieben dargestellt, dadabei möchte ich bleiben, aber äh das man eben da aufpassen soll, wie man eben die Wörter wählt, das sind immernoch Schüler, die ähm natürlich nichts dafür können, wie sie was verstehen teilweise, das eben Schüler sich auch darauf einlassen müssen und das hängt natürlich vom Tag der Schüler ab, wie sie was mitnehmen und ähm das man da eben nicht so einen harten Druck drauf geben soll. #00:15:06-3#

41 I: Okay. Ähm dann wäre meine nächste Frage, ob das Material aus deiner Frage einen Mehrwert bietet und warum. #00:15:15-3#

42 B2: Ähm ja schon ähm wie gesagt, es ist halt sehr hilfreich, wenn man sich mit (dieser?) Situation auseinander setzt äh wenn man (.) nochmal (.) äh vielleicht auch nochmal ins Auge gesetzt bekommen möchte, wo man eher aufpassen soll, wo man ja ähm was nachvoll- also wo man eher darauf achten soll, wie man mit Schülern umgeht, äh was man vielleicht verbessern sollte, wo es Probleme geben könnte und wenn man erst mit dieser Situation konfrontiert wird, versteht man auch vielleicht, äh wo man dann später dran arbeiten muss. Und allein schon deswegen fand ich das sehr hilfreich. Und dann noch mit der Rückmeldung der Situation ähm, ob man eben die Situation genauso wertet, denn je- von Mensch zu Mensch versteht man ja (immer?) auch nochmal was unterschiedlich. Aber dann ähm, sage ich mal, auf einen gleichen Nenner zu kommen und zu sehen, okay, das ist jetzt eigentlich die Priorität dieser Situation, das ist eigentlich jetzt das Wichtige ähm das hilft (halt?) nochmal ganz doll, da nochmal die Situation richtig zu einzuschätzen und nicht einfach ähm die anders zu werten oder auch falsch zu interpretieren. #00:16:22-8#

43 I: Mhm[Ja]. Und gab es in dem Material Inhalte, die besonders hilfreich oder interessant für dich waren und welche wären das? #00:16:30-8#

44 B2: Ähm jetzt direkt fällt mir glaube ich nichts soweit ein. Ich fand das- also ich habe das Material im Allgemeinen ähm (.) erstmal bewerten können überhaupt und ich fand das halt im Allgemeinen, wenn man sich damit intensiv auseinander setzt und natürlich in jeder Situation dann nochmal intensiv drauf eingeht und ähm sich damit auseinander setzt und das auch alles nachvollzieht und vielleicht auch (.) im Nachhinein hätte ich da nochmal ein bisschen anders gemacht. ich hätte meine oder mir Notizen nebenbei gemacht und äh wäre gerne ins Gespräch gekommen mit meinen Kommilitonen, aber so im Allgemeinen ähm (..) so einzelne Situationen nicht, aber wie gesagt im Allgemeinen fand ich die ganze Arbeit hilfreich. #00:17:13-7#
45 I: Mhm[Ja]. Und gab es da andersrum vielleicht Inhalte, die überflüssig sind aus deiner Perspektive? #00:17:19-8#
46 B2: Ähm (5s) äh teilweise war es ja so, dass es erstmal geschrieben wurde, was in der Situation zu sehen ist und dann war nochmal die Lehrkraft, die das nochmal erzählt hat, was niedergeschrieben wurde, also da fand ich das halt immer so ich habe es mir schon durchgelesen, dann höre ich das nochmal von ihm und dann sehe ich die Situation. Ähm vielleicht ist es auch ganz gut, wenn man die ähm (.) die- das Geschriebene erstmal nicht gesehen hätte, (.) weil (das?) dann das irgendwie so doppelt ist, aber irgendwie wiederum kann ich es auch nachvollziehen, dass es halt doppelt ist, damit man eben das gelesen hat, gesehen hat und dann wieder- also praktisch war das ja auch so eine Einführung ins Video. Ähm (.) naja, ich sag mal so, wenn man sich dran gewöhnt hat, dass es erstmal das geschrieben ist, dann die Situation nochmal erklärt wurde und dann das in dem Video direkt darauf ange- daran angeschlossen hat, dann konnte man halt damit, sage ich mal, umgehen und damit arbeiten. Aber am Anfang war es ein bisschen nervig, wenn man sich das ein bisschen durchgelesen hat, das äh verstanden hat die Situation und dann hört man diese Situation, das waren dann so die ersten zwanzig Sekunden ach, jetzt erzählt er genau das, was ich eben gelesen habe, aber sonst, wie gesagt, wenn man sich darauf eingelassen hat, dann ähm war es eigentlich ganz okay. #00:18:35-3#
47 I: Okay. Und gibt es Fragen, die für dich offen geblieben sind aus der Arbeit mit dem Material? #00:18:40-7#
48 B2: Ähm (.) soweit nicht, glaube ich. (.) Ähm nein, also ich fand, wie gesagt, das Material hilfreich, um überhaupt die Situation einzuschätzen, darauf einzugehen und auf den gleichen Nenner zu kommen und selber nochmal zu gucken, habe ich das richtig verstanden, kann ich das nachvollziehen, aber sonst ähm (..) nein, ich glaube nicht. #00:19:03-4#
49 I: Okay. Dann wäre ich durch mit meinen Fragen. Wenn du noch irgendwelche Fragen, Anmerkungen hast zum Test, zur Rückmeldung, zu dem Material, (.) gerne jetzt. #00:19:12-6#
50 B2: Ähm nein, nicht soweit. #00:19:15-2#
51 I: Okay. Dann mache ich die Aufnahme einmal aus. #00:19:18-3#

P.2.3 Interview 3: B3

1 I: Okay. Also dann nochmal ganz offiziell vielen Dank, dass du das Interview mit mir führst. Ähm genau, wie eben angekündigt, ich brauche zunächst das Einverständnis für die Verwendung deiner Daten. Dafür lese ich die Einverständniserklärung nochmal vor und bitte danach einfach um Einverständnis. (.) Die Einverständniserklärung lautet ich bin damit einverstanden, dass dieses Interview aufgezeichnet wird. Dieses Interview wird für eine Verschriftlichung, Verwendung oder Weitergabe anonymisiert. Die Einverständniserklärung zur Nutzung der erhobenen Daten kann jederzeit mit Wirkung für die Zukunft widerrufen werden. (.) Bist du damit einverstanden? #00:00:32-6#
2 B3: Ja. #00:00:33-0#
3 I: Super. Ähm genau. Ihr habt vor ein paar Wochen diesen Reflexionstest bearbeitet und danach eine Rückmeldung bekommen und in der letzten Woche auch mit diesem Onlinematerial gearbeitet. Ähm (.) und ich würde gleich gerne ein bisschen was darüber erfahren, wie du diese Rückmeldung wahrnimmst beziehungsweise wahrgenommen hast. Ähm mir geht es überhaupt nicht darum, dich irgendwie zu bewerten oder das, was du sagst, sondern diese Rückmeldung oder die Gestaltung und das Format zu verbessern. Das heißt auch, alles, was du hier sagst, hat nichts mit der Veranstaltung zu tun, keinen Einfluss auf Note, Bestehen, was auch immer, das ist unabhängig davon. (.) Hast du noch Fragen bevor wir anfangen? #00:01:12-4#
4 B3: Nein. #00:01:13-3#
5 I: Okay. (.) Äh dann würde ich vorschlagen, wir machen zum Einstieg äh ja gucken wir einmal- oder was heißt gucken wir. Also es geht zum Einstieg um deine Rückmeldung. So. Kannst du dich noch an die Rückmeldung erinnern? #00:01:27-6#
6 B3: Äh vom Reflexionstest? #00:01:29-3#
7 I: Ja. #00:01:30-0#
8 B3: Ja. Doch. #00:01:31-0#
9 I: Okay. Welche Schlüsse hast denn du daraus gezogen? #00:01:34-2#
10 B3: Mhm aus der Rückmeldung? #00:01:36-9#
11 I: Mhm[Ja]. #00:01:37-3#
12 B3: Einmal habe ich- Also ich fand die- den Reflexionstest insoweit gut, weil ich mir vorher für die das Modul auch gar nicht so sicher war, wie so eine Prüfungsleistung aussehen kann oder wie Fragen formuliert werden können. Und daran habe ich mich so ein bisschen orientiert jetzt, (.) (wo?) ich sagen kann, okay wie kann denn Reflexion so getestet werden und wie wird das vielleicht auch später für das Referendariat bewertet oder welche Fragen kann man sich stellen. Was- wie sieht das aus, wenn man, sage ich mal, etwas nicht ganz so gut macht. (.) Genau und dann war das ganz gut auch, si- mich auch ein bisschen einzuschätzen, wie gut das denn gelaufen ist, oder wie schlecht. #00:02:13-8#
13 I: Okay. Äh ich habe hier einmal, damit wir nicht so abstrakt bleiben, eine Beispielrückmeldung mitgebracht. Das ist mit ziemlicher Sicherheit nicht deine. (unv.) einfach kurz einen Moment, da rein einmal zu schauen und dann (.) machen wir weiter. #00:02:28-3#
14 B3: (103s) #00:04:11-1#
15 I: Okay. Ähm meine erste Frage wäre, ob du vor dieser Rückmeldung schonmal mit solchen Boxplots gearbeitet hast. #00:04:19-4#
16 B3: Nein. #00:04:20-5#
17 I: Nie gesehen. #00:04:21-4#
18 B3: Nein. #00:04:22-1#
19 I: Okay. Ähm und dann ist in dieser- ich mach das mal ganz kurz auf. In dem ersten Plot ist auch (..) so ist hier auf der rechten Seite auch die Effektstärke eingetragen. Hast du mit der schonmal gearbeitet? #00:04:37-5#

20 B3: Ja, in Statistik. (.) Also ich habe vorher schonmal BWL studiert und da hatten wir ein Modul, das Statistik hieß. #00:04:43-9#
21 I: Ah okay. Ähm (.) gut. Dann wäre meine Frage, wenn du dir diese Rückmeldung jetzt angeschaut hast, welche Schlüsse würdest du aus dieser Rückmeldung ziehen? #00:04:54-4#
22 B3: (...) Mhm naja im Großen und Ganzen konnte ich mich ja selber da ein bisschen vergleichen im statistischen Mittel wie ich da schon stehe. (..) Da wäre- also wir hatten Reflexion ja noch nicht in der Vorlesung, deswegen war es irgendwie eher etwas, okay, was bringe ich schon mit im Vergleich. Ähm (..) ja also (...) an den oberen oder unteren Kästchen habe ich mich jetzt nicht so gemessen, es kann ja auch sein, dass Leute das irgendwie nebenbei gemacht haben, oder weiß ich nicht. (.) Ja. #00:05:31-3#
23 I: Also verstehe ich richtig, du hast dich am Mittelwert // orientiert? #00:05:34-9#
24 B3: (Ja?) genau, ich habe mich mit dem Mittelwert orientiert. #00:05:37-0#
25 I: Okay. (.) Ähm (..) und (.) wenn wir jetzt nochmal auf die Effektstärke gucken, wie würdest du die hier interpretieren (.) in diesem (.) Kontext? (..) Oder welche Informationen ziehst du daraus? #00:05:55-5#
26 B3: (4s) Darauf habe ich ehrlich gesagt damals noch gar nicht geachtet, ich habe immer nur das- auf das Ergebnis in Prozent gerech- äh geachtet. (4s) Naja, das müsste ja- (5s) Also die Effektstärke null ist ja hier bei dem Median, das heißt, wenn man schaut, okay, wie weit liegt man da drüber, dann (..) müsste das ja- Also null Komma fünf, hat äh eben Herr N gesagt, ist eine hohe Effektstärke, also ab einer Effektstärke von null Komma fünf wird es als relevant oder als hoch eingeschätzt. Kann man schauen, okay gut, wie hoch ist der Unterschied vielleicht? Ich weiß n- Nein, da bin ich mir nicht ganz sicher. (lacht) #00:06:43-2#
27 I: Okay. Ähm (..) und in dieser Rückmeldung insgesamt, welche Informationen sind da besonders hilfreich oder besonders interessant? #00:06:52-5#
28 B3: Mhm also ich fand das Diagramm- Also das Gesamtergebnis natürlich- Also ich habe mich- Also vor allem fand ich die Graphen ganz gut, weil die halt auf einen Blick irgendwie das (direkt?) also rübergebracht haben. Ich habe- muss auch sagen, den Text habe ich damals relativ schnell überflogen. Ich dachte auch am Anfang, dass das ähm Infomaterial ist, das gar nicht auf mich zugeschnitten war ähm und d- das nur die Graphen da meine Ergebnisse ähm wiederspiegeln. Ähm die fand ich ziemlich gut [Anm: zeigt auf Teilergebnisse], weil- #00:07:26-8#
29 I: Die Teilergebnisse, oder- #00:07:28-7#
30 B3: Die- Genau, die Teilergebnisse. Die fand ich sehr interessant. Die habe ich mir auch länger angeschaut, weil da eben auch noch dann im Vergleich mir gezeigt wurde, wo ich dann nochmal dran arbeiten kann für später. (.) Genau. #00:07:40-6#
31 I: Okay. Und gibt es andersrum Inhalte, die aus deiner Perspektive überflüssig sind? #00:07:46-7#
32 B3: Mhm (...) also ich habe es mir- ich habe es mir ja gerade durchgelesen, da fand ich es eigentlich ganz gut, also sowei- so, wie es war. Ähm einige Sachen, zum Beispiel jetzt über die Schülervorstellung, das hatten wir im Modul schon, das (..) finde ich aber gut, dass es da nochmal drin steht. Also (..) das hätte ich mir jetzt nicht, glaube ich, ähm (.) ja nochmal so intensiv durchgelesen. Aber es ist ganz gut, dass (so?) manchmal so Sachen noch erläutert sind. #00:08:17-9#
33 I: Okay. Äh und gibt es Fragen, die für dich offen bleiben? #00:08:22-7#
34 B3: Mhm so spontan nicht, nein. #00:08:25-2#
35 I: Nein. Oder weitere Infos, die du gerne gehabt hättest oder hättest? #00:08:29-4#

36 B3: Ich glaube tatsächlich, dass da alles drin gewesen wäre, was ich gebraucht hätte. Ich bin fälschlicherweise davon ausgegangen, dass es irgendwie ein zehn Minuten Test war und äh habe dann- Ich glaube, auf der ersten Seite stand das wahrscheinlich irgendwie, dass das ein bisschen länger dauert, aber da ja da wurde ich überrascht von, dass das dann ein bisschen- das ich da mehr Zeit einplanen musste. Genau. #00:08:51-5#
37 I: Okay. (.) Ähm genau. Dann wären das meine Fragen zu der Rückmeldung. Jetzt nochmal zu diesem Onlinmaterial, mit dem habt ihr euch in der letzten Seminarsitzung beschäftigt. Ich frage vorweg einmal warst du da da? Hast du das gemacht? #00:09:04-8#
38 B3: Ja. Ich bin mir aber gerade nicht mehr sicher, was das genau war. #00:09:08-5#
39 I: Ähm. #00:09:09-8#
40 B3: Ach das war das, wo wir diskutieren sollten über die- über- da waren nochmal- war nochmal das gleiche, also die gleiche Person, die das Video vorgetragen hat und da waren dann nochmal Lösungsvorschläge und so weiter. #00:09:19-5#
41 I: Genau. #00:09:20-3#
42 B3: Ja. #00:09:20-6#
43 I: Okay. Dann äh genau die folgenden Fragen natürlich beziehen sich da drauf und zwar wie hast du mit diesem Material gearbeitet? #00:09:29-0#
44 B3: Äh ich habe das tatsächlich als Einzelauftrag verstanden und ich glaube, das- also es gab nur eine Zweiergruppe, die das ähm ja die diskutiert hat. Ich bin tatsächlich da rein gegangen und habe (.) mir die Frage durchgelesen, habe mich meine eigenen Gedanken gemacht und dann aber die Lösungsvorschläge durchgeklickt. (..) Genau. #00:09:49-4#
45 I: Okay. Und was nimmst du aus der Arbeit mit dem Material für dich mit? #00:09:54-7#
46 B3: Mhm (.) ich fand das nochmal ganz gut auch gerade diese Lösungsvorschläge zu schauen und zu sehen okay, was wird denn erwartet. (.) Ähm (.) wie könnte ich denn auf so eine Frage antworten oder wie geht man da an so eine Situation ran. Das (ist?) war schon ganz gut. #00:10:09-6#
47 I: Okay. Und ähm so ganz allgemein gefragt, bietet das Material aus deiner Perspektive einen Mehrwert und warum? #00:10:17-7#
48 B3: Äh also jetzt in diesem Modul bietes- bietet es mir vor allem einen Mehrwert, weil ich eben genauso wie bei dem Test auch eben (.) mich Frage, wie eine Prüfungsleistung so aussehen kann ähm und das wäre ja so ein Schritt oder eben auch auf das Referendariat dann bezogen, gerade didaktisch, was bemängelt werden kann, wie wie analysiert wird und daran werden sich ja später die Prüfer auch halten. (.) Genau. #00:10:44-2#
49 I: Okay. und gab es in dem Material Inhalte, die da besonders hilfreich oder besonders interessant waren? #00:10:50-3#
50 B3: (...) Mir fällt gerade nichts ein. Es ist wahrscheinlich zu lange her, tut mir leid. #00:10:58-5#
51 I: Kein Problem. Ähm andersrum fällt dir dann vielleicht auch nicht ein, aber welche- ob da Inhalte überflüssig waren und welche das waren? #00:11:06-9#
52 B3: (..) Nein (unv.). (...) Ich muss gerade- ich überlege gerade, aber- also negativ nicht, aber positiv überlege ich gerade noch. (...) Nein, es ist zu lange her. #00:11:22-1#
53 I: Ähm und gibt es Fragen, die für dich offen geblieben sind aus dem Material oder aus der Arbeit mit dem Material? #00:11:28-6#
54 B3: Nein. #00:11:30-4#
55 I: Okay. (.) Dann wären das meine Fragen. Wenn du noch irgendwie Fragen, Anmerkungen hast zum Test, zur Rückmeldung, zu diesem Onlinematerial (.) gerne. #00:11:39-5#
56 B3: Nein, es war gut. Also es ist eine coole coole Sache gewesen, (da?) man ein bisschen interaktiv auch ein bisschen mitarbeiten konnte. Das war ganz gut. #00:11:47-6#
57 I: (..) Okay. Dann mache ich mal die Aufnahme aus. #00:11:52-1#

P.2.4 Interview 4: B4

1 I: Ähm das [Anm: die Aufnahme] ist nur, damit ich mir nicht alles merken muss, genau. (.) Genau. Dann erstmal Danke, dass du das Interview mit machst. Ich habe es eben angekündigt, ich brauche zunächst das Einverständnis für die Verwendung der Daten. Ich lese dafür einmal die Einverständniserklärung vor und bitte dich dann danach, dass du, wenn du einverstanden bist, das sagst. Die Einverständniserklärng ist ich bin damit einverstanden, dass dieses Interview aufgezeichnet wird. Das Interview wird für eine Verschriftlichung, Verwendung oder Weitergabe anonymisiert. Die Einverständniserklärung zur Nutzung der erhobenen Daten kann jederzeit mit Wirkung für die Zukunft widerrufen werden. Bist du einverstanden? #00:00:33-5#
2 B4: Ja. // Das bin ich. #00:00:34-2#
3 I: Perfekt. (.) Ähm genau. Ich habe es vorhin schon angekündigt, ihr habt vor ein paar Wochen diesen Reflexionstest bearbeitet, eine Rückmeldung bekommen und letzte Woche mit diesem Material gearbeitet und ich würde gerne gleich mehr darüber erfahren, wie du die Rückmeldung wahrnimmst oder wahrgenommen hast. Ähm es geht mir überhaupt nicht darum, dich zu bewerten oder deine Aussagen, sondern das Format und die Gestaltung der Materialien zu verbessern. (.) Und alles was du hier sagst hat natürlich überhaupt keinen Einfluss auf die Veranstaltung, Benotung, Bestehen, das ist komplett unabhängig davon. Hast du noch Fragen bevor wir anfangen? #00:01:08-6#
4 B4: Nein, habe ich nicht. #00:01:09-9#
5 I: Okay. Dann äh geht es zum Einstieg um dein eigene Rückmeldung und zwar mit der Frage kannst du dich noch an deine Rückmeldung zum Reflexionstest erinnern? #00:01:20-0#
6 B4: Mhm (.) ja. Also zum prozentuellen Teil schon. #00:01:26-7#
7 I: Okay. (.) Und welche Schlüsse hast du aus der Rückmeldung gezogen? #00:01:30-1#
8 B4: Ähm dass ich (.) entweder beim nächsten Test mehr Arbeit reinstecken sollte in den Test, oder ähm mein didaktisches Verständnis äh verbessern sollte. (.) Eins von beidem. Also ich hatte halt eine ähm Rückmeldung unter dem Schnitt, also knapp unter dem Schnitt. Der Schnitt lag zum Beispiel bei einer Rückmeldung bei fünfzig Prozent und ich war bei siebenundvierzig. Und (.) ja, das habe ich so daraus gezogen. #00:02:02-2#
9 I: Mhm[Ja]. (.) // Und wovon- #00:02:04-1#
10 B4: Wenn das die Frage beantwortet. #00:02:04-8#
11 I: Genau. Wovon machst du quasi abhängig, ob das quasi mehr Arbeit reinstecken ist oder- #00:02:11-7#
12 B4: Achso äh genauer durchlesen, mehr Gedanken machen auf jeden Fall. Ähm also ich muss sagen, ich habe mir- (piepen) Irgendwas piept hier gerade. #00:02:20-2#
13 I: Ich befürchte, da ist irgendwo im Schrank // eine Stoppuhr. #00:02:22-4#
14 B4: Oh mann. Egal. Ähm ne also dass man sich halt mehr Zeit dafür nimmt, die Aufgaben zu überdenken und sich richtig Gedanken zu machen. Äh das habe ich zwar oft gemacht, aber nicht immer. Also vor allem am Ende habe ich eher okay durchgeschaut, was kommt jetzt spontan infrage für mich am meisten und dann habe ich das gewählt. So. Ähm und ich habe mich alles durchgelesen, ich habe nichts geskipped oder sowas, aber ähm ja (.) wie schon gesagt. #00:02:50-0#
15 I: Okay. Ähm genau. Damit das ein bisschen konkreter wird und nicht so abstrakt über deine Rückmeldung. Ich habe hier eine Beispielrückmeldung mitgebracht, das ist höchstwahrscheinlich nicht deine. Ähm es ist genau vom Text das gleiche, es ist nur ein anderes Ergebnis, schau gern kurz rein. #00:03:04-9#
16 B4: Nein, ist nicht meine #00:03:01-9# #00:03:13-0#
17 B4: (9s, liest) Achso ja. (20s, liest) Okay. (8s, liest) #00:03:44-1#
18 I: Okay. (.) Ähm dann wäre meine erste Frage hast du vor dieser Rückmeldung schonmal mit Boxplots gearbeitet. (.) Also mit dieser Diagrammart? #00:03:55-5#

19 B4: Boxplots? Also- #00:03:56-9#
20 I: Also diese Art von Diagrammen. #00:03:58-8#
21 B4: Mhm nicht, das ich wüsste. #00:04:01-5#
22 I: (...) Okay. #00:04:05-0#
23 B4: Wieso? #00:04:05-7#
24 I: Nein, alles gut. Es ist einfach nur eine Frage, um zu schauen, ob die Diagramme passen, oder ob ob ihr da Vorerfahrungen mit habt. #00:04:13-7#
25 B4: Achso. Das- ich finde das Diagramm ähm habe ich noch nie gesehen so die- die Vorstellungsart, also speziell diese hier. #00:04:19-5#
26 I: Mhm[Ja]. #00:04:20-2#
27 B4: Aber ich fand sie auch nicht schwierig zu lesen oder zu deuten oder sowas. #00:04:23-8#
28 I: (..) Okay. #00:04:25-4#
29 B4: Falls das die- egal. Weiter gehts. #00:04:27-1#
30 I: Nein, alles gut. Kannst gern alles sagen. Ähm hast du vorher schonmal mit der Effektstärke gearbeitet? #00:04:33-4#
31 B4: (..) Sagt mir nichts. #00:04:35-9#
32 I: Alles klar. Ähm (.) wen du dir die Rückmeldung jetzt so anschaust, welche Schlüsse würdest du daraus ziehen? #00:04:43-0#
33 B4: Mhm (7s) mhm ich würde eine persönliche Rückmeldung- Wenn das jetzt meine Bewertung wäre, würde ich daraus eine kurze ähm grobe Rückmeldung zu meinem didaktischen Verständnis vielleicht ähm (.) erfahren. Also so eine so eine allgemeine Rückmeldung wie es bei mir aussieht. #00:05:09-6#
34 I: Mhm[Ja]. #00:05:10-1#
35 B4: So in den bestimmten Themenbereichen. Ähm ich würde mir aber auch nicht zu viele Gedanken machen, wenn ich jetzt schlecht abschneiden würde im Test, ähm da das halt nur sehr grob ist. So da- heißt nicht, dass sie dass es jetzt äh irgendwie schlecht gemacht ist oder so, aber es ist halt (.) ja (.) äh nur eine grobe Rückmeldung, weil man- weil man da ankreuzen kann uns sowas. ja. #00:05:31-7#
36 I: Mhm[Ja]. Ähm du hast jetzt gesagt, du kriegst irgendwie Informationen so über dein fachdidaktisches (.) // ne Abschneiden #00:05:40-5#
37 B4: Ja. (.) Mhm[Ja]. #00:05:40-7#
38 I: Ähm welche Informationen kriegst du denn? #00:05:43-4#
39 B4: (.) Äh (.) Vo- Schülervorstellungen, die ich selber noch habe auf jeden Fall, ähm über das faire Bewerten wahrscheinlich würde ich sagen, also ob ich ob ich ähm gut bewerte sozusagen, ob ich vernünftig bewerte und neutral, heterogen. (.) Ähm (..) was gabs dann noch. Alternativen. (.) Ja über diese beiden Hauptpunkte. würde ich sagen. #00:06:08-9#
40 I: Okay. Also diese drei Teilbereiche, die da- #00:06:11-1#
41 B4: Mhm Alternativen weiß ich jetzt nicht, was das waren, aber mir würden die äh die beiden hier. #00:06:14-7#
42 I: Mhm[Ja]. // Bewertung und Schüler- #00:06:15-5#
43 B4: Bewertung und Schülervorstellungen. #00:06:16-6#
44 I: Okay. Und ähm wie interpretierst du das quasi die Lage der Ergebnisse, also wo jetzt die die Ergebnisse eingezeichnet sind. Also wie bewertest du das Abschneiden (.) jetzt von dieser Rückmeldung oder von dieser Person, der die Rückmeldung gehört? #00:06:33-1#
45 B4: Du meinst, ob ich zufrieden damit wäre? #00:06:35-0#
46 I: Zum Beispiel. Oder wo- #00:06:37-1#
47 B4: Achso. #00:06:37-5#
48 I: Woran du das festmachst, ob es jetzt okay oder- #00:06:40-4#

49 B4: Eindeutig am am Schnitt würde ich sagen. (.) Also das ist der Haupt- ähm punkt wo man das- wo ich das jetzt in Relation setzen würde, (.) ob ich jetzt gut abgeschnitten bin oder nicht. Oder vor allem vor allem hier kann man das halt gut sehen. (.) Also das ist ja so der der Bereich, wo die meisten abschneiden. #00:06:56-4#
50 I: Bei den Schülervorstellungen jetzt, oder- #00:06:58-2#
51 B4: Genau. Also bei den bei den drei Klötzen hier. #00:07:00-2#
52 I: Mhm[Ja]. #00:07:00-1#
53 B4: Ähm ja und dann würde ich hier beispielsweise sagen, ja ich könnte mich verbessern und hier bin ich schon äh auf einem ganz guten Weg. #00:07:06-1#
54 I: Ja #00:07:07-0#
55 B4: So. #00:07:07-5#
56 I: Okay. (.) Ähm jetzt ist hier bei diesem ersten Gesamtergebnis die Effektstärke eingetragen. #00:07:15-4#
57 B4: // Mhm[Ja]. #00:07:16-0#
58 I: Wenn du dir die anschaust, (..) wie würdest du das interpretieren? #00:07:20-1#
59 B4: (..) Die Effektstärke war dieser Abstand hier? Oder (.) was war das? #00:07:25-2#
60 I: (4s) Nein. Aber ist okay. Also das ist total in Ordnung, wenn du das nicht kennst und einfach sagst, du weißt nicht, wie du es interpretieren sollst. #00:07:33-3#
61 B4: Okay. #00:07:33-6# #00:07:35-1#
62 B4: Nein, würde ich jetzt nicht wissen. #00:07:36-8#
63 I: Okay. Ähm (.) gut. Dann wieder zur Rückmeldung quasi insgesamt. Welche Informationen sind da besonders hilfreich oder besonders interessant für dich? #00:07:47-4#
64 B4: (..) Mhm (...) bei diesen Informationen hier? #00:07:54-5#
65 I: Mhm[Ja]. #00:07:55-0#
66 B4: Ähm (4s) naja die maximal erreichten Punktzahlen sagen ja auch schon viel darüber aus, wie ich da bei den einzelnen Aufgaben abgeschnitten bin, aber ansonsten würde ich auf jeden Fall sagen, bei diesen drei Punkten hier. Vor allem, weil ich ja vorhin gelernt habe, dass speziell die Schülervorstellungen und die Bewertungen für mich persönlich eine ganz gute Reflexion gebildet haben. (.) Und dann nochmal den statistischen Vergleich sozusagen zu sehen, ist, glaube ich, ganz gut. #00:08:22-2#
67 I: Okay. (.) Und gibt es andersrum Inhalte, die aus deiner Perspektive überflüssig sind? #00:08:28-2#
68 B4: (...) Jetzt in der- Also hier bei der Rückmeldung oder allgemein in der Umfrage? #00:08:33-9#
69 I: Äh jetzt gerade für die Rückmeldung. #00:08:36-5#
70 B4: (..) Ähm (.) also um ehrlich zu sein, habe ich mir den ganzen Text hier nicht durchgelesen, weil ich dachte, das sind so allgemeine Informationen // zum Test und so. Von daher kann ich dazu nicht viel sagen, aber (...) nein. Also würde mir jetzt spontan nicht einfallen. #00:08:53-8#
71 I: Mhm[Ja]. #00:08:44-8#
72 I: Okay. Und gibt es Fragen, die für dich offen geblieben sind? #00:08:57-7#
73 B4: (4s) Mhm (..) spontan nicht, nein. #00:09:04-4#
74 I: Oder Informationen, die du gerne gehabt hättest noch zusätzlich oder (was?)? #00:09:08-5#
75 B4: (4s) Unter Umständen was die (.) bestmöglichen Lösungen da wären, weil die weiß man ja jetzt ja nicht. Man weiß nur was, also wo man also prozentuell steht, aber nicht, was man genau falsch gemacht hat. Also falsch gemacht hat. #00:09:27-0#
76 I: Also du meinst quasi eine Musterlösung, // für den Test. #00:09:30-3#
77 B4: Richtig. Ja, genau. #00:09:31-4#
78 I: Mhm[Ja]. Okay. Ähm genau. Das wären meine Fragen zu der Rückmeldung. #00:09:36-3#
79 B4: Mhm[Ja]. #00:09:36-7#

80 I: Dann habt ihr ja in der letzten Woche dieses Onlinematerial ähm bearbeitet. Ich frage mal vorweg warst du da da, hast du das mitgemacht? #00:09:43-4#
81 B4: Ähm (.) ich war da, ja. Aber ich (..) kann mich gerade nicht- Bestimmt. Ja. ich muss wieder rein kommen. #00:09:51-6#
82 I: Das war- das war so ein Onlinematerial, wo auch wieder solche Videos waren und danach gab es aber nicht diese Fragen zum ankreuzen, // sondern so Beispiel- #00:10:01-5#
83 B4: Ah. (.) Ja, ja. Wir haben- ich habe aber nur eine Aufgabe geschafft, wie die meisten die meisten anderen. Also wir hatten nur zehn Minuten davon. Ich glaube, das waren drei Aufgaben und ich habe nur die zweite Aufgabe geschafft. #00:10:07-1#
84 I: Ja. #00:10:03-8#
85 I: Ja, okay. (.) Ähm genau darauf beziehen sich quasi meine nächsten Fragen und zwar wie hast du mit diesem Material gearbeitet? #00:10:14-2#
86 B4: (4s) Das waren doch Videos, oder? (.) Ich habe mir die Videos angeschaut. (.) Ähm (7s) weiß ich jetzt nicht mehr. Also ich habe- wir haben (ja?) die Videos angeschaut und dann analysiert, aber wie genau (.) habe ich jetzt nicht mehr im Kopf (.) gerade. #00:10:35-8#
87 I: Okay. Ähm (..) was nimmst du aus der Arbeit mit dem Material vielleicht mit für dich? #00:10:43-7#
88 B4: Ähm ich würde sagen, dass man auf jeden Fall (.) weil es ging ja darum, dass eine Person da gelehrt hat und man sollte dann die Situation analysieren mit der Klasse, ähm auf jeden Fall so eine dritte Perspektive hat. Also wenn man später ja (irgendwann?) selber unterrichten wird, das man sozusagen weiß, was für Fehler man machen kann aus der dritten Perspektive, weil wenn man die Lehrperson ist, die die Fehler macht, ist es schwer, die Fehler zu erkennen. Und ähm ja bei der Analyse würde ich sagen, dass es ganz hilfreich war. #00:11:13-8#
89 I: Mhm[ja]. Würdest du dann auch sagen, dass das Material einen Mehrwert bietet? Und wenn ja, warum? #00:11:20-1#
90 B4: (.) Inwiefern? #00:11:21-8#
91 I: Für dich. Einen Mehrwert, // dass du dieses Material- #00:11:25-1#
92 B4: Achso. (.) Ah okay. (.) Ja, ja klar, wie gerade eben beschrieben, wenn das rübergekommen ist. Ich meine halt das man quasi Fehler die man selber vielleicht als Lehrer gar nicht sehen würde, die man gemacht hat halt bei einer anderen Person sieht. Zum Beispiel ich glaube, ich hätte vielleicht (.) ohne irgendwie böse Absichten zu haben aber auch gesagt, ja, hier, du hast doch so eine schöne Schrift also zu einem Mädchen. (.) Vielleicht kannte er das Mädchen (auch?), ich weiß nicht, was in diesem Video war. Aber vielleicht hätte ich das auch so sagen können und dann bringt mich so ein Video doch auch zum Nachdenken, ob das überhaupt das richtig ist. Auch, wenn es überhaupt nicht böse gemeint ist oder so. (.) Aber ähm (.) ja, das hat schon einen Mehrwert für mich in dem Sinne, dass man halt reflektieren kann. #00:12:01-9#
93 I: Alles klar. Und gab es in dem Material oder in dem Teil, den du bearbeitet hast Inhalte, die da besonders interessant oder hilfreich für dich waren? #00:12:09-7#
94 B4: (6s) Naja, die Sache, die ich gerade eben schon angesprochen habe und das, worüber wir gerade ja auch gesprochen haben. #00:12:18-9#
95 I: Das // mit dem- #00:12:19-6#
96 B4: (Im?) Unterricht. #00:12:20-1#
97 I: Ja, okay. #00:12:20-9#
98 B4: Und das man da halt- Dass das so kleine Dinge sind, äh die aber einen großen Einfluss haben können, so // wie zum Beispiel sowas. #00:12:26-8#
99 I: Ja. #00:12:26-1#
100 B4: (.) Mhm weil das hätte man ja auch so interpre- interpretieren können, ja, ist so eine kleine Sache, ne, ähm sagt man halt eben nebenbei, dass äh dass bei Mädchen so die Schrift besser ist. (.) Ähm (.) aber ja. (..) Was war die Frage nochmal? #00:12:41-7#

101 I: Ob da irgendwas besonders interessant // oder besonders hilfreich war. #00:12:45-8#
102 B4: Achso. Ja, das- das war besonders // interessant. (.) Äh mir würde bestimmt noch was einfallen, aber ich will das nicht so so nicht hinauszögern.
103 I: Ja. #00:12:46-3#
104 I: Gab es andersrum Dinge, die aus deiner Perspektive überflüssig waren? #00:12:56-0#
105 B4: (4s) Ähm für mich persönlich vielleicht das (..) Thema (.) mhm (4s) nein, das kann man auch nicht so sagen. Ich wollte sagen, dass dass man äh dass man die Heterogenität und die Gleichberechtigung im Physikunterricht halt na sehr sehr äh (.) hervorgehoben hat, aber ähm und dass äh also dass wir halt relativ aufgeklärt schon sind und sowas. Aber gerade in dem Beispiel, was ich genannt habe, sieht man ja, dass es wohl nicht so ist. (.) Von daher weiß ich nicht. Fä- Fällt mir nicht spontan so ein. #00:13:31-9#
106 I: Okay. Und gibt es zu dem Material Fragen, die für dich offen geblieben sind? (.) Oder aus der Arbeit mit dem Material? #00:13:39-5#
107 B4: Mhm (...) Nein. #00:13:43-5#
108 I: Okay. Äh dann wäre ich soweit durch mit meinen Fragen. Hast du noch irgendwie Fragen, Anmerkungen zum Test, zur Rückmeldung, zum Material? Du hattest vorhin irgendwie, glaube ich, gefragt, ob das Inhalte überflüssig sich auf den Test bezog oder so oder- #00:13:58-6#
109 B4: Nochmal. (..) Inhalte (unv.) #00:14:02-0#
110 I: Vergiss die zweite Frage und wir gehen zurück zu hast du noch irgendwelche Fragen oder Anmerkungen? #00:14:07-1#
111 B4: (lacht) Äh nein. Habe ich nicht. Also (.) ich glaube, ich hatte geschrieben, ich hatte in meiner Antwort da geschrieben, dass ich die Videos cool fand. Also cool gemacht. (.) Ähm ich weiß nicht, bist du da beteiligt an der Videoproduktion oder? #00:14:20-7#
112 I: Ähm ich war beteiligt, die ist abgeschlossen, aber da war ich noch Studentin. #00:14:26-5#
113 B4: Ah okay. #00:14:27-5#
114 I: Genau. #00:14:27-9#
115 B4: Nein. #00:14:29-0#
116 I: Ich habe die nicht nicht selber geschrieben. #00:14:30-4#
117 B4: Aber du hast die Fragen geschrieben? #00:14:31-6#
118 I: Ja. #00:14:32-1#
119 B4: Ah okay. (.) Äh (..) ich hätte mich wahrscheinlich doch vorbereiten sollen, dann hätte ich das alles wieder- Nein, ich fand ich fand es gut, wie es ist. (.) Ähm wie macht ihr das, wieso macht ihr diese Umfrage jetzt hier eigentlich? #00:14:45-8#
120 I: Das, was ich jetzt hier mache mit dir? #00:14:47-2#
121 B4: Ja. #00:14:47-8#
122 I: Äh das ist, um zu schauen, ob die Rückmeldung für Studierende angemessen ist, also ob die euch was bringt oder ob wir die irgendwie anders umgestalten müssen. // (.) Ähm was ihr da draus mitnehmt, ne, wie versteht ihr das oder worauf legt ihr wert? Es geht darum, die quasi besser zu machen, damit im nächsten Jahrgang, im nächsten Semester // die besser ist. #00:15:09-6#
123 B4: Achso. #00:14:56-5#
124 B4: Ja. (.) Man könnte halt noch wie gesagt spezifizieren, (.) aber- #00:15:12-8#
125 I: Was meinst du mit // spezifizieren? #00:15:14-6#

126 B4: Ähm dass man halt genau weiß, wa- welche Punkte man quasi wo- wo- was (ich?) vorher schon erzählt habe, // äh was also dass man direkt sehen kann, wo habe ich (.) quasi was falsch gemacht // bei welchem Experiment, weil das waren ja echt viele Experiment, die man da beobachtet hat. Und wo man sich verbessern könnte. Aber dann kann man dagegen halt sagen, es ist halt ein riesen Aufwand, sowas zu machen, ne? Und wenn man wenn man hä die Genauigkeit da erhöhen wollen würde, müsste man noch mehr Experimente machen und noch mehr Analysen, dann würde man noch mehr dran sitzen, also ich als Student, der das äh- Egal, ich glaube ich glaube, du weißt, was ich meine so dabei. #00:15:49-1#
127 I: Ja. #00:15:20-8#
128 I: Ja. #00:15:25-9#
129 I: Mhm[Ja]. #00:15:30-2#
130 I: Eine detailliertere // Ergebnisaufschlüsselung quasi, ne? // Welche Frage wie- #00:15:53-6#
131 B4: Genau. #00:15:50-8#
132 B4: Ja. (.) Ja, ja. (.) Richtig. Dass dass ich so genau weiß, (.) wo ich meine Fehler habe, aber man könnte eigentlich schon sagen, dass das hier auch reicht. Also (.) das ist ja schon eine sehr gute Darstellung von den äh Themen. #00:16:03-6#
133 I: Die Teil- Teilergebnisse meinst // du? #00:16:05-1#
134 B4: Richtig. Ja, genau, was man hier sieht in diesem Graphen (.) ähm ich glaube, das muss man nicht unbedingt nochmal genauer machen, aber f- für mich persönlich wäre es ganz cool gewesen. Aber wie gesagt, so eine Musterlösung zu haben. Oder das so- #00:16:16-9#
135 I: Ja. #00:16:17-5#
136 B4: Ja. (.) Aber ansonsten nichts. #00:16:19-8#
137 I: Nein. Okay. Vielen Dank. // Dann mache ich die Aufnahme- #00:16:21-5#
138 B4: Okay. #00:16:22-5#

P.2.5 Interview 5: B5

1 I: Okay. Dann (.) vielen Dank, dass du das hier mitmachst mit dem Interview. Ähm (.) ich möchte zunächst um dein Einverständnis für die Verwendung der Daten bitten. Dazu lese ich einmal die Einverständniserklärung vor und bitte danach, (.) dass du einmal sagst, ob du einverstanden bist. (.) Ähm die Einverständniserklärung lautet ch bin damit einverstanden, dass dieses Interview aufgezeichnet wird. Das Interview wird für eine Verschriftlichung, Verwendung oder Weitergabe anonymisiert. Die Einverständniserklärung zur Nutzung der erhobenen Daten kann jederzeit mit Wirkung für die Zukunft widerrufen werden. Bist du damit einverstanden? #00:00:33-0#
2 B5: Ja. #00:00:33-6#
3 I: Super. (.) Ähm genau. Ihr habt vor ein paar Wochen diesen Reflexionstest bearbeitet und eine Rückmeldung dazu bekommen und mit letzter Woche in dies- mit diesem Onlinematerial gearbeitet. Und ich würde gleich gerne so ein bisschen mehr darüber erfahren, wie du diese Rückmeldung wahrnimmst oder wahrgenommen hast ähm (.) und (.) genau und eben auch über das die Arbeit mit dem Onlinematerial. Es geht mir überhaupt nicht darum, dich zu bewerten oder deine Aussagen, sondern ich möchte eben die Forma- äh das Format, die Gestaltung der Materialien verbessern. (.) Und, ganz wichtig, alles, was du hier sagst hat natürlich überhaupt gar nichts mit der Veranstaltung, also (.) nichts mit Bestehen, nichts mit Note zu tun. Ähm genau. (.) Welche Fragen hast du noch bevor wir anfangen? #00:01:16-0#
4 B5: Mhm ich habe erstmal keine Fragen. #00:01:18-0#
5 I: Alles klar. (.) Ähm dann geht es zum Einstieg einmal um deine eigene Rückmeldung. Erinnerst du dich noch (.) an die? #00:01:25-3#
6 B5: Ja. Ich erinnere mich daran, dass ich besonders äh den Schülervorstellungen wohl im äh unteren Bereich gelandet bin und mich dann im Nachhinein gefragt habe, inwiefern das tatsächlich äh (.) richtig auf die Antworten, die ich gegeben hätte, hätte ich das anders zu Gemüte geführt, diesen Test äh also repräsentativ war. Einfach aus dem Grund da dem da der Test einmal bei mir abgestürzt ist und zwar kurz vor Ende und dann ähm habe ich beim zweiten Durchlaufen, (in?) dem ich das Ganze nochmal ausgefüllt habe, äh das nur durchgeklickt und mir die Sachen nur überflogen. Das war natürlich ein bisschen äh ja schlecht für das das Ergebnis, nehme ich an, weil ich einfach keinen Nerv hatte, das Ganze nochmal zu machen. Das war irgendwie ein bisschen doof von mir, dass ich den Test quasi angefangen hatte und dann äh kurz vor Ende ange- äh was mir was zu Essen gemacht habe und als ich zurück gekommen bin, meinte die Seite leider, dass sie (.) jetzt äh neu geladen werden musste und dann waren alle meine Daten weg. #00:02:22-6#
7 I: Ärgerlich. #00:02:23-7#
8 B5: Dementsprechend war ich mir so ein bisschen unsicher und habe mich gefragt okay, hätte ich dann, wenn ich mir das Ganze beim zweiten Mal nochmal genau richtig durchgelesen hätte dann das gleiche Ergebnis bekommen. #00:02:33-6#
9 I: Mhm[Ja]. Du hast es jetzt schon angeschnitten. Ich frage es trotzdem nochmal. Welche Schlüsse hast du denn aus deiner Rückmeldung dann gezogen? #00:02:40-3#
10 B5: Ja also ähm die Schlüsse, die ich gezogen hätte, wenn ich gesagt hätte gut, das Ergebnis ist auf jeden Fall repräsentativ für meinen Wissensstand äh wären gewesen, ich muss mir definitiv Schülervorstellungen nochmal äh zu Gemüte führen, weil dort bin ich im weit unteren Bereich gewe- äh gelandet, während ich in den anderen Bereichen ziemlich äh nahe an dem (.) dem erwünschten Wert gewesen war, also im oberen Bereich. Ähm genau deswegen, das wäre mein mein Schluss // gewesen danach, (.) weil es ja schon ganz schön graphisch aufgestellt war, das war ganz angenehm.
11 I: Ja. #00:03:08-8#
12 I: Du meintest gerade äh erwünschter Wert. Was meinst du damit? #00:03:16-4#

13 B5: Ja also ich habe bei der graphischen äh Darstellung der Ergebnisse waren ja so äh Balken gegeben, in denen man liegen hätte sollen #00:03:25-2#
14 I: Mhm[Ja]. #00:03:25-8#
15 B5: am besten also wenn man jetzt (.) besonders gut reflektiert hätte. Ähm und da lag ich eben in den zwei anderen Gebieten aus- aus- außerhalb von Schülervorstellungen lag ich dort in diesen (..) beiden (.). #00:03:40-6#
16 I: Magst du einmal zeigen, was du meinst? [Anm: Nutzt Beispielrückmeldung] #00:03:43-4#
17 B5: Ja. Genau, hier. Hier drin. #00:03:44-9#
18 I: Achso in dieser in diesem Kasten. Quasi in der Box. #00:03:47-6#
19 B5: In dem Kasten, genau. Ja, ja. Genau. #00:03:48-6#
20 I: Okay. (.) Dann habe ich es verstanden. #00:03:49-8#
21 B5: Ja. #00:03:50-9#
22 I: Äh super. (.) Dann genau. Ich habe es jetzt schon hier geöffnet. Damit es nicht so abstrakt ist, ich habe eine Beispielrückmeldung mitgebracht. Das sind ziemlich sicher nicht deine Ergebnisse. Ähm der Text und so weiter ist natürlich identisch, nur das Ergebnis ist unterschiedlich. Du kannst gerne kurz reinschauen, wenn du magst und dann äh hätte ich da ein paar Fragen zu. #00:04:10-6#
23 B5: (73s, liest) Oh jetzt verstehe ich auch, dass ich das falsch verstanden hatte (...) während ich mir die meine Rückmeldung angeguckt habe. Ich ich war davon ausgegangen, dass der Median der gewünschte Wert ist und nicht, das di- das Ergebnis in Prozent, also je besser je höher, desto besser // tatsächlich äh damit dargestellt werden sollte. #00:05:41-2#
24 I: Mhm[Ja]. #00:05:37-6#
25 B5: Aber es macht auch irgendwie beim zweiten Betrachten mehr Sinn. Ich habe das tatsächlich, als ich mein- meinen Feedbackbogen angeguckt habe nur sehr stark überflogen und bin dann zum falschen Schluss gekommen. #00:05:52-8#
26 I: (...) Okay. Dann wäre meine erste Frage, ob du vor dieser Rückmeldung schonmal mit dieser Art von Diagramm gearbeitet hast, mit solchen Boxplots. #00:06:02-1#
27 B5: Nein (.) noch nicht. #00:06:04-4#
28 I: Okay. #00:06:05-1#
29 B5: Also noch nicht bewusst, wenn wenn dann habe ich es vergessen. #00:06:07-9#
30 I: (.) Ähm (.) und mit der Effektstärke, (.) hast du damit schonmal gearbeitet // vorher? #00:06:14-4#
31 B5: Das habe ich gerade heute das erste Mal gehört. #00:06:16-0#
32 I: Okay. #00:06:16-8#
33 B5: In der Vorlesung (lacht). #00:06:17-5#
34 I: Ähm wenn du dir die Rückmeldung jetzt anschaust, welche Schlüsse (.) ziehst du aus dieser Rückmeldung? #00:06:25-0#
35 B5: (..) Ähm (.) ich glaube, sobald ich da gelesen hätte, dass ich im äh oberen Teil der Vergleichsgruppe liege, hätte ich mir erstmal gedacht okay, ja dann reicht das ja. (.) Also das wäre mein erster Schluss gewesen tatsächlich. Dann hätte ich mir vielleicht das hier nochmal angeguckt und gesehen okay gut, Schülervorstellungen, da kann ich mich auf jeden Fall nochmal verbessern. (.) Ähm (..) und bei Alternativen ist auch nochmal Luft nach oben, aber äh Schülervorstellungen würde ich mir dann wahrscheinlich nochmal angucken, hätte ich die Zeit. #00:06:55-0#
36 I: Okay. Und ähm jetzt ist hier in diesem ersten Diagramm die Effekstärke auf der rechten y-Achse eingetragen. #00:07:03-4#
37 B5: Mhm[Ja]. #00:07:04-0#
38 I: Wenn du dir das so anschaust, welche Schlüsse würdest du daraus ziehen oder wie würdest du das interpretieren? #00:07:10-5#

39 B5: (..) Meinen Wert bei der Effektstärke? (..) Ähm (5s) ja also ich habe das wie gesagt dieses Wort überhaupt in dieser Vorlesung das erste Mal gehört und da bin ich mir noch nicht so richtig sicher, was ich damit anfangen soll tatsächlich. (.) Also mir fehlt da so ein bisschen Kontext dazu okay, was sagt mir denn jetzt eine Effektstärke von eins. #00:07:35-8#
40 I: Okay. (..) Ähm dann gehen wir von der Effektstärke einfach wieder weg und zurück zur ganzen Rückmeldung. Und zwar welche Informationen in der Rückmeldung sind besonders hilfreich oder besonders interessant für dich? #00:07:48-9#
41 B5: Mhm zum Beispiel die Erläuterung zur Effektstärke, die ich gerade gesehen habe. Die würden mir nämlich Aufschluss darüber geben, was mir denn die Effektstärke tatsächlich sagt. (.) Ähm (...) ja und da ich diese Rückmeldung wirklich einfach nur überflogen habe und das jetzt auch gerade wieder gemacht habe, diese prägnanten Aussagen, wie ich bin im oberen Bereich, dann die graphischen Aus- äh ähm (.) ja Auswertungen, die äh geben natürlich eigentlich wenn man sie versteht erstmal schneller guten Überblick. (.) Ähm (.) na also in meinem in meinem Fall habe ich mir das wirklich angeguckt äh mit wenig Zeit. Also nicht so, dass ich mich jetzt hingesetzt habe, okay das habe ich gemacht (.) äh und ich lese mir das jetzt alles durch, sondern ähm da so die graphischen Auftragungen sind auf jeden Fall sehr hilfreich und eben die dick gedruckte äh Rückmeldung, in welchem Bereich ich liege. (.) Ja. #00:08:41-7#
42 I: (.) Okay. Ähm gibt es andersum Informationen, die aus deiner Perspektive überflüssig sind? #00:08:47-6#
43 B5: (6s) Dafür müsste ich mir wahrscheinlich alles durchlesen, (.) äh dann könnte ich das eher sagen. (..) Mhm (..) das sind wahrscheinlich jetzt auch gar nicht so wenig Seiten (unv.). (4s) Mhm (19s) nein eigentlich nicht. Ich glaube, ich bräuchte einfach mehr Zeit, (.) um mir das wirklich- weil wenn ich das alles durchgelesen hätte, dann hätte ich überhaupt keine Fragen zu den Diagrammen gehabt und äh hätte sie auch richtig verstanden. Aber ich glaube, bei mir war einfach der Fehler, dass ich mir nicht genug Zeit dafür genommen habe. #00:09:41-1#
44 I: Okay. Ähm gibt es denn Fragen, die sonst noch für dich offen geblieben sind aus dieser Rückmeldung? #00:09:48-9#
45 B5: (10s) Naja also hier ist ja jetzt keine Rückmeldung dazu, wie ich die einzelnen Fragen richtig oder falsch beantwortet habe, aber das ist ja im Onlinematerial möglich, die richtigen oder mögliche richtige Antworten (.) äh sich äh zu den einzelnen Fragen anzuschauen, deswegen da eigentlich nicht. Ich glaube, ähm niederschwelliger, (.) (aber?) wahrscheinlich ungenauer wäre es wahrscheinlich gewesen, so ein einfaches Balkendiagramm am Anfang hinzusetzen, wo steht okay, das ist das gewollte Ergebnis, da warst du. (.) Das hätte ich direkt verstanden, weil ich kenne das schon. Und jetzt hier das musste ich natürlich erstmal äh hätte ich mir einmal durchlesen müssen. Gut, das hätte auch nur eine fünf Minuten gedauert, aber die habe ich mir halt nicht genommen. #00:10:36-0#
46 I: Also quasi die die gleichen Werte also deine Ergebnisse, die hier- also sind jetzt nicht deine, // aber die Ergebnisse, die als Gesamtergebnis und Teilergebnisse sind, die einfach als Balkendiagramm, oder? #00:10:45-7#
47 B5: Ja. #00:10:42-4#
48 B5: Genau ja. Als äh auf die erste Seite so als Schnellübersicht // und wenn einen das dann äh wenn man das genauer haben will, dann den Rest nochmal lesen. Also so wäre es für mich zugänglicher gewesen, ich hatte- kannte halt eben dieses äh diese Art der Darstellung noch nicht. Aber hier ist sie auch erklärt. Also ist halt die Frage, ne, also ist jetzt der Fehler bei mir, weil ich halt einfach die Zeit nicht investiert habe, die ich auch einfach mal hätte investieren können oder (.) ähm ja muss man davon ausgehen, dass alle nur dreißig Sekunden Zeit haben, um sich ihre Ergebnisse anzugucken. #00:11:15-0#
49 I: Mhm[Ja]. #00:10:49-6#
50 I: Okay. Gibt es noch andere Informationen, die du gern gehabt hättest? #00:11:19-8#

51 B5: Nein, ich glaube nicht. Also ich würde sagen, das sind auf jeden Fall genug Informationen und wenn ich dann die äh ja diese Reflexion nochmal einzeln aufgeschlüsselt haben will, war das ja online verfügbar in dem Material. Also ich habe auch ich habe mit meinem äh mit einem äh Kommilitonen dann auch mir das Ganze äh sehr intensiv angeguckt und auch wirklich nur mit dem ersten Videoausschnitt gearbeitet und das allein hat schon sehr viel äh Gesprächs- äh bedpotenzial gegeben, einfach nur zu gucken okay, was würden wir dazu sagen. #00:11:48-6#
52 I: Mhm[Ja]. Da sind wir auch quasi schon bei meiner nächsten Frage. Und zwar (.) wäre ich jetzt auch zu diesem Onlinematerial, was ihr in der letzten Woche angeschaut habt (.) äh übergegangen. Und zwar ähm wäre meine erste Frage, wie du mit dem Material gearbeitet hast. #00:12:05-4#
53 B5: Ja wir haben uns tatsächlich das Video zusammen angeguckt, also wir saßen da zu zweit vor und dann haben wir dann uns die Fragen angesehen und direkt angefangen, so ein bisschen darüber zu diskutieren. Wir hatten dann oft auch unterschiedliche Standpunkte und haben uns dann Tipps die Tipps dazu angeguckt und daraufhin nochmal weiter diskutiert und dann die äh Antworten und sind da tatsächlich auf sehr ähm also wir sind beide auf äh also ich glaube beim ersten Video war es irgendwie zwei Möglichkeiten, das anzugehen und si- jeder ist auf eine davon gekommen. Also das war dann, wenn man viel Zeit darein steckt und wenn wir auch äh, in dem Fall haben wir das ja gemacht, dann kommen wir auch zu den richtigen Ergebnissen, aber das war nicht ansatzweise die Zeit, die wir äh investiert haben, als wir den Test selber gemacht haben. Das war einfach (.) weniger. Das war so eine Zwischendurchsache. Man hat das so als Aufgabe bekommen im Fachdidaktik, das war bis jetzt noch nie irgendwie eine Veranstaltung, wo man zu viel Zeit investieren musste und äh auch konnte, weil der Rest schon so zeitintensiv ist und äh dann dementsprechend hat man sich da eine halbe Stunde für freigenommen und äh das war aber lange nicht genug. Also da müsste man also mhm ich müsste, glaube ich, zwei Stunden mindestens investieren, um das dann äh wirklich (.) lange darüber nachzudenken und das dann äh richtig- #00:13:17-8#
54 I: Für das Material? Also oder den Test? #00:13:20-1#
55 B5: Den Test. #00:13:20-2#
56 I: Den Test. // Mhm[Ja]. #00:13:20-8#
57 B5: Ja. Damit ich da wirklich (.) dann auch Antworten gebe, die ich auch äh über die ich lange nachgedacht habe. Ja. #00:13:26-5#
58 I: (..) Okay. Ähm was nimmst du für dich aus der Arbeit mit diesem Onlinematerial mit? #00:13:35-0#
59 B5: (...) Mhm (...) ja also so ein so ein ich weiß jetzt ich habe jetzt keine zentrale Erkenntnis, aber ich finde es sehr hilfreich, um äh sein Reflektieren zu reflektieren irgendwie und die also um (.) ähm mögliche Reflexionsansätze (.) äh neben den Reflexionen, die ich selber hatte einfach dort austesten zu können oder ja (.) daneben betrachten zu können. #00:14:04-4#
60 I: Mhm[Ja]. (.) Ähm und ganz ähnlich eigentlich. Würdest du sagen, dass das Material einen Mehrwert bietet und wenn ja, warum? Oder wenn nein, warum nicht, natürlich auch. #00:14:15-0#
61 B5: I- Ich finde auf jeden Fall einen Mehrwert, weil ich dann dadurch eben diese Rückmeldung hatte okay, das wäre jetzt die richtige Antwort gewesen ähm oder genau äh eine richtige Antwort und ähm das habe ich ja jetzt hier aus diesem Material nicht rauslesen können. Ich wusste jetzt nicht okay, zu Frage eins hatte ich diese Antwort und die war falsch oder richtig und deswegen also das war ja schon sehr ausführlich dann dort äh (.) beantwortet und mit Tipps versehen, das fand ich gut. #00:14:42-8#
62 I: Und gab es in dem Material Inhalte, die da besonders hilfreich oder besonders interessant waren? #00:14:48-8#

63 B5: Ich fand es gut, dass man einen Tipp bekommen hat, bevor man die Lösung hatte, weil dann konnte man selber nochmal seine eigentlichen Gedanken reflektieren und gucken okay, passt das auch mit dem Tipp. (.) Und ja (.) also wi- wir sind wirklich auch nur über diesen einen Abschnitt gekommen, also nicht weiter darüber hinaus, deswegen weiß ich nicht, inwiefern das jetzt (.) ja, aber in dem Fall finde ich es gut, also so besonders den Tipp. #00:15:10-4#
64 I: Und gibt es Inhalte, die da überflüssig waren oder sind oder deiner Perspektive? #00:15:15-8#
65 B5: Mhm (.) ist mir nichts aufgefallen tatsächlich. #00:15:20-2#
66 I: Okay. (.) Und Fragen, die offen geblieben sind oder die offen bleiben? #00:15:25-2#
67 B5: (...) Auch nicht. Also keine, die mir jetzt gerade in den Sinn kommt. #00:15:31-2#
68 I: (..) Okay. Ähm dann habe ich noch eine Frage. Und zwar hast du (.) vorhin gesagt, du hättest gerne eine Aufschlüsselung, wie du bei den einzelnen Aufgaben abgeschnitten hast. #00:15:42-7#
69 B5: Mhm[Ja]. #00:15:43-4#
70 I: Beziehungsweise was da die richtige Lösung gewesen wäre. Und dann hast du gesagt, das bekommt man ja in dem Onlinematerial. #00:15:48-9#
71 B5: Mhm[Ja]. #00:15:49-4#
72 I: In dem Onlinematerial sind das ja jetzt aber nicht genau (.) also das ist ja nicht das gleiche Format mit Fragen und Antworten, sondern das ist freier. #00:15:58-4#
73 B5: Mhm[Ja]. #00:15:58-8#
74 I: Reicht dir das trotzdem oder fändst du es trotzdem besser, zusätzlich noch eine Aufschlüsselung über alle Aufgaben zu bekommen? #00:16:06-0#
75 B5: Also ich hätte es auf jeden Fall spannend gefunden, in meinem Fall nochmal eine Aufschlüsselung aller Antworten und meine- also aller richtigen Antworten und meiner Antworten da gegenübergestellt. #00:16:14-0#
76 I: Mhm[Ja]. #00:16:14-7#
77 B5: Ähm einfach um zu sehen okay, liegt das jetzt dieses Ergebnis daran, dass ich immernoch so denke, dass ich immernoch diese Antwort gegeben hätte oder lag das jetzt daran, dass ich mir das einfach beim zweiten Durchgang einfach nur durchgeklickt habe, weil ich mir das nicht nochmal alles genau durchlesen wollte. Ähm in dem Fall hätte ich das also wenn das einfach zu machen wäre, hätte ich das mir angeguckt wahrscheinlich einfach nur, um das gegenüberzustellen meine Antwort und das das, was richtig war. Ja. (.) Aber ich fand es auch äh spannend, das äh auf diese Weise, wie es dort dargestellt war einfach zu äh nochmal zu diskutieren, aber genau ja. Ich hätte es auch spannend gefunden eben so aufgeschlüsselt. Ja. #00:16:56-5#
78 I: Ja. Okay. Ähm dann wären das meine Fragen gewesen. Ähm hast du nach Fragen, Anmerkungen zum Test, zur Rückmeldung, zu dem Material? #00:17:05-9#
79 B5: Mhm das Einzige ist, das äh also das stand ja tatsächlich vor dem Test, das habe ich danach gehört, ich habe mir das aber gar nicht durchgelesen, sondern einfach auf starten gedrückt. Und da stand ja tatsächlich, dass man fünfundvierzig Minuten mindestens einplanen soll (oder so?) irgendwieso. Ähm das äh ich das vorher nochmal gewusst hätte gut, es müssen wirklich mindestens fünfundvierzig Minuten sein und eigentlich eher mehr. #00:17:29-5#
80 I: Mhm[Ja]. #00:17:30-3#

81 B5: Und ähm (.) für mich war das jetzt an so einem Zeitpunkt um zwanzig Uhr äh kurz vor dem Abendbrot so und das ist halt auch äh leider oft so, dass solche Aufgaben dann in solche Randzeiten rutschen und dafür, finde ich, ist der Zeit- äh Test schon sehr umfangreich. Also sich da nochmal äh eineinhalb Stunden, glaube ich, hätte ich eigentlich mindestens gebraucht, ähm (.) oder zwei, ich weiß es nicht, sind da zu konzentrieren, ist wirklich gar nicht so wenig. (.) Vor allem, wenn man so eine Reflexion noch nie gemacht hat. (.) Also ich finde, der ist schon sehr umfangreich der (.) äh der die ja dieser Test. (.) Also ich für mich hätte es besser gepasst, hätte ich den in zwei Teilen machen können. Also hätte ich gesagt gut, erster Teil (.) eine dreiviertel Stunde, zweiter Teil eine dreiviertel Stunde, dann hätte ich länger diese Konzentration aufrecht erhalten können. Aber das wäre alles. #00:18:16-3#
82 I: (..) Okay. Dann danke. Dann mache ich mal Aufnahme aus. #00:18:21-8#

P.2.6 Interview 6: B6

1 I: Genau. Nochmal offiziell vielen Dank, dass du das Interview mit mir führst. (.) Ich brauche, wie eben angekündigt zuerst das Einverständnis für die Verwendung deiner Daten. Dazu lese ich die Einverständniserklärung einmal vor (.) und bitte dann danach um Einverständnis vorausgesetzt, du bist einverstanden. Äh die Einverständniserklärung lautet ich bin damit einverstanden, dass dieses Interview aufgezeichnet wird. Das Interview wird für eine Verschriftlichung, Verwendung oder Weitergabe anonymisiert. Die Einverständniserklärung zur Nutzung der erhobenen Daten kann jederzeit mit Wirkung für die Zukunft widerrufen werden. Bist du damit einverstanden? #00:00:33-3#
2 B6: Ich bin einverstanden, ja. #00:00:34-6#
3 I: Super. (.) Ähm (.) genau. Ihr habt vor einiger Zeit, vor ein paar Wochen inzwischen schon, diesen Reflexionstest bearbeitet und dann danach die Rückmeldung bekommen und (.) inzwischen vorletzte Woche auch mit diesem Onlinematerial gearbeitet. #00:00:48-8#
4 B6: Ja. #00:00:49-6#
5 I: Ähm genau. Mir geht es gleich darum oder ich würde gleich gerne ein bisschen mehr darüber erfahren, wie du die Rückmeldung wahrnimmst oder wahrgenommen hast (.) und auch das Onlinematerial. Es geht explizit nicht darum, dich zu bewerten oder deine Aussagen, sondern darum, Format, Gestaltung äh der Materialien zu verbessern. Das heißt insbesondere natürlich auch, dass das, was du hier sagst, überhaupt keinen Einfluss auf die Bewertung (.) in der Veranstaltung X hat. Also in keiner anderer Veranstaltung, aber- genau. #00:01:21-1#
6 B6: Ja. #00:01:21-7#
7 I: Ähm welche Fragen hast du noch bevor wir anfangen? #00:01:24-8#
8 B6: (..) Eigentlich kannst du so anfangen. #00:01:27-4#
9 I: Gut, dann machen wir das. Äh wir fangen vielleicht zum Einstieg einmal an mit deiner eigenen Rückmeldung. Kannst du dich an die noch erinnern? #00:01:35-3#
10 B6: (..) Was jetzt beim Test bei rausgekommen ist? #00:01:39-0#
11 I: Mhm[Ja]. #00:01:38-9#
12 B6: Ja. #00:01:39-5#
13 I: Weißt du noch, welche Schlüsse du daraus gezogen hast? #00:01:42-9#
14 B6: (..) Das ich noch äh Potenzial habe, mich zu verbessern, würde ich mal so beschreiben. (.) Aber ich hatte ehrlich gesagt auch (..) mir erschien der Test recht lang zum Schluss, fand ich und ich muss dann selbst eingestehen, was ich auch quasi jetzt für mich reflektiert dabei rausgefunden habe, dass ich die Lehrkraft, ich weiß jetzt nicht, ob das jetzt eine offizielle Lehrkraft war, ist ja auch egal, eigentlich mehr danach bewertet habe, wie er die Schüler behandelt hat und nicht, wie er seine Aufgaben gestaltet hat. Weil zum Schluss, muss ich ehrlich sagen, hat mich seine Art ein bisschen genervt und dann habe ich ein bisschen den Fokus verloren, muss ich ehrlich sagen beim bei dem Test. Und dann habe ich (.) ja wenn er beispielsweise gesagt hat, was mir so ein bisschen hängen geblieben ist, Marie an die Tafel, das ist dann so ein Aufhänger wo man darauf hängen bleibt und dann (.) (auf?) den Fokus von den Aufgaben, sage ich mal, die er gestaltet hat, verloren und dann (.) ja das hat sich so ein bisschen in der Reflexion wiedergespiegelt, fand ich. (.) Wobei ich das dann (.) die Fragen wurden ja recht genau aufgeschlüsselt dann in in der (.) in der Bewertung, da konnte man das dann eigentlich auch wiederfinden. #00:02:43-6#
15 I: Ähm in welcher Bewertung mein- oder welche Fragen wurden // aufgelistet? #00:02:48-2#
16 B6: In dem Feedback. (.) (unv.) #00:02:47-7#
17 I: Achso. In diese drei Bereiche meinst du. #00:02:51-7#
18 B6: Ja genau. Das ist ja recht (.) ja separat alles gewesen, man konnte alles genau nachlesen (.) oder diese ganzen Unterkategorien öffnen, sage ich mal, wo eigentlich jede Frage dann ja nachlesbar, wenn nicht jede, aber jeder Block. #00:03:05-3#
19 I: Ach in dem in diesem Onlinematerial dann. Mhm[Ja]. #00:03:08-7#

20 B6: Ja genau. #00:03:08-4#
21 I: Okay. (..) Gut dann habe ich hier einmal mitgebracht, damit es ein bisschen konkreter wird eine Beispielrückmeldung. Die ist natürlich vom Aufbau genau wie deine, ist aber nicht deine. #00:03:20-9#
22 B6: Mhm[Ja]. #00:03:21-4#
23 I: Ähm also soweit ich weiß. (.) Und genau, nimm dir ruhig einen Moment, das einmal anzugucken und dann (.) stelle ich meine Fragen. #00:03:31-5#
24 B6: (Ja?). (38s) Mhm ja ist noch soweit im Hinterkopf, glaube ich. So lange ist das ja noch nicht her. (12s) Ja. #00:04:27-8#
25 I: Okay. Ähm dann wäre meine erste Frage, ob du vorher schonmal mit solchen Boxplots gearbeitet hast. (.) Also mit dieser Diagrammart. #00:04:37-2#
26 B6: Für eine Reflexion oder? #00:04:38-7#
27 I: Überhaupt. #00:04:39-6#
28 B6: Nur in der Mathematik. #00:04:42-1#
29 I: Äh in welchem Kontext da? #00:04:44-8#
30 B6: Irgendwann im Abitur hatten wir das mal. Da hatten wir dann auch in der Statistik damit ja Werte dann eben dargestellt oder irgendwelche (.) Auswertungen eben graphisch visu- visualisiert. #00:04:56-9#
31 I: (..) Okay. #00:04:58-9#
32 B6: Also ganz fremd war es nicht. (.) Man kommt ja auch al- also die Erklärung ist ja noch daneben, man kommt ja eigentlich gut damit klar. #00:05:05-3#
33 I: Okay. (.) Ähm das zweite wäre also hier im oberen Plot ist ja die Effektstärke noch eingezeichnet. Hast du vorher mit der Effektstärke schonmal gearbeitet? #00:05:15-3#
34 B6: Nein. #00:05:16-5#
35 I: (.) Nein. Okay. Ähm (.) gut. Dann weg von diesen konkreten Begriffen wieder zur ganzen Rückmeldung. Wenn du die jetzt so anguckst, welche Schlüsse ziehst du aus dieser Rückmeldung? #00:05:29-1#
36 B6: (...) Die ist sehr umfangreich. (.) Ähm (.) was ich gut fand war eben jetzt zum Beispiel dieser Link, wenn man jetzt jetzt (.) sein Ergebnis hat und dann auch nicht damit zufrieden ist, dass man noch eben Alternativen hat, wo man nochmal (.) sich informieren kann, was man eben auch besser machen kann, oder wo man etwas finden kann, wie man es besser machen kann. (..) Dafür war der doch da, ne? #00:05:53-1#
37 I: // Ja. #00:05:54-4#
38 B6: Dass man eben Alternativen findet. Ja. (.) Und äh (..) wo ich mich, glaube ich, auch selbst ertappt habe mit den Schülervorstellungen, da bin ich nicht (.) genug drauf eingegangen, glaube ich. #00:06:06-0#
39 I: (...) Okay. Und welche Schlüsse würden Sie aus den Ergebnissen ziehen hier in dem also in dieser Rückmeldung? #00:06:14-1#
40 B6: (5s) Ja diese Rückmeldung ist ja sehr positiv erstmal zum Bewerten. #00:06:21-6#
41 I: (...) Warum? #00:06:24-7#
42 B6: (.) Weil das Ergebnis hier im oberen Mittel liegt. #00:06:28-5#
43 I: (..) Okay. #00:06:30-4#
44 B6: So hatte ich das jetzt verstanden. (.) Er hat ja jetzt irgendwie achtundsechtzig Prozent, das entspricht ja irgendwie dem Median, ne, ist es, (unv.) müsste doch, oder? Ja. #00:06:40-2#
45 I: (..) Okay. Und ähm (.) jetzt hatte ich eben schonmal auf diese Effektstärke angesprochen. (.) Welche Schlüsse würden sie aus der Information aus der Ef- aus den Informationen der Effektstärke ziehen (.) oder über die Effektstärke? #00:06:57-5#
46 B6: (12s) Wie ist denn diese Effektstärke (.) wie ist denn die gemeint, wie ist die definiert? #00:07:14-3#

47 I: (...) Ähm (.) ich würde vorschlagen, das klären wir später und ich nehme jetzt einfach hin, dass Ihnen die nichts sagt, okay? #00:07:23-4#
48 B6: Ja. #00:07:23-4#
49 I: Okay. Können wir aber gerne später noch klären. Ähm (.) dann (.) wäre die nächste Frage, welche Informationen der Rückmeldung insgesamt besonders hilfreich oder besonders interessant für Sie sind. #00:07:36-6#
50 B6: (7s) Äh für mich persönlich? (.) Ja was ich eingangs schon gesagt hatte und insbesondere die Schülervorstellungen. (..) (unv.) Seite (unv.). (.) Ähm ja eben dass man noch (.) also für mich selbst (da?) habe ich jetzt herausgefunden, dass ich noch ein bisschen daran arbeiten muss an meiner Reflex- Reflexionsfähigkeit (.) und dass man eben sich nicht verleiten lassen sollte, (.) die Situation zu objektiv betrachten. (4s) Gerade, wenn jetzt ja das war jetzt ja ein Beispiel, wo man eben eine ganze Stunde beurteilen soll und das ist halt auch ein bisschen länger und das ist ja nicht immer die Situation, dass man eben (.) fünf Minuten eine Situation beurteilen sollte, son- (.) oder soll, sondern dass es auch mal über einen längeren Zeitraum sein kann und dass man dann eben trotzdem noch fokussiert bleibt (.) und eben alle Aspekte eben berücksichtigt. (..) Das war für mich so das Wesentliche, (.) was ich daraus mitgenommen habe. #00:08:43-8#
51 I: Okay. Und sind andersum Inhalte der Rückmeldung überflüssig? #00:08:50-3#
52 B6: (...) Mhm (...) mir ist bis jetzt nichts aufgefallen, was mich jetzt gestört hätte oder so. #00:09:00-3#
53 I: (...) Okay. Und (.) welche Fragen bleiben für Sie offen? #00:09:06-8#
54 B6: (8s) Naja wenn ich mir jetzt (.) mir nochmal Gedanken machen würde und das Material nochmal durcharbeiten würde und den Test nochmal machen würde und es wäre vielleicht immernoch schlecht, dann würde ich mir doch Fragen stellen, warum das nicht besser geworden ist, aber (.) so kann ich jetzt erstmal nur (..) mich weiter informieren oder versuchen, reflektierter zu sein und dann hoffen, dass es besser wird. #00:09:34-8#
55 I: Mhm[Ja]. Und gibt es vielleicht Informationen, also zusätzliche Informationen, die Sie gerne gehabt hätten? #00:09:40-4#
56 B6: (5s) Mhm (4s) zum Test jetzt oder oder zu der Rückmeldung? #00:09:51-9#
57 I: Äh vorrangig zur Rückmeldung. Aber // ich nehme auch natürlich Infos zum Test. Also gerne auch. #00:09:53-5#
58 B6: Achso. (.) Achso ja. #00:09:54-3#
59 B6: (7s) Ich glaube nicht. Nein. Nein. #00:10:02-8#
60 I: Nein. (.) Okay. Dann wäre es das quasi ganz konkret zu dieser Rückmeldung. Dann würde ich weiter gehen zu dem Onlinematerial, also das, was Sie vor zwei Wochen in der Veranstaltung noch kurz (.) bearbeitet haben und äh genau hätte dazu noch folgende Fragen. Und zwar zuerst (.) ich frage zuerst mal waren Sie da und haben Sie das überhaupt mitgemacht? #00:10:23-9#
61 B6: Ja. #00:10:23-5#
62 I: Okay. Ähm (.) genau. Wie- Ach jetzt bin ich in das Sie gerutscht. Naja. // Wie hast du mit dem Material gearbeitet? #00:10:31-7#
63 B6: Macht nichts. #00:10:30-0#
64 B6: (.) Wir haben uns da angemeldet, dann habe ich mir das angeguckt. (.) Dann habe ich die Videosequenzen übersprungen, weil das erst zwei Wochen her war und ich das für nötig, sage ich mal, gehalten habe, weil das noch im Hinterkopf war. #00:10:44-5#
65 I: Diese einleitenden Videos // zur Lerngruppe und den Kram, oder? #00:10:48-6#
66 B6: Ja genau dann- #00:10:46-4#
67 B6: Die Videos, die den Unterricht nochmal beschrieben haben, // dass man weiß, wo man ist, in welchem Teil der Bewertung oder der Rückmeldung. #00:10:55-4#
68 I: Achso ja. #00:10:52-7#

69 B6: Ja und dann habe ich mir im wesentlichen die Unterpunkte angeguckt, die man immer weiter aufklappen konnte. (.) Da wurde ja dann auch sehr genau beschrieben, was jetzt (..) ja, ich sage mal, was da gefordert war oder was die mögliche beste Antwort war. (.) Ja. (.) Dann (.) nach dem spätestens nach dem dritten (.) Block, ich weiß nicht, wie das jetzt- (oder?) Gruppe war hat sich das ja auch recht schnell wiederholt immer die besonders die Anfangs- (.) fragen, die man immer wieder aufklappen musste. (..) Das hat mich jetzt nicht gestört, ist halt vom Umfang her okay, also (.) man weiß halt dann, dass die Frage da kommt, dann klappt man sie nicht immer wieder auf. (.) Dann guckt man sich eher nur die möglichen Lösungsvorschläge an. (.) Ähm (.) ich hatte aber nicht immer im Hinterkopf, was ich jetzt (.) da als Antwort gegeben habe. Ich weiß nicht, ob das da irgendwo stand und ich das übersehen habe, oder (.) also man hat dann die Situation hatte man im Hinterkopf, aber ich wusste halt nicht mehr genau, (.) habe ich das jetzt so gelöst, (.) weil alle Details wusste man ja auch nicht mehr. #00:11:57-1#
70 I: Ja. #00:11:58-5#
71 B6: Das ist hätte ich vielleicht noch so (..) ich weiß nicht, ob man das in der Bewertung, wenn das dann anonymisiert ist, ob man das dann trotzdem daneben aufrufen kann irgendwie, dass (.) damit man sieht, was man selber geantwortet hat eben. #00:12:10-8#
72 I: (..) Okay. (.) Ähm und was nimmst du aus der Arbeit mit dem Material mit? #00:12:18-2#
73 B6: (...) Mit den? #00:12:21-1#
74 I: (.) Also aus aus diesem Onlinematerial mit aus der Arbeit damit quasi. #00:12:26-0#
75 B6: (...) Also ich muss sagen, wir hatten ja nur, nur in Anführungszeichen, zwanzig Minuten Zeit. (..) Ich habe mich da erstmal (.) den ersten Block habe ich recht umfangreich gelesen und dann habe ich gemerkt, dass sind noch zwei Kategorien, die eigentlich ähnlich aufgebaut sind. Dann habe ich mich da ein bisschen schneller durchgeklickt. (..) Mhm (..) ja wenn ich jetzt beispielsweise nochmal in den Unterricht gehe, würde ich mir das vielleicht nochmal genauer angucken, um zu gucken eben, wie man auf bestimmte Sachverhalte am besten reagiert oder wie man auf Vorstellungen eingeht (..) (so?) dafür. Also der Link ist ja noch verfügbar, oder? Kann ich ja (noch?) aufrufen. #00:13:02-4#
76 I: Ja der bleibt verfügbar. #00:13:02-8#
77 B6: Ja dann könnte ich darauf nochmal zurückgreifen. Das würde ich auf jeden Fall mitnehmen, dass man nochmal die Möglichkeit hat, (.) eben darauf zurückzugreifen. #00:13:09-5#
78 I: Ja. Okay. Und ähm bietet das Material aus deiner Perspektive dann auch einen Mehrwert und warum? #00:13:17-5#
79 B6: Also einen Mehrwert auf jeden Fall. Hätte ich jetzt diesen Test gemacht und hätte jetzt keine Rückmeldung bekommen, das wäre ja eigentlich (.) dann wäre es eben nur was für die Wissenschaft gewesen. Ja und jetzt habe ich da kann ich ja persönlich auch noch was draus mitnehmen. #00:13:30-3#
80 I: (..) Okay auch aus dem aus diesem Onlinematerial. #00:13:34-9#
81 B6: (..) Ja wenn ich jetzt ge- spezielle Situationen mitnehmen wollte, dann müsste ich halt nochmal nachgucken, was ich jetzt geantwortet habe. Aber sonst könnte ich halt allgemein gucken, in welche Richtung das gehen soll. #00:13:47-9#
82 I: Ja. (.) Ähm (.) bei diesem Onlinematerial, welche Inhalte waren da besonders hilfreich oder besonders interessant (.) für dich? #00:14:00-0#
83 B6: (4s) besonders hilfreich im Wesentlichen die die (.) Lösungsvorschläge dann, also wie die Situation eben aufgelöst hätte werden sollen oder im besten Fall ja. (.) (Dem?) (.) Der Rest eigentlich, die Videos waren noch im Hinterkopf und diese (.) ich weiß jetzt nicht mehr genau wie das hieß, am Anfang diese Buttons, die man ja aufklappen sollte, ja die waren halt da und die hat man dann irgendwann nicht mehr angeguckt, weil die Fragen sich ja dann wiederholten. #00:14:28-8#

84 I: Ja. (.) Ähm gab es in diesem Onlinematerial Inhalte, die aus deiner Perspektive überflüssig waren oder sind? #00:14:35-8#
85 B6: Ja. Das, was ich gerade erwähnt habe, also dass sich das gedoppelt hat dann. #00:14:39-0#
86 I: Ja. #00:14:39-5#
87 B6: Also es ni- also es ist gut gestaltet, dass es nicht alles immer aufgeklappt ist, dass man quasi runterscrollen muss, dass man das optional aufklappen kann, aber (..) ja, für mich war es überflüssig, aber man könnte es auch so lassen (.) im Prinzip. Weil es halt eingeklappt war, (hat?) es halt nicht gestört dann. #00:14:55-3#
88 I: (..) Okay. Und bleiben aus der Arbeit mit diesem Material Fragen für dich offen? #00:15:00-9#
89 B6: (...) Bis jetzt sind keine aufgekommen, nein.. #00:15:05-2#
90 I: Nein. Okay. Ähm dann wäre ich soweit durch mit den Fragen. Ich habe mir aufgeschrieben, dass wir gleich noch über die Effekstärke sprechen, das machen wir auch. Aber äh hast du sonst noch irgendwie Fragen, Anmerkungen zum Test, zur Rückmeldung, zum Onlinematerial? #00:15:20-8#
91 B6: Das war eine gestellte Unterrichtssituation, oder? #00:15:23-6#
92 I: Ja. #00:15:24-7#
93 B6: Achso. Ja. Das war es eigentlich. #00:15:26-2#
94 I: Ja das ist genau das ist gescripted, ähm- #00:15:30-1#
95 B6: Ja. #00:15:30-6#
96 I: Damit man die Punkte findet, die man finden- also die Punkte darstellen kann, die gefunden werden sollen und andere eben nicht noch also zusätzliche so Störfaktoren in Anführungszeichen mit drin sind. // Ja. #00:15:41-8#
97 B6: Ja. #00:15:41-8#
98 I: Okay. #00:15:42-7#
99 B6: Nein, sonst zum Test eigentlich- #00:15:44-6#
100 I: Du hattest ähm (.) bei dieser Frage, welche weiteren Infos man gerne gehabt vorhin noch was zum Test gesagt. War das das die Frage, ob es gestellt ist, oder- #00:15:53-5#
101 B6: Ich glaube schon, ja. #00:15:54-8#
102 I: Ja. Okay. Nur, dass da nichts untergeht. Ähm genau dann die Effektstärke. (.) Äh die ist zu verstehen (.) quasi als Maß ja für die Unterschiede. Also es gibt ja hier diesen den in blau den Mittelwert eingezeichnet, in rot das Ergebnis und die Effektstärke ist quasi einfach ein Maß, was angibt, wie stark ein Unterschied zu werten ist. #00:16:18-0#
103 B6: Mhm[Ja]. #00:16:19-0#
104 I: Das ist so die Idee dahinter. Ähm (..) mir ist aber auch klar, dass viele von euch da vorher noch nicht mit gearbeitet haben, deswegen frage ich das nämlich hier explizit auch noch nach, ob das überhaupt Sinn ergibt, (.) diese Effektstärke da drin aufzuschlüsseln. #00:16:36-4#
105 B6: Also ich hatte das jetzt bei meiner Bewertung gelesen, aber dann eher überlesen, sage ich mal. (.) Ja auch aus dem Grund, weil ich das vielleicht irgendwann mal in Mathe gehört hatte, aber dann auch jetzt ni- damit nichts mehr anfangen kann. (.) Ja, ich hätte dann halt diese Einordnung fand ich ganz nett, dann wusste man ungefähr, wo man steht. #00:16:54-4#
106 I: Mhm[Ja]. #00:16:55-2#
107 B6: Ja. Aber mit der Effektstärke konnte ich jetzt nichts anfangen. #00:16:58-4#
108 I: Okay. (.) Ähm gut. Dann bin ich durch. Gibt es von dir inzwischen noch was Neues? #00:17:04-3#
109 B6: (.) Ich wüsste nichts, nein. #00:17:07-0#
110 I: Okay. Dann mache ich mal die Aufnahme aus. #00:17:09-0#

P.3 Interviews zur 2. Evaluation (Kohorte C)

P.3.1 Interview 1: C1

1. I: So. Okay. Genau, ich habe eben schon gesagt, ich ähm- ne, erstmal danke, das du mitmachst. Vielen Dank. Ähm ich habe eben schon gesagt, ich muss zuerst einmal das Einverständnis einholen. Ich lese einmal vor und dann sagst du, ob du einverstanden // bist. #00:00:13-6#
2. C1: Ja. #00:00:13-7#
3. I: Ich bin damit einverstanden, dass das Interview aufgezeichnet wird. Es wird für eine Verschriftlichung, Verwendung oder Weitergabe anonymisiert. Die Einverständniserklärung zur Nutzung der erhobenen Daten kann jederzeit mit Wirkung für die Zukunft widerrufen werden. Bist du damit einverstanden? #00:00:29-9#
4. C1: Ja, das bin ich. #00:00:30-9#
5. I: Super. Genau, also du hast, wir haben eben schon darüber gesprochen, du hast vor ein paar Wochen den Reflexionstest bearbeitet und eine Rückmeldung bekommen und ich würde gern bisschen was darüber erfahren, wie du die wahrgenommen hast oder wahrnimmst. Ich habe es eben auch schon gesagt, es geht überhaupt nicht darum, dich irgendwie zu bewerten, sondern für mich darum, das Material zu verbessern. Es gibt auch keinen Einfluss auf irgendeine Bewertung, Benotung. Genau. Hast du noch eine Frage, irgendwas vorweg? (C1 schüttelt den Kopf) Ne. // Okay. #00:00:56-7#
6. C1: Nein. #00:00:55-9#
7. I: Äh dann würde ich zum Anfang einmal zu deiner eigenen Rückmeldung fragen, kannst du dich noch dran erinnern an die Rückmeldung, die du bekommen hast? #00:01:04-7#
8. C1: Ja, das kann ich. #00:01:05-8#
9. I: Ähm weißt du, welche Schlüsse du daraus gezogen hast? #00:01:09-2#
10. C1: Ja, das weiß ich. Also ich ähm (...) habe verstanden, dass ich noch mehr an ähm Schülervorstellungen arbeiten muss äh und an der (.) äh (.) an Bewertungen und äh (.) einigermaßen gut ausgefallen bei mir ist äh sind Alternativen, also Vorschl- äh Vorschläge unterbreiten äh für Alternativen, das war okay, aber ansonsten Schülervorstellungen und Bewertungen, das das war eben das äh im Mittleren äh ganz unten in dem äh in dem in der Box // das heißt. nicht so gut. #00:01:48-8#
11. I: Mhm(ja) (..) alles klar. (..) Okay. Genau, ich habe hier zur Erinnerung äh einfach nochmal- das ist jetzt ein Beispiel, das ist nicht deine, // von irgendwem einfach nur zur- äh ach so sorum, ne, äh genau, kannst du, wenn du willst reingucken, nur zur Erinnerung. (.) Vielleicht schaust du einmal kurz ähm einmal kurz auf die Ergebnisse, ich mache das mal einmal auf. (.) So. (.) Ähm erstmal eine Frage, hast du davor schonmal mit solchen Boxplots gearbeitet? #00:02:18-6#
12. C1: Okay. #00:01:56-3#
13. C1: Mhm ja, das kenne ich noch ähm aus dem Physikdidaktikunterricht bei dem Herrn O. #00:02:26-9#
14. I: Mhm(ja). #00:02:28-0#
15. C1: Also wir haben das äh (.) da gewisse Statistiken ausgewertet, äh was Schülervorstellungen angeht und äh ich kann mich erinnern, dass wir äh auch äh solche Boxplots äh ich glaube auch in Bezug auf PISA-Studien und so weiter (.) äh besprochen haben. #00:02:44-3#
16. I: Alles klar. (..) Ähm und wenn du dir jetzt diese einfach als Beispiel mal // diese Rückmeldung, diese Ergebnisse anschaust, welche Schlüsse würdest du da draus ziehen? #00:02:52-6#
17. C1: Ja. #00:02:49-2#
18. C1: Mhm (5s) ähm Bezug worauf? Ähm auf äh darauf, wo ich mich verbessern // müsste? #00:03:05-6#

19. I: Genau, also wenn du dir jetzt vorstellen würdest, dass wäre deine Rückmeldung. #00:03:07-9#
20. C1: Ja. #00:03:08-5#
21. I: Was würdest du dann daraus // mitnehmen so für dich? Was wäre deine Schlussfolgerung? #00:03:13-3#
22. C1: Okay. #00:03:10-8#
23. C1: (...) Mhm ähm (4s) dass ich äh unterhalb des äh Medians bin, also unterhalb des Durchschnitts (.) ähm (..) dass ich in zwei Bereichen ähm (...) schlecht abgeschnitten habe äh und in einem Bereich einigermaßen mittelmäßig, also wenn man sich das hier anschaut (C1 zeigt auf Teilergebnis zu Schülervorstellungen). #00:03:44-9#
24. I: Also bei // den Schülervorstellungen mittelmäßig, meinst du, oder-? #00:03:48-8#
25. C1: Äh das ist- genau bei- (..) Also für mich ist es äh sieht es so aus, als äh eigentlich- ist es wirklich meins? Also (weil?) // ich meine micht zu (erinnern?)- Das ist nicht meins, genau, okay. #00:03:55-0#
26. I: Nein, das ist nicht nicht nicht deins, das ist ein Beispiel. #00:03:57-2#
27. C1: Alles klar, weil bei mir war das umgekehrt äh- // ähm genau- #00:04:01-0#
28. I: Ich weiß ja gar nicht, was deins ist, ist ja anonym. #00:04:01-4#
29. C1: Ja, also ich würde sagen, ähm hier die Schülervorstellungen, das ist ja im oberen Bereich. #00:04:06-9#
30. I: Mhm(ja). #00:04:07-3#
31. C1: Äh das äh ähm der fünfzig Prozent sozusagen und äh die Alternativen und die Bewertung, also diese beiden Aspekte, also die Bewertung sind unterhalb der fünfzig Prozent der Ergebnisse (.) und die Alternativen sind ganz unten, also äh (..) ich würde sagen, dass in den beiden Bereichen muss man sich verbessern auf jeden Fall und der der ist eigentlich einigermaßen gut ausgefallen. #00:04:37-1#
32. I: Okay. (.) Und wenn // du jetzt quasi die Rückmeldung so insgesamt nimmst, welche Informationen sind da jetzt besonders interessant oder besonders hilfreich? #00:04:47-2#
33. C1: Ja. #00:04:40-9# #00:04:47-4#
34. C1: (..) Äh Rückmeldung welche besonders interessant, (besonders?) hilfreich- (.) Ähm (...) dass ähm Reflexion äh wichtig ist, um den eigenen Unterricht zu verbessern. #00:05:04-3#
35. I: Mhm(ja). #00:05:05-0#
36. C1: Dass man äh (7s) dass vor allem auch Schülervorstellungen ähm (...) ähm wichtig sind, um ähm (..) die ähm (.) Unterrichtsgestaltung ähm entsprechend äh (.) den ähm Vorkenntnissen der Schüler besser ähm ähm besser umsetzen zu können (..) ähm, dass die ähm (.) also vor allem, dass es überhaupt diese drei Bereiche wichtig sind, also Bewertung, Schülervorstellung als Beschreibung äh Bewertung und und Alternativen, also das sind sozusagen ähm (..) die sind ja nicht einfach so entstanden, sondern ich glaube, du beziehst dich da auf Nowak // meine ich? #00:06:03-2#
37. I: Ja. #00:06:03-1#
38. C1: Genau und äh (...) ja was gibt es denn noch ähm o- kannst du die Frage noch ein bisschen präzisieren, also ähm (.) was meinst du jetzt, was besonders wichtig bei Rückmeldungen // ähm- #00:06:18-7#
39. I: Ne- was für dich jetzt, wenn du dir das- wenn du die Rückmeldung bekommen hast, was davon war für dich- das ist ja einfach mehrere Seiten- #00:06:25-7#
40. C1: Okay. #00:06:26-3#
41. I: Einfach was davon für dich- und wenn du sagst, die drei Bereiche, dann ist das auch in Ordnung. #00:06:30-0#
42. C1: Ist in Ordnung. #00:06:30-6#
43. I: Ja. // Es gibt hier kein richtig und falsch. #00:06:33-0#

44. C1: Okay. (.) Alles klar. #00:06:33-1#
45. I: (Okay?). Gibt es denn anders rum was, wo du sagst, das ist überflüssig an Information oder so, irgendwas? #00:06:40-1#
46. C1: Mhm (.) würde ich so nicht sagen. Ähm (..) was (kann?) überflüssig? (8s) Kann ich nichts äh ist mir nichts aufgefallen. #00:06:58-0#
47. I: Okay. (.) Äh und gibt es Fragen, die für die offen geblieben sind, Sachen, die unklar sind noch? #00:07:04-0#
48. C1: Ähm (..) ja ich wollte fragen, also man sollte ja nicht so oft zurückschauen, das heißt, sich das Video nicht so oft // nochmal anschauen. #00:07:16-0#
49. I: Mhm(ja). #00:07:15-6#
50. C1: Äh ist es denn so gedacht, dass man äh auch im richtigen Unterricht äh viele Dinge auch nicht äh so richtig äh protokollieren kann und dann sich äh auf das eigene Gedächtnis verlassen muss äh hat es diesen Hintergrund oder äh wa was was hast du dir dabei gedacht, // (I: Mhm(ja).) äh als du dir äh diese diesen Hinweis gegeben hast, dass man nicht so oft sich das Video nochmal anschauen soll. #00:07:40-5#
51. I: Ja, also das hat ein- das ist natürlich so, wenn ich mir Unterricht anschaue, kann ich nicht alles immer alles sehen und äh das ist nicht realistisch für später- also für für ne reale Situation in der Schule, dass man dann eben ganz ganz viel Zeit hat und ähm genau das ist natürlich die eine Komponente und dann gibt es noch einen ganz platten Grund, der ist, dass das einfach sehr lange dauert und äh wenn man- also dass dass man diese realistische Situation so ein bisschen nachstellt und nicht, wenn man jetzt jedes Video mehrmals guckt, dann dauert dieses äh diese Lernumgebung unglaublich lange äh und dass das dann einfach auch nicht zu lange wird, das so in // Kombination, genau. #00:08:22-8#
52. C1: Okay (.) alles klar. #00:08:23-1#
53. I: Aber das ist eben genau das, also in der Schule oder wenn du jetzt auch im Praktikum zum Beispiel Nachbesprechungen mit deinem Mentor machst, dann habt ihr ja auch kein Video, was ihr immer anschauen könnt. Ja. #00:08:32-8#
54. C1: Okay. #00:08:33-6#
55. I: Genau. Äh ja gibt es sonst noch weitere Informationen auch zur Rückmeldung, die du gern gehabt hättest oder auch Sachen, die in der Rückmeldung fehlen für dich? #00:08:43-0#
56. C1: Mhm (...) kann ich nicht sagen // (unv.) habe ich nicht #00:08:48-2#
57. I: Ne. Okay. (.) Ähm dann ist hier in dieser Rückmeldung ist ja so ein Material verlinkt. Hattest du da mal reingeschaut? #00:08:56-5#
58. C1: Ja, das habe ich. Ähm d- wir haben ja eine E-Mail von Frau äh P // bekommen. #00:09:01-4#
59. I: Ja. #00:09:01-8#
60. C1: Äh ich bin aber nicht ganz fertig geworden, ich fand es sehr hilfreich, dass man das auch mit äh von einem Smartphone aus machen konnte. #00:09:07-7#
61. I: Mhm(ja). #00:09:08-3#
62. C1: Ähm ich äh wie weit bin ich denn gekommen? Also ich habe mir die Literaturhinweise angeschaut und äh die- ich glaube, das waren die wichtigen äh Fragen, die man sich stellen sollte äh // bei der- #00:09:20-5#
63. I: So Leitfragen, ne? Ja. #00:09:21-1#
64. C1: Genau, die Leitfragen, aber ich bin noch nicht so weit, dass ich das äh mit der Bearbeitung fertig bin. #00:09:26-1#
65. I: Mhm(ja). #00:09:26-6#
66. C1: Genau. #00:09:27-7#
67. I: Okay, also das heißt, du hast es erstmal geöffnet und ein bisschen // gelesen so. Ja. #00:09:32-0#

68. C1: Äh ich habe es geöffnet, ich habe es ein bisschen äh genau überflogen, aber ich bin noch nicht fertig, also ich hab fest vor äh mir fest vorgenommen, da fertig zu werden. #00:09:37-7#
69. I: Mhm(ja). Ähm okay das ist jetzt natürlich wir // (C1: Ja.) sprechen jetzt nur über das, was du schon gesehen hast, über den Rest // (C1: Okay.) kannst du ja natürlich nichts sagen. Ähm was nimmst du daraus mit für dich? #00:09:49-1#
70. C1: Ähm. #00:09:51-8#
71. I: Oder kannst du da überhaupt schon was zu sagen? Wenn nicht, dann ist auch okay. #00:09:56-5#
72. C1: (5s) Im Moment kann ich dazu nichts // sagen. #00:10:02-9#
73. I: Okay. (.) Ähm und würdest du sagen, dass es einen Mehrwert hat, also ein- (.) ja etwas positives bietet, das Material? #00:10:10-5#
74. C1: Ähm ja, also der Mehrwert besteht darin, dass man ähm äh dass da bereits ähm Literaturhinweise bestehen, das heißt, man weiß, äh wo man nachschlagen könnte Bezug auf die- auf Reflexion. Ähm wenn ich das so richtig verstanden habe, sind die Leitfragen der Literatur entnommen, das heißt, ähm ähm man kann sich an den Leitfragen orientieren bei der Reflexion, das ist doch auch ein Mehrwert. (.) Und äh (..) äh ja ansonsten ähm (..) die Auswertung an sich mit den äh mit den drei Bereichen äh man hat ja sozusagen, es wird einem aufgezeigt, wo (.) sagen mal so, woran man äh besonders arbeiten sollte, äh um sich zu verbessern, auch das ist ein Mehrwert. (.) Ähm und das- das- da gab es ja ein ein Hinweis, also Sie sollten dann an den in dem in dem Bereich arbeiten // und äh (.) dann eben diesen Link und mit Literaturhinweisen und so weiter. Also das sind alles Dinge, die einen Mehrwert haben, finde ich. #00:11:20-8#
75. I: Mhm(Ja). #00:11:12-4#
76. I: Okay. Und aus dem Bereich, wo du jetzt schon in dieses // Material geschaut hast, gibt es da Sachn, die du irgendwie für besonders hilfreich empfindest oder auch andersrum Sachen, die vielleicht überflüssig sind oder sowas? #00:11:33-9# #00:11:45-3#
77. C1: Ja. #00:11:25-0#
78. C1: Ähm ich habe die Literatur noch nicht komplett äh durchgelesen, ähm also da kann ich- (..) noch nichts sagen, ich habe das nur überflogen, also äh zur zu genuaen Inhalten kann ich äh noch nichts sagen. #00:11:51-4#
79. I: Okay. Und gibt es noch was- also gibt es Fragen, die für dich noch offen sind zu dem Material? #00:11:57-0#
80. C1: Mhm nein. #00:11:59-6#
81. I: Ne. #00:12:00-5#
82. C1: Noch nicht. #00:12:01-2#
83. I: Okay. (.) Falls da nochmal was aufkommt, // ne, du kannst dich immer auch melden, aber wie es dir passt, genau. #00:12:07-6#
84. C1: Okay. Ja, alles klar. #00:12:05-6#
85. I: Gut. Dann wäre ich durch mit meinen Fragen. Hast du sonst noch irgendwie Anmerkungen, Kommentare, Fragen, fällt dir noch was ein, was du sagen möchtest? #00:12:16-5#
86. C1: Ähm (...) Nein. (..) Alles okay. Also ich ähm äh wen ich fertig bin mit der Reflexion, dann äh ähm habe ich kann ich vielleicht qualitativ bessere Fragen stellen, also im Moment kann ich äh keine guten Fragen formulieren, deshalb lasse ich das lieber. #00:12:34-3#
87. I: Äh // kannst auch- #00:12:35-7#
88. C1: Also nein, nein, nein, also ich äh muss mich erstmal einarbeiten äh und dann schauen, ähm, also ich habe die Literatur ja noch gar nicht durchgesehen, äh deswegen kann ich auch keine Fragen formulieren. #00:12:46-2#
89. I: Okay. Dann mache ich einmal schonmal aus. So. #00:12:52-2#

P.3.2 Interview 2: C2

1. I: So, okay. Ähm dann danke, dass du das mit mir machst hier, das Interview. Ich brauche zunächst das Einverständnis für die Verwendung deiner Daten, dafür lese ich nochmel die Einverständniserklärung vor und bitte danach mündlich um Einverständnis. (.) Die Einverständniserklärung lautet (.) ich bin damit einverstanden, dass dieses Interview aufgezeichnet wird. Das Interview wird für eine Verschriftlichung, Verwendung oder Weitergabe anonymisiert. Diese Einverständniserklärung zur Nutzung der erhobenen Daten kann jederzeit mit Wirkung für die Zukunft widerrufen werden. Bist du damit einverstanden? #00:00:33-6#
2. C2: Das bin ich. #00:00:34-7#
3. I: Super. (.) Genau. Du hast vor einer Zeit diesen Reflexionstest jetzt zum, ich meine, zweiten Mal bearbeitet? #00:00:41-3#
4. C2: Zum zweiten Mal, ja. #00:00:42-0#
5. I: Mhm(Ja) und dann auch eben wieder eine Rückmeldung bekommen. Ich würde gleich gerne eben ein bisschen was darüber erfahren, wie du die jetzt wahrgenommen hast oder wahrnimmst, es geht, ich habe es eben schon gesagt, überhaupt nicht darum, dich zu bewerten, sondern eben das Material zu verbessern, das ist völlig unabhängig von einer Bewertung hier in der Veranstaltung. (.) Irgendwelche Fragen von dir? #00:01:01-7#
6. C2: Ne. #00:01:02-2#
7. I: Ne. Okay. (.) Dann würde ich einmal nochmal zum kurzen Recap mit deiner Rückmeldung anfangen, kannst du dich an die noch erinnern? #00:01:09-9#
8. C2: Meine Rückmeldung zu dem Test? #00:01:12-9#
9. I: Mhm(Ja). #00:01:12-9#
10. C2: Zu dem Quiz? (..) War beim Ankreuzen positiv, aber ich habe nichts besonderes dazu geschrieben #00:01:19-8#
11. I: Okay. (.) Und ähm die Rückmeldung, die du jetzt bekommen hast, hast du da- hattest du da schon reingeschaut? #00:01:27-6#
12. C2: Ach so, ja, hatte ich. #00:01:28-5#
13. I: Und welche Schlussfolgerungen hast du daraus gezogen? #00:01:31-7#
14. C2: Mhm ich war zufrieden mit mir beziehungsweise froh, dass ich mich verbessert hatte, weil beim ersten Mal, wo ich diesen Test gemacht hatte, hatte ich, glaube ich, im Schnitt lag ich bei fünfzig Prozent und diesmal lag ich bei fünfundsiebzig, meine ich. #00:01:47-1#
15. I: Mhm(Ja). (.) Ähm dann, du hast es eben schon so ein bisschen angesprochen, (.) inwiefern hat jetzt die Beschäftigung mit- also oder inwiefern hast du jetzt, wenn du dir die Rückmeldung angeschaut hast, deine Rückmeldung oder die Bearbeitung vom letzten Semester berücksichtigt? #00:02:02-9#
16. C2: (...) Ich würde sagen, fast gar nicht. Also ich habe jetzt nicht gedacht, okay, das habe ich beim letzten Mal so und so gemacht, das mache ich jetzt anders, sondern ich bin sehr neutral dran gegangen, würde ich sagen. #00:02:13-5#
17. I: Beim- bei der Bearbeitung von dem-? #00:02:15-2#
18. C2: Ja. #00:02:15-7#
19. I: Ja. Okay. Und jetzt als du die Rückmeldung bekommen hast, da hast du dann die alte nochmal rausgeholt, oder wie hast // du das-? #00:02:22-1#
20. C2: Nein. #00:02:21-8#
21. I: Nö. Also has- hattest du noch im Kopf. #00:02:23-7#
22. C2: Ja. #00:02:24-1#
23. I: Ah, okay. #00:02:24-8#
24. C2: Äh oder wie meinst du das? (.) Die alte Rückmeldung? #00:02:28-5#
25. I: Ja. (.) Hattest du die- oder kannst du dich an die- also hast du die irgendwie // berücksichtigt? #00:02:33-7#

26. C2: Nein, gar nicht. #00:02:33-9#
27. I: Ne, die ist einfach weg. Okay. Auch gut. #00:02:36-6#
28. C2: Ich weiß nur noch die Prozentzahl. #00:02:38-2#
29. I: Ja, okay. Also von diesem- von diesem Gesamt- // (.) Ja. #00:02:42-1#
30. C2: Overall, genau, waren das ja einundfünzig Prozent, glaube ich, bei mir und diesmal halt mehr. Also // das weiß ich noch. #00:02:46-7#
31. I: Ja. (.) Ja, okay. Super. Und dann hast du quasi dich an die Zahl erinnert, die Zahl // verglichen und nicht detailliert. #00:02:52-2#
32. C2: Genau, ja. Die Prozentzahl. Ja, richtig. Ja. #00:02:52-7#
33. I: Ja, okay, dann habe ich es richtig verstanden. Super. (.) Okay. Ich habe, damit wir wieder irgendwas haben, worüber wir sprechen, natürlich wieder eine Beispielrückmeldung mit. (.) Ähm (.) ich habe es letztes Mal schon gefragt, ich frage es jetzt nochmal, hast du schonmal mit solchen Boxplots gearbeitet. #00:03:10-1#
34. C2: (..) Naja, beim letzten Mal halt // einmal kurz. #00:03:14-5#
35. I: Aber nicht in anderen- #00:03:15-4#
36. C2: Bisher noch nicht, // nein. #00:03:16-7#
37. I: Ne, okay. Ähm genau, wenn du dir die Rückmeldung anschaust, welche Schlussfolgerungen würdest du aus dieser Rückmeldung ziehen? #00:03:22-5#
38. C2: (7s) Ich würde sagen, dass die Person sich in- (..) als- als- aus Sicht der Person, oder-? #00:03:36-5#
39. I: Ja, also (mit?) wenn du dir vorstellst, es wäre deine. #00:03:39-3#
40. C2: Achso (unv.) #00:03:40-0#
41. I: Ich habe es jetzt nur als Beispiel, damit man die Anonymität // beibehält. #00:03:43-7#
42. C2: Ähm das würde ich dann schade finden, (.) weil ich, ähm ja, unter den fünfundzwanzig Prozent der Leute bin, die (.) die Aufgaben am schlechtesten ausgeführt hat, also abgegeben hat (.) und äh ich würde mich verbessern wollen. #00:03:58-3#
43. I: Mhm(Ja). #00:03:59-4#
44. C2: Genau, aber dazu muss man sagen, fällt mir gerade auf, äh man weiß ja gar nicht, wie die anderen abgeschnitten haben. Es kann (ja?) auch sein, dass alle sehr gut abgeschnitten haben, oder? #00:04:07-7#
45. I: // Joa. #00:04:08-9#
46. C2: Ja. Das ist meine Meinung dazu. #00:04:11-1#
47. I: Okay. Genau und das würde dann für dich einen Einfluss machen, also // wenn alle sehr gut, oder- #00:04:16-2#
48. C2: Ja. (..) Doch, bestimmt, also ich würde mir (ja?) schon überlegen, was- wieso habe ich es halt nicht geschafft. (.) Doch, bestimmt, ja. #00:04:22-4#
49. I: Okay. (.) Und ähm (.) welche Infos oder ja- Informationen oder Inhalte aus dieser Rückmeldung sind besonders hilfreich oder besonders interessant für dich? #00:04:34-0#
50. C2: (4s) Ähm nur aus dieser Grafik hier oder // auch insgesamt? #00:04:41-0#
51. I: Nö, insgesamt jetzt diese ganzen Zettel, gibt es da irgendwas, was eben besonders relevant raussticht? #00:04:46-6#
52. C2: Oh, Abbildung zwei ist ganz interessant für mich. (.) Die stellt (dann?) auch nochmal die äh (.) einzelnen Stärken sozusagen, in welchen Bereichen man gut abgeschnitten hat und in welchen man äh nicht so gut abgeschnitten hat äh dar, das finde ich ganz hilfreich. #00:05:01-7#
53. I: Mhm(Ja). Und äh gibt es andersrum was, was du überflüssig findest? #00:05:06-4#
54. C2: (3s) Nö. #00:05:10-7#
55. I: Ne, okay. Gibt es, ja, Fragen, die für dich offen geblieben sind, oder Informationen, die du zusätzlich gerne gehabt hättest? #00:05:17-3#
56. C2: (.) Mhm ich finde, dass (.) die wichtigsten Informationen dabei sind. #00:05:22-9#

57. I: (..) Okay. (.) Ähm genau in dieser Rückmeldung ist ja auch so ein äh Material verlinkt. (..) Wenn- also du hast die Rückmeldung irgendwie, glaube ich, am Wochenende bekommen. #00:05:35-4#
58. C2: Ja, // sehr früh schon. #00:05:37-0#
59. I: Ähm (.) genau, hattest du da nochmal reingeschaut bis jetzt quasi? #00:05:40-7#
60. C2: Nein, bisher noch nicht. #00:05:41-7#
61. I: Okay. #00:05:42-2#
62. C2: Ich hatte nur den, die ganzen Ergebnisse angeschaut. #00:05:44-4#
63. I: Mhm(Ja). Ähm was mich da jetzt interessieren würde ist, ist das quasi eine (..) bewusste Entscheidung, das zu- dafür beziehungsweise dagegen, das Fördermaterial anzuschauen oder planst du, noch reinzuschauen, oder- also hast du da- genau, hast du da irgendwie drüber nachgedacht? #00:06:03-5#
64. C2: (.) Nein, momentan nicht. Also ich habe mir- es hat mich schon interessiert, wie ich abgeschnitten habe, #00:06:09-1#
65. I: Mhm(Ja). #00:06:09-6#
66. C2: aber die Fördermaterialien hatte ich mir persönlich jetzt nicht angeschaut oder nicht vor, anzuschauen. #00:06:15-9#
67. I: Okay. Hat das irgendeinen konkreten Grund? #00:06:18-5#
68. C2: Das hat keinen Grund bezüglich der Qualität der Materialien oder so oder der Seite, ähm ich habe nur gerade viel zu tun, deswegen. #00:06:26-7#
69. I: Ist (auch?) kein Problem, also- #00:06:28-6#
70. C2: Ne, ne, (äh?) es hat nichts mit der Seite zu tun. #00:06:30-1#
71. I: Okay. Wenn ich mich richtig erinnere, kennst du das Material ja auch schon so ein bisschen, weil ihr im letzten Semester reingeschaut habt. #00:06:37-0#
72. C2: Ja. #00:06:37-6#
73. I: (.) Ähm hat jetzt die Tatsache, dass du das Material schon kennst, hat das deine Entscheidung irgendwie beeinflusst? #00:06:43-0#
74. C2: Nein. #00:06:43-7#
75. I: Das (nicht?). Okay. (.) Ähm gibt es irgendwelche Rahmenbedingungen, eine bestimmte Situation, die du benötigen würdest, damit du in das Material schaust, oder was was müssten so die Rahmenbedingungen sein, damit du das machen // würdest? #00:06:57-2#
76. C2: Mehr Zeit (.) hauptsächlich. #00:06:58-8#
77. I: Okay. #00:06:59-4#
78. C2: Und dann hätte ich auch bestimmt das mehr- also ich habe das Interesse ja so oder so, aber mehr die Möglichkeit, da rein zu schauen zeitlich gesehen. Dann würde ich das auch bestimmt machen. Ja. #00:07:09-7#
79. I: Okay. Ähm genau dann würde ich sagen, wir kommen nochmal weg von der Rückmeldung, dem Material, einmal hin zu diesem Online-Instrument, (.) äh das hast du ja jetzt das zweite Mal, meine // ich, bearbeitet. #00:07:21-0#
80. C2: Das zweite Mal, genau. #00:07:21-4#
81. I: Genau. Hast du dich (.) ähm oder woran hast du dich noch dran erinnert, als du jetzt das zweite Mal das Instrument bearbeitet hast? #00:07:28-4#
82. C2: (.) An (..) gravierende Fehler, die die Lehrkraft (da?) meiner Meinung nach gemacht hat in den Videos. (.) Daran habe ich mich noch erinnert, // also daran konnte ich mich noch erinnern. #00:07:41-6#
83. I: Was (.) was wäre das? #00:07:42-2#

84. C2: Also das war eine Sache mit- über die wir uns schon in der äh Freundesgruppe hier ausgetauscht haben, die- mit den Leuten, die den Test auch gemacht haben, dass die Lehrkraft da zum Beispiel (.) extra immmer die Mädchen nach vorne geholt hat, um (.) ähm irgendwelche Werte aufzuschreiben, und die Jungs (.) nach vorne geholt hat, um zu rechnen und physikalisch nachzudenken. (..) Genau, und natürlich habe ich mich noch an die Person erinnert (..) und (..) an meine Prozentzahl. So, das waren die Dinge, die mir auf jeden Fall hängen geblieben sind vom letzten Mal. #00:08:21-6#
85. I: (.) Okay. (.) Und äh inwiefern hat das die Bearbeitung, also die jetztige Bearbeitung beeinflusst, dass du das vorher schonmal bearbeitet hattest? #00:08:31-6#
86. C2: (4s) Wie gesagt, ich würde sagen, nicht wirklich. Also (..) spezifische Fragen, Punkte, Aspekte da sind mir nicht im Kopf geblieben und ich habe das relativ neutral, relativ neu ähm bearbeitet, deswegen war ich auch so froh, dass ich da äh, ich glaube, einundsiebzig Prozent oder so bekommen habe am Ende, weil das halt von mir aus kam so, nicht, weil ich mir irgendwelche Lösungen gemerkt habe oder sowas. #00:09:01-1#
87. I: Ja. #00:09:02-1#
88. C2: Ja. #00:09:02-4#
89. I: (Gut?). Gut. Dann wäre ich durch // mit meinen Fragen. #00:09:05-7#
90. C2: Okay. #00:09:05-1#
91. I: Hast du noch irgendwas, (.) was du anmerken, sagen- Kommentare, Fragen, alles (lacht). #00:09:12-0#
92. C2: Nö. #00:09:12-8#
93. I: Nö. #00:09:13-3#
94. C2: Keine weiteren Fragen. #00:09:14-3#
95. I: Okay, dann stoppe ich einmal. #00:09:16-5#

P.3.3 Interview 3: C3

1. I: Gut. Dann vielen Dank, dass du das mit machst. (.) Äh ich, genau, möchte zunächst um dein Einverständnis für die Verwendung der Daten bitten, dazu lese ich gleich die Einverständniserklärung einmal vor und bitte danach um dein mündliches Einverständnis. (.) Die Einverständniserklärung lautet ich bin damit einverstanden, dass dieses Interview aufgezeichnet wird. Dieses Interview wird für eine Verschriftlichung, Verwendung oder Weitergabe anonymisiert. Diese Einverständniserklärung zur Nutzung der erhobenen Daten kann jederzeit mit Wirkung für die Zukunft widerrufen werden. Bist du damit einverstanden? #00:00:30-4#
2. C3: Damit bin ich einverstanden. #00:00:31-7#
3. I: Super. (.) Ähm genau, du hast vor einer Zeit, ich glaube, letzte Woche, ne, den Reflexionstest, genau, bearbeitet und eine Rückmeldung erhalten am Samstag, // meine ich. #00:00:43-5#
4. C3: Ja. // Genau so. #00:00:45-1#
5. I: Plus minus. (.) Okay. Und ich würde gleich gerne bisschen mehr darüber erfahren, wie du die Rückmeldung wahrnimmst oder wahrgenommen hast. Es geht mir überhaupt nicht darum, dich zu bewerten, sondern eben äh das Format und die Gestaltung vom Material zu verbessern und das Interview oder deine Aussagen hier nehmen natürlich überhaupt keinen Einfluss auf die Bewertung in universitären Veranstaltungen. #00:01:03-9#
6. C3: Alles // klar. #00:01:05-0#
7. I: Hast du noch irgendwelche Fragen bevor wir anfangen? #00:01:07-1#
8. C3: Ne, erstmal nicht. #00:01:07-9#

9. I: Okay. Dann würde ich zum Einstieg einmal (.) ähm ja mit deiner eigenen Rückmeldung anfangen. Kannst du dich an die noch erinnern? #00:01:16-6#
10. C3: Also die Rückmeldung, die ich jetzt bekommen habe // am Samstag? #00:01:19-8#
11. I: Ja. #00:01:19-6#
12. C3: Genau, ja, also habe ich natürlich durchge- also überflogen mehr oder weniger ähm war so zwischen Tür und Angel, Mail halt, zwischen anderen Physikveranstaltungen. Ähm ansich, joa, positives Ergebnis, bin ich zufrieden mit, kann ich mit leben. #00:01:33-8#
13. I: Okay, also das wäre deine Schlussfolgerung, // die du draus gezogen hast. #00:01:37-4#
14. C3: Das (.) ja, das ist jetzt erstmal meine Schlussfolgerung, ähm dass ich anscheinend den Test ähm erfolgreich absolviert habe oder zumindest äh die relevanten Dinge gesehen habe, die mir jetzt auch wich- persönlich wichtig gewesen wären, die ich hätte erkennen sollen. (.) Ja. #00:01:52-7#
15. I: Mhm(Ja). Ähm du hast ja letztes Semester auch eine Rückmeldung bekommen? #00:01:57-2#
16. C3: Ja. #00:01:57-6#
17. I: (.) Inwiefern hast du, als du dir jetzt die Rückmeldung angeschaut hast, inwiefern hast du dabei die Rückmeldung aus dem letzten Semester irgendwie berücksichtigt? #00:02:07-7#
18. C3: (..) Äh ich habe die alte Rückmeldung halt nochmal hervorgekramt, hatte ich halt noch im Mailverteiler und ähm oder ähm nicht im Mailverteiler, ich habe sie (ja auch?) runtergeladen und nochmal verglichen und halt festgestellt, dass ich mich verbessert habe und (.) (ne?) sonst habe ich jetzt auf die alte- auf den alten Test nicht so viel (.) mehr gegeben, weil es ist auch (eigentlich?) einfach ein Jahr her, ich wusste nicht mehr, was ich genau angekreuzt habe. #00:02:31-4#
19. I: Ja, klar. (.) Okay. Dann habe ich, damit wir was haben, worüber wir reden können // und es anonym bleibt, einmal eine Beispielrückmeldung mitgebracht, ich (packe?) die einmal auf. Also von der Struktur her genau die gleiche, wie deine, vom Ergebnis wahrscheinlich nicht, es sei denn, es ist durch Zufall identisch. (.) Ähm (.) genau, also ich habe hier das Beispiel von einer Rückmeldung, wenn du möchtest, kannst du dir auch ruhig einen Moment Zeit nehmen, das einmal anzuschauen, zu überfliegen ähm und dann habe ich danach so ein paar (.) Fragen. #00:03:02-8#
20. C3: Mhm(Ja). #00:02:38-0#
21. C3: Okay. Ja, ich würde noch einmal kurz // überfliegen wollen. #00:03:06-1#
22. I: Ja, klar. #00:03:06-1#
23. C3: (37s) (Joa?) äh glaube ich. (..) Fangen wir mal an, (ne?). #00:03:48-0#
24. I: Okay. (.) Ähm hast du vorher schonmal mit solchen Boxplots gearbeitet? #00:03:54-1#
25. C3: Also jetzt um Unterricht zu bewerten oder // generell mit solchen Diagrammen? #00:03:58-0#
26. I: Überhaupt. (.) Ja, mit den Diagrammen. #00:03:59-4#
27. C3: Äh (.) eher nicht. Also Diagramme generell ja, aber mit solchen Blockdiagrammen jetzt eher seltener, aber sie sind ja relativ klar nachzuvollziehen, wenn man sie einmal (.) genau angeguckt hat. (.) Also ich kann mir schon dran- ich kann mir drunter vorstellen, was was sie bedeuten. #00:04:16-8#
28. I: Okay. (.) Ähm wenn du dir jetzt vorstellst, das wäre deine Rückmeldung. #00:04:22-2#
29. C3: Mhm(Ja). #00:04:23-1#
30. I: Welche Schlüsse ziehst du aus dieser Rückmeldung? #00:04:25-6#

31. C3: (...) Äh da muss ich nochmal eben kurz ganz kurz gucken. (..) (Ja?), dann haben wir ja einmal hier das Gesamtergebnis (.) ähm, da würde ich (halt?) diesen Rückschluss draus ziehen, dass ich noch unter dem Median bin, also unter dem Durchschnitt, (.) aber halt nah dran, also (.) wenn ich das jetzt ja richtig verstanden habe, ist dieser Block ja so der der durchschnittliche Bereich, plus, minus, mit einer kleinen Kohärenz (.) und ähm (.) ja, also dass es jetzt erstmal nicht so dramatisch weit weg ist (..) und äh wenn (wir?) zu den anderen gehen, dann könnte ich ja halt aufgeschlüsselt auf die einzelnen Kompetenzen, die abgefragt worden sind im Test, nochmal ein bisschen genauer nachvollziehen, äh in welcher Kompetenz ich jetzt nacharbeiten müsste, (.) wie das Gesamtergebnis halt aufgeschlüsselt ist und ähm welchen (.) ja, genau. #00:05:16-9#
32. I: Mhm(Ja). Was würdest du // sagen, wo müsstest du da nacharbeiten jetzt zum Beispiel oder müsstest du da überhaupt nacharbeiten. #00:05:23-6#
33. C3: Also (in?)- #00:05:18-8#
34. C3: Ich persönlich jetzt- wenn ich den z- den kleinen, also den die Abbildung zwei als Ergebnis kriegen würde, würde ich auf jeden Fall sagen, dass ich in der Bewertungskompetenz von Unterricht nochmal ein bisschen mich konzen- fokussieren müsste. (.) In Alternativensuche auch ähm das Schülervorstellungen würde in dem Fall ja- ist ja ziemlich gut, da würde ich das erstmal so beibehalten, dass ich sagen würde, okay, ich lege (mir?) jetzt mal meinen Fokus, um mich zu verbessern erstmal auf die Bewertung meines Unterrichts, vielleicht auch nochmal bei Kollegen in Unterricht gehen, wenn ich Zeit habe und die das dulden (unv.) vielleicht nochmal andere Eindrücke kriegt, einfach um auch nochmal zu üben, wie ich Unterricht bewerte, um vielleicht auch Muster v- von Kollegen an mir selber wiederzuerkennen. (.) Ja. (.) // Das wäre so mein Vorgehen. #00:06:05-2#
35. I: Ja. (.) Okay und ähm aus dieser Rückmeldung, welche Informationen sind da besonders hilfreich oder besonders interessant für dich? #00:06:13-7#
36. C3: (..) Mhm (.) (wie?) also inwiefern kann ich jetzt diese Frage verstehen? #00:06:20-9#
37. I: Also wenn du- Ich meine, du kriegst hier äh im- inklusive Deckblatt fünf Seiten. #00:06:25-5#
38. C3: Ja. #00:06:26-2#
39. I: Ähm was davon ist quasi das besonders interessante oder bestonders relevant für dich das- an Informationen quasi. #00:06:33-0#
40. C3: (.) Also das relevanteste wäre tatsächlich für mich die Abbildung zwei ähm (.) Deckblatt immer- ja, gut, ist immer schön und gut und e- Erklärung und Einführung auch immer wunderbar, (.) äh ich weiß jetzt gar nicht mehr, inwiefern die Einführungstexte auch in dem Test vorkamen, kamen glaube ich nicht, ne? #00:06:51-3#
41. I: Ne, // so nicht. #00:06:52-4#
42. C3: Ne, ne, so nicht in der Form, aber (.) ähm auf jeden Fall für die Einordnung nochmal gut, dass man nochmal selber weiß rückwirkend, wie setzt sich das zusammen, wie setzt sich auch das Bewertungsschema, glaube ich, zusammen, wird zumindest (so?) einigermaßen grob angeschnitten. (.) Aber das wirklich der zentrale Punkt wär- wäre für mich jetzt Abbildung zwei, weil da kann ich die meisten Rückschlüsse draus ziehen. S- Aus Abbildung eins (.) ist ja eher für die Vergleichsgruppe, wobei ich kenne die Vergleichsgruppe ja nicht, außer jetzt irgendein Kollege von mir, dass ich weiß, der hat das auch gemacht, aber dann- selbst dann könnte ich es ja nur (.) einordnen, wenn ich sein Ergebnis kennen würde. Das ist nice-to-have, aber eher relevant, glaube ich, für die Forschung, als für mich selber dann. Für mich wäre Abbildung zwei, da kann ich persönliche Rückschlüsse draus ziehen. #00:07:36-9#
43. I: Okay. (.) Ähm und gibt es andersum Inhalte in dieser Rückmeldung, die aus deiner Perspektive überflüssig sind? #00:07:44-6#

44. C3: Mhm (...) an sich (..) ne, es hat alles irgendwie schon seine Bewandnis. Also ich muss es ja auch einordnen können, ähm weil der Test an sich ist ja erstmal ja nur der Test (.) und (.) mit dem mit dem Beschreibungstext kriege ich ja denn erstmal die Einordnung, äh was überhaupt das Ziel des Tests war, also (na?) gut, vorher weiß man das auch schon ungefähr, man kann sich an den Fragen langhangeln und ich weiß ja, was ungefähr gefordert ist, aber (.) um selber nochmal die am Ende die Einordnung und (auch?) das eigene Ergebnis selber einordnen zu können, finde ich das schon relativ wichtig. (.) Ähm (..) ja (.) von daher (.) finde ich es so eigentlich in Ordnung so. #00:08:28-7#
45. I: (.) Gibt es äh Fragen, die für dich offen bleiben? #00:08:33-1#
46. C3: Mhm (...) Fragen, die offen bleiben. (..) Ne, an sich eigentlich nicht. Wüsste ich jetzt nicht. #00:08:44-0#
47. I: Oder Informationen, die du drüber hinaus gern hättest zum Beispiel. #00:08:48-0#
48. C3: (...) Mhm (...) (unv.) muss ich mal kurz überlegen. (8s) Ne, an sich eigentlich auch nicht unbedingt. (..) Eine genauere Aufschlüsselung der Aufgaben, wo man irgendwas hatte, ist ist eigentlich auch gar nicht so (.) wichtig. (..) Ne, eigentlich nicht, ne. #00:09:20-4#
49. I: Okay. (.) Ähm dann ist in dieser Rückmeldung ist hier ja so ein Material verlinkt. #00:09:27-1#
50. C3: // Mhm(Ja). #00:09:28-3#
51. I: Ähm (..) genau, hast du da mal reingeschaut, das irgendwie teilweise oder auch ganz bearbeitet oder planst du, damit zu arbeiten? #00:09:38-9#
52. C3: Muss ich jetzt ehrlich gestehen, habe ich noch nicht reingeguckt (lacht), kann ich also jetzt noch nicht zu viel sagen. Mhm (...) werde ich aber mal reingucken, weil kann ja helfen, oder wird wahrscheinlich helfen. Ähm (..) ist ja auf jeden Fall sinnvoll, auf jeden Fall nochmal Material zu bekommen, um Unterricht selber bewerten zu können, oder zu reflektieren, ist (.) ist halt nicht einfach, muss man halt üben, ähm ohne Übung wird das halt nicht und (.) im Referendariat kriegt man vielleicht später noch ein bisschen Rückmeldung, aber darüber hinaus wird es dann irgendwann schwierig, weil dann ist man halt auf sich allein gestellt, kriegt vielleicht nochmal Rückmeldung von den Schülern oder Schülerinnen, aber sonst (..) dann muss es halt sitzen und dann muss ich mir selber irgendwie Möglichkeiten finden, das zu reflektieren und dann ist so, ist so Material, glaube ich, nicht verkehrt, (wenn?) man nochmal Hilfestellung bekommt. #00:10:27-0#
53. I: Mhm(Ja). (.) Ähm (.) ich weiß nicht ganz genau, hast du im letzten Semester in der Veranstaltung bei Herrn N, wo du den Test schonmal gemacht (hattest?), da habt ihr, glaube ich, weiß ich natürlich nicht, ob du da da warst, in einer Sitzung mal in dieses Material reingeschaut. (..) Wenn du dich nicht erinnerst, dann // (.) nicht. #00:10:45-5#
54. C3: Dann erinnere ich mich, glaube ich, nicht // dran. #00:10:45-3#
55. I: Ne, okay. Gut, das wär sonst nur einmal als Hint, das ist das immernoch das gleiche quasi. Okay. Ähm (.) aber dann hat- also dann hat das jetzt auch nicht deine Entscheidung beeinflusst, dass du schonmal in das Material geschaut hast, wenn du dich- (lacht) #00:11:00-3#
56. C3: Äh // ne. (.) Ja, auf jeden Fall nicht. (lacht) #00:11:01-9#
57. I: Ne. #00:11:02-5#
58. C3: Das Problem auch, dass in den Veranstaltungen- Kann sein, dass ich wirklich nicht da war, weil ich oft doppelt Veranstaltungen noch mit anderen Fachbereichen (habe?). Ja (unv.) #00:11:09-0#
59. I: Ja, okay, das ist unglücklich. (.) Äh ne, ist auch kein Problem, ich hätte es nur als Hint // einmal mitgegeben. #00:11:14-5#
60. C3: Ja. #00:11:14-1#

61. I: Ähm okay, dann kommen wir vielleicht noch einmal weg von der Rückmeldung oder diesem Material zurück zu dem Test selber, zu dem Diagnoseinstrument. Du hast das jetzt das zweite Mal, meine ich, bearbeitet? #00:11:25-9#
62. C3: Mhm(Ja). #00:11:26-5#
63. I: Äh woran hast du dich bei der jetzigen Bearbeitung beziehungsweise davor noch erinnert? #00:11:32-1#
64. C3: (...) Woran ich mich noch erinnert (habe?). (.) Äh ich wusste halt, dass da im Test das eine Videosequenzen kommen und ich wusste (dann?) natürlich, dass äh (.) ähm dass ich Sachen- dass ich halt äh meinen Fokus (darauf?) legen musste, wie wie wird gesprochen, wie wie wie agieren die Lehrkraft mit den Schülerinnen etc, das wusste ich noch so ungefähr. Ich ich wusste natürlich nicht mehr nach einem Jahr, was unge- was die Fehler waren, (.) das nicht, aber ich wus- ich- Im Gegensatz zum ersten Mal wusste ich schon ein bisschen, worauf ich vielleicht zu achten habe, (.) was natürlich auch ein Vorteil vielleicht später bei der Beantwortung der Fragen ist, aber ist ja auch nicht verkehrt, ist ja eine Übung. (.) Ähm (..) was wusste ich sonst noch? Äh dass halt viele Fragen waren (lacht), dass es ein bisschen dauert, aber (.) gut, und dass sich manche Sachen vielleicht auch gefühlt irgendwie wiederholt haben, aber (.) ist aber auch nicht verkehrt, ähm (.) muss halt auch mal sein. #00:12:28-4#
65. I: Okay und inwiefern hat das deine Bearbeitung jetzt beeinflusst, dass du das Instrument vorher schonmal bearbeitet hast? #00:12:35-6#
66. C3: (..) Mhm (.) an sich das Vorgehen war gleich, da hat sich gar nicht so viel dran geändert, ähm (..) da war halt genügend Zeitraum zwischen, dass ich jetzt nicht hätte sagen können, ich kann mich da jetzt so durchklicken ähm das- dafür hats (nicht?)- das heißt, ich musste (.) das halt wieder gucken. Wäre halt genauso gewesen, als wenn ich beim Kollegen Unterricht geguckt hätte und der sagt ein Jahr später, ich habe die gleiche Einheit jetzt wieder mit einer neuen Klasse, kannst du nochmal gucken, (.) hat sich was verbessert, hat sich nichts- das wäre für mich die gleiche Situation gewesen. #00:13:07-0#
67. I: (.) Ja, okay. Ähm dann wäre ich durch mit meinen Fragen hier. Hast du noch Fragen, Anmerkungen? #00:13:13-9#
68. C3: (.) Ne, an sich nicht. #00:13:16-5#
69. I: Ne, okay. Dann stoppe ich mal eben die Aufnahme. #00:13:19-6#

P.3.4 Interview 4: C4

1. I: So, okay. Dann erstmal vielen Dank, dass du das mitmachst. Ich möchte zunächst um dein Einverständnis für die Verwendung der Daten bitten, dazu lese ich die Einverständniserklärung vor und bitte danach um dein mündliches Einverständnis, wenn du einverstanden bist. (.) Äh die Einverständniserklärung lautet ich bin damit einverstanden, dass dieses Interview aufgezeichnet wird. Dieses Interview wird für eine Verschriftlichung, Verwendung oder Weitergabe anonymisiert. Die Einverständniserklärung zur Nutzung der erhobenen Daten kann jederzeit mit Wirkung für die Zukunft widerrufen werden. Bist du damit einverstanden? #00:00:30-0#
2. C4: Mhm(Ja), ja. #00:00:30-7#
3. I: Supi. (.) Genau. Also ich hatte es eben schon gesagt, du hast den Test bearbeitet oder dieses Diagnoseinstrument und eine Rückmeldung bekommen (.) äh und mir geht es gleich vor allem so ein bisschen um die Rückmeldung, darum, zu erfahren, wie du die wahrnimmst oder wahrgenommen hast. (.) Es geht explizit nicht darum, dich oder deine Aussagen irgendwie zu bewerten, sondern darum, das Format oder die Gestaltung dieser Materialien zu verbessern. Das heißt, das, was du jetzt sagst, hat überhaupt keinen Einfluss auf irgendwelche Bewertungen in universitären Veranstaltungen. #00:01:01-0#
4. C4: Mhm(Ja). #00:01:01-7#
5. I: (.) Hast du noch irgendwelche Fragen, bevor wir anfangen? #00:01:04-7#
6. C4: Ja, ich habe mir die Rückmeldung gar n- noch gar nicht so ganz genau durchgelesen, aber- #00:01:08-2#
7. I: Aber einmal // grob reingeguckt? #00:01:10-2#
8. C4: Ich habe mir die Graphen angeguckt. Ja, das auf jeden Fall. Ich habe mir die Graphen angeguckt, ich habe mir irgendwie diese Prozentzahlen kurz angeguckt, (.) ähm ja, genau. #00:01:16-5#
9. I: Ist in Ordnung. #00:01:17-4#
10. C4: Mhm(Ja). #00:01:18-0#
11. I: (.) Kriegen wir hin. Okay. Dann äh fangen wir da vielleicht direkt einmal an. Kannst du dich noch so grob dran erinnern an die Rückmeldung? #00:01:25-6#
12. C4: Ja, so ganz grob, ja. #00:01:27-4#
13. I: Okay. Welche Schlüsse hast du daraus gezogen? #00:01:30-6#
14. C4: (..) Mhm also für mich (..) ich habe mir quasi angeguckt, irgendwie, da war ja ganz- da stand irgendwie, Sie haben so und so viel Prozent, damit liegen in dem und dem Drittel (.) oder in dem und dem und bei mir war es das obere und das war für mich dann irgendwie (auch ein?) Schluss, dass das jetzt ähm trotz meiner (.) nicht sehr konzentrierten Bearbeitung dieser Aufgaben anscheinend ganz gut war und ganz gut geklappt hat und ja, genau. Mehr habe ich da eigentlich auch gar nicht als Schluss draus gezogen. Ich war dann doch eigentlich zufrieden. (.) Punkt. (lacht) #00:02:00-6#
15. I: Okay. (.) Äh jetzt bin ich nicht ganz sicher, hast du den das erste Mal bearbeitet? #00:02:05-2#
16. C4: Ne, das zweite Mal. #00:02:05-8#
17. I: Im letzten Semester bei Herrn N? #00:02:08-1#
18. C4: Genau zu Beginn bei Herr N letztes Semester habe ich das- haben wir das schonmal machen dürfen #00:02:12-3#
19. I: Okay. Dann ähm wäre meine Frage, ihr habt da ja auch eine Rückmeldung zu bekommen. #00:02:18-1#
20. C4: Mhm(Ja). #00:02:18-7#
21. I: Inwiefern (.) das deine jetzige Beschäftigung mit der Rückmeldung (.) unter Umständen beeinflusst hat, (.) dass du damals schon das Instrument bearbeitet hast beziehungsweise damals schon eine Rückmeldung bekommen hast, hast du das irgendwie berücksichtigt? #00:02:32-9#

22. C4: (..) Mhm nicht so ganz. Also ich habe die Rückmeldung von früher mir auch irgendwo gespeichert gehabt (.) äh und weiß aber, dass sie schlechter war. Also schlechter im- natürlich im objektiven Sinne, aber ähm auf jeden Fall hatte ich weniger Prozente, ich war auf jeden Fall im Mittel (..) ähm (.) ja und sonst hat das aber eigentlich über- gar nichts gemacht, (so?), ich habe nur geguckt, okay, ja anscheinend hat sich das irgendwo- in (irgendeiner?) Weise verbessert, ähm was aber auch hoff- ja, zu erwarten war, denke ich mal, wenn man sich noch gar nicht mit dem Thema beschäftigt hatte. #00:03:04-4#
23. I: (.) Ja. Okay. Ich habe, damit wir etwas haben, worüber wir sprechen können, einmal eine Beispielrückmeldung. Das ist ähm (.) nicht deine, es sei denn, die Ergebnisse sind durch Zufall gleich, davon gehe ich aber nicht aus. (..) Wenn du willst, kannst du nochmal kurz drauf schauen, einmal reinschauen, ähm danach hätte ich ein paar Fragen dazu. #00:03:24-8#
24. C4: Mhm(Ja). (97s) Ja. (...) Ja. #00:05:05-9#
25. I: Okay. (.) Ähm hast du (.) zuvor schonmal mit solchen Boxplots gearbeitet, also mit diesen Diagrammformen? #00:05:14-4#
26. C4: Mhm nein. #00:05:16-5#
27. I: (.) (Nein?). #00:05:17-7#
28. C4: Ich glaube a- wobei, doch wahrscheinlich schon, weiß ich nicht, man hat doch alles irgendwie schonmal gemacht. ja, ich weiß nicht. // Doch, bestimmt. #00:05:24-2#
29. I: Aber nicht präsent? #00:05:24-4#
30. C4: Nicht präsent, nein. #00:05:25-5#
31. I: Okay. (.) Ähm wenn du dir die Rückmeldung amschaust, oder dir vorstellst für den Moment, das wäre deise- äh deine, welche Schlussfolgerunden würdest du dadraus ziehen? #00:05:35-3#
32. C4: (.) (unv.) puh also äh mhm (.) erstmal würde ich schlussfolgern, dass, wenn das jetzt meine Rückmeldung wäre, dass ich auf jeden Fall noch an mir arbeiten muss in vielen Sachen, einfach nur, weil ich unter dem Durchschnitt liege und hast ist halt (unv.) akzeptabel, meiner Meinung nach, allerdings schon sehr, sehr gute Kenntnisse über Schülervorstellungen habe. (.) Also das ist ein Punkt, wo man quasi oder wo die Person oder ich dann in dem Fall (.) schonmal geglänzt hat oder glänzen kann, aber alle anderen Aspekte der- ähm die erhoben wurden, auf jeden Fall noch ausbaufähig sind (.) und ich weiß nicht, wie das dann (.) oder wie das jeder für sich handhabt, aber ma- natürlich versucht man selber ja schon mindestens mal im Durchschnitt aller anderen zu sein (.) und äh nicht dadrunter. (.) Genau. #00:06:15-7#
33. I: (.) Okay. #00:06:17-3#
34. C4: Also zum Beispiel Bewertung- Sorry. #00:06:18-6#
35. I: Ne, (unv.) #00:06:19-5#
36. C4: Also zum Beispiel bei Bewertungen und Alternativen, (.) wie das hier steht in Abbildung zwei, (.) ähm sieht man ja, dass- also Bewertung ist ja relativ schlecht, (.) also relativ weit unter dem Durchschnitt und Alternativen (.) quasi am unteren Ende des Durchschnitts, diese beiden Aspekte, vor allen Dingen aber die Bewertung, sollte man auf jeden Fall verbessern, sollte man auf jeden Fall nochmal schulen, (.) sich dazu vielleicht nochmal was durchlesen, (.) genau. (.) Ja. #00:06:42-0#
37. I: (...) Okay. Ähm wenn du (.) diese Rückmeldung jetzt quasi bekommst oder dir anschaust. #00:06:48-3#
38. C4: Mhm(Ja). #00:06:49-0#
39. I: Welche Informationen sind da besonders hilfreich oder besonders, ja, interessant, relevant für dich? #00:06:53-6#
40. C4: (...) Mhm (4s) Naja (...) meinst du jetzt auf die Graphen bezogen, oder auf die wirkliche- auf die wirklichen Daten, die du zurück bekommst? #00:07:07-3#
41. I: Also es geht quasi jetzt drum, wenn du diese Rückmeldung kriegst, dann kriegst du // so ein fünfseitiges Dokument. #00:07:12-3#

42. C4: Mhm(Ja) (.) mhm(ja). #00:07:12-2#
43. I: Was davon ist für dich aus diesem kompletten Dokument besonders interessant // besonders relevant? #00:07:17-6#
44. C4: Mhm(Ja) mhm(ja) mhm(ja). Ab dem Punkt Rückmeldung zu Ihrer äh Testbearbeitung, (.) wo quasi ein- eingangs so ein bisschen erklärt wird, äh wie (diesen?) Graphen lesen kann, eher nicht so interessant, kriegt, glaube ich, jeder hin. Dann der Graph an sich (.) ist natürlich sehr interessant, vor allem das Gesamtergebnis finde ich tatsächlich viel interessanter, als die Teilergebnisse (.) ähm und danach natürlich noch die Erklärung der jeweiligen äh aufgeschlüsselten Teilergebnisse, was ist die Bewertung, was sind Alternativen, was sind äh Schülervorstellungen. (hustet) Sorry. Wichtig, aber weniger interessant, finde ich tatsächlich aber auch danach direkt den Punkt äh wie kann man das schulen. (.) Wie kann man dran gehen, ähm wobei der ja doch relativ kurz gehalten ist ähm ja genau. (.) Wenn ich jetzt das lesen würde und sehen würde, äh ich bin bei Bewertung bei fünfzig Prozent und alle anderen sind da irgendwie (.) bestenfalls bei fünfundsiebzig Prozent, würde ich mich tatsächlich fragen, wie kann man das denn noch besser schulen. Ähm genau. Wahrscheinlich wäre dann ein ein größerer Absatz über wie kann man diese- die Reflexionsfähigkeit schulen oder so (.) nochmal interessanter gewesen. Genau. Ja. #00:08:21-5#
45. I: (.) Ähm du hast gerade gesagt, du findest das Gesamtergebnis interessanter, also Abbildung eins, als // Abbildung zwei. #00:08:27-4#
46. C4: Ja. (.) Ja. #00:08:27-8#
47. I: Warum? #00:08:28-4#
48. C4: (..) Naja, weil am Ende kommt es ja auf jeden der Aspekte an und ähm wie sagt man immer so schön, die Kette ist nur so stark, wie ihr schwächstes Glied und äh das ist natürlich dann- das schlägt sich im Gesamtergebnis eher nieder, als äh in den Teilergebnissen. Man sich kann ja so ein bisschen einreden, ja, gut, ich äh bin jetzt in der Bewertung nicht so gut, aber dafür in den Schülervorstellungen mega, (.) was gut ist, allerdings vielleicht nicht immer unbedingt zielführend, weil am Ende ist man dan trotzdem in der Bewertung nicht gut und ein Aspekt des Unterrichts ist dann halt immer schlechter oder nach diesem Test schlechter und demnach ist das Gesamtergebnis vielleicht aussagekräftiger, (.) was die Unterrichtsqualität dieser Person angeht, (.) weil sie halt alles berücksichtigt. (.) Außerdem ist es f- also für jemanden, der den Test gemacht hat und der dauert ja auch ein Weilchen, ähm ist es vielleicht auch einfach, sehr sehr schön, akkurat ganz schnell zu sehen auf einen Blick, ich bin hier, die anderen sind da, das ist das Gesamtergebnis und nicht- sich- sich nicht noch Gedanken machen zu müssen über Aufschlüsselung der Teilergebnisse, was ist überhaupt Bewertung, was ist über- (unv.) Alternativen, was zählt hier alles rein. Einfach (auch?) einfacher zu lesen, also, ne, (.) kommt auch (drauf?) (unv.) #00:09:33-4#
49. I: (..) Okay. Ähm gibt es Inhalte in dieser Rückmeldung, die aus deiner Perspektive überflüssig sind. #00:09:41-3#
50. C4: Mhm (4s) das ist eine sehr gute Frage. (...) Ne, eigentlich nicht, muss ich sagen. Also w- ah ich weiß nicht, also so ein bisschen diese Diagrammerklärung (Hinweise zur Interpretation der Diagramme), die horizontale Linie markiert den Median und so weiter, ähm das steht ja auch hier in der Legende, das ist alles so ein bisschen doppelt gemoppelt. (.) Oder auch die Erklärung, die hier in dem Diagramm (Abbildung 1) steht, genau, das steht ja hier oben auch nochmal drin, eigentlich genau das gleiche. Am Ende richtet sich der Test äh ja an Lehrende äh Lernende der Physik (.) und ich weiß nicht, ob man das dann alles noch acht Mal erklären (muss?), ich glaube, Graphen haben wir alle schonmal gesehen, ähm anderer- andererseits würde ich auch nicht sagen, dass es überflüssig ist, vielleicht hilft es schon einigen Leuten weiter. Demnach, ne, ich glaube, eigentlich ist nichts überflüssig. #00:10:30-1#
51. I: Okay. #00:10:30-9#
52. C4: (Ja?). #00:10:31-3#

53. I: Gibt es Fragen, die für dich offen bleiben? #00:10:34-3#
54. C4: (.) Mhm (..) ich habe mich so ein bisschen gefragt tatsächlich, die Fehlerbereiche, also die jeweils angegeben sind, sind ja sehr sehr groß, ähm also irre groß und (.) ja, zum Beispiel bei Schülervorstellungen ist der Fehlerbereich von Null bis Hundert (lacht), was ist da denn los? Ähm irgendwie finde ich das- also das das hat mich einfach gewundert als- als ich auch meine Sache bekommen habe, habe ich mich (.) da so ein bisschen hinterfragt und mich auch gefragt, woher diese überhaupt diese Fehler kommen. (.) Ob die gewählt sind und gesagt wird, das ist der Fehler, den wir bei dieser Fragestellung ha- annehmen müssen, oder ob die sich auch irgendwie so ein bisschen ergeben aus den Antworten, wenn ja- weil das wird ja, denke ich mal, auch immer so ein bisschen doppelt gefragt und, ne, und äh die t- der Test bestätigt sich (ja?) sicherlich sicherlich auch in der Fragestellung selber. Die Fehler werden irgendwie gefühlt gar nicht thematisiert, ich weiß nicht, wo die herkommen, das fand ich ein bisschen äh k- tatsächlich komisch. Die Frage bleibt für mich auch offen, ob es normal ist, dass der Fehlerbereich so groß ist. Weil wenn ich mir das so angucke, eigentlich kann ich da ja nichts draus lesen, also zumindest sehr wenig und das äh genau, das bleibt auf jeden Fall offen für mich. #00:11:43-7#
55. I: Mhm(Ja). Ich kann das vielleicht ganz kurz klären. #00:11:45-4#
56. C4: Ja. // Gerne. Sehr gerne. #00:11:47-4#
57. I: Ähm (.) dieser Bereich zwischen oberem Ende der Box bis zum oberen Ende vom Balken. #00:11:53-4#
58. C4: (.) Du bist hier, ne? (zeigt auf oberen Teil der Box in Beispielrückmeldung) #00:11:55-3#
59. I: Äh ne, ich bin da drüber, ich bin hier (zeigt auf obere Antenne in Beispielrückmeldung). #00:11:59-5#
60. C4: Mhm(Ja), ja. #00:12:00-4#
61. I: Das sind das ist der Bereich, wo die besten fünfundzwanzig Prozent der Ergebnisse liegen. (.) Der Bereich unter der Box bis ganz unten, das ist der Bereich, wo die niedrigsten fünfundzwanzig Prozent der Ergebnisse liegen. #00:12:13-3#
62. C4: Mhm(Ja) mhm(ja). #00:12:14-2#
63. I: Das heißt jetzt zum Beispiel Schülervorstellungen, wo wir wirklich einen Bereich von Null bis Hundert haben, heißt das, es gibt Studierende, die im Bereich von Null bis Hundert Punkte- also hundert Prozent der // Punkte kriegen. #00:12:23-9#
64. C4: Mhm(Ja). Also (unv.) Null? #00:12:25-0#
65. I: Genau. #00:12:25-7#
66. C4: Eieiei, // okay. #00:12:26-3#
67. I: Gibt es auch und es gibt eben auch welche, die in diesem Bereich, ähm genau, hundert Prozent der Punkte kriegen. Und das ist eben, ne, ein relativ großer Bereich, weil sich der zusammensetzt aus eben allen, also, ich glaube, in dieser Rückmeldung sind das, genau, hundert Studierende und // drei Referendare insgesamt. Also bei hundertdrei Personen jeweils der maximale Bereich, in dem irgendjemand Punkte erreicht, das sind diese Balken. #00:12:47-5#
68. C4: Mhm(Ja). #00:12:40-1#
69. C4: Mhm(Ja). #00:12:48-0#
70. I: Und deswegen sind die da // eben so groß, weil es tatsächlich vorkommt. #00:12:52-9#
71. C4: Ja. (.) (Genau?) das- das- das wäre tatsächlich eine Erklärung gewesen, die total interessant gewesen wäre. Also steht halt hier- hier nicht drin, aber vor allen Dingen bei den Schüler- innenvorstellungen, dass wirklich von Null bis Hundert alles vertreten ist, ist ja unglaublich krass. (.) Ähm das ist ja interessant, ja. #00:13:06-3#
72. I: Mhm(Ja). Ähm das ist versucht, so ein bisschen aufzufangen in dieser (.) // in dieser Beschriftung, untere fünfundzwanzig Prozent der Ergebnisse und obere fünfundzwanzig Prozent der Ergebnisse hier oben (zeigt jeweils auf die Beschriftung in Abbildung 1). Ähm (..) #00:13:20-2#

73. C4: Mhm(Ja) mhm(ja) (4s) mhm(ja) mhm(ja) mhm(ja) mhm(ja) mhm(ja) #00:13:19-1#
74. C4: Wahrscheinlich (unv.) habe ich // habe ich mir jetzt nicht gut überflogen wahrscheinlich. (.) // Oder- #00:13:24-5#
75. I: Hast du ein- (.) Ja oder hast du eine Idee, wie man das anders oder oder irgendwie ein Hinweis, an welcher Stelle es stehen müsste, oder (.) irgendwie irgendeinen Gedanken dazu? #00:13:33-8#
76. C4: Mhm ja, lass mich (mal?) ganz kurz gucken. #00:13:35-7#
77. I: Klar #00:13:36-3#
78. C4: (6s) Also wenn, dann- also sinnhaftig wäre es natürlich vor dem ersten vor der ersten Abbildung, das irgendwie noch mit reinzunehmen. Wie gesagt, es kann natürlich auch jetzt einfach sein, ich habe das jetzt gerade nochmal überflogen gehabt und das habe ich jetzt vielleicht übersehen, das passiert ja auch. (.) Ähm aber wenn man eh schon einen Text schreibt zur oder wie man diesen wie man dieses Diagramm liest, dann kann man das vielleicht noch bei Rückmeldung zu ihren Test- äh Rückmeldung zu ihrer Testbearbeitung noch irgendwie mit ein einbauen. #00:14:07-2#
79. I: Also in den Text, // nicht (unv.) #00:14:09-5#
80. C4: In den in den Text, genau. Also, (ich meine?), hier hat man es ja auch, oder ob man es hier nochmal mit einbaut, dann hätte- hätte man es zwei Mal (.) ähm oder vielleicht das nochmal so ein bisschen ausführen. Ähm (..) ja, doch, also vielleicht das- also so, wie du es mir gerade eben erklärt hast, es ist ja mega verständlich, dass man das da noch so ähm, ja, doch, das wäre, glaube ich, cool gewesen. (.) Weil ich habe mich tatsächlich bei meinen Rückmeldungen gefragt, ob es ob die Fehlerbereiche gar nicht die Gesamtgruppe darstellen, sondern meine Ergebnisse. #00:14:37-1#
81. I: Mhm(Ja). #00:14:37-7#
82. C4: Also ob- man kann ja auch, keine Ahnung, du kannst ja einmal fragen, ähm (.) wo äh- also könnte- es sind ja auch teilweise Fragen, die sich doppeln, ich habe gerade kein- mir fällt jetzt gerade kein Beispiel ein, aber du hast ja teilweise auch Fragen o- man kann- es kann ja auch sein, dass ich bei der einen dann Ja ankreuze und bei der gedoppelten Frage Nein und dann ist es ja auch irgendwie dann wächst ja auch mein eigener Fehlerbereich und ich dachte erst, dass das mein eigener Fehlerbereich ist, der halt davon kommt, dass ich nicht konsistent in meinen eigenen Anklickungen war. #00:15:02-7#
83. I: Mhm(Ja). #00:15:03-1#
84. C4: Ähm aber so ist es natürlich auch viel sinniger, steht auch hier, aber habe ich mir nicht genau genug durchgelesen anscheinen dafür. Ja. #00:15:08-6#
85. I: Ja. Okay. (.) Ähm gibt es ansonste noch (.) weitere Informationen, die du gern gehabt hättest? #00:15:14-9#
86. C4: (...) Nö, eigentlich nicht. (.) Eigentlich nicht, ne. #00:15:20-5#
87. I: Okay. (.) Äh dann ist hier unten ist, also in diesem Absatz wie kann ich meine Reflexionsfähigkeit schulen, da ist so ein Material verlinkt. Ich glaube, es äh heißt Fördermaterial da. Hast du das Material bearbeitet oder da mal reingeschaut ganz oder teilweise beziehungsweise planst du das? #00:15:40-5#

88. C4: Ich habe es nicht getan, nein. Plane ich das ist, finde ich, immer so ein bisschen schwierig äh schwierig zu sagen. Äh definitiv äh ich kann nicht sagen, ich plane das jetzt, // ich kann aber sagen ähm wir hier im Lehramtsstudium kriegen unglaublich viel Material an der Hand an die Hand, du wirst es wissen. Man kann sich nicht alles angucken, aber man kann sich vielleicht ähm die Sachen speichern wir uns ja alle ab und man kann vielleicht einfach wissen, dass da was ist und wenn man dann äh merkt oder in die Situation kommt, dass das vielleicht helfen könnte, kann man dadrauf zurück kommen. Also ich habe jetzt erstmal nicht geplant, das zu machen, außer du sagst, das ist richtig cooles Material, dann gucke ich mir das natürlich noch an, ähm aber ähm (.) der Wi- das das das Wissen darüber ist auf jeden Fall vorhanden und vielleicht kommt man ja irgendwann darauf zurück und wenn es in drei Jahren ist. #00:16:24-2#
89. I: Mhm(Ja). #00:15:50-6#
90. I: Mhm(Ja). Ähm (.) jetzt weiß ich natürlich nicht, ob du da letztes Semester da warst, aber es gab äh ihr habt in einer Sitzung bei Hernn N mal reingeschaut in das Material. #00:16:35-6#
91. C4: Mhm das- #00:16:37-2#
92. I: Okay, zumindest erinnerst du dich nicht. Ne, dann ist kein Problem. #00:16:40-3#
93. C4: Ich war sicherlich da, aber ich habe- ich erinnere mich auf jeden Fall nicht mehr dran, ne. #00:16:43-3#
94. I: Das ist kein Problem, das wäre sonst nur ein äh quasi Hint gewesen. Okay äh gibt es irgendwie bestimmte, ja, Rahmenbedingungen, die du quasi benötigen würdest, damit du ins Material schauen würdest. Irgendwas, was es dafür bräuchte oder Rahmenbedingungen, die man schaffen könnte, was es, ja, quasi einfacher macht für dich? #00:17:00-5#
95. C4: (.) Einfacher, ich meine, es ist doch relativ einfach, ich muss nur auf einen Link klicken, ähm oder oder also (.) oder meinst du attraktiver, so dass man da eher drauf klickt, oder? #00:17:09-8#
96. I: Nehme ich auch, ja. #00:17:10-7#
97. C4: Mhm(Ja) mhm(ja). (4s) Ne, (.) da gibt es, glaube ich, keine Möglichkeiten. (.) Also bin ich ehrlich, weil da ist ein Link, ich glaube, dass (.) mehr kann man nicht machen an eurer Stelle jetzt, ne. #00:17:24-9#
98. I: Okay. (.) Dann ähm kommen wir nochmal weg von der Rückmeldung und diesem (.) // also Material, was hier liegt einmal zum Instrument zurück. Du hast es jetzt das zweite Mal bearbeitet, meine ich. Woran hast du dich vor und bei der Bearbeitung jetzt noch dran erinnert? #00:17:43-3#
99. C4: Mhm(Ja). #00:17:31-6# #00:17:45-1#
100. C4: (...) Vor? Also du meinst bevor ich das- #00:17:47-7#
101. I: Also als du es jetzt bearbeitet hast- #00:17:50-3#
102. C4: Ja. #00:17:50-8#
103. I: Woran hast du dich erinnert vom Mal davor? #00:17:53-3#
104. C4: (.) Äh ü- äh Rahmenbedingungen? Auf jeden Fall, also dass ähm ich quasi einem Kollegen einen Kol- einen Kollegen Rückmeldung geben soll, ähm das war definitiv noch präsent. (.) Ähm mir war auch so, ich sage mal, so (.) das Szenario auch noch irgendwie präsent und auch noch sehr präsent, dass dieser Kollege (diese Kollegin?) im Umgang mit ähm (.) geschlechtsspezifischen Rollen sehr (.) interessant war. Das war mir auch noch präsent. (.) // Ja. #00:18:24-5#
105. I: Okay. Und inwiefern hat das deine Bearbeitung jetzt beeinflusst, dass du das Instrument vorher schonmal bearbeitet hattest? #00:18:30-2#

106. C4: (..) Ich glaube, gar nicht so stark, ehrlich gesagt. Ich glaube tatsächlich, dass ich mit einem anderen Blick (.) da drauf gegangen bin, als damals. Ähm (...) ich musste mir eh nochmal alles angucken und anhören, weil das ist natürlich dann, ne, man merkt sich das ja nicht ganz genau, was der da gemacht hat und gesagt hat. Ich glaube, das hat sich gar nicht so stark ähm geändert. Ich habe allerdings tatsächlich, muss ich sagen, äh mehr hinterfragt, auch die Fragestellungen mehr hinterfragt, (und?) auch (.) drüber nachgedacht, ob ich diese Fragestellungen so irgendwie gut finde, oder ob ich auch ähm jetzt eine andere Möglichkeit finde, die mir jetzt gefällt oder so, ne, also das ähm das habe ich mehr, aber im Vorfeld überlegt oder anders rangegangen bin ich auf jeden Fall nicht. Ne, glaube ich nicht, nein. #00:19:11-9#
107. I: Okay. (.) Ähm ja dan wäre ich durch mit meinen Fragen. Hast du irgendwie Fragen, Anmerkungen (.) zum Instrument, zur Rückmeldung? #00:19:21-0#
108. C4: Mhm (...) nicht so ganz. Anmerkung von mir wäre jetzt eher, also äh dass ich finde es relativ lang, (.) das ist auch okay, also ich (habe?) andere Studien sind auch so lang, das weiß ich auch, ne, ähm (...) ja. Punkt. Man kann jetzt irgendwie noch was das weiter ausführen und sagen, es ist halt so, dass man irgendwann vielleicht auch fertig werden will und dann nimmt ganz ganz schnell die Motivation ab, sich jede Antwortmöglichkeit immer ganz genau durchzulesen und man- da steht dann immer Ja, Spiegelstrich, er kann so weitermachen, Nein, Spiegelstrich, und dann klickt man ja, okay, ja, komm. Ja, nächstes. (.) Ähm einfach nur, weil es vielleicht ein bisschen zu lang ist und weil ähm es auch sehr sehr subtil ist in den verschiedenen Antwortmöglichkeiten, was sich jetzt ändert. Aber insgesamt ähm habe ich eigentlich nichts mehr, ne. #00:20:07-8#
109. I: (..) Okay. Dann mache ich hier mal aus. #00:20:12-0#

P.3.5 Interview 5: C5

1. I: So dann danke, dass du das Interview mit mir führst. (.) Ich brauche gleich zunächst dein Einverständnis äh für die Verwendung der Daten. Dazu lese ich die Einverständniserklärung vor und bitte danach um dein mündliches Einverstand- äh Einverständnis, wenn du einverstanden bist. (.) Die Einverständniserklärung lautet ich bin damit einverstanden, dass dieses Interview aufgezeichnet wird. Dieses Interview wird für eine Verschriftlichung, Verwendung oder Weitergabe anonymisiert. Diese Einverständniserklärung zur Nutzung der erhobenen Daten kann jederzeit mit Wirkung für die Zukunft wiederrufen werden. (.) Bist du damit einverstanden? #00:00:33-6#
2. C6: Ja. #00:00:34-0#
3. I: Gut. (.) Äh genau. Ihr habt vor ein paar Wochen den- dieses Diagnoseinstrument, diesen Reflexionstest nochmal bearbeitet und danach auch wieder eine Rückmeldung bekommen. Und mir geht es darum, dass ich gleich gerne was überfahren würde, wie du die Rückmeldung wahrnimmst oder wahrgenommen hast in erster Linie. Es geht explizit nicht darum, dich in irgendeiner Form zu bewerten oder deine Aussagen, sondern das Format und die Gestaltung der Rückmeldung zu verbessern. (.) Das heißt natürlich auch, dass äh alle Aussagen in diesem Interview keinen Einfluss auf Bewertung nehmen in universitären Veranstaltungen, egal welche. #00:01:11-3#
4. C6: Mhm(Ja). #00:01:12-0#
5. I: (.) Fragen bevor wir anfangen? #00:01:14-9#
6. C6: Nein. #00:01:15-6#
7. I: Nein. Okay. (.) Äh dann (.) fangen wir (.) direkt an mit (der?) deiner Rückmeldung, kannst du dich an die noch erinnern? #00:01:22-9#
8. C6: Ja. (.) Aber ich kann die leider nicht mehr- (..) Also ich habe den Test ja schon mal gemacht und ich weiß nicht mehr, was ich bei- (.) bei der vorigen Rückmeldung hatte. (Sonst?) hätte ich das ja gut vergleichen können. (.) Ich weiß wohl, dass ich damals, (..) ich müsste jetzt lügen, zwischen fünfzig und siebzig Prozent war, genauer kann ich das nicht mehr einordnen. (.) Also bei dem ersten Test. (.) Und jetzt war ich bei einundsechzig. Um das jetzt aber einzuordnen, ich habe den Test quasi eineinhalb Mal gemacht, weil ich einmal, ich glaube das hatte ich auch bei der- beim Feedback geschrieben beim Test, (.) quasi bei dreiundvierzig Prozent Fortschritt war, (.) dann habe ich aber eine Pause gemacht und dann war der Fortschritt gelöscht, weil ich zu lange nichts gemacht habe. (.) Und dann muss ich gestehen, dass ich diese dreiundvierzig Prozent recht intuitiv durchgeklickt habe, weil ich keine Lust hatte, das alles nochmal zu gucken, deshalb könnte es sein, dass ich den (.) ersten Teil quasi (.) nicht so gut habe. (lacht) Und das kann man jetzt, also- #00:02:15-6#
9. I: Also das war bei der jetzigen Testdurchführung? #00:02:17-5#
10. C6: Genau, // bei der ersten nicht. Ne. #00:02:19-2#
11. I: Nee, okay. #00:02:19-8#
12. C6: Und das habe ich dann recht schnell durchgeklickt, weil ich dann auch nicht so viel Lust hatte, das alles nochmal zu machen (..) und da habe ich dann- (.) also was ich noch so im Hinterkopf hatte, habe ich wieder angekreuzt, aber es könnte natürlich sein, dass man da auch noch was verkehrt angekreuzt hat. (.) Ja, um das so einzuordnen. #00:02:35-1#
13. I: Ja, okay. #00:02:36-5#
14. C6: Und das kann man jetzt ja, (.) so wie ich das bei der Rückmeldung nicht gesehen hatte- Also man kann ja nicht sehen, ob man in der ersten Phase besonders schlecht war oder in- oder prozentzahlmäßig gesehen und dann in der letzten Phase besser war, das ist ja ein Gesamtergebnis. #00:02:49-9#
15. I: Ja. (.) Genau, das bezieht sich auf alle Videos, ja. (.) Ähm genau, du hast jetzt schon so ein bisschen angefangen, aber ich frage trotzdem manchmal, welche Schlüsse hast du aus deiner Rückmeldung gezogen? #00:02:59-9#

16. C6: (...) Ja, tro- trotzdem ist das Ergebnis ja nicht sehr gut, also dass ich jetzt noch (..) A Reflexionsbedarf für Unterricht habe (.) und (...) ja an- Also wir haben den Test ja letztes Mal auch besprochen, A hier und auch bei Frau Q oder bei Herrn N, eins von beiden, (.) und dass ich ja trotzdem noch nicht so genau weiß, worauf es jetzt ankommt, anscheinend. (..) // (Also?)- #00:03:23-9#
17. I: Bei der bei der Durchführung quasi, bei der Reflexion, oder? #00:03:28-6#
18. C6: Ja, bei der Reflexion von Unterricht. #00:03:30-0#
19. I: (.) Okay. (.) Ähm (..) du hast eben auch gesagt, du hattest die alte Rückmeldung jetzt nicht mehr, ne? Aber // du hast (unv.)- #00:03:39-7#
20. C6: Nee, ich habe extra nochmal geguckt, ob die Datei ähnlich hieß, also ich habe noch mal (am?) Rechner gesucht, aber ich habe die nicht wieder gefunden. #00:03:45-4#
21. I: okay. Ja, also // kurz off topic, sie hätte eigentlich genauso heißen müssen, aber okay. #00:03:51-1#
22. C6: Aber- (...) Ja, hatte ich nämlich auch danach gesucht, aber dann habe ich sie vielleicht nicht gespeichert, oder- #00:03:54-4#
23. I: Ja. (.) Okay, aber so so ein ganz bisschen Erinnerung, meintest du, hattest du ja // quasi noch, ne? #00:04:00-9#
24. C6: Ich würde so sagen, zwischen fünfzig und siebzig Prozent, also irgendwie im, (..) das wurde ja irgendwie in oberes und unteres Drittel oder so eingeteilt, ne? #00:04:08-0#
25. I: Ja. (..) Und- #00:04:10-6#
26. C6: Ja. (..) Ich meine, es war eher unteres (.) oder irgendwie so (..) schlechter als die (.) andere Hälfte, oder irgendwie so, wurde das ja doch- (.) Irgendwie so war das, ne? #00:04:18-4#
27. I: Ja. #00:04:18-9#
28. C6: Der Wortlaut. #00:04:19-5#
29. I: Genau, es war äh in Viertel (.) aufgeteilt. #00:04:22-4#
30. C6: In Viertel, ja. #00:04:23-4#
31. I: (unv.) Ähm und inwiefern hat dich jetzt, also als du jetzt quasi deine Rückmeldung bekommen hast, inwiefern hast du dabei eben das berücksichtigt, was beim letzten Mal- (.) also was du beim letzten Mal erfahren hast, bzw. hat sich quasi deine Perspektive dadrauf verändert, dass du es jetzt ein zweites Mal gemacht hast auf die Rückmeldung? #00:04:42-7#
32. C6: (...) Also wir haben ja mal da in der Vorlesung oder im Plenum darüber gesprochen, (.) und im Prinzip habe ich meine- (.) oder das, was wir quasi seitdem gelernt haben, habe ich versucht darauf anzuwenden, (..) also worauf der, hier, ich weiß nicht, wie er hieß, achten sollte, beziehungsweise den Umgang und (.) wie er den Unterricht gestaltet. (...) Ja und beim ersten Mal hatte man eigentlich nur so, (..) ich sage mal, die, (...) ein paar Schülervorstellungen hatte man gehört, was man daraus- dazu beachten sollte und eben, wie man mit Geschlechterrollen umgehen soll. (..) Darauf konnte man sich so ein bisschen beziehen und jetzt hatte man ja schon ein bisschen mehr, (..) auch bei Herrn N, gehört wie man einen guten Unterricht plant. (..) Und (.) ja im Grunde fand ich es immer noch schwierig, weil (..) dieses- also das ist natürlich eine gestellte Situation, haben wir dann ja auch irgendwann gehört, aber das gesamte Bild ist irgendwie nicht stimmig und da irgendwie Hin- explizit Hinweise zu geben, was besser gemacht werden kann, wobei das Ganze (.) irgendwie schwierig ist, (lacht) ist dann doch trotzdem noch schwierig gewesen. Also wenn ich jetzt irgendwie Ausbilder wäre, wäre mir das, wäre mir das- würde mir das schwer fallen, irgendwie zu sagen, das hast du schon gut gemacht, (.) aber da würde ich jetzt erstmal anfangen, weil das sind ja viele Punkte, wo (.) wo es nicht gut läuft, sag ich mal. (.) Und dann auch- es gibt ja viele Auswahlmöglichkeiten, ich weiß nicht, wie das dann gewichtet ist, mit was da richtig ist und was eben falsch ist und dann einen Abzug gibt, (.) aber es gibt eben viele Punkte, wo man anfangen könnte. #00:06:13-5#
33. I: (...) Ja. #00:06:15-6#

34. C6: (.) War trotzdem noch irgendwie schwierig, das (..) zu reflektieren so. #00:06:20-6#
35. I: Ja. (.) Okay. Ähm ich habe auch wieder eine Beispielrückmeinung mitgebracht, damit man wieder irgendwie was Greifbares hat, (zum?) Erinnern. (.) (Ich?) packe die einmal aus. (.) Ähm genau. Willst du einmal kurz einen Moment Zeit haben, die anzuschauen? #00:06:37-0#
36. C6: Ja. #00:06:37-7#
37. I: (Dann?) (unv.) #00:06:38-3#
38. C6: (16s) Ich habe meine überhaupt nicht mehr im Hinterkopf, soll ich die mal (.) aufrufen oder brauchen wir das nicht? #00:06:58-8#
39. I: Also wir können im Folgenden auch (.) äh einfach bei der bleiben, das ist kein // Problem. #00:07:03-9#
40. C6: Okay. (.) Ach ja, du wolltest ja Rückmeldung dazu haben, (..) wie die- Zwecks Aufbau und- #00:07:09-3#
41. I: Ja, genau. Also ich habe quasi so ein paar Fragen, // die können wir gleich einfach durchgehen, aber wenn dir was einfällt, kannst du gerne alles sagen einfach. #00:07:16-0#
42. C6: Ja. #00:07:11-9#
43. C6: Ja. (14s) Ja. #00:07:30-3#
44. I: Okay. Ähm gut. Dann meine erste Frage. Hast du (.) schonmal mit solchen Boxplots, also mit so einer Art von Diagramm gearbeitet? #00:07:39-0#
45. C6: Ja, ziemlich genau, als ich letzte Mal die Rückmeldung erhalten habe. #00:07:42-4#
46. I: Äh aber abgesehen davon? #00:07:44-4#
47. C6: (.) Mhm ich glaube, letztes Mal habe ich nein gesagt, deswegen sollte ich jetzt auch nein sagen. #00:07:48-9#
48. I: Hätte ja in der Zwischenzeit- #00:07:50-3#
49. C6: Und seitdem in der Zwischenzeit auch nicht. Nein. #00:07:54-2#
50. I: Okay. (.) Ähm wenn du dir die Rückmeldung anschaust, (.) für den Moment quasi so tust, als wäre es deine. #00:08:02-0#
51. C6: Mh(Ja). #00:08:02-8#
52. I: Ähm welche Schlüsse würdest du daraus ziehen, wenn du die bekommst? #00:08:05-9#
53. C6: (5s) Dass noch Reflektionsbedarf besteht (5s) für die Person? #00:08:19-0#
54. I: (..) Mhm(Ja). #00:08:20-8#
55. C6: (...) Weil (er?) eben im unteren Bereich der fünfundzwanzig- (..) äh ja, im unteren Bereich der fünfundzwanzig hier, der Vergleichsgruppe ist. #00:08:30-7#
56. I: Äh beim bei dem Gesamtergebnis bist du da, ne? #00:08:33-2#
57. C6: Achso, // ja. #00:08:33-8#
58. I: Ja, // ja. #00:08:34-2#
59. C6: Hier auf mich bezogen. #00:08:34-9#
60. I: Achso, im Text vorne. // Ja. #00:08:36-4#
61. C6: Ja. #00:08:36-7#
62. I: Okay. (4s) Mhm(Ja). Und ähm welche Inhalte dieser Rückmeldung oder welche welche ähm Informationen sind da besonders hilfreich oder besonders interessant für dich? #00:08:51-7#
63. C6: (4s) Ja, wie man gegenüber den anderen sich- (.) Oder wie man gegenüber anderen abgeschnitten hat. (..) Und (.) was ich auch, glaube ich, letztes Mal schon gesagt habe, (was?) ich ganz gut fand, dass man eben guckt, wie weit (also?) wie weit ist man und wie weit sind die anderen. Also wie weit sind die anderen schon (.) im Vergleich zu dir? #00:09:14-0#
64. I: Also die diese // Fachsemester #00:09:16-8#
65. C6: Bezogen aufs Fachsemester, genau. #00:09:17-1#
66. I: Mhm(Ja). (.) // Okay. #00:09:18-3#

67. C6: Also man kann natürlich auch erste Semester mit achten Semestern vergleichen, aber (.) das bringt einem selber hier nicht so viel. (..) Weil es gibt ja so ein- (.) so eine Art Rückmeldung, wie weit man eigentlich sein sollte schon. (.) Denke ich mal, ne? (.) Also- (4s) (Ja?), und sonst äh (7s) hätte ich noch irgendwas sehen sollen, was mir das bringt? #00:09:46-4#
68. I: Nein, nein, es // gibt- #00:09:48-1#
69. C6: Okay. #00:09:48-4#
70. I: Also es gibt kein richtig und falsch, es geht um deine deine Perspektive einfach. #00:09:53-6#
71. C6: Ich glaube nicht, sonst habe ich daraus nicht so viel mitgenommen, sonst- #00:09:57-7#
72. I: Gut. (.) Ähm würdest du sagen, dass Inhalte aus dieser Rückmeldung überflüssig sind? #00:10:03-2#
73. C6: (10s) Ja, was ich vorab gut fand, dass erstmal so die Prozentzahl, die man erreicht hat und wie das dann eingeordnet ist am Anfang steht, das wäre nicht so gut, wenn das irgendwie am Ende stehen würde, weil man das natürlich als erstes wissen möchte. #00:10:26-5#
74. I: Mhm(Ja). #00:10:27-4#
75. C6: (..) Und sonst- Sie ist recht- Also (ist?) eine recht umfangreiche Rückmeldung. Sind ja auch fünf Sei- vier Seiten mit Deckblatt fünf. (..) Das hat wahrscheinlich alles seine Berechtigung. Wenn man sich das jetzt anguckt, ist natürlich alles genau beschrieben, (.) ähm (..) was man darunter zu verstehen hat, also es wird ja alles sehr gut eingeordnet. (...) Und ja, dadurch kann man eben auch ohne Vor- äh Vorwissen das wissen, wie man abgeschnitten hat. (...) Weiß ich nicht, kann ich jetzt jetzt nichts zu sagen, also wird ja alles (..) seine Daseinsberichtigung haben, (.) oder? #00:11:05-3#
76. I: Okay. (.) Ähm gibt es (.) Fragen, die für dich offen bleiben bei der Rückmeldung? #00:11:11-1#
77. C6: (7s) Mhm ja, was ich jetzt genau zum Beispiel tun kann, um besser zu werden, wird jetzt nicht beantwortet. (.) (Bleibe?) jetzt erstmal so bei meinem (.) Prozentsatz, bei meinem Wissenstand stehen. (..) Klar, ich könnte den Test jetzt nochmal machen, vielleicht wenn ich da einfach nochmal drauf drücke auf den, und das funktioniert. (..) Aber was ich jetzt genau falsch beantwortet habe, wird ja jetzt nicht- (also?) ist ja jetzt nicht ersichtlich, oder was ich richtig beantwortet habe. #00:11:42-5#
78. I: (..) Ja. (..) Also würde dich interessieren quasi eine Aufschlüsselung auf die (.) Aufgaben, verstehe ich das richtig, oder? (.) Meinst du (das?)- #00:11:55-1#
79. C6: Ja, vielleicht irgendwie sowas, oder wo man jetzt selbst gravierende Fehlvorstellungen hat. Also wenn man jetzt immer, kann ja sein, dass ich das seit letzten Mal immer noch nicht, (.) auch nicht besser weiß, und immer wieder das falsche ankreuze als Antwort, (.) weil ich denke, das ist das Richtige, aber das ist gerade so eine krasse Fehlvorstellung, die man irgendwie selber hat, (.) dass man das so machen muss, oder irgendwie so auslegen sollte in Unterricht (.) und das wird ja jetzt (so?) nicht aufgeschlüsselt. Das würde ich ja beim nächsten Mal wahrscheinlich dann (..) auch wieder so ankreuzen, denke ich mal, und auch vielleicht so selber machen. #00:12:26-4#
80. I: (...) Okay. Äh gibt es, ja, sonst noch andere Informationen, die du gerne gehabt hättest, (...) also darüber hinaus? #00:12:37-4#
81. C6: (...) Ich könnte mir auch vorstellen, wenn man jetzt (..) deutlich- also noch schlechter ist, dass man vielleicht- obwohl, mich würde es vielleicht auch interessieren, wie denn eben (.) so ein idealer Unterricht aussehen würde, oder eine mögliche Form des idealen Unterrichts, also wie so eine Art (.) Musterlösung wäre. #00:12:59-4#
82. I: (.) Also eine eine Musterlösung für den Test, was man wo ankreuzt, oder einfach so ein Best-Practice-Beispiel für guten Unterricht? #00:13:09-2#

83. C6: Eher so ein Best-Practice-Beispiel. Ich glaube, die Lösungs-, die liest man sich vielleicht die ersten drei, vier durch und dann (.) skippt man da eher doch nur so durch, (.) dann eher so, wie man es eigentlich (.) machen könnte, wo es gut funktioniert, vielleicht. #00:13:22-9#
84. I: (...) Okay. (..) Ähm (.) dann ist in der Meldung hier auf der vorletzten Seite, da ist so ein Fördermaterial verlinkt. (...) Hast du da mal äh reingeschaut beziehungsweise das bearbeitet? #00:13:41-7#
85. C6: Nein. #00:13:42-5#
86. I: Nein. #00:13:43-6#
87. C6: Ehrlich gesagt nur bis zur zweiten Seite gescrollt. #00:13:45-6#
88. I: Bis zur- Also zweiten- (.) also inklusive Diagramm für das // Gesamtergebnis. #00:13:51-5#
89. C6: Ja. #00:13:50-9#
90. I: Okay. (.) Ähm (.) das heißt, du hast es gar nicht wahrgenommen? #00:13:55-5#
92. I: Nein. Okay. (...) Dann (..) äh können wir wegkommen von der Rückmeldung, dem ganzen Material, einmal zurück zu dem Instrument selber. Du hast das ja jetzt mindestens das zweite Mal, ich weiß nicht, // wahrscheinlich das zweite Mal, ne? #00:14:13-1#
93. C6: Ja, zwei Mal. #00:14:12-5#
94. I: Äh bearbeitet. Woran hast du dich (.) vor der Bearbeitung des-, oder bei der Bearbeitung jetzt noch erinnert, (.) aus dem letzten Mal quasi? #00:14:21-9#
95. C6: (..) An die Personen. (...) Und hauptsächlich (.) den Umgang mit den Schülern und insbesondere die Mädchen. (..) Vorab. #00:14:33-6#
96. I: Mhm(Ja). #00:14:34-8#
97. C6: (.) Ja und als man dann- (.) Lag dann doch, glaube ich, ein oder anderthalb Jahre dazwischen, (..) ja, (also?) irgendwann kam es wieder, ah okay, wir machen heute Impuls und (.) die physikalischen Inhalte kamen dann wieder, (.) also was was er vorhat. (..) Ja, und die- diese Tabellen habe ich dann beim ersten Mal deutlich genauer angeguckt, als- also diese Unterrichtsplanung, was er wie vorhat, als beim zweiten Mal. (.) Weil es dann doch noch irgendwo im Hinterkopf war, so, irgendwann kam es dann wieder. #00:15:04-5#
98. I: Ja, also es ist quasi so ein- Ja, sorry- #00:15:07-5#
99. C6: Oder (.) ging es auch um den Test? (..) Ja. #00:15:11-4#
100. I: Ja, alles, gerne. #00:15:12-8#
101. C6: (...) Und dass es eben sehr viele Antwortmöglichkeiten gibt, wovon man nicht weiß, was richtig und falsch ist. #00:15:19-2#
102. I: Ja, okay. Ähm (du hast?) eben schon so ein bisschen angefangen, dass du gesagt hast, man liest vielleicht nicht mehr alles so detailliert durch, gerade die Tabellen. Ähm (.) inwiefern würdest du ansonsten sagen, hat das deine jetzige Bearbeitung beeinflusst, dass du das Instrument vorher schon mal bearbeitet hast? #00:15:35-7#
103. C6: (.) Ja, man wird ja ein bisschen- (..) man bearbeitet das ein bisschen schneller, ob das dann besser ist, anscheinend ja nicht, (..) sei mal dahingestellt, also man liest halt nur so ein bisschen flüchtiger und oberflächlicher. #00:15:49-3#
104. I: (..) Okay. #00:15:51-3#
105. C6: Weil man eben denkt, man weiß, was er vorhat und was er machen will. #00:15:54-7#
106. I: (.) Ja. #00:15:56-0#
107. C6: (.) Und worauf man zu achten hat, eigentlich. #00:15:58-3#
108. I: (...) Orientiert an den Sachen, an die man sich noch erinnert, meinst du, oder das, was dann wiederkommt? #00:16:05-0#
109. C6: Ja, man ist so ein bisschen voreingenommen, würde ich mal sagen, ne, // auch. #00:16:08-0#
110. I: Okay. (...) In Bezug auf bestimmte Aspekte oder (..) ist das eher so eine Grundhaltung? #00:16:15-5#

111. C6: (..) Mhm beides, vielleicht, würde ich sagen. Also die Grundhaltung war voreingenommen und dann, (..) man (.) man denkt schon so, okay, darauf sollte ich vielleicht achten, wenn er das jetzt macht, oder wie er das geplant hat. #00:16:29-0#
112. I: (..) Okay. Ähm dann (.) wäre werde ich durch mit meinen Fragen, hast du noch irgendwie Anmerkungen, Fragen, Kommentare? #00:16:37-9#
113. C6: Nein, bis zum nächsten Mal. #00:16:39-6#
114. I: (lacht) Genau. Super, dann mach ich einmal raus.

Q *R*-Skript für die Bestimmung der Beurteilerübereinstimmung zur Prüfung der Anforderung an Teil 2 des INA

Q.1 Vorbereitend: Pakete einfügen und Einlesen der Datei mit der Übersicht über die Kodierungen

Die eingelesene Datei enthält für beide Kodierungen die Häufigkeit des Auftretens jeder Kategorie (Spalten) für jedes Segment (Zeilen).

```
1 library(haven)
2 library(DescTools)
3 doppelkodierung <- read_sav("Uebereinstimmung.sav")
```

Q.2 Vorbereitung der Kreuztabellen zur Bestimmung von κ_{Cohen}

```
1  VL_K1 <- doppelkodierung$K1Vorlesen
2  VL_K2 <- doppelkodierung$K2Vorlesen
3  VL_l <- union(VL_K1, VL_K2)
4  VL_tab <- table(factor(VL_K1, VL_l), factor(VL_K2, VL_l))
5
6  Besch_K1 <- doppelkodierung$K1Beschreibung
7  Besch_K2 <- doppelkodierung$K2Beschreibung
8  Besch_l <- union(Besch_K1, Besch_K2)
9  Besch_tab <- table(factor(Besch_K1, Besch_l), factor(Besch_
     K2, Besch_l))
10
11 Bew_K1 <- doppelkodierung$K1Bewertung
```

```
12 Bew_K2 <- doppelkodierung$K2Bewertung
13 Bew_l <- union(Bew_K1, Bew_K2)
14 Bew_tab <- table(factor(Bew_K1, Bew_l), factor(Bew_K2, Bew_l
     ))
15
16 BewBegr_K1 <- doppelkodierung$K1BewertungemitBegründung
17 BewBegr_K2 <- doppelkodierung$K2BewertungemitBegründung
18 BewBegr_l <- union(BewBegr_K1, BewBegr_K2)
19 BewBegr_tab <- table(factor(BewBegr_K1, BewBegr_l), factor(
     BewBegr_K2, BewBegr_l))
20
21 Alt_K1 <- doppelkodierung$K1Alternative
22 Alt_K2 <- doppelkodierung$K2Alternative
23 Alt_l <- union(Alt_K1, Alt_K2)
24 Alt_tab <- table(factor(Alt_K1, Alt_l), factor(Alt_K2,Alt_l)
     )
25
26 AltBegr_K1 <- doppelkodierung$K1AlternativemitBegründung
27 AltBegr_K2 <- doppelkodierung$K2AlternativemitBegründung
28 AltBegr_l <- union(AltBegr_K1, AltBegr_K2)
29 AltBegr_tab <- table(factor(AltBegr_K1, AltBegr_l), factor(
     AltBegr_K2, AltBegr_l))
30
31 ZuAb_K1 <- doppelkodierung$K1Zustimmung
32 ZuAb_K2 <- doppelkodierung$K2Zustimmung
33 ZuAb_l <- union(ZuAb_K1, ZuAb_K2)
34 ZuAb_tab <- table(factor(ZuAb_K1, ZuAb_l), factor(ZuAb_K2,
     ZuAb_l))
35
36 ZuAbBegr_K1 <- doppelkodierung$K1ZustimmungmitBegründung
37 ZuAbBegr_K2 <- doppelkodierung$K2ZustimmungmitBegründung
38 ZuAbBegr_l <- union(ZuAbBegr_K1, ZuAbBegr_K2)
39 ZuAbBegr_tab <- table(factor(ZuAbBegr_K1, ZuAbBegr_l),
     factor(ZuAbBegr_K2, ZuAbBegr_l))
40
41 Ueberl_K1 <- doppelkodierung$K1ÜberlegungzumUnterricht
42 Ueberl_K2 <- doppelkodierung$K2ÜberlegungzumUnterricht
43 Ueberl_l <- union(Ueberl_K1, Ueberl_K2)
44 Ueberl_tab <- table(factor(Ueberl_K1, Ueberl_l), factor(
     Ueberl_K2, Ueberl_l))
45
46 Konstruktrel_K1 <- doppelkodierung$K1Konstruktrelevant
47 Konstruktrel_K2 <- doppelkodierung$K2Konstruktrelevant
48 Konstruktrel_l <- union(Konstruktrel_K1, Konstruktrel_K2)
49 Konstruktrel_tab <- table(factor(Konstruktrel_K1,
     Konstruktrel_l), factor(Konstruktrel_K2, Konstruktrel_l))
50
```

```
Sonstiges_K1 <- doppelkodierung$K1Sonstiges
Sonstiges_K2 <- doppelkodierung$K2Sonstiges
Sonstiges_l <- union(Sonstiges_K1, Sonstiges_K2)
Sonstiges_tab <- table(factor(Sonstiges_K1, Sonstiges_l),
   factor(Sonstiges_K2, Sonstiges_l))
```

Q.3 Bestimmung von κ_{Cohen} für die Kategorien

```
CohenKappa(VL_tab, conf.level = 0.95)
CohenKappa(Besch_tab, conf.level = 0.95)
CohenKappa(Bew_tab, conf.level = 0.95)
CohenKappa(BewBegr_tab, conf.level = 0.95)
CohenKappa(Alt_tab, conf.level = 0.95)
CohenKappa(AltBegr_tab, conf.level = 0.95)
CohenKappa(ZuAb_tab, conf.level = 0.95)
CohenKappa(ZuAbBegr_tab, conf.level = 0.95)
CohenKappa(Ueberl_tab, conf.level = 0.95)
CohenKappa(Konstruktrel_tab, conf.level = 0.95)
CohenKappa(Sonstiges_tab, conf.level = 0.95)
```

R *R*-Skript für die statischen Auswertungen zur Prüfung der Anforderung an Teil 3 des INA

R.1 Vorbereitend: Pakete einfügen und Aufbereitung der Daten aus dem Reflexionstest

```
1 library(haven)
2 library(tidyverse)
3 library(psych)
4 library(car)
5 library(rstatix)
6 library(rosetta)
```

Importieren der Daten des Reflexionstests sowie u. a. Bestimmung der Scores auf Aufgabenebene (R1Bew1 bis RGAlt2, ScoreMCNo1 bis ScoreMCNoG) bzw. Testebene (als Absolut- und Prozentwerte: ScoreMC bis ScoreMCSVProz) und Entfernung einer bekanntermaßen nicht ernsthaften Bearbeitung

```
1 results <- read.csv("./results-survey516586.csv", header=
    TRUE, sep="\t", encoding="UTF-8")
2 results[is.na(results)]<-0
3 results <- mutate(results,
4                   R1Bew1 = (results$Refl02bIIIBew.SQ002.+
                       results$Refl02bIIIBew.SQ003.+ results$
                       Refl02bIIIBew.SQ004.+ results$
                       Refl02bIIIBew.SQ007.)/4,
5                   R1Alt1 = (results$Refl02bIIIAlt.SQ002.+
                       results$Refl02bIIIAlt.SQ004.+ results$
                       Refl02bIIIAlt.SQ005.+ results$
                       Refl02bIIIAlt.SQ006.)/4,
6                   R2Bew1 = (results$Refl03bIVBeob.SQ003.+
                       results$Refl03bIVBeob.SQ004.+ results$
                       Refl03bIVBeob.SQ006.)/3,
```

```
7                   R2Alt1 = (results$Refl03IVUmgang.SQ003.+
                       results$Refl03IVUmgang.SQ004.+ results$
                       Refl03IVUmgang.SQ005.)/3,
8                   R3Bew1 = (results$Refl06bIVBew.SQ001.+
                       results$Refl06bIVBew.SQ002.+ results$
                       Refl06bIVBew.SQ003.+ results$Refl06bIVBew
                       .SQ005.)/4,
9                   R3Alt1 = (results$Refl06bIVAlt.SQ001.+
                       results$Refl06bIVAlt.SQ002.+ results$
                       Refl06bIVAlt.SQ003.+ results$Refl06bIVAlt
                       .SQ005.)/4,
10                  R3Bew2 = (results$Refl06bVBew.SQ001.+
                       results$Refl06bVBew.SQ005.+ results$
                       Refl06bVBew.SQ006.+ results$Refl06bVBew.
                       SQ007.)/4,
11                  R3Alt2 = (results$Refl06bVAlt.SQ002.+
                       results$Refl06bVAlt.SQ005.+ results$
                       Refl06bVAlt.SQ006.)/3,
12                  R4Bew1 = (results$Refl07bIVBew.SQ001.+
                       results$Refl07bIVBew.SQ002.+ results$
                       Refl07bIVBew.SQ003.+ results$Refl07bIVBew
                       .SQ004.)/4,
13                  R4Alt1 = (results$Refl07bIVAlt.SQ001.+
                       results$Refl07bIVAlt.SQ003.+ results$
                       Refl07bIVAlt.SQ005.)/3,
14                  R4Bew2 = (results$Refl07bVBew.SQ001.+
                       results$Refl07bVBew.SQ003.+ results$
                       Refl07bVBew.SQ005.+ results$Refl07bVBew.
                       SQ007.)/4,
15                  R4Alt2 = (results$Refl07bVAlt.SQ004.+
                       results$Refl07bVAlt.SQ005.+ results$
                       Refl07bVAlt.SQ006.+ results$Refl07bVAlt.
                       SQ007.)/4,
16                  R5Bew1 = (results$Refl09bIBew.SQ001.+
                       results$Refl09bIBew.SQ002.+ results$
                       Refl09bIBew.SQ004.)/3,
17                  R5Alt1 = (results$Refl09bIAlt.SQ003.+
                       results$Refl09bIAlt.SQ005.+ results$
                       Refl09bIAlt.SQ006.+ results$Refl09bIAlt.
                       SQ007.)/4,
18                  R5Bew2 = (results$Refl09bIVBew.SQ001.+
                       results$Refl09bIVBew.SQ002.+ results$
                       Refl09bIVBew.SQ004.+ results$Refl09bIVBew
                       .SQ005.)/4,
19                  R5Alt2 = (results$Refl09bIVAlt.SQ002.+
                       results$Refl09bIVAlt.SQ003.+ results$
```

```
                    Refl09bIVAlt.SQ005.+ results$Refl09bIVAlt
                    .SQ006.)/4,
20                  R5Bew3 = (results$Refl09bVIBew.SQ001.+
                    results$Refl09bVIBew.SQ004.+ results$
                    Refl09bVIBew.SQ007.)/3,
21                  R5Alt3 = (results$Refl09bVIAlt.SQ001.+
                    results$Refl09bVIAlt.SQ005.+ results$
                    Refl09bVIAlt.SQ007.)/3,
22                  R6Bew1 = (results$Refl12bIBew.SQ002.+
                    results$Refl12bIBew.SQ003.+ results$
                    Refl12bIBew.SQ004.)/3,
23                  R6Alt1 = (results$Refl12bIAlt.SQ002.+
                    results$Refl12bIAlt.SQ003.+ results$
                    Refl12bIAlt.SQ006.+ results$Refl12bIAlt.
                    SQ007.)/4,
24                  R6Bew2 = (results$Refl12bIIBew.SQ004.+
                    results$Refl12bIIBew.SQ006.+ results$
                    Refl12bIIBew.SQ007.)/3,
25                  R6Alt2 = (results$Refl12bIIAlt.SQ005.+
                    results$Refl12bIIAlt.SQ006.+ results$
                    Refl12bIIAlt.SQ007.)/3,
26                  R6Bew3 = (results$Refl12bIIIBew.SQ001.+
                    results$Refl12bIIIBew.SQ002.+ results$
                    Refl12bIIIBew.SQ005.)/3,
27                  R6Alt3 = (results$Refl12bIIIAlt.SQ002.+
                    results$Refl12bIIIAlt.SQ003.+ results$
                    Refl12bIIIAlt.SQ005.)/3,
28                  R7Bew1 = (results$Refl13bIBew.SQ003.+
                    results$Refl13bIBew.SQ004.)/2,
29                  R7Alt1 = (results$Refl13bIAlt.SQ002.+
                    results$Refl13bIAlt.SQ004.+ results$
                    Refl13bIAlt.SQ006.+ results$Refl13bIAlt.
                    SQ007.)/4,
30                  R7Bew2 = (results$Refl13bIIBew.SQ002.+
                    results$Refl13bIIBew.SQ003.+ results$
                    Refl13bIIBew.SQ004.+ results$Refl13bIIBew
                    .SQ007.)/4,
31                  R7Alt2 = (results$Refl13bIIAlt.SQ002.+
                    results$Refl13bIIAlt.SQ004.+ results$
                    Refl13bIIAlt.SQ007.)/3,
32                  RGBew1 = (results$GesbVIBew.SQ002.+ results
                    $GesbVIBew.SQ003.+ results$GesbVIBew.
                    SQ005.+ results$GesbVIBew.SQ007.+ results
                    $GesbVIBew.SQ008.)/5,
33                  RGAlt1 = (results$GesVIAlt.SQ002.+ results$
                    GesVIAlt.SQ004.+ results$GesVIAlt.SQ005.+
                     results$GesVIAlt.SQ007.)/4,
```

```
34                  RGBew2 = (results$GesVIIBew.SQ001.+ results
                       $GesVIIBew.SQ002.+ results$GesVIIBew.
                       SQ005.)/3,
35                  RGAlt2 = (results$GesVIIAlt.SQ002.+ results
                       $GesVIIAlt.SQ003.+ results$GesVIIAlt.
                       SQ005.)/3,
36                  ScoreMCRefl1 = (Refl02bIIIBew.SQ002.+
                       Refl02bIIIBew.SQ003.+ Refl02bIIIBew.SQ004
                       .+ Refl02bIIIBew.SQ007.)/4+ (
                       Refl02bIIIAlt.SQ002.+ Refl02bIIIAlt.SQ004
                       .+ Refl02bIIIAlt.SQ005.+ Refl02bIIIAlt.
                       SQ006.)/4,
37                  ScoreMCRefl2 = (Refl03bIVBeob.SQ003.+
                       Refl03bIVBeob.SQ004.+ Refl03bIVBeob.SQ006
                       .)/3+ (Refl03IVUmgang.SQ003.+
                       Refl03IVUmgang.SQ004.+ Refl03IVUmgang.
                       SQ005.)/3,
38                  ScoreMCRefl3 = (Refl06bIVBew.SQ001.+
                       Refl06bIVBew.SQ002.+ Refl06bIVBew.SQ003.+
                        Refl06bIVBew.SQ005.)/4+ (Refl06bIVAlt.
                       SQ001. +Refl06bIVAlt.SQ002. +Refl06bIVAlt
                       .SQ003. +Refl06bIVAlt.SQ005.)/4+ (
                       Refl06bVBew.SQ001.+ Refl06bVBew.SQ005.+
                       Refl06bVBew.SQ006.+ Refl06bVBew.SQ007.)/
                       4+ (Refl06bVAlt.SQ002.+ Refl06bVAlt.SQ005
                       .+ Refl06bVAlt.SQ006.)/3,
39                  ScoreMCRefl4 = (Refl07bIVBew.SQ001.+
                       Refl07bIVBew.SQ002.+ Refl07bIVBew.SQ003.+
                        Refl07bIVBew.SQ004.)/4+ (Refl07bIVAlt.
                       SQ001.+ Refl07bIVAlt.SQ003.+ Refl07bIVAlt
                       .SQ005.)/3+ (Refl07bVBew.SQ001.+
                       Refl07bVBew.SQ003.+ Refl07bVBew.SQ005.+
                       Refl07bVBew.SQ007.)/4+ (Refl07bVAlt.SQ004
                       .+ Refl07bVAlt.SQ005.+ Refl07bVAlt.SQ006
                       .+ Refl07bVAlt.SQ007.)/4,
40                  ScoreMCRefl5 = (Refl09bIBew.SQ001.+
                       Refl09bIBew.SQ002.+ Refl09bIBew.SQ004.)/
                       3+ (Refl09bIAlt.SQ003.+ Refl09bIAlt.SQ005
                       .+ Refl09bIAlt.SQ006.+ Refl09bIAlt.SQ007
                       .)/4+ (Refl09bIVBew.SQ001.+ Refl09bIVBew.
                       SQ002.+ Refl09bIVBew.SQ004.+ Refl09bIVBew
                       .SQ005.)/4+ (Refl09bIVAlt.SQ002.+
                       Refl09bIVAlt.SQ003.+ Refl09bIVAlt.SQ005.+
                        Refl09bIVAlt.SQ006.)/4+ (Refl09bVIBew.
                       SQ001.+ Refl09bVIBew.SQ004.+ Refl09bVIBew
                       .SQ007.)/3+ (Refl09bVIAlt.SQ001.+
```

```
                    Refl09bVIAlt.SQ005.+ Refl09bVIAlt.SQ007.)
                    /3,
41                  ScoreMCRefl6 = (Refl12bIBew.SQ002.+
                    Refl12bIBew.SQ003.+ Refl12bIBew.SQ004.)/
                    3+ (Refl12bIAlt.SQ002.+ Refl12bIAlt.SQ003
                    .+ Refl12bIAlt.SQ006.+ Refl12bIAlt.SQ007
                    .)/4+ (Refl12bIIBew.SQ004.+ Refl12bIIBew.
                    SQ006.+ Refl12bIIBew.SQ007.)/3+ (
                    Refl12bIIAlt.SQ005.+ Refl12bIIAlt.SQ006.+
                     Refl12bIIAlt.SQ007.)/3+ (Refl12bIIIBew.
                    SQ001.+ Refl12bIIIBew.SQ002.+
                    Refl12bIIIBew.SQ005.)/3+ (Refl12bIIIAlt.
                    SQ002.+ Refl12bIIIAlt.SQ003.+
                    Refl12bIIIAlt.SQ005.)/3,
42                  ScoreMCRefl7 = (Refl13bIBew.SQ003.+
                    Refl13bIBew.SQ004.)/2+ (Refl13bIAlt.SQ002
                    .+ Refl13bIAlt.SQ004.+ Refl13bIAlt.SQ006
                    .+ Refl13bIAlt.SQ007.)/4+ (Refl13bIIBew.
                    SQ002.+ Refl13bIIBew.SQ003.+ Refl13bIIBew
                    .SQ004.+ Refl13bIIBew.SQ007.)/4+ (
                    Refl13bIIAlt.SQ002.+ Refl13bIIAlt.SQ004.+
                     Refl13bIIAlt.SQ007.)/3,
43                  ScoreMCReflG = (GesbVIBew.SQ002.+ GesbVIBew
                    .SQ003.+ GesbVIBew.SQ005.+ GesbVIBew.
                    SQ007.+ GesbVIBew.SQ008.)/5+ (GesVIAlt.
                    SQ002.+ GesVIAlt.SQ004.+ GesVIAlt.SQ005.+
                     GesVIAlt.SQ007.)/4+ (GesVIIBew.SQ001.+
                    GesVIIBew.SQ002.+ GesVIIBew.SQ005.)/3+ (
                    GesVIIAlt.SQ002.+ GesVIIAlt.SQ003.+
                    GesVIIAlt.SQ005.)/3,
44                  ScoreMC = ScoreMCRefl1+ ScoreMCRefl2+
                    ScoreMCRefl3+ ScoreMCRefl4+ ScoreMCRefl5+
                     ScoreMCRefl6+ ScoreMCRefl7+ ScoreMCReflG
                    ,
45                  ScoreMCProz = ScoreMC/32*100,
46                  ScoreMCBew = (Refl02bIIIBew.SQ002.+
                    Refl02bIIIBew.SQ003.+ Refl02bIIIBew.SQ004
                    .)/4+ (Refl06bIVBew.SQ001.+ Refl06bIVBew.
                    SQ002.+ Refl06bIVBew.SQ003.+ Refl06bIVBew
                    .SQ005.)/4+ (Refl06bVBew.SQ001.+
                    Refl06bVBew.SQ005.+ Refl06bVBew.SQ006.+
                    Refl06bVBew.SQ007.)/4+ (Refl07bIVBew.
                    SQ001.+ Refl07bIVBew.SQ002.+ Refl07bIVBew
                    .SQ003.+ Refl07bIVBew.SQ004.)/4+ (
                    Refl07bVBew.SQ001.+ Refl07bVBew.SQ003.+
                    Refl07bVBew.SQ005.+ Refl07bVBew.SQ007.)/
                    4+ (Refl09bIBew.SQ001.+ Refl09bIBew.SQ002
```

```
                .+ Refl09bIBew.SQ004.)/3+ (Refl09bIVBew.
                SQ001.+ Refl09bIVBew.SQ002.+ Refl09bIVBew
                .SQ004.+ Refl09bIVBew.SQ005.)/4+ (
                Refl09bVIBew.SQ001.+ Refl09bVIBew.SQ004.+
                 Refl09bVIBew.SQ007.)/3+ (Refl12bIBew.
                SQ002.+ Refl12bIBew.SQ003.+ Refl12bIBew.
                SQ004.)/3+ (Refl12bIIBew.SQ004.+
                Refl12bIIBew.SQ006.+ Refl12bIIBew.SQ007.)
                /3+ (Refl12bIIIBew.SQ001.+ Refl12bIIIBew.
                SQ002.+ Refl12bIIIBew.SQ005.)/3+ (
                Refl13bIBew.SQ003.+ Refl13bIBew.SQ004.)/
                2+ (Refl13bIIBew.SQ002.+ Refl13bIIBew.
                SQ003.+ Refl13bIIBew.SQ004.+ Refl13bIIBew
                .SQ007.)/4+ (GesbVIBew.SQ002.+ GesbVIBew.
                SQ003.+ GesbVIBew.SQ005.+ GesbVIBew.SQ007
                .+ GesbVIBew.SQ008.)/5+ (GesVIIBew.SQ001
                .+ GesVIIBew.SQ002.+ GesVIIBew.SQ005.)/3,
47              ScoreMCBewProz = ScoreMCBew/15*100,
48              ScoreMCAlt = (Refl02bIIIAlt.SQ002.+
                Refl02bIIIAlt.SQ004.+ Refl02bIIIAlt.SQ005
                .+ Refl02bIIIAlt.SQ006.)/4+ (Refl06bIVAlt
                .SQ001.+ Refl06bIVAlt.SQ002.+
                Refl06bIVAlt.SQ003.+ Refl06bIVAlt.SQ005.)
                /4+ (Refl06bVAlt.SQ002.+ Refl06bVAlt.
                SQ005.+ Refl06bVAlt.SQ006.)/3+ (
                Refl07bIVAlt.SQ001.+ Refl07bIVAlt.SQ003.+
                 Refl07bIVAlt.SQ005.)/3+ (Refl07bVAlt.
                SQ004.+ Refl07bVAlt.SQ005.+ Refl07bVAlt.
                SQ006.+ Refl07bVAlt.SQ007.)/4+ (
                Refl09bIAlt.SQ003.+ Refl09bIAlt.SQ005.+
                Refl09bIAlt.SQ006.+ Refl09bIAlt.SQ007.)/
                4+ (Refl09bIVAlt.SQ002.+ Refl09bIVAlt.
                SQ003.+ Refl09bIVAlt.SQ005.+ Refl09bIVAlt
                .SQ006.)/4+ (Refl09bVIAlt.SQ001.+
                Refl09bVIAlt.SQ005.+Refl09bVIAlt.SQ007.)/
                3+ (Refl12bIAlt.SQ002.+ Refl12bIAlt.SQ003
                .+ Refl12bIAlt.SQ006.+ Refl12bIAlt.SQ007
                .)/4+ (Refl12bIIAlt.SQ005.+ Refl12bIIAlt.
                SQ006.+ Refl12bIIAlt.SQ007.)/3+ (
                Refl12bIIIAlt.SQ002.+ Refl12bIIIAlt.SQ003
                .+ Refl12bIIIAlt.SQ005.)/3+ (Refl13bIAlt.
                SQ002.+ Refl13bIAlt.SQ004.+ Refl13bIAlt.
                SQ006.+ Refl13bIAlt.SQ007.)/4+ (
                Refl13bIIAlt.SQ002.+ Refl13bIIAlt.SQ004.+
                 Refl13bIIAlt.SQ007.)/3+ (GesVIAlt.SQ002
                .+ GesVIAlt.SQ004.+ GesVIAlt.SQ005.+
```

```
                    GesVIAlt.SQ007.)/4+ (GesVIIAlt.SQ002.+
                    GesVIIAlt.SQ003.+ GesVIIAlt.SQ005.)/3,
49                  ScoreMCAltProz = ScoreMCAlt/15*100,
50                  ScoreMCSV = (Refl03bIVBeob.SQ003.+
                    Refl03bIVBeob.SQ004.+ Refl03bIVBeob.SQ006
                    .)/3+ (Refl03IVUmgang.SQ003.+
                    Refl03IVUmgang.SQ004.+ Refl03IVUmgang.
                    SQ005.)/3,
51                  ScoreMCSVProz = ScoreMCSV/2*100,
52                  ScoreMCNo1 = (Refl02b.SQ001.+ Refl02b.SQ002
                    .+ Refl02b.SQ003.+ Refl02b.SQ004.+
                    Refl02b.SQ005.)/5,
53                  ScoreMCNo2 = (Refl03b.SQ001.+ Refl03b.SQ002
                    .+ Refl03b.SQ003.)/3,
54                  ScoreMCNo3 = (Refl06b.SQ001.+ Refl06b.SQ003
                    .+ Refl06b.SQ004.+ Refl06b.SQ005.)/4,
55                  ScoreMCNo4 = (Refl07b.SQ001.+ Refl07b.SQ002
                    .+ Refl07b.SQ004.+ Refl07b.SQ005.)/4,
56                  ScoreMCNo5 = (Refl09b.SQ001.+ Refl09b.SQ002
                    .+ Refl09b.SQ003.+ Refl09b.SQ004.+
                    Refl09b.SQ006.+ Refl09b.SQ007.)/6,
57                  ScoreMCNo6 = (Refl12b.SQ001.+ Refl12b.SQ002
                    .+ Refl12b.SQ003.)/3,
58                  ScoreMCNo7 = (Refl13b.SQ001.+ Refl13b.SQ002
                    .)/2,
59                  ScoreMCNoG = (Gesb.SQ001.+ Gesb.SQ002.+
                    Gesb.SQ003.+ Gesb.SQ006.+ Gesb.SQ007.)/5,
60                  ScoreMCNo = ScoreMCNo1 + ScoreMCNo2 +
                    ScoreMCNo3 + ScoreMCNo4 + ScoreMCNo5 +
                    ScoreMCNo6 + ScoreMCNo7 + ScoreMCNoG,
61                  PC = paste(toupper(results$PersCode.PC1.),
                    toupper(results$PersCode.PC2.), toupper(
                    results$PersCode.PC3.), toupper(results$
                    PersCode.PC4.), sep=""))
62  results = results[-21,]
```

Vereinheitlichen der Schreibweise der Persönlichen Codes sowie Umbenennung zweier falscher Codes

```
1  PC1 <- toupper(results$PersCode.PC1.)
2  PC1[69] <- "X"
3  PC1[78] <- "X"
4  PC2 <- toupper(results$PersCode.PC2.)
5  PC2[69] <- "X"
6  PC2[78] <- "X"
7  PC3 <- toupper(results$PersCode.PC3.)
8  PC3[69] <- "X"
```

```
PC3[78] <- "X"
PC4 <- toupper(results$PersCode.PC4.)
PC4[69] <- "1"
PC4[78] <- "2"
```

Erstellung von reduzierten Dataframes für Bachelor- und Masterstudierende

```
results_BA <- subset(results[(results$Phase=="A1"),])
results_MA <- subset(results[(results$Phase=="A2"),])
```

R.2 Zu Indikator 3.1: Untersuchung der Verteilung und Lösungshäufigkeiten

R.2.1 Prüfung der Normalverteilung

Plot: Histogramm der erreichten Scores (Gesamtscore sowie Teilscores für Bewertungen und Alternativen) in Prozent mit Normalverteilung

```
ybreaks = seq(0,30,5)
ggplot(results, aes(x=ScoreMCProz))+
  geom_histogram(aes(y=..density..), binwidth=5, colour = "
     black", fill = "white")+
  xlim(0,100)+
  stat_function(fun=dnorm, args=list(mean=mean(results$
     ScoreMCProz), sd=sd(results$ScoreMCProz)), color="black"
     , size=1)+
  xlab("Testwerte␣in␣%␣(zusammengefasst␣in␣Intervalle␣mit␣
     einer␣Breite␣von␣5%)")+
  theme_bw()+
  theme(axis.text=element_text(size=12))+
  scale_y_continuous(Häufigkeit, breaks = round(ybreaks/(5*
     192), 3), labels = ybreaks)
```

Ergänzende Berechnung der Kolmogorow-Smirnow-Tets und beschreibender Statistiken der Verteilung für die Gesamtstichprobe sowie Bachelor- und Masterstudierende

```
ks.test(results$ScoreMCProz, "pnorm", mean=mean(results$
   ScoreMCProz), sd=sd(results$ScoreMCProz))
ks.test(results$ScoreMCBewProz, "pnorm", mean=mean(results$
   ScoreMCBewProz), sd=sd(results$ScoreMCBewProz))
ks.test(results$ScoreMCAltProz, "pnorm", mean=mean(results$
   ScoreMCAltProz), sd=sd(results$ScoreMCAltProz))
```

```
4
5  describe(results$ScoreMCProz)
6  describe(results$ScoreMCBewProz)
7  describe(results$ScoreMCAltProz)
8
9  ks.test(results_BA$ScoreMCProz, "pnorm", mean=mean(results_
       BA$ScoreMCProz), sd=sd(results_BA$ScoreMCProz))
10 ks.test(results_BA$ScoreMCBewProz, "pnorm", mean=mean(
       results_BA$ScoreMCBewProz), sd=sd(results_BA$
       ScoreMCBewProz))
11 ks.test(results_BA$ScoreMCAltProz, "pnorm", mean=mean(
       results_BA$ScoreMCAltProz), sd=sd(results_BA$
       ScoreMCAltProz))
12
13 describe(results_BA$ScoreMCProz)
14 describe(results_BA$ScoreMCBewProz)
15 describe(results_BA$ScoreMCAltProz)
16
17 ks.test(results_MA$ScoreMCProz, "pnorm", mean=mean(results_
       MA$ScoreMCProz), sd=sd(results_MA$ScoreMCProz))
18 ks.test(results_MA$ScoreMCBewProz, "pnorm", mean=mean(
       results_MA$ScoreMCBewProz), sd=sd(results_MA$
       ScoreMCBewProz))
19 ks.test(results_MA$ScoreMCAltProz, "pnorm", mean=mean(
       results_MA$ScoreMCAltProz), sd=sd(results_MA$
       ScoreMCAltProz))
20
21 describe(results_MA$ScoreMCProz)
22 describe(results_MA$ScoreMCBewProz)
23 describe(results_MA$ScoreMCAltProz)
```

R.2.2 Betrachtung der Lösungshäufigkeiten für die Gesamtstichprobe sowie Bachelor- und Masterstudierende

```
1  LH <- c(mean(results$R1Bew1), mean(results$R1Alt1), mean(
       results$R2Bew1), mean(results$R2Alt1), mean(results$R3Bew1
       ), mean(results$R3Alt1), mean(results$R3Bew2), mean(
       results$R3Alt2), mean(results$R4Bew1), mean(results$R4Alt1
       ), mean(results$R4Bew2), mean(results$R4Alt2), mean(
       results$R5Bew1), mean(results$R5Alt1), mean(results$R5Bew2
       ), mean(results$R5Alt2), mean(results$R5Bew3), mean(
       results$R5Alt3), mean(results$R6Bew1), mean(results$R6Alt1
       ), mean(results$R6Bew2), mean(results$R6Alt2), mean(
```

```
   results$R6Bew3), mean(results$R6Alt3), mean(results$R7Bew1
   ), mean(results$R7Alt1), mean(results$R7Bew2), mean(
   results$R7Alt2), mean(results$RGBew1), mean(results$RGAlt1
   ), mean(results$RGBew2), mean(results$RGAlt2))
LH
mean(LH)
sd(LH)
min(LH)
max(LH)

LH_BA <- c(mean(results_BA$R1Bew1), mean(results_BA$R1Alt1),
    mean(results_BA$R2Bew1), mean(results_BA$R2Alt1), mean(
   results_BA$R3Bew1), mean(results_BA$R3Alt1), mean(results_
   BA$R3Bew2), mean(results_BA$R3Alt2), mean(results_BA$
   R4Bew1), mean(results_BA$R4Alt1), mean(results_BA$R4Bew2),
    mean(results_BA$R4Alt2), mean(results_BA$R5Bew1), mean(
   results_BA$R5Alt1), mean(results_BA$R5Bew2), mean(results_
   BA$R5Alt2), mean(results_BA$R5Bew3), mean(results_BA$
   R5Alt3), mean(results_BA$R6Bew1), mean(results_BA$R6Alt1),
    mean(results_BA$R6Bew2), mean(results_BA$R6Alt2), mean(
   results_BA$R6Bew3), mean(results_BA$R6Alt3), mean(results_
   BA$R7Bew1), mean(results_BA$R7Alt1), mean(results_BA$
   R7Bew2), mean(results_BA$R7Alt2), mean(results_BA$RGBew1),
    mean(results_BA$RGAlt1), mean(results_BA$RGBew2), mean(
   results_BA$RGAlt2))
LH_BA
mean(LH_BA)
sd(LH_BA)
min(LH_BA)
max(LH_BA)

LH_MA <- c(mean(results_MA$R1Bew1), mean(results_MA$R1Alt1),
    mean(results_MA$R2Bew1), mean(results_MA$R2Alt1), mean(
   results_MA$R3Bew1), mean(results_MA$R3Alt1), mean(results_
   MA$R3Bew2), mean(results_MA$R3Alt2), mean(results_MA$
   R4Bew1), mean(results_MA$R4Alt1), mean(results_MA$R4Bew2),
    mean(results_MA$R4Alt2), mean(results_MA$R5Bew1), mean(
   results_MA$R5Alt1), mean(results_MA$R5Bew2), mean(results_
   MA$R5Alt2), mean(results_MA$R5Bew3), mean(results_MA$
   R5Alt3), mean(results_MA$R6Bew1), mean(results_MA$R6Alt1),
    mean(results_MA$R6Bew2), mean(results_MA$R6Alt2), mean(
   results_MA$R6Bew3), mean(results_MA$R6Alt3), mean(results_
   MA$R7Bew1), mean(results_MA$R7Alt1), mean(results_MA$
   R7Bew2), mean(results_MA$R7Alt2), mean(results_MA$RGBew1),
    mean(results_MA$RGAlt1), mean(results_MA$RGBew2), mean(
   results_MA$RGAlt2))
LH_MA
```

```
mean(LH_MA)
sd(LH_MA)
min(LH_MA)
max(LH_MA)
```

R.3 Zu Indikator 3.2: Untersuchung der Reliabilität und Trennschärfen

Berechnung von Cronbachs Alpha (über alle Aufgaben bzw. die Aufgaben zu Bewertungen und Alternativen) für die Gesamtstichprobe sowie Bachelor- und Masterstudierende

R.3.1 Über alle Aufgaben

```
alpha_ges <- psych::alpha(subset(results, select = c("R1Bew1","R1Alt1","R2Bew1", "R2Alt1", "R3Bew1", "R3Alt1", "R3Bew2", "R3Alt2", "R4Bew1", "R4Alt1", "R4Bew2", "R4Alt2", "R5Bew1", "R5Alt1", "R5Bew2", "R5Alt2", "R5Bew3", "R5Alt3", "R6Bew1", "R6Alt1", "R6Bew2", "R6Alt2", "R6Bew3", "R6Alt3", "R7Bew1", "R7Alt1", "R7Bew2", "R7Alt2", "RGBew1", "RGAlt1", "RGBew2", "RGAlt2")), check.keys =TRUE)
alpha_ges
mean(alpha_ges$item.stats$r.drop)
min(alpha_ges$item.stats$r.drop)
max(alpha_ges$item.stats$r.drop)

alpha_ges_BA <- psych::alpha(subset(results_BA, select = c("R1Bew1","R1Alt1","R2Bew1", "R2Alt1", "R3Bew1", "R3Alt1", "R3Bew2", "R3Alt2", "R4Bew1", "R4Alt1", "R4Bew2", "R4Alt2", "R5Bew1", "R5Alt1", "R5Bew2", "R5Alt2", "R5Bew3", "R5Alt3", "R6Bew1", "R6Alt1", "R6Bew2", "R6Alt2", "R6Bew3", "R6Alt3", "R7Bew1", "R7Alt1", "R7Bew2", "R7Alt2", "RGBew1", "RGAlt1", "RGBew2", "RGAlt2")), check.keys =TRUE)
alpha_ges_BA
mean(alpha_ges_BA$item.stats$r.drop)
min(alpha_ges_BA$item.stats$r.drop)
max(alpha_ges_BA$item.stats$r.drop)

alpha_ges_MA <- psych::alpha(subset(results_MA, select = c("R1Bew1","R1Alt1","R2Bew1", "R2Alt1", "R3Bew1", "R3Alt1", "R3Bew2", "R3Alt2", "R4Bew1", "R4Alt1", "R4Bew2", "R4Alt2", "R5Bew1", "R5Alt1", "R5Bew2", "R5Alt2", "R5Bew3", "R5Alt3
```

```
   ", "R6Bew1", "R6Alt1", "R6Bew2", "R6Alt2", "R6Bew3", "
   R6Alt3", "R7Bew1", "R7Alt1", "R7Bew2", "R7Alt2", "RGBew1",
    "RGAlt1", "RGBew2", "RGAlt2")), check.keys =TRUE)
alpha_ges_MA
mean(alpha_ges_MA$item.stats$r.drop)
min(alpha_ges_MA$item.stats$r.drop)
max(alpha_ges_MA$item.stats$r.drop)
```

R.3.2 Jeweils über die Aufgaben zu Bewertungen und Alternativen

```
alpha_Bew <- psych::alpha(subset(results, select = c("R1Bew1
   ", "R3Bew1", "R3Bew2", "R4Bew1", "R4Bew2", "R5Bew1", "
   R5Bew2", "R5Bew3", "R6Bew1", "R6Bew2", "R6Bew3", "R7Bew1",
    "R7Bew2",  "RGBew1", "RGBew2")), check.keys =TRUE)
alpha_Bew
mean(alpha_Bew$item.stats$r.drop)
min(alpha_Bew$item.stats$r.drop)
max(alpha_Bew$item.stats$r.drop)

alpha_Bew_BA <- psych::alpha(subset(results_BA, select = c("
   R1Bew1", "R3Bew1", "R3Bew2", "R4Bew1", "R4Bew2", "R5Bew1",
    "R5Bew2", "R5Bew3", "R6Bew1", "R6Bew2", "R6Bew3", "R7Bew1
   ", "R7Bew2",  "RGBew1", "RGBew2")), check.keys =TRUE)
alpha_Bew_BA
mean(alpha_Bew_BA$item.stats$r.drop)
min(alpha_Bew_BA$item.stats$r.drop)
max(alpha_Bew_BA$item.stats$r.drop)

alpha_Bew_MA <- psych::alpha(subset(results_MA, select = c("
   R1Bew1", "R3Bew1", "R3Bew2", "R4Bew1", "R4Bew2", "R5Bew1",
    "R5Bew2", "R5Bew3", "R6Bew1", "R6Bew2", "R6Bew3", "R7Bew1
   ", "R7Bew2",  "RGBew1", "RGBew2")), check.keys =TRUE)
alpha_Bew_MA
mean(alpha_Bew_MA$item.stats$r.drop)
min(alpha_Bew_MA$item.stats$r.drop)
max(alpha_Bew_MA$item.stats$r.drop)
```

```
alpha_Alt <- psych::alpha(subset(results, select = c("R1Alt1
   ", "R3Alt1", "R3Alt2", "R4Alt1", "R4Alt2", "R5Alt1", "
   R5Alt2", "R5Alt3", "R6Alt1", "R6Alt2", "R6Alt3", "R7Alt1",
    "R7Alt2", "RGAlt1", "RGAlt2")), check.keys =TRUE)
alpha_Alt
mean(alpha_Alt$item.stats$r.drop)
```

```
4  min(alpha_Alt$item.stats$r.drop)
5  max(alpha_Alt$item.stats$r.drop)
6
7  alpha_Alt_BA <- psych::alpha(subset(results_BA, select = c("
     R1Alt1", "R3Alt1", "R3Alt2", "R4Alt1", "R4Alt2", "R5Alt1",
      "R5Alt2", "R5Alt3", "R6Alt1", "R6Alt2", "R6Alt3", "R7Alt1
     ", "R7Alt2", "RGAlt1", "RGAlt2")), check.keys =TRUE)
8  alpha_Alt_BA
9  mean(alpha_Alt_BA$item.stats$r.drop)
10 min(alpha_Alt_BA$item.stats$r.drop)
11 max(alpha_Alt_BA$item.stats$r.drop)
12
13 alpha_Alt_MA <- psych::alpha(subset(results_MA, select = c("
     R1Alt1", "R3Alt1", "R3Alt2", "R4Alt1", "R4Alt2", "R5Alt1",
      "R5Alt2", "R5Alt3", "R6Alt1", "R6Alt2", "R6Alt3", "R7Alt1
     ", "R7Alt2", "RGAlt1", "RGAlt2")), check.keys =TRUE)
14 alpha_Alt_MA
15 mean(alpha_Alt_MA$item.stats$r.drop)
16 min(alpha_Alt_MA$item.stats$r.drop)
17 max(alpha_Alt_MA$item.stats$r.drop)
```

R.4 Zu Indikator 3.3: Explorative Faktorenanalyse

```
1  results_FA_Aufgaben <- subset(results, select = c("R1Bew1","
     R1Alt1","R2Bew1", "R2Alt1", "R3Bew1", "R3Alt1", "R3Bew2",
     "R3Alt2", "R4Bew1", "R4Alt1", "R4Bew2", "R4Alt2", "R5Bew1"
     , "R5Alt1", "R5Bew2", "R5Alt2", "R5Bew3", "R5Alt3", "
     R6Bew1", "R6Alt1", "R6Bew2", "R6Alt2", "R6Bew3", "R6Alt3",
      "R7Bew1", "R7Alt1", "R7Bew2", "R7Alt2", "RGBew1", "RGAlt1
     ", "RGBew2", "RGAlt2"))
2  cortest.bartlett(results_FA_Aufgaben)
3  KMO(results_FA_Aufgaben)
```

Untersuche Anzahl an Faktoren

```
1  nfactors(results_FA_Aufgaben, rotate="varimax", fm="mle")
2  fa.parallel(results_FA_Aufgaben, fa="both")
3  principal(results_FA_Aufgaben, nfactors=2, rotate="varimax")
4  scree(results_FA_Aufgaben, factors=TRUE, pc=FALSE, hline=-2,
       main="")
```

R.5 Zu Indikator 3.4: Vergleich mit dem Reflexionsperformanztest

R.5.1 Aufbereiten der Daten des Reflexionsperformanztests

Importieren der Daten aus dem Performanztest sowie Bestimmung der Scores auf Videoebene (ScorePRefl1 bis ScorePReflG), bzw. Testebene (ab Gesamtscore (ScoreP))

```
1  Performanz <- read_sav("Ergebnisse-Performanztest.sav")
2  Performanz[is.na(Performanz)]<-0
3  Performanz <- {mutate(Performanz,
4                        ScorePRefl1 = R_pre_1.1.1+ R_pre_1.1.2+
                             R_pre_1.1.3+ R_pre_1.1.4,
5                        ScorePRefl2 = R_pre_2.1.1+ R_pre_2.2.1+
                             R_pre_2.2.2+ R_pre_2.3.1+ R_pre_
                            2.3.2+ R_pre_2.3.4,
6                        ScorePRefl3 = R_pre_3.1.1a+ R_pre_3.1.1
                            aU+ R_pre_3.1.1b+ R_pre_3.1.1bU+ R_
                            pre_3.1.1c+ R_pre_3.1.1cU+ R_pre_
                            3.1.1d+ R_pre_3.2.1+ R_pre_3.2.2+ R_
                            pre_3.2.3,
7                        ScorePRefl4 = R_pre_4.1.1+ R_pre_4.2.1+
                             R_pre_4.2.2,
8                        ScorePRefl5 = R_pre_5.1.1+ R_pre_5.2.1+
                             R_pre_5.2.2,
9                        ScorePRefl6 = R_pre_6.1.1+ R_pre_6.2.1a
                            + R_pre_6.2.1b+ R_pre_6.2.1c+ R_pre_
                            6.3.1,
10                       ScorePRefl7 = R_pre_7.1.1+ R_pre_7.2.1+
                             R_pre_7.2.2+ R_pre_7.3.1+ R_pre_
                            7.4.1+ R_pre_7.5.1+ R_pre_7.5.2+ R_
                            pre_7.6.1+ R_pre_7.6.2+ R_pre_7.6.3,
11                       ScorePRefl8 = R_pre_8.1.1+ R_pre_8.1.2+
                             R_pre_8.2.1+ R_pre_8.3.1+ R_pre_
                            8.3.2+ R_pre_8.3.3+ R_pre_8.3.4+ R_
                            pre_8.4.1,
12                       ScorePRefl9 = R_pre_9.1.1+ R_pre_9.1.2+
                             R_pre_9.1.3+ R_pre_9.1.4+ R_pre_
                            9.1.5+ R_pre_9.2.1+ R_pre_9.3.1+ R_
                            pre_9.4.1+ R_pre_9.5.1+ R_pre_9.6.1,
13                       ScorePRefl10 = R_pre_10.1.1+ R_pre_
                            10.1.2+ R_pre_10.2.1+ R_pre_10.3.1,
14                       ScorePRefl11 = R_pre_11.1.1+ R_pre_
                            11.2.1+ R_pre_11.2.2+ R_pre_11.3.1+ R
```

```
                    _pre_11.4.1+ R_pre_11.4.2+ R_pre_
                    11.5.1,
15                  ScorePRefl12 = R_pre_12.1.1+ R_pre_
                    12.2.1+ R_pre_12.2.2+ R_pre_12.3.1+ R
                    _pre_12.4.1+ R_pre_12.5.1,
16                  ScorePRefl13 = R_pre_13.1.2+ R_pre_
                    13.2.1+ R_pre_13.3.1+ R_pre_13.4.1+ R
                    _pre_13.5.1+ R_pre_13.6.1,
17                  ScorePReflG = R_pre_Ges.1.1+ R_pre_Ges
                    .1.2+ R_pre_Ges.1.3+ R_pre_Ges.2.1+ R
                    _pre_Ges.2.2+ R_pre_Ges.3.1+ R_pre_
                    Ges.4.1+ R_pre_Ges.5.1+ R_pre_Ges
                    .6.1+ R_pre_Ges.6.2+ R_pre_Ges.7.1+ R
                    _pre_Ges.8.1,
18                  ScoreP = ScorePRefl1+ ScorePRefl2+
                    ScorePRefl3+ ScorePRefl4+ ScorePRefl5
                    + ScorePRefl6+ ScorePRefl7+
                    ScorePRefl8+ ScorePRefl9+
                    ScorePRefl10+ ScorePRefl11+
                    ScorePRefl12+ ScorePRefl13+
                    ScorePReflG,
19                  ScorePVid = ScorePRefl2+ ScorePRefl3+
                    ScorePRefl6+ ScorePRefl7+ ScorePRefl9
                    + ScorePRefl12+ ScorePRefl13+
                    ScorePReflG,
20                  ScorePfd = R_pre_1.1.1+ R_pre_1.1.4+
21                    R_pre_2.1.1+
22                    R_pre_3.1.1a+ R_pre_3.1.1aU+ R_pre_
                      3.1.1b+ R_pre_3.1.1bU+ R_pre_3.1.1c
                      + R_pre_3.1.1cU+ R_pre_3.1.1d+
23                    R_pre_4.1.1+
24                    R_pre_6.2.1a+ R_pre_6.2.1c+ R_pre_
                      6.3.1+
25                    R_pre_7.2.2+ R_pre_7.3.1+ R_pre_
                      7.4.1+ R_pre_7.6.1+ R_pre_7.6.2+ R_
                      pre_7.6.3+
26                    R_pre_8.1.1+ R_pre_8.1.2+ R_pre_
                      8.3.1+ R_pre_8.3.2+ R_pre_8.3.3+ R_
                      pre_8.3.4+ R_pre_8.4.1+
27                    R_pre_9.1.1+ R_pre_9.1.5+ R_pre_
                      9.2.1+ R_pre_9.4.1+ R_pre_9.6.1+
28                    R_pre_10.1.1+ R_pre_10.1.2+ R_pre_
                      10.3.1+
29                    R_pre_11.1.1+ R_pre_11.3.1+ R_pre_
                      11.4.2+
30                    R_pre_12.1.1+ R_pre_12.3.1+
```

```
31                          R_pre_13.2.1+ R_pre_13.4.1+ R_pre_
                               13.5.1+ R_pre_13.6.1+
32                          R_pre_Ges.1.1+ R_pre_Ges.1.2+ R_pre_
                               Ges.1.3+ R_pre_Ges.2.1+ R_pre_Ges
                               .2.2+ R_pre_Ges.5.1+ R_pre_Ges.6.1+
                                R_pre_Ges.6.2+ R_pre_Ges.8.1,
33                        ScorePf = R_pre_2.2.1+ R_pre_2.2.2+
34                          R_pre_5.1.1+
35                          R_pre_6.1.1+ R_pre_6.2.1b+
36                          R_pre_7.1.1+ R_pre_7.5.1+ R_pre_
                               7.5.2+
37                          R_pre_8.2.1+
38                          R_pre_9.1.3+ R_pre_9.1.4+ R_pre_
                               9.5.1+
39                          R_pre_11.4.1+
40                          R_pre_12.2.1+ R_pre_12.2.2+
41                          R_pre_13.3.1+
42                          R_pre_Ges.7.1,
43                        ScorePap = R_pre_1.1.2+ R_pre_1.1.3+
44                          R_pre_2.3.1+ R_pre_2.3.2+ R_pre_
                               2.3.4+
45                          R_pre_3.2.1+ R_pre_3.2.2+ R_pre_
                               3.2.3+
46                          R_pre_4.2.1+ R_pre_4.2.2+
47                          R_pre_5.2.1+ R_pre_5.2.2+
48                          R_pre_7.2.1+
49                          R_pre_9.1.2+ R_pre_9.3.1+
50                          R_pre_10.2.1+
51                          R_pre_11.2.1+ R_pre_11.2.2+ R_pre_
                               11.5.1+
52                          R_pre_12.4.1+ R_pre_12.5.1+
53                          R_pre_13.1.2+
54                          R_pre_Ges.3.1+ R_pre_Ges.4.1,
55                        )}
```

Zusammenfügen der Ergebnisse von Performanz- und MC-Test für die Personen, die beide Instrumente bearbeitet haben, in einen Dataframe (für die folgenden Berechnungen lediglich Berücksichtigung der Personen, die beide Instrumente bearbeitet haben)

```
1  Vergleich <- data.frame(
2    Performanz[-c(13,23),],
3    rbind(results[111,], results[113,], results[121,], results
       [117,], results[132,], results[127,], results[118,],
       results[122,], results[145,], results[126,], results
       [119,], results[114,], results[136,], results[146,],
       results[133,], results[139,], results[148,], results
```

```
    [140,], results[150,], results[134,], results[137,],
    results[147,], results[141,], results[144,], results
    [143,], results[149,], results[135,], results[138,],
    results[142,], results[156,], results[157,], results
    [167,], results[162,], results[158,], results[164,],
    results[165,], results[166,], results[160,], results
    [168,])
)
```

R.5.2 Prüfen von Voraussetzungen und Skalengüte

Prüfen der Normalverteilung für beide Instrumente und die Gesamtstichprobe sowie die Teilstichproben (Masterstudierende (1), Grundschullehramtsstudierende (2) und Bachelorstudierende (3))

Performanztest

Gesamtscore

```
ks.test(Vergleich$ScoreP, "pnorm", mean=mean(Vergleich$
    ScoreP), sd=sd(Vergleich$ScoreP))
ks.test(Vergleich[Vergleich$Gruppe=="3",]$ScoreP, "pnorm",
    mean=mean(Vergleich[Vergleich$Gruppe=="3",]$ScoreP), sd=sd
    (Vergleich[Vergleich$Gruppe=="3",]$ScoreP))
ks.test(Vergleich[Vergleich$Gruppe=="1",]$ScoreP, "pnorm",
    mean=mean(Vergleich[Vergleich$Gruppe=="1",]$ScoreP), sd=sd
    (Vergleich[Vergleich$Gruppe=="1",]$ScoreP))
ks.test(Vergleich[Vergleich$Gruppe=="2",]$ScoreP, "pnorm",
    mean=mean(Vergleich[Vergleich$Gruppe=="2",]$ScoreP), sd=sd
    (Vergleich[Vergleich$Gruppe=="2",]$ScoreP))
```

Gesamtscore im Performanztest nur unter Berücksichtigung der in beiden Instrumenten vorkommenden Videos (nur Gesamtstichprobe)

```
ks.test(Vergleich$ScorePVid, "pnorm", mean=mean(Vergleich$
    ScorePVid), sd=sd(Vergleich$ScorePVid))
```

Teilscores zum fachlichen, fachdidaktischen und pädagogischen Wissen

```
ks.test(Vergleich$ScorePf, "pnorm", mean=mean(Vergleich$
    ScorePf), sd=sd(Vergleich$ScorePf))
ks.test(Vergleich$ScorePfd, "pnorm", mean=mean(Vergleich$
    ScorePfd), sd=sd(Vergleich$ScorePfd))
```

```
3 ks.test(Vergleich$ScorePap, "pnorm", mean=mean(Vergleich$
    ScorePap), sd=sd(Vergleich$ScorePap))
4
5 ks.test(Vergleich[Vergleich$Gruppe=="3",]$ScorePf, "pnorm",
    mean=mean(Vergleich[Vergleich$Gruppe=="3",]$ScorePf), sd=
    sd(Vergleich[Vergleich$Gruppe=="3",]$ScorePf))
6 ks.test(Vergleich[Vergleich$Gruppe=="3",]$ScorePfd, "pnorm",
     mean=mean(Vergleich[Vergleich$Gruppe=="3",]$ScorePfd), sd
    =sd(Vergleich[Vergleich$Gruppe=="3",]$ScorePfd))
7 ks.test(Vergleich[Vergleich$Gruppe=="3",]$ScorePap, "pnorm",
     mean=mean(Vergleich[Vergleich$Gruppe=="3",]$ScorePap), sd
    =sd(Vergleich[Vergleich$Gruppe=="3",]$ScorePap))
8
9 ks.test(Vergleich[Vergleich$Gruppe=="1",]$ScorePf, "pnorm",
    mean=mean(Vergleich[Vergleich$Gruppe=="1",]$ScorePf), sd=
    sd(Vergleich[Vergleich$Gruppe=="1",]$ScorePf))
10 ks.test(Vergleich[Vergleich$Gruppe=="1",]$ScorePfd, "pnorm",
     mean=mean(Vergleich[Vergleich$Gruppe=="1",]$ScorePfd), sd
    =sd(Vergleich[Vergleich$Gruppe=="1",]$ScorePfd))
11 ks.test(Vergleich[Vergleich$Gruppe=="1",]$ScorePap, "pnorm",
     mean=mean(Vergleich[Vergleich$Gruppe=="1",]$ScorePap), sd
    =sd(Vergleich[Vergleich$Gruppe=="1",]$ScorePap))
12
13 ks.test(Vergleich[Vergleich$Gruppe=="2",]$ScorePf, "pnorm",
    mean=mean(Vergleich[Vergleich$Gruppe=="2",]$ScorePf), sd=
    sd(Vergleich[Vergleich$Gruppe=="2",]$ScorePf))
14 ks.test(Vergleich[Vergleich$Gruppe=="2",]$ScorePfd, "pnorm",
     mean=mean(Vergleich[Vergleich$Gruppe=="2",]$ScorePfd), sd
    =sd(Vergleich[Vergleich$Gruppe=="2",]$ScorePfd))
15 ks.test(Vergleich[Vergleich$Gruppe=="2",]$ScorePap, "pnorm",
     mean=mean(Vergleich[Vergleich$Gruppe=="2",]$ScorePap), sd
    =sd(Vergleich[Vergleich$Gruppe=="2",]$ScorePap))
```

Reflexionstest

```
1 ks.test(Vergleich$ScoreMC, "pnorm", mean=mean(Vergleich$
    ScoreMC), sd=sd(Vergleich$ScoreMC))
2 ks.test(Vergleich[Vergleich$Gruppe=="3",]$ScoreMC, "pnorm",
    mean=mean(Vergleich[Vergleich$Gruppe=="3",]$ScoreMC), sd=
    sd(Vergleich[Vergleich$Gruppe=="3",]$ScoreMC))
3 ks.test(Vergleich[Vergleich$Gruppe=="1",]$ScoreMC, "pnorm",
    mean=mean(Vergleich[Vergleich$Gruppe=="1",]$ScoreMC), sd=
    sd(Vergleich[Vergleich$Gruppe=="2",]$ScoreMC))
```

```
4 ks.test(Vergleich[Vergleich$Gruppe=="2",]$ScoreMC, "pnorm",
    mean=mean(Vergleich[Vergleich$Gruppe=="2",]$ScoreMC), sd=
    sd(Vergleich[Vergleich$Gruppe=="2",]$ScoreMC))
```

Untersuchung der Reliabilität für beide Instrumente (Skalenanalyse mit Chronbachs-Alpha) für den Gesamtscore im Performanztest und den Gesamtscore im Performanztest nur unter Berücksichtigung der in beiden Tests vorkommenden Videos (in dieser Reihenfolge) auf Ebene der Reflexionspunkte

```
1 psych::alpha(subset(Vergleich, select = c("R_pre_1.1.1","R_
    pre_1.1.2","R_pre_1.1.3","R_pre_1.1.4","R_pre_2.1.1","R_
    pre_2.2.1","R_pre_2.2.2","R_pre_2.3.1","R_pre_2.3.2","R_
    pre_2.3.4","R_pre_3.1.1a","R_pre_3.1.1aU","R_pre_3.1.1b","
    R_pre_3.1.1bU","R_pre_3.1.1c","R_pre_3.1.1cU","R_pre_3.1.1
    d","R_pre_3.2.1","R_pre_3.2.2","R_pre_3.2.3","R_pre_4.1.1"
    ,"R_pre_4.2.1","R_pre_4.2.2","R_pre_5.1.1","R_pre_5.2.1","
    R_pre_5.2.2","R_pre_6.1.1","R_pre_6.2.1a","R_pre_6.2.1b","
    R_pre_6.2.1c","R_pre_6.3.1","R_pre_7.1.1","R_pre_7.2.1","R
    _pre_7.2.2","R_pre_7.3.1","R_pre_7.4.1","R_pre_7.5.1","R_
    pre_7.5.2","R_pre_7.6.1","R_pre_7.6.2","R_pre_7.6.3","R_
    pre_8.1.1","R_pre_8.1.2","R_pre_8.2.1","R_pre_8.3.1","R_
    pre_8.3.2","R_pre_8.3.3","R_pre_8.3.4","R_pre_8.4.1","R_
    pre_9.1.1","R_pre_9.1.2","R_pre_9.1.3","R_pre_9.1.4","R_
    pre_9.1.5","R_pre_9.2.1","R_pre_9.3.1","R_pre_9.4.1","R_
    pre_9.5.1","R_pre_9.6.1","R_pre_10.1.1","R_pre_10.1.2","R_
    pre_10.2.1","R_pre_10.3.1","R_pre_11.1.1","R_pre_11.2.1","
    R_pre_11.2.2","R_pre_11.3.1","R_pre_11.4.1","R_pre_11.4.2"
    ,"R_pre_11.5.1","R_pre_12.1.1","R_pre_12.2.1","R_pre_
    12.2.2","R_pre_12.3.1","R_pre_12.4.1","R_pre_12.5.1","R_
    pre_13.1.2","R_pre_13.2.1","R_pre_13.3.1","R_pre_13.4.1","
    R_pre_13.5.1","R_pre_13.6.1","R_pre_Ges.1.1","R_pre_Ges
    .1.2","R_pre_Ges.1.3","R_pre_Ges.2.1","R_pre_Ges.2.2","R_
    pre_Ges.3.1","R_pre_Ges.4.1","R_pre_Ges.5.1","R_pre_Ges
    .6.1","R_pre_Ges.6.2","R_pre_Ges.7.1","R_pre_Ges.8.1")),
    check.keys =TRUE)
2 psych::alpha(subset(Vergleich, select = c("R_pre_2.1.1","R_
    pre_2.2.1","R_pre_2.2.2","R_pre_2.3.1","R_pre_2.3.2","R_
    pre_2.3.4","R_pre_3.1.1a","R_pre_3.1.1aU","R_pre_3.1.1b","
    R_pre_3.1.1bU","R_pre_3.1.1c","R_pre_3.1.1cU","R_pre_3.1.1
    d","R_pre_3.2.1","R_pre_3.2.2","R_pre_3.2.3","R_pre_6.1.1"
    ,"R_pre_6.2.1a","R_pre_6.2.1b","R_pre_6.2.1c","R_pre_6.3.1
    ","R_pre_7.1.1","R_pre_7.2.1","R_pre_7.2.2","R_pre_7.3.1",
    "R_pre_7.4.1","R_pre_7.5.1","R_pre_7.5.2","R_pre_7.6.1","R
    _pre_7.6.2","R_pre_7.6.3","R_pre_9.1.1","R_pre_9.1.2","R_
    pre_9.1.3","R_pre_9.1.4","R_pre_9.1.5","R_pre_9.2.1","R_
    pre_9.3.1","R_pre_9.4.1","R_pre_9.5.1","R_pre_9.6.1","R_
    pre_12.1.1","R_pre_12.2.1","R_pre_12.2.2","R_pre_12.3.1","
```

```
   R_pre_12.4.1","R_pre_12.5.1","R_pre_13.1.2","R_pre_13.2.1"
   ,"R_pre_13.3.1","R_pre_13.4.1","R_pre_13.5.1","R_pre_
   13.6.1","R_pre_Ges.1.1","R_pre_Ges.1.2","R_pre_Ges.1.3","R
   _pre_Ges.2.1","R_pre_Ges.2.2","R_pre_Ges.3.1","R_pre_Ges
   .4.1","R_pre_Ges.5.1","R_pre_Ges.6.1","R_pre_Ges.6.2","R_
   pre_Ges.7.1","R_pre_Ges.8.1")), check.keys =TRUE)
```

Untersuchung der Reliabilität für beide Instrumente (Skalenanalyse mit Chronbachs-Alpha) für den Gesamtscore im MC-Test auf Aufgabenebene

```
1 psych::alpha(subset(Vergleich, select = c("R1Bew1","R1Alt1",
   "R2Bew1", "R2Alt1", "R3Bew1", "R3Alt1", "R3Bew2", "R3Alt2"
   , "R4Bew1", "R4Alt1", "R4Bew2", "R4Alt2", "R5Bew1", "
   R5Alt1", "R5Bew2", "R5Alt2", "R5Bew3", "R5Alt3", "R6Bew1",
    "R6Alt1", "R6Bew2", "R6Alt2", "R6Bew3", "R6Alt3", "R7Bew1
   ", "R7Alt1", "R7Bew2", "R7Alt2", "RGBew1", "RGAlt1", "
   RGBew2", "RGAlt2")), check.keys =TRUE)
```

R.5.3 Berechnung von Korrelationen zwischen den Tests

Untersuchung der Korrelationen zwischen dem Reflexionstest und dem Reflexionsperformanztest (Gesamtscore sowie Score nur unter Berücksichtigung der in beiden Tests vorkommenden Videos)

```
1 cor.test(Vergleich$ScoreP, Vergleich$ScoreMC, method = "
   pearson")
2 cor.test(Vergleich$ScorePVid, Vergleich$ScoreMC, method = "
   pearson")
```

R.5.4 Untersuchung von Gruppenunterschieden (Masterstudierende (1), Grundschullehramtsstudierende (2), Bachelorstudierende (3))

Voraussetzung für die Untersuchung von Mittelwertsunterschieden zwischen den Kohorten mit einer ANOVA ist neben der Normalverteilung (s.o.) Varianzhomogenität mit dem Levene-Test für Gesamtscores beider Instrumente und Teilscores im Performanztest

```
1 Vergleich$Gruppef <- as.factor(Vergleich$Gruppe)
2
3 leveneTest(Vergleich$ScoreP, Vergleich$Gruppef)
4 leveneTest(Vergleich$ScorePfd, Vergleich$Gruppef)
```

```
leveneTest(Vergleich$ScorePf, Vergleich$Gruppef)
leveneTest(Vergleich$ScorePap, Vergleich$Gruppef)
leveneTest(Vergleich$ScoreMC, Vergleich$Gruppef)
```

Berechnung der ANOVAs

```
Anova(aov(Vergleich$ScoreP ~ Vergleich$Gruppef), type="3")
Anova(aov(Vergleich$ScorePfd ~ Vergleich$Gruppef), type="3")
Anova(aov(Vergleich$ScorePf ~ Vergleich$Gruppef), type="3")
Anova(aov(Vergleich$ScorePap ~ Vergleich$Gruppef), type="3")
Anova(aov(Vergleich$ScoreMCProz ~ Vergleich$Gruppef), type="
   3")
```

Post-hoc-Analyse zu den Mittelwertsunterschieden bzgl. der Testwerte zum Fachwissen im Reflexionsperformanztest

```
pairwise.t.test(Vergleich$ScorePf, Vergleich$Gruppe, p.
   adjust="bonferroni")
Vergleich %>%
  cohens_d(ScorePf ~ Gruppe) %>%
  as.data.frame
```

R.6 Indikator 3.5: Zusammenhänge zu Lerngelegenheiten/demografischen Variablen

R.6.1 Quasi-längsschnittliche Betrachtungen: Zusammenhänge zwischen Gesamtscore/Teilergebnissen und Indikatoren für Lerngelegenheiten

Nachfolgend werden jeweils Voraussetzungstests sowie passend dazu Tests auf Zusammenhänge/Unterschiede zu den verschiedenen Variablen berechnet.

Zusammenhang zur Ausbildungsphase

Unterschiede zwischen Teilnehmenden im Bachelorstudium, Masterstudium und Referendariat (Ausschluss der Phasen Staatsexamen und Sonstiges, da keine entsprechende Zuordnung vorgenommen werden kann))

```
results_Phase <- subset(results[(results$Phase!="A3" &
   results$Phase!="-oth-"),])
FS_Phase <- as.numeric(gsub("A", "", results_Phase$FSPhy))

```

```
4 ks.test(results_Phase[results_Phase$Phase=="A1",]$ScoreMC, "
    pnorm", mean=mean(results_Phase[results_Phase$Phase=="A1"
    ,]$ScoreMC), sd=sd(results_Phase[results_Phase$Phase=="A1"
    ,]$ScoreMC))
5 ks.test(results_Phase[results_Phase$Phase=="A2",]$ScoreMC, "
    pnorm", mean=mean(results_Phase[results_Phase$Phase=="A2"
    ,]$ScoreMC), sd=sd(results_Phase[results_Phase$Phase=="A2"
    ,]$ScoreMC))
6 ks.test(results_Phase[results_Phase$Phase=="A4",]$ScoreMC, "
    pnorm", mean=mean(results_Phase[results_Phase$Phase=="A4"
    ,]$ScoreMC), sd=sd(results_Phase[results_Phase$Phase=="A4"
    ,]$ScoreMC))
7 leveneTest(results_Phase$ScoreMC, results_Phase$Phase)
8
9 welch_anova_test(results_Phase, ScoreMC ~ as.factor(Phase))
10 posthocTGH(results_Phase$ScoreMC, as.factor(results_Phase$
    Phase), method="games-howell", p.adjust = "bonferroni")
```

Zusammenhang zum Fachsemester

Ausschluss der Phasen Sonstiges und Referendariat, da kein Semester zugeornet werden kann

```
1 results_FS <- subset(results[(results$Phase!="A4" & results$
    Phase!="-oth-"),])
2 FS_FS <- as.numeric(gsub("A", "", results_FS$FSPhy))
3
4 ks.test(results_FS$ScoreMC, "pnorm", mean=mean(results_FS$
    ScoreMC), sd=sd(results_FS$ScoreMC))
5 ks.test(FS_FS, "pnorm", mean=mean(FS_FS), sd=sd(FS_FS))
6
7 cor.test(results_FS$ScoreMC, FS_FS, method = "spearman")
```

Zusätzliche Betrachtung des Zusammenhangs zum Fachsemester unter Ausschluss der Sachunterrichtsstudierenden

```
1 results_FS_ISSU <- subset(results[(results$Phase!="A4" &
    results$Phase!="-oth-" & results$id!="300" & results$id!="
    302" & results$id!="303" & results$id!="305" & results$id!
    ="306" & results$id!="307" & results$id!="308" & results$
    id!="309" & results$id!="310" & results$id!="311" &
    results$id!="312" & results$id!="314" & results$id!="316"
    & results$id!="317" & results$id!="318" & results$id!="319
    " & results$id!="320" & results$id!="342" & results$id!="
    343"),])
```

```
FS_FS_ISSU <- as.numeric(gsub("A", "", results_FS_ISSU$FSPhy))

ks.test(results_FS_ISSU$ScoreMC, "pnorm", mean=mean(results_FS_ISSU$ScoreMC), sd=sd(results_FS_ISSU$ScoreMC))
ks.test(FS_FS_ISSU, "pnorm", mean=mean(FS_FS_ISSU), sd=sd(FS_FS_ISSU))

cor.test(results_FS_ISSU$ScoreMC, FS_FS_ISSU, method = "spearman")
```

Zusammenhang zwischen Scores und Praktika

Für alle vorliegenden Datensätze

```
results_Praktika <- subset(results[(results$Phase!="A4" & results$Phase!="-oth-"),])
FS_Praktika <- as.numeric(gsub("A", "", results_Praktika$FSPhy))
```

Gruppenvergleich: Absolvierte Pflichtpraktika ja(1)/nein(0)

```
ks.test(results_Praktika[results_Praktika$Pflichtpraktika==1,]$ScoreMC, "pnorm", mean=mean(results_Praktika[results_Praktika$Pflichtpraktika==1,]$ScoreMC), sd=sd(results_Praktika[results_Praktika$Pflichtpraktika==1,]$ScoreMC))
ks.test(results_Praktika[results_Praktika$Pflichtpraktika==0,]$ScoreMC, "pnorm", mean=mean(results_Praktika[results_Praktika$Pflichtpraktika==0,]$ScoreMC), sd=sd(results_Praktika[results_Praktika$Pflichtpraktika==0,]$ScoreMC))

\#\# parametrisch
leveneTest(results_Praktika$ScoreMC ~ as.factor(results_Praktika$Pflichtpraktika))
t.test(results_Praktika$ScoreMC ~ results_Praktika$Pflichtpraktika, var.equal=TRUE)
cohens_d(results_Praktika, ScoreMC ~ Pflichtpraktika)
\#\# non-parametrisch
wilcox.test(ScoreMC ~ as.factor(Pflichtpraktika), data = results_Praktika)
```

Untersuchung des Zusammenhangs zum Umfang der absolvierten Pflichtpraktika (Anzahl der Wochen)

```
ks.test(results_Praktika[results_Praktika$Pflichtpraktika
   ==1,]$WoPflichtpraktika, "pnorm", mean=mean(results_
   Praktika[results_Praktika$Pflichtpraktika==1,]$
   WoPflichtpraktika), sd=sd(results_Praktika[results_
   Praktika$Pflichtpraktika==1,]$WoPflichtpraktika))
ks.test(results_Praktika$WoPflichtpraktika, "pnorm", mean=
   mean(results_Praktika$WoPflichtpraktika), sd=sd(results_
   Praktika$WoPflichtpraktika))

\#\# non-parametrisch
cor.test(results_Praktika[results_Praktika$Pflichtpraktika
   ==1,]$ScoreMC, results_Praktika[results_Praktika$
   Pflichtpraktika==1,]$WoPflichtpraktika, method = "spearman
   ")
```

Gruppenvergleich: Absolvierte freiwillige Schulpraktika ja(1)/nein(0)

```
ks.test(results_Praktika[results_Praktika$SonstigePraktika
   ==1,]$ScoreMC, "pnorm", mean=mean(results_Praktika[results
   _Praktika$SonstigePraktika==1,]$ScoreMC), sd=sd(results_
   Praktika[results_Praktika$SonstigePraktika==1,]$ScoreMC))
ks.test(results_Praktika[results_Praktika$SonstigePraktika
   ==0,]$ScoreMC, "pnorm", mean=mean(results_Praktika[results
   _Praktika$SonstigePraktika==0,]$ScoreMC), sd=sd(results_
   Praktika[results_Praktika$SonstigePraktika==0,]$ScoreMC))

\#\# parametrisch
leveneTest(results_Praktika$ScoreMC ~ as.factor(results_
   Praktika$SonstigePraktika))
t.test(results_Praktika$ScoreMC ~ results_Praktika$
   SonstigePraktika, var.equal=TRUE)
\#\# non-parametrisch
wilcox.test(ScoreMC ~ as.factor(SonstigePraktika), data =
   results_Praktika)
```

Zusammenhang zwischen Scores und anderen schulischen Lehrerfahrungen

Gruppenvergleich: Arbeit als Vertretungslehrkraft ja(1)/nein(0)

```
ks.test(results_Praktika[results_Praktika$VertretungsLK==1,]
   $ScoreMC, "pnorm", mean=mean(results_Praktika[results_
   Praktika$VertretungsLK==1,]$ScoreMC), sd=sd(results_
   Praktika[results_Praktika$VertretungsLK==1,]$ScoreMC))
```

```
ks.test(results_Praktika[results_Praktika$VertretungsLK==0,]
   $ScoreMC, "pnorm", mean=mean(results_Praktika[results_
   Praktika$VertretungsLK==0,]$ScoreMC), sd=sd(results_
   Praktika[results_Praktika$VertretungsLK==0,]$ScoreMC))

\#\# parametrisch
leveneTest(results_Praktika$ScoreMC, results_Praktika$
   VertretungsLK)
t.test(results_Praktika$ScoreMC ~ results_Praktika$
   VertretungsLK, var.equal=FALSE)
\#\# non-parametrisch
wilcox.test(ScoreMC ~ VertretungsLK, data = results_Praktika
   )
```

Sonstiges Lehrerfahrung

Gruppenvergleich: Außerschulische Lehrerfahrung ja(1)/nein(0)

```
ks.test(results_Praktika[results_Praktika$SonstigeLehre==1,]
   $ScoreMC, "pnorm", mean=mean(results_Praktika[results_
   Praktika$SonstigeLehre==1,]$ScoreMC), sd=sd(results_
   Praktika[results_Praktika$SonstigeLehre==1,]$ScoreMC))
ks.test(results_Praktika[results_Praktika$SonstigeLehre==0,]
   $ScoreMC, "pnorm", mean=mean(results_Praktika[results_
   Praktika$SonstigeLehre==0,]$ScoreMC), sd=sd(results_
   Praktika[results_Praktika$SonstigeLehre==0,]$ScoreMC))

\#\# parametrisch
leveneTest(results_Praktika$ScoreMC, results_Praktika$
   SonstigeLehre)
t.test(results_Praktika$ScoreMC ~ results_Praktika$
   SonstigeLehre, var.equal = FALSE)
cohens_d(results_Praktika, ScoreMC ~ SonstigeLehre, var.
   equal = FALSE)
\#\# non-parametrisch
wilcox.test(ScoreMC ~ SonstigeLehre, data = results_Praktika
   )
```

Gruppenvergleich: Universitäre Lehrerfahrung ja(1)/nein(0)

```
ks.test(results_Praktika[results_Praktika$Unilehre==1,]$
   ScoreMC, "pnorm", mean=mean(results_Praktika[results_
   Praktika$Unilehre==1,]$ScoreMC), sd=sd(results_Praktika[
   results_Praktika$Unilehre==1,]$ScoreMC))
ks.test(results_Praktika[results_Praktika$Unilehre==0,]$
   ScoreMC, "pnorm", mean=mean(results_Praktika[results_
```

```
   Praktika$Unilehre==0,]$ScoreMC), sd=sd(results_Praktika[
   results_Praktika$Unilehre==0,]$ScoreMC))
3
4 \#\# parametrisch
5 leveneTest(results_Praktika$ScoreMC, results_Praktika$
   Unilehre)
6 t.test(results_Praktika$ScoreMC ~ results_Praktika$Unilehre,
    var.equal=TRUE)
7 cohens_d(results_Praktika, ScoreMC ~ Unilehre)
```

Gruppenvergleich: Lehramtstyp

Physik-Lehramtsstudierende(0)/Sachunterrichtsstudierende(1)

```
1 results_Demographie <- subset(results[(results$Phase!="A4" &
    results$Phase!="-oth-"),])
2 results_Demographie_ISSU <- mutate(results_Demographie,
3                                    ISSU = case_when(id=="300
                                       " | id=="302" | id=="
                                       303" | id=="305" | id==
                                       "306" | id=="307" | id
                                       =="308" | id=="309" |
                                       id=="310" | id=="311" |
                                        id=="312" | id=="314"
                                       | id=="316" | id=="317"
                                        | id=="318" | id=="319
                                       " | id=="320" | id=="
                                       342" | id=="343" ~ 1,
4                                                    TRUE ~
                                                       0))
5
6 ks.test(results_Demographie_ISSU[results_Demographie_ISSU$
   ISSU==1,]$ScoreMC, "pnorm", mean=mean(results_Demographie_
   ISSU[results_Demographie_ISSU$ISSU==1,]$ScoreMC), sd=sd(
   results_Demographie_ISSU[results_Demographie_ISSU$ISSU
   ==1,]$ScoreMC))
7 ks.test(results_Demographie_ISSU[results_Demographie_ISSU$
   ISSU==0,]$ScoreMC, "pnorm", mean=mean(results_Demographie_
   ISSU[results_Demographie_ISSU$ISSU==0,]$ScoreMC), sd=sd(
   results_Demographie_ISSU[results_Demographie_ISSU$ISSU
   ==0,]$ScoreMC))
8
9 leveneTest(results_Demographie_ISSU$ScoreMCProz ~ as.factor(
   results_Demographie_ISSU$ISSU))
10 t.test(results_Demographie_ISSU$ScoreMCProz ~ results_
   Demographie_ISSU$ISSU, var.equal=TRUE)
```

Demografische Daten (jeweils mit max. vorhandenem N)

Zusamenhang zum Alter (N = 192)

```
ks.test(results$ScoreMC, "pnorm", mean=mean(results$ScoreMC), sd=sd(results$ScoreMC))
ks.test(results$Alter, "pnorm", mean=mean(results$Alter), sd=sd(results$Alter))

\#\# non-parametrisch
cor.test(results$ScoreMC, results$Alter, method = "spearman")
```

Gruppenvergleich: Geschlecht weiblich(1)/männlich(2) (N = 192)

```
ks.test(results[results$Geschlecht=="A1",]$ScoreMC, "pnorm", mean=mean(results[results$Geschlecht=="A1",]$ScoreMC), sd=sd(results[results$Geschlecht=="A1",]$ScoreMC))
ks.test(results[results$Geschlecht=="A2",]$ScoreMC, "pnorm", mean=mean(results[results$Geschlecht=="A2",]$ScoreMC), sd=sd(results[results$Geschlecht=="A2",]$ScoreMC))

\#\# parametrisch
leveneTest(results$ScoreMC ~ as.factor(results$Geschlecht))
t.test(results$ScoreMC ~ results$Geschlecht)
cohens_d(results, ScoreMC ~ Geschlecht)

\#\# non-parametrisch
wilcox.test(ScoreMC ~ Geschlecht, data=results, conf.int=TRUE)
```

Zusammenhang zu Abiturnote (Ausländische Abschlüsse wurden nicht berücksichtigt, da die Umrechnung unklar; N = 175)

```
results_Abinote <- subset(results[(results$Phase!="A4" & results$Phase!="-oth-" & results$OrtHSReife != "0"),])

ks.test(results_Abinote$ScoreMC, "pnorm", mean=mean(results_Abinote$ScoreMC), sd=sd(results_Abinote$ScoreMC))
ks.test(results_Abinote$HSReifeD1, "pnorm", mean=mean(results_Abinote$HSReifeD1), sd=sd(results_Abinote$HSReifeD1))

cor.test(results_Abinote$ScoreMC, results_Abinote$HSReifeD1, method = "pearson")
cor.test(results_Abinote$ScoreMC, results_Abinote$HSReifeD1, method = "spearman")
```

Gruppenvergleich G8(A1)/G9(A2) (N = 165)

```
results_G8G9 <- subset(results[(results$Phase!="A4" &
   results$Phase!="-oth-" & results$OrtHSReife != "0" &
   results$HSReifeD2!="-oth-"),])
ks.test(results_G8G9[results_G8G9$HSReifeD2=="A1",]$ScoreMC,
    "pnorm", mean=mean(results_G8G9[results_G8G9$HSReifeD2=="
   A1",]$ScoreMC), sd=sd(results_G8G9[results_G8G9$HSReifeD2
   =="A1",]$ScoreMC))
ks.test(results_G8G9[results_G8G9$HSReifeD2=="A2",]$ScoreMC,
    "pnorm", mean=mean(results_G8G9[results_G8G9$HSReifeD2=="
   A2",]$ScoreMC), sd=sd(results_G8G9[results_G8G9$HSReifeD2
   =="A2",]$ScoreMC))

leveneTest(results_G8G9$ScoreMC ~ as.factor(results_G8G9$
   HSReifeD2))
t.test(results_G8G9$ScoreMC ~ results_G8G9$HSReifeD2)

wilcox.test(ScoreMC ~ HSReifeD2, results_G8G9)
```

Zusammenhang zu letzter (a) Physiknote, (b) Mathematiknote, (c) Deutschnote (hier und im Folgenden N = 178)

```
ks.test(results_Demographie$ScoreMC, "pnorm", mean=mean(
   results_Demographie$ScoreMC), sd=sd(results_Demographie$
   ScoreMC))

ks.test(results_Demographie$Schulnoten.SQ001., "pnorm", mean
   =mean(results_Demographie$Schulnoten.SQ001.), sd=sd(
   results_Demographie$Schulnoten.SQ001.))
cor.test(results_Demographie$ScoreMC, results_Demographie$
   Schulnoten.SQ001., method = "spearman")

ks.test(results_Demographie$Schulnoten.SQ002., "pnorm", mean
   =mean(results_Demographie$Schulnoten.SQ002.), sd=sd(
   results_Demographie$Schulnoten.SQ002.))
cor.test(results_Demographie$ScoreMC, results_Demographie$
   Schulnoten.SQ002., method = "spearman")

ks.test(results_Demographie$Schulnoten.SQ003., "pnorm", mean
   =mean(results_Demographie$Schulnoten.SQ003.), sd=sd(
   results_Demographie$Schulnoten.SQ003.))
cor.test(results_Demographie$ScoreMC, results_Demographie$
   Schulnoten.SQ003., method = "pearson")
cor.test(results_Demographie$ScoreMC, results_Demographie$
   Schulnoten.SQ003., method = "spearman")
```

Gruppenvergleich: Phyisk als Prüfungsfach im Abitur ja(1)/nein(0)

```
ks.test(results_Demographie[results_Demographie$PhyPruef
   ==1,]$ScoreMC, "pnorm", mean=mean(results_Demographie[
   results_Demographie$PhyPruef==1,]$ScoreMC), sd=sd(results_
   Demographie[results_Demographie$PhyPruef==1,]$ScoreMC))

ks.test(results_Demographie[results_Demographie$PhyPruef
   ==0,]$ScoreMC, "pnorm", mean=mean(results_Demographie[
   results_Demographie$PhyPruef==0,]$ScoreMC), sd=sd(results_
   Demographie[results_Demographie$PhyPruef==0,]$ScoreMC))

\#\# parametrisch
leveneTest(results_Demographie$ScoreMC ~ as.factor(results_
   Demographie$PhyPruef))
t.test(results_Demographie$ScoreMC ~ results_Demographie$
   PhyPruef, var.equal=FALSE)
```

Gruppenvergleich abhängig vom Zweitfach (Naturwissenschaftlich (2), Mathe (1) oder Sonstiges (0))

```
results_Demographie_ZF <- mutate(results_Demographie,
                                 ZF_MNW = c(2, 1, 1, 1, 0,
                                    0, 1, 1, 1, 0, 1, 2, 1,
                                    1, 0, 1, 1, 1, 2, 1,
1, 1, 0, 2, 1, 1, 0, 0, 1, 1, 0, 1, 0, 1, 2, 0, 1, 1, 1, 1,
1, 2, 1, 0, 1, 2, 1, 0, 0, 1, 1, 1, 2, 1, 1, 0, 0, 0, 1, 1,
1, 1, 2, 2, 0, 1, 2, 0, 1, 0, 0, 1, 0, 1, 1, 0, 1, 1, 0, 1,
2, 1, 1, 1, 0, 1, 2, 1, 1, 2, 0, 1, 1, 0, 0, 1, 0, 1, 2, 1,
2, 1, 1, 2, 2, 1, 1, 0, 2, 1, 1, 1, 0, 0, 2, 1, 0, 1, 0, 1,
0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 1, 1, 1, 1, NA, 0,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 1, 1, 1, 0, NA, 1, 1, 1, 1,
0, 2, 2, 2, 2, 1, 0, 1, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1),
                                 ZF_MNW_Ges = case_when(ZF_
                                    MNW == "0" ~ 0, ZF_MNW ==
                                     "NA" ~ NA, TRUE ~ 1),
                                 ZF_NW = case_when(ZF_MNW ==
                                     "2" ~ 1, ZF_MNW == "NA"
                                    ~ NA, TRUE ~ 0))
```

Ausschluss von NA (N = 176)

```
results_Demographie_ZF_o_NA <- subset(results_Demographie_ZF
   [-c(139, 156),])
```

Berechnung der ANOVA nach Prüfung der Voraussetzungen

```
ks.test(results_Demographie_ZF_o_NA[results_Demographie_ZF_o_NA$ZF_MNW==0,]$ScoreMC, "pnorm", mean=mean(results_Demographie_ZF_o_NA[results_Demographie_ZF_o_NA$ZF_MNW==0,]$ScoreMC), sd=sd(results_Demographie_ZF_o_NA[results_Demographie_ZF_o_NA$ZF_MNW==0,]$ScoreMC))

ks.test(results_Demographie_ZF_o_NA[results_Demographie_ZF_o_NA$ZF_MNW==1,]$ScoreMC, "pnorm", mean=mean(results_Demographie_ZF_o_NA[results_Demographie_ZF_o_NA$ZF_MNW==1,]$ScoreMC), sd=sd(results_Demographie_ZF_o_NA[results_Demographie_ZF_o_NA$ZF_MNW==1,]$ScoreMC))

ks.test(results_Demographie_ZF_o_NA[results_Demographie_ZF_o_NA$ZF_MNW==2,]$ScoreMC, "pnorm", mean=mean(results_Demographie_ZF_o_NA[results_Demographie_ZF_o_NA$ZF_MNW==2,]$ScoreMC), sd=sd(results_Demographie_ZF_o_NA[results_Demographie_ZF_o_NA$ZF_MNW==2,]$ScoreMC))

leveneTest(results_Demographie_ZF_o_NA$ScoreMCProz ~ as.factor(results_Demographie_ZF_o_NA$ZF_MNW))
Anova(aov(results_Demographie_ZF_o_NA$ScoreMCProz ~ as.factor(results_Demographie_ZF_o_NA$ZF_MNW)), type="3")
```

R.6.2 Längsschnittliche Betrachtungen

Aufbereitung der Datensätze

Erstellung der beiden Dataframes zu echten Bachelor-Längsschnitten

Kohorte mit Code 1005/1012

```
results_1005 <- subset(results[results$Code=="1005",])
results_1005 <- mutate(results_1005,
                       PP = "Prae")
results_1012 <- subset(results[results$Code=="1012",])
results_1012 <- mutate(results_1012,
                       PP = "Post")

colnames(results_1005) <- paste(colnames(results_1005), ".Prae", sep="")
N_1005 <- length(results_1005$id)

colnames(results_1012) <- paste(colnames(results_1012), ".Post", sep="")
```

```
N_1012 <- length(results_1012$id)


results_1005_1012_sort <- data.frame()
for (i in 1:N_1005){
  j <- which((results_1005$PC[i]==results_1012$PC)==TRUE)
  if (length(j)>0) {
    results_1005_1012_sort <- rbind(results_1005_1012_sort,
      c(results_1005[i,],
                    results_1012[j,]))
    colnames(results_1005_1012_sort) <- c(colnames(results_
      1005), colnames(results_1012))

    }
  else {}
}
```

Kohorte mit Code 1021/1027

```
(Dataframe: results_1021_1027_sort)
results_1021 <- subset(results[results$Code=="1021",])
results_1021 <- mutate(results_1021,
                       PP = "Prae")
results_1027 <- subset(results[results$Code=="1027",])
results_1027 <- mutate(results_1027,
                       PP = "Post")

colnames(results_1021) <- paste(colnames(results_1021), ".
   Prae", sep="")
N_1021 <- length(results_1021$id)

colnames(results_1027) <- paste(colnames(results_1027), ".
   Post", sep="")
N_1027 <- length(results_1027$id)


results_1021_1027_sort <- data.frame()
for (i in 1:N_1021){
  j <- which((results_1021$PC[i]==results_1027$PC)==TRUE)
  if (length(j)>0) {
    results_1021_1027_sort <- rbind(results_1021_1027_sort,
      c(results_1021[i,],
                    results_1027[j,]))
    colnames(results_1021_1027_sort) <- c(colnames(results_
      1021), colnames(results_1027))

    }
```

```
25   else {}
26 }
```

Erstellung des Dataframes für die Prä-Post-PS-Studierenden

```
1  results_nur_PS <- subset(results[(
2    results$Code=="1009" | results$Code=="1024" |
3      results$Code=="1016" | results$Code=="1025" |
4      results$Code=="1029" | results$Code=="1035" |
5      results$Code=="1030" | results$Code=="1036"),])
6
7  results_nur_PS <- mutate(results_nur_PS,
8                                  PP = case_when(results_nur_
                                     PS$Code=="1009" ~ "Prae",
9                                   results_nur_PS$Code=="1024"
                                       ~ "Post",
10                                  results_nur_PS$Code=="1016"
                                       ~ "Prae",
11                                  results_nur_PS$Code=="1025"
                                       ~ "Post",
12                                  results_nur_PS$Code=="1029"
                                       ~ "Prae",
13                                  results_nur_PS$Code=="1035"
                                       ~ "Post",
14                                  results_nur_PS$Code=="1030"
                                       ~ "Prae",
15                                  results_nur_PS$Code=="1036"
                                       ~ "Post",
16                                  ))
17
18 results_nur_PS_Prae <- subset(results_nur_PS[results_nur_PS$
     PP=="Prae",])
19 colnames(results_nur_PS_Prae) <- paste(colnames(results_nur_
     PS_Prae), ".Prae", sep="")
20 N_Prae <- length(results_nur_PS_Prae$id)
21
22 results_nur_PS_Post <- subset(results_nur_PS[results_nur_PS$
     PP=="Post",])
23 colnames(results_nur_PS_Post) <- paste(colnames(results_nur_
     PS_Post), ".Post", sep="")
24 N_Post <- length(results_nur_PS_Post$id)
25
26
27 results_PS_sort=data.frame()
28 for (i in 1:N_Prae){
29   j <- which((results_nur_PS_Prae$PC[i]==results_nur_PS_Post
       $PC)==TRUE)
```

```
30   if (length(j)>0) {
31     results_PS_sort <- rbind(results_PS_sort, c(results_nur_
         PS_Prae[i,],
32                         results_nur_PS_Post[j,]))
33     colnames(results_PS_sort) <- c(colnames(results_nur_PS_
         Prae), colnames(results_nur_PS_Post))
34
35     }
36   else {}
37 }
```

Zusammenfügen aller Prä-Post-Gruppen in einen Dataframe

```
1 results_Laengsschnitte_BA <- rbind(results_1005_1012_sort,
2                                    results_1021_1027_sort)
3
4 results_Laengsschnitte <- rbind(results_1005_1012_sort,
5                                 results_1021_1027_sort,
6                                 results_PS_sort)
```

Untersuchungen zur Entwicklung im Längsschnitt über ein Semester

Gesamtstichprobe N = 41

```
1 ks.test((results_Laengsschnitte$ScoreMC.Post-results_
    Laengsschnitte$ScoreMC.Prae), "pnorm", mean=mean((results_
    Laengsschnitte$ScoreMC.Post-results_Laengsschnitte$ScoreMC
    .Prae)), sd=sd((results_Laengsschnitte$ScoreMC.Post-
    results_Laengsschnitte$ScoreMC.Prae)))
2
3 mean(results_Laengsschnitte$ScoreMCProz.Prae)
4 sd(results_Laengsschnitte$ScoreMCProz.Prae)
5
6 mean(results_Laengsschnitte$ScoreMCProz.Post)
7 sd(results_Laengsschnitte$ScoreMCProz.Post)
8
9 t.test(results_Laengsschnitte$ScoreMC.Prae, results_
    Laengsschnitte$ScoreMC.Post, paired = TRUE)
```

Berechnung der angepassten Effektstärke $d_{\text{Cohen},d}$

```
1 -3.6961/sqrt(41)
```

Bachelor N = 16

```
ks.test((results_Laengsschnitte_BA$ScoreMC.Post-results_
   Laengsschnitte_BA$ScoreMC.Prae), "pnorm", mean=mean((
   results_Laengsschnitte_BA$ScoreMC.Post-results_
   Laengsschnitte_BA$ScoreMC.Prae)), sd=sd((results_
   Laengsschnitte_BA$ScoreMC.Post-results_Laengsschnitte_BA$
   ScoreMC.Prae)))

mean(results_Laengsschnitte_BA$ScoreMCProz.Prae)
sd(results_Laengsschnitte_BA$ScoreMCProz.Prae)

mean(results_Laengsschnitte_BA$ScoreMCProz.Post)
sd(results_Laengsschnitte_BA$ScoreMCProz.Post)

t.test(results_Laengsschnitte_BA$ScoreMC.Prae, results_
   Laengsschnitte_BA$ScoreMC.Post, paired = TRUE)
```

Berechnung der angepassten Effektstärke $d_{\text{Cohen},d}$

```
-2.8766/sqrt(16)
```

Master (PS) N = 25

```
ks.test((results_PS_sort$ScoreMC.Post-results_PS_sort$
   ScoreMC.Prae), "pnorm", mean=mean((results_PS_sort$ScoreMC
   .Post-results_PS_sort$ScoreMC.Prae)), sd=sd((results_PS_
   sort$ScoreMC.Post-results_PS_sort$ScoreMC.Prae)))

mean(results_PS_sort$ScoreMCProz.Prae)
sd(results_PS_sort$ScoreMCProz.Prae)

mean(results_PS_sort$ScoreMCProz.Post)
sd(results_PS_sort$ScoreMCProz.Post)

t.test(results_PS_sort$ScoreMC.Prae, results_PS_sort$ScoreMC
   .Post, paired = TRUE)
```

Berechnung der angepassten Effektstärke $d_{\text{Cohen},d}$

```
-2.3796/sqrt(25)
```

R.7 Indikator 3.6: Zusammenhang zum fachdidaktischen Wissen

R.7.1 Aufbereiten der Daten aus FDW-Test

Importieren der Daten zum FDW-Test für beide MZPs

```
FDW_all <- read.csv("./FDW_24.08.23_Datensätze PraePost.csv", header=TRUE, sep=";", encoding="UTF-8")
FDW <- FDW_all[(FDW_all$Bearbeitung=="vollstaendig" & is.na(FDW_all$Einsatzcode)==FALSE),]
FDW <- mutate(FDW,
              ProbCode = gsub("r", "", Probandencode),
                  PP = case_when(FDW$Einsatzcode=="1014" ~ "Prae",
                                 FDW$Einsatzcode=="1016" ~ "Prae",
                                 FDW$Einsatzcode=="1024" ~ "Post",
                                 FDW$Einsatzcode=="1025" ~ "Post",
                                 FDW$Einsatzcode=="1034" ~ "Prae",
                                 FDW$Einsatzcode=="1035" ~ "Post",
                                 FDW$Einsatzcode=="1036" ~ "Post",
                                 ))

FDW_nur_PS_Prae <- subset(FDW[FDW$PP=="Prae",])
colnames(FDW_nur_PS_Prae) <- paste(colnames(FDW_nur_PS_Prae), ".Prae", sep="")
FDW_nur_PS_Post <- subset(FDW[FDW$PP=="Post",])
colnames(FDW_nur_PS_Post) <- paste(colnames(FDW_nur_PS_Post), ".Post", sep="")
```

R.7.2 Zusammenfügen der Daten aus dem Reflexions- und FDW-Test für beide MZPs

Erstellung eines Datensatzes mit Prä-Post-Daten der Personen

```
results_Prae_R_FDW <- merge(results_nur_PS_Prae, FDW_nur_PS_Prae, by.x = "PC.Prae", by.y = "ProbCode.Prae")
```

```
results_Post_R_FDW <- merge(results_nur_PS_Post, FDW_nur_PS_Post, by.x = "PC.Post", by.y = "ProbCode.Post")

results_PS_FDW <- rbind(subset(FDW[FDW$PP=="Prae",]), subset(FDW[FDW$PP=="",]))
results_PS_FDW_sort=data.frame()
for (i in 1:N_Prae){
  j <- which((results_Prae_R_FDW$PC[i]==results_Post_R_FDW$PC)==TRUE)
  if (length(j)>0) {
    results_PS_FDW_sort <- rbind(results_PS_FDW_sort, c(results_Prae_R_FDW[i,],
                    results_Post_R_FDW[j,]))
    colnames(results_PS_FDW_sort) <- c(colnames(results_Prae_R_FDW), colnames(results_Post_R_FDW))

    }
  else {}
}
```

Erstellung eines Datensatzes mit den Prä + Post Bearbeitungen ohne Zuordnung zu den Personen (N = 41) (für nachfolgende Berechnung von Cronbachs-Alpha)

```
results_Post_R_FDW_colnames_Prae <- results_Post_R_FDW
colnames(results_Post_R_FDW_colnames_Prae) <- colnames(results_Prae_R_FDW)
results_R_FDW <- rbind(results_Prae_R_FDW, results_Post_R_FDW_colnames_Prae)
```

R.7.3 Berechnung der Kennzahlen (M, SD, Test auf Normalverteilung, Cronbachs-Alpha)

Für die Gesamtstichprobe (Prä +Post)

Reflexionstest

```
ks.test(c(results_Prae_R_FDW$ScoreMC.Prae, results_Post_R_FDW$ScoreMC.Post), "pnorm", mean=mean(c(results_Prae_R_FDW$ScoreMC.Prae, results_Post_R_FDW$ScoreMC.Post)), sd=sd(c(results_Prae_R_FDW$ScoreMC.Prae, results_Post_R_FDW$ScoreMC.Post)))
mean(c(results_Prae_R_FDW$ScoreMC.Prae, results_Post_R_FDW$ScoreMC.Post))
sd(c(results_Prae_R_FDW$ScoreMC.Prae, results_Post_R_FDW$ScoreMC.Post))
```

```
4 psych::alpha(subset(results_R_FDW, select = c("R1Bew1.Prae",
     "R1Alt1.Prae","R2Bew1.Prae", "R2Alt1.Prae", "R3Bew1.Prae",
      "R3Alt1.Prae", "R3Bew2.Prae", "R3Alt2.Prae", "R4Bew1.Prae
     ", "R4Alt1.Prae", "R4Bew2.Prae", "R4Alt2.Prae", "R5Bew1.
     Prae", "R5Alt1.Prae", "R5Bew2.Prae", "R5Alt2.Prae", "
     R5Bew3.Prae", "R5Alt3.Prae", "R6Bew1.Prae", "R6Alt1.Prae",
      "R6Bew2.Prae", "R6Alt2.Prae", "R6Bew3.Prae", "R6Alt3.Prae
     ", "R7Bew1.Prae", "R7Alt1.Prae", "R7Bew2.Prae", "R7Alt2.
     Prae", "RGBew1.Prae", "RGAlt1.Prae", "RGBew2.Prae", "
     RGAlt2.Prae")), check.keys =FALSE)
```

FDW-Test

```
1 ks.test(c(results_Prae_R_FDW$Summe_Probanden.Prae, results_
     Post_R_FDW$Summe_Probanden.Post), "pnorm", mean=mean(c(
     results_Prae_R_FDW$Summe_Probanden.Prae, results_Post_R_
     FDW$Summe_Probanden.Post)), sd=sd(c(results_Prae_R_FDW$
     Summe_Probanden.Prae, results_Post_R_FDW$Summe_Probanden.
     Post)))
2 mean(c(results_Prae_R_FDW$Summe_Probanden.Prae, results_Post
     _R_FDW$Summe_Probanden.Post)/62*100)
3 sd(c(results_Prae_R_FDW$Summe_Probanden.Prae, results_Post_R
     _FDW$Summe_Probanden.Post)/62*100)
4 psych::alpha(subset(results_R_FDW, select = c(736:766)),
     check.keys = FALSE)
```

Für die Prä-Stichprobe

Reflexionstest

```
1 ks.test(results_Prae_R_FDW$ScoreMC.Prae, "pnorm", mean=mean(
     results_Prae_R_FDW$ScoreMC.Prae), sd=sd(results_Prae_R_FDW
     $ScoreMC.Prae))
2 mean(results_Prae_R_FDW$ScoreMC.Prae)
3 sd(results_Prae_R_FDW$ScoreMC.Prae)
4 psych::alpha(subset(results_Prae_R_FDW, select = c("R1Bew1.
     Prae","R1Alt1.Prae","R2Bew1.Prae", "R2Alt1.Prae", "R3Bew1.
     Prae", "R3Alt1.Prae", "R3Bew2.Prae", "R3Alt2.Prae", "
     R4Bew1.Prae", "R4Alt1.Prae", "R4Bew2.Prae", "R4Alt2.Prae",
      "R5Bew1.Prae", "R5Alt1.Prae", "R5Bew2.Prae", "R5Alt2.Prae
     ", "R5Bew3.Prae", "R5Alt3.Prae", "R6Bew1.Prae", "R6Alt1.
     Prae", "R6Bew2.Prae", "R6Alt2.Prae", "R6Bew3.Prae", "
     R6Alt3.Prae", "R7Bew1.Prae", "R7Alt1.Prae", "R7Bew2.Prae",
      "R7Alt2.Prae", "RGBew1.Prae", "RGAlt1.Prae", "RGBew2.Prae
     ", "RGAlt2.Prae")), check.keys =FALSE)
```

FDW-Test

```
1 ks.test(results_Prae_R_FDW$Summe_Probanden.Prae, "pnorm",
     mean=mean(results_Prae_R_FDW$Summe_Probanden.Prae), sd=sd(
     results_Prae_R_FDW$Summe_Probanden.Prae))
2 mean(results_Prae_R_FDW$Summe_Probanden.Prae/62*100)
3 sd(results_Prae_R_FDW$Summe_Probanden.Prae/62*100)
4 psych::alpha(subset(results_Prae_R_FDW, select = c(736:766))
     , check.keys = FALSE)
```

Für die Post-Stichprobe

Reflexionstest

```
1 ks.test(results_Post_R_FDW$ScoreMC.Post, "pnorm", mean=mean(
     results_Post_R_FDW$ScoreMC.Post), sd=sd(results_Post_R_FDW
     $ScoreMC.Post))
2 mean(results_Post_R_FDW$ScoreMC.Post)
3 sd(results_Post_R_FDW$ScoreMC.Post)
4 psych::alpha(subset(results_Post_R_FDW, select = c("R1Bew1.
     Post","R1Alt1.Post","R2Bew1.Post", "R2Alt1.Post", "R3Bew1.
     Post", "R3Alt1.Post", "R3Bew2.Post", "R3Alt2.Post", "
     R4Bew1.Post", "R4Alt1.Post", "R4Bew2.Post", "R4Alt2.Post",
      "R5Bew1.Post", "R5Alt1.Post", "R5Bew2.Post", "R5Alt2.Post
     ", "R5Bew3.Post", "R5Alt3.Post", "R6Bew1.Post", "R6Alt1.
     Post", "R6Bew2.Post", "R6Alt2.Post", "R6Bew3.Post", "
     R6Alt3.Post", "R7Bew1.Post", "R7Alt1.Post", "R7Bew2.Post",
      "R7Alt2.Post", "RGBew1.Post", "RGAlt1.Post", "RGBew2.Post
     ", "RGAlt2.Post")), check.keys =FALSE)
```

FDW-Test

```
1 ks.test(results_Post_R_FDW$Summe_Probanden.Post, "pnorm",
     mean=mean(results_Post_R_FDW$Summe_Probanden.Post), sd=sd(
     results_Post_R_FDW$Summe_Probanden.Post))
2 mean(results_Post_R_FDW$Summe_Probanden.Post/62*100)
3 sd(results_Post_R_FDW$Summe_Probanden.Post/62*100)
4 psych::alpha(subset(results_Post_R_FDW, select = c(736:766))
     , check.keys = FALSE)
```

R.7.4 Berechnung der Korrelationen für die Gesamt-, Prä- und Post-Stichprobe

```
1 cor.test(results_R_FDW$ScoreMC.Prae, results_R_FDW$Summe_
    Probanden.Prae, method = "pearson")
2 cor.test(results_Prae_R_FDW$ScoreMC.Prae, results_Prae_R_FDW
    $Summe_Probanden.Prae, method = "pearson")
3 cor.test(results_Post_R_FDW$ScoreMC.Post, results_Post_R_FDW
    $Summe_Probanden.Post, method = "pearson")
```

R.7.5 Berechnung der Tests auf Normalverteilung sowie der Korrelationen zur Reflexionsfähigkeit für die Facetten des FDW-Tests

Für die Gesamtstichprobe

```
1 ks.test(c(results_Prae_R_FDW$Instruktionsstrategien.Prae,
    results_Post_R_FDW$Instruktionsstrategien.Post), "pnorm",
    mean=mean(c(results_Prae_R_FDW$Instruktionsstrategien.Prae
    , results_Post_R_FDW$Instruktionsstrategien.Post)), sd=sd(
    c(results_Prae_R_FDW$Instruktionsstrategien.Prae, results_
    Post_R_FDW$Instruktionsstrategien.Post)))
2 cor.test(results_R_FDW$ScoreMC.Prae, results_R_FDW$
    Instruktionsstrategien.Prae, method = "pearson")
3
4 ks.test(c(results_Prae_R_FDW$Schuelervorstellungen.Prae,
    results_Post_R_FDW$Schuelervorstellungen.Post), "pnorm",
    mean=mean(c(results_Prae_R_FDW$Schuelervorstellungen.Prae,
     results_Post_R_FDW$Schuelervorstellungen.Post)), sd=sd(c(
    results_Prae_R_FDW$Schuelervorstellungen.Prae, results_
    Post_R_FDW$Schuelervorstellungen.Post)))
5 cor.test(results_R_FDW$ScoreMC.Prae, results_R_FDW$
    Schuelervorstellungen.Prae, method = "pearson")
6
7 ks.test(c(results_Prae_R_FDW$Experimente.Prae, results_Post_
    R_FDW$Experimente.Post), "pnorm", mean=mean(c(results_Prae
    _R_FDW$Experimente.Prae, results_Post_R_FDW$Experimente.
    Post)), sd=sd(c(results_Prae_R_FDW$Experimente.Prae,
    results_Post_R_FDW$Experimente.Post)))
8 cor.test(results_R_FDW$ScoreMC.Prae, results_R_FDW$
    Experimente.Prae, method = "spearman")
9
```

```
10 ks.test(c(results_Prae_R_FDW$Fachdidaktische_Konzepte.Prae,
      results_Post_R_FDW$Fachdidaktische_Konzepte.Post), "pnorm"
      , mean=mean(c(results_Prae_R_FDW$Fachdidaktische_Konzepte.
      Prae, results_Post_R_FDW$Fachdidaktische_Konzepte.Post)),
      sd=sd(c(results_Prae_R_FDW$Fachdidaktische_Konzepte.Prae,
      results_Post_R_FDW$Fachdidaktische_Konzepte.Post)))
11 cor.test(results_R_FDW$ScoreMC.Prae, results_R_FDW$
      Fachdidaktische_Konzepte.Prae, method = "pearson")
```

Für die Prä-Stichprobe

```
1  ks.test(results_Prae_R_FDW$Instruktionsstrategien.Prae, "
      pnorm", mean=mean(results_Prae_R_FDW$
      Instruktionsstrategien.Prae), sd=sd(results_Prae_R_FDW$
      Instruktionsstrategien.Prae))
2  cor.test(results_Prae_R_FDW$ScoreMC.Prae, results_Prae_R_FDW
      $Instruktionsstrategien.Prae)
3
4  ks.test(results_Prae_R_FDW$Schuelervorstellungen.Prae, "
      pnorm", mean=mean(results_Prae_R_FDW$Schuelervorstellungen
      .Prae), sd=sd(results_Prae_R_FDW$Schuelervorstellungen.
      Prae))
5  cor.test(results_Prae_R_FDW$ScoreMC.Prae, results_Prae_R_FDW
      $Schuelervorstellungen.Prae)
6
7  ks.test(results_Prae_R_FDW$Experimente.Prae, "pnorm", mean=
      mean(results_Prae_R_FDW$Experimente.Prae), sd=sd(results_
      Prae_R_FDW$Experimente.Prae))
8  cor.test(results_Prae_R_FDW$ScoreMC.Prae, results_Prae_R_FDW
      $Experimente.Prae)
9
10 ks.test(results_Prae_R_FDW$Fachdidaktische_Konzepte.Prae, "
      pnorm", mean=mean(results_Prae_R_FDW$Fachdidaktische_
      Konzepte.Prae), sd=sd(results_Prae_R_FDW$Fachdidaktische_
      Konzepte.Prae))
11 cor.test(results_Prae_R_FDW$ScoreMC.Prae, results_Prae_R_FDW
      $Fachdidaktische_Konzepte.Prae)
```

Für die Post-Stichprobe

```
ks.test(results_Post_R_FDW$Instruktionsstrategien.Post, "pnorm", mean=mean(results_Post_R_FDW$Instruktionsstrategien.Post), sd=sd(results_Post_R_FDW$Instruktionsstrategien.Post))
cor.test(results_Post_R_FDW$ScoreMC.Post, results_Post_R_FDW$Instruktionsstrategien.Post)

ks.test(results_Post_R_FDW$Schuelervorstellungen.Post, "pnorm", mean=mean(results_Post_R_FDW$Schuelervorstellungen.Post), sd=sd(results_Post_R_FDW$Schuelervorstellungen.Post))
cor.test(results_Post_R_FDW$ScoreMC.Post, results_Post_R_FDW$Schuelervorstellungen.Post)

ks.test(results_Post_R_FDW$Experimente.Post, "pnorm", mean=mean(results_Post_R_FDW$Experimente.Post), sd=sd(results_Post_R_FDW$Experimente.Post))
cor.test(results_Post_R_FDW$ScoreMC.Post, results_Post_R_FDW$Experimente.Post)

ks.test(results_Post_R_FDW$Fachdidaktische_Konzepte.Post, "pnorm", mean=mean(results_Post_R_FDW$Fachdidaktische_Konzepte.Post), sd=sd(results_Post_R_FDW$Fachdidaktische_Konzepte.Post))
cor.test(results_Post_R_FDW$ScoreMC.Post, results_Post_R_FDW$Fachdidaktische_Konzepte.Post)
```

S *R*-Skript für die statistischen Auswertungen der explorativen Untersuchungen zur Reflexionsfähigkeit

S.1 Vorbereitend: Pakete einfügen und Aufbereitung der Daten aus dem Reflexionstest

```
1 library(tidyverse)
2 library(psych)
3 library(car)
4 library(rstatix)
5 library(lsr)
6 library(tidyLPA)
7 library(ggalluvial)
8 library(ppcor)
```

Importieren der Daten aus dem Reflexionstest sowie u. a. Bestimmung der Scores auf Aufgabenebene (R1Bew1 bis RGAlt2, ScoreMCNo1 bis ScoreMCNoG) bzw. Testebene (als Absolut- und Prozentwerte: ScoreMC bis ScoreMCSVProz) und Entfernung einer bekanntermaßen nicht ernsthaften Bearbeitung

```
1 results <- read.csv("./results-survey516586.csv", header=
    TRUE, sep="\t", encoding="UTF-8")
2 results[is.na(results)]<-0
3 results <- mutate(results,
4                   R1Bew1 = (results$Refl02bIIIBew.SQ002.+
                       results$Refl02bIIIBew.SQ003.+ results$
                       Refl02bIIIBew.SQ004.+ results$
                       Refl02bIIIBew.SQ007.)/4,
5                   R1Alt1 = (results$Refl02bIIIAlt.SQ002.+
                       results$Refl02bIIIAlt.SQ004.+ results$
                       Refl02bIIIAlt.SQ005.+ results$
                       Refl02bIIIAlt.SQ006.)/4,
```

```
6                   R2Bew1 = (results$Refl03bIVBeob.SQ003.+
                       results$Refl03bIVBeob.SQ004.+ results$
                       Refl03bIVBeob.SQ006.)/3,
7                   R2Alt1 = (results$Refl03IVUmgang.SQ003.+
                       results$Refl03IVUmgang.SQ004.+ results$
                       Refl03IVUmgang.SQ005.)/3,
8                   R3Bew1 = (results$Refl06bIVBew.SQ001.+
                       results$Refl06bIVBew.SQ002.+ results$
                       Refl06bIVBew.SQ003.+ results$Refl06bIVBew
                       .SQ005.)/4,
9                   R3Alt1 = (results$Refl06bIVAlt.SQ001.+
                       results$Refl06bIVAlt.SQ002.+ results$
                       Refl06bIVAlt.SQ003.+ results$Refl06bIVAlt
                       .SQ005.)/4,
10                  R3Bew2 = (results$Refl06bVBew.SQ001.+
                       results$Refl06bVBew.SQ005.+ results$
                       Refl06bVBew.SQ006.+ results$Refl06bVBew.
                       SQ007.)/4,
11                  R3Alt2 = (results$Refl06bVAlt.SQ002.+
                       results$Refl06bVAlt.SQ005.+ results$
                       Refl06bVAlt.SQ006.)/3,
12                  R4Bew1 = (results$Refl07bIVBew.SQ001.+
                       results$Refl07bIVBew.SQ002.+ results$
                       Refl07bIVBew.SQ003.+ results$Refl07bIVBew
                       .SQ004.)/4,
13                  R4Alt1 = (results$Refl07bIVAlt.SQ001.+
                       results$Refl07bIVAlt.SQ003.+ results$
                       Refl07bIVAlt.SQ005.)/3,
14                  R4Bew2 = (results$Refl07bVBew.SQ001.+
                       results$Refl07bVBew.SQ003.+ results$
                       Refl07bVBew.SQ005.+ results$Refl07bVBew.
                       SQ007.)/4,
15                  R4Alt2 = (results$Refl07bVAlt.SQ004.+
                       results$Refl07bVAlt.SQ005.+ results$
                       Refl07bVAlt.SQ006.+ results$Refl07bVAlt.
                       SQ007.)/4,
16                  R5Bew1 = (results$Refl09bIBew.SQ001.+
                       results$Refl09bIBew.SQ002.+ results$
                       Refl09bIBew.SQ004.)/3,
17                  R5Alt1 = (results$Refl09bIAlt.SQ003.+
                       results$Refl09bIAlt.SQ005.+ results$
                       Refl09bIAlt.SQ006.+ results$Refl09bIAlt.
                       SQ007.)/4,
18                  R5Bew2 = (results$Refl09bIVBew.SQ001.+
                       results$Refl09bIVBew.SQ002.+ results$
                       Refl09bIVBew.SQ004.+ results$Refl09bIVBew
                       .SQ005.)/4,
```

```
19                  R5Alt2 = (results$Refl09bIVAlt.SQ002.+
                       results$Refl09bIVAlt.SQ003.+ results$
                       Refl09bIVAlt.SQ005.+ results$Refl09bIVAlt
                       .SQ006.)/4,
20                  R5Bew3 = (results$Refl09bVIBew.SQ001.+
                       results$Refl09bVIBew.SQ004.+ results$
                       Refl09bVIBew.SQ007.)/3,
21                  R5Alt3 = (results$Refl09bVIAlt.SQ001.+
                       results$Refl09bVIAlt.SQ005.+ results$
                       Refl09bVIAlt.SQ007.)/3,
22                  R6Bew1 = (results$Refl12bIBew.SQ002.+
                       results$Refl12bIBew.SQ003.+ results$
                       Refl12bIBew.SQ004.)/3,
23                  R6Alt1 = (results$Refl12bIAlt.SQ002.+
                       results$Refl12bIAlt.SQ003.+ results$
                       Refl12bIAlt.SQ006.+ results$Refl12bIAlt.
                       SQ007.)/4,
24                  R6Bew2 = (results$Refl12bIIBew.SQ004.+
                       results$Refl12bIIBew.SQ006.+ results$
                       Refl12bIIBew.SQ007.)/3,
25                  R6Alt2 = (results$Refl12bIIAlt.SQ005.+
                       results$Refl12bIIAlt.SQ006.+ results$
                       Refl12bIIAlt.SQ007.)/3,
26                  R6Bew3 = (results$Refl12bIIIBew.SQ001.+
                       results$Refl12bIIIBew.SQ002.+ results$
                       Refl12bIIIBew.SQ005.)/3,
27                  R6Alt3 = (results$Refl12bIIIAlt.SQ002.+
                       results$Refl12bIIIAlt.SQ003.+ results$
                       Refl12bIIIAlt.SQ005.)/3,
28                  R7Bew1 = (results$Refl13bIBew.SQ003.+
                       results$Refl13bIBew.SQ004.)/2,
29                  R7Alt1 = (results$Refl13bIAlt.SQ002.+
                       results$Refl13bIAlt.SQ004.+ results$
                       Refl13bIAlt.SQ006.+ results$Refl13bIAlt.
                       SQ007.)/4,
30                  R7Bew2 = (results$Refl13bIIBew.SQ002.+
                       results$Refl13bIIBew.SQ003.+ results$
                       Refl13bIIBew.SQ004.+ results$Refl13bIIBew
                       .SQ007.)/4,
31                  R7Alt2 = (results$Refl13bIIAlt.SQ002.+
                       results$Refl13bIIAlt.SQ004.+ results$
                       Refl13bIIAlt.SQ007.)/3,
32                  RGBew1 = (results$GesbVIBew.SQ002.+ results
                       $GesbVIBew.SQ003.+ results$GesbVIBew.
                       SQ005.+ results$GesbVIBew.SQ007.+ results
                       $GesbVIBew.SQ008.)/5,
```

```
33                  RGAlt1 = (results$GesVIAlt.SQ002.+ results$
                       GesVIAlt.SQ004.+ results$GesVIAlt.SQ005.+
                        results$GesVIAlt.SQ007.)/4,
34                  RGBew2 = (results$GesVIIBew.SQ001.+ results
                       $GesVIIBew.SQ002.+ results$GesVIIBew.
                       SQ005.)/3,
35                  RGAlt2 = (results$GesVIIAlt.SQ002.+ results
                       $GesVIIAlt.SQ003.+ results$GesVIIAlt.
                       SQ005.)/3,
36                  ScoreMCRefl1 = (Refl02bIIIBew.SQ002.+
                       Refl02bIIIBew.SQ003.+ Refl02bIIIBew.SQ004
                       .+ Refl02bIIIBew.SQ007.)/4+ (
                       Refl02bIIIAlt.SQ002.+ Refl02bIIIAlt.SQ004
                       .+ Refl02bIIIAlt.SQ005.+ Refl02bIIIAlt.
                       SQ006.)/4,
37                  ScoreMCRefl2 = (Refl03bIVBeob.SQ003.+
                       Refl03bIVBeob.SQ004.+ Refl03bIVBeob.SQ006
                       .)/3+ (Refl03IVUmgang.SQ003.+
                       Refl03IVUmgang.SQ004.+ Refl03IVUmgang.
                       SQ005.)/3,
38                  ScoreMCRefl3 = (Refl06bIVBew.SQ001.+
                       Refl06bIVBew.SQ002.+ Refl06bIVBew.SQ003.+
                        Refl06bIVBew.SQ005.)/4+ (Refl06bIVAlt.
                       SQ001. +Refl06bIVAlt.SQ002. +Refl06bIVAlt
                       .SQ003. +Refl06bIVAlt.SQ005.)/4+ (
                       Refl06bVBew.SQ001.+ Refl06bVBew.SQ005.+
                       Refl06bVBew.SQ006.+ Refl06bVBew.SQ007.)/
                       4+ (Refl06bVAlt.SQ002.+ Refl06bVAlt.SQ005
                       .+ Refl06bVAlt.SQ006.)/3,
39                  ScoreMCRefl4 = (Refl07bIVBew.SQ001.+
                       Refl07bIVBew.SQ002.+ Refl07bIVBew.SQ003.+
                        Refl07bIVBew.SQ004.)/4+ (Refl07bIVAlt.
                       SQ001.+ Refl07bIVAlt.SQ003.+ Refl07bIVAlt
                       .SQ005.)/3+ (Refl07bVBew.SQ001.+
                       Refl07bVBew.SQ003.+ Refl07bVBew.SQ005.+
                       Refl07bVBew.SQ007.)/4+ (Refl07bVAlt.SQ004
                       .+ Refl07bVAlt.SQ005.+ Refl07bVAlt.SQ006
                       .+ Refl07bVAlt.SQ007.)/4,
40                  ScoreMCRefl5 = (Refl09bIBew.SQ001.+
                       Refl09bIBew.SQ002.+ Refl09bIBew.SQ004.)/
                       3+ (Refl09bIAlt.SQ003.+ Refl09bIAlt.SQ005
                       .+ Refl09bIAlt.SQ006.+ Refl09bIAlt.SQ007
                       .)/4+ (Refl09bIVBew.SQ001.+ Refl09bIVBew.
                       SQ002.+ Refl09bIVBew.SQ004.+ Refl09bIVBew
                       .SQ005.)/4+ (Refl09bIVAlt.SQ002.+
                       Refl09bIVAlt.SQ003.+ Refl09bIVAlt.SQ005.+
                        Refl09bIVAlt.SQ006.)/4+ (Refl09bVIBew.
```

```
                    SQ001.+ Refl09bVIBew.SQ004.+ Refl09bVIBew
                    .SQ007.)/3+ (Refl09bVIAlt.SQ001.+
                    Refl09bVIAlt.SQ005.+ Refl09bVIAlt.SQ007.)
                    /3,
41                  ScoreMCRefl6 = (Refl12bIBew.SQ002.+
                    Refl12bIBew.SQ003.+ Refl12bIBew.SQ004.)/
                    3+ (Refl12bIAlt.SQ002.+ Refl12bIAlt.SQ003
                    .+ Refl12bIAlt.SQ006.+ Refl12bIAlt.SQ007
                    .)/4+ (Refl12bIIBew.SQ004.+ Refl12bIIBew.
                    SQ006.+ Refl12bIIBew.SQ007.)/3+ (
                    Refl12bIIAlt.SQ005.+ Refl12bIIAlt.SQ006.+
                     Refl12bIIAlt.SQ007.)/3+ (Refl12bIIIBew.
                    SQ001.+ Refl12bIIIBew.SQ002.+
                    Refl12bIIIBew.SQ005.)/3+ (Refl12bIIIAlt.
                    SQ002.+ Refl12bIIIAlt.SQ003.+
                    Refl12bIIIAlt.SQ005.)/3,
42                  ScoreMCRefl7 = (Refl13bIBew.SQ003.+
                    Refl13bIBew.SQ004.)/2+ (Refl13bIAlt.SQ002
                    .+ Refl13bIAlt.SQ004.+ Refl13bIAlt.SQ006
                    .+ Refl13bIAlt.SQ007.)/4+ (Refl13bIIBew.
                    SQ002.+ Refl13bIIBew.SQ003.+ Refl13bIIBew
                    .SQ004.+ Refl13bIIBew.SQ007.)/4+ (
                    Refl13bIIAlt.SQ002.+ Refl13bIIAlt.SQ004.+
                     Refl13bIIAlt.SQ007.)/3,
43                  ScoreMCReflG = (GesbVIBew.SQ002.+ GesbVIBew
                    .SQ003.+ GesbVIBew.SQ005.+ GesbVIBew.
                    SQ007.+ GesbVIBew.SQ008.)/5+ (GesVIAlt.
                    SQ002.+ GesVIAlt.SQ004.+ GesVIAlt.SQ005.+
                     GesVIAlt.SQ007.)/4+ (GesVIIBew.SQ001.+
                    GesVIIBew.SQ002.+ GesVIIBew.SQ005.)/3+ (
                    GesVIIAlt.SQ002.+ GesVIIAlt.SQ003.+
                    GesVIIAlt.SQ005.)/3,
44                  ScoreMC = ScoreMCRefl1+ ScoreMCRefl2+
                    ScoreMCRefl3+ ScoreMCRefl4+ ScoreMCRefl5+
                     ScoreMCRefl6+ ScoreMCRefl7+ ScoreMCReflG
                    ,
45                  ScoreMCProz = ScoreMC/32*100,
46                  ScoreMCBew = (Refl02bIIIBew.SQ002.+
                    Refl02bIIIBew.SQ003.+ Refl02bIIIBew.SQ004
                    .)/4+ (Refl06bIVBew.SQ001.+ Refl06bIVBew.
                    SQ002.+ Refl06bIVBew.SQ003.+ Refl06bIVBew
                    .SQ005.)/4+ (Refl06bVBew.SQ001.+
                    Refl06bVBew.SQ005.+ Refl06bVBew.SQ006.+
                    Refl06bVBew.SQ007.)/4+ (Refl07bIVBew.
                    SQ001.+ Refl07bIVBew.SQ002.+ Refl07bIVBew
                    .SQ003.+ Refl07bIVBew.SQ004.)/4+ (
                    Refl07bVBew.SQ001.+ Refl07bVBew.SQ003.+
```

```
                    Refl07bVBew.SQ005.+ Refl07bVBew.SQ007.)/
                    4+ (Refl09bIBew.SQ001.+ Refl09bIBew.SQ002
                    .+ Refl09bIBew.SQ004.)/3+ (Refl09bIVBew.
                    SQ001.+ Refl09bIVBew.SQ002.+ Refl09bIVBew
                    .SQ004.+ Refl09bIVBew.SQ005.)/4+ (
                    Refl09bVIBew.SQ001.+ Refl09bVIBew.SQ004.+
                     Refl09bVIBew.SQ007.)/3+ (Refl12bIBew.
                    SQ002.+ Refl12bIBew.SQ003.+ Refl12bIBew.
                    SQ004.)/3+ (Refl12bIIBew.SQ004.+
                    Refl12bIIBew.SQ006.+ Refl12bIIBew.SQ007.)
                    /3+ (Refl12bIIIBew.SQ001.+ Refl12bIIIBew.
                    SQ002.+ Refl12bIIIBew.SQ005.)/3+ (
                    Refl13bIBew.SQ003.+ Refl13bIBew.SQ004.)/
                    2+ (Refl13bIIBew.SQ002.+ Refl13bIIBew.
                    SQ003.+ Refl13bIIBew.SQ004.+ Refl13bIIBew
                    .SQ007.)/4+ (GesbVIBew.SQ002.+ GesbVIBew.
                    SQ003.+ GesbVIBew.SQ005.+ GesbVIBew.SQ007
                    .+ GesbVIBew.SQ008.)/5+ (GesVIIBew.SQ001
                    .+ GesVIIBew.SQ002.+ GesVIIBew.SQ005.)/3,
47                  ScoreMCBewProz = ScoreMCBew/15*100,
48                  ScoreMCAlt = (Refl02bIIIAlt.SQ002.+
                    Refl02bIIIAlt.SQ004.+ Refl02bIIIAlt.SQ005
                    .+ Refl02bIIIAlt.SQ006.)/4+ (Refl06bIVAlt
                    .SQ001.+ Refl06bIVAlt.SQ002.+
                    Refl06bIVAlt.SQ003.+ Refl06bIVAlt.SQ005.)
                    /4+ (Refl06bVAlt.SQ002.+ Refl06bVAlt.
                    SQ005.+ Refl06bVAlt.SQ006.)/3+ (
                    Refl07bIVAlt.SQ001.+ Refl07bIVAlt.SQ003.+
                     Refl07bIVAlt.SQ005.)/3+ (Refl07bVAlt.
                    SQ004.+ Refl07bVAlt.SQ005.+ Refl07bVAlt.
                    SQ006.+ Refl07bVAlt.SQ007.)/4+ (
                    Refl09bIAlt.SQ003.+ Refl09bIAlt.SQ005.+
                    Refl09bIAlt.SQ006.+ Refl09bIAlt.SQ007.)/
                    4+ (Refl09bIVAlt.SQ002.+ Refl09bIVAlt.
                    SQ003.+ Refl09bIVAlt.SQ005.+ Refl09bIVAlt
                    .SQ006.)/4+ (Refl09bVIAlt.SQ001.+
                    Refl09bVIAlt.SQ005.+Refl09bVIAlt.SQ007.)/
                    3+ (Refl12bIAlt.SQ002.+ Refl12bIAlt.SQ003
                    .+ Refl12bIAlt.SQ006.+ Refl12bIAlt.SQ007
                    .)/4+ (Refl12bIIAlt.SQ005.+ Refl12bIIAlt.
                    SQ006.+ Refl12bIIAlt.SQ007.)/3+ (
                    Refl12bIIIAlt.SQ002.+ Refl12bIIIAlt.SQ003
                    .+ Refl12bIIIAlt.SQ005.)/3+ (Refl13bIAlt.
                    SQ002.+ Refl13bIAlt.SQ004.+ Refl13bIAlt.
                    SQ006.+ Refl13bIAlt.SQ007.)/4+ (
                    Refl13bIIAlt.SQ002.+ Refl13bIIAlt.SQ004.+
                     Refl13bIIAlt.SQ007.)/3+ (GesVIAlt.SQ002
```

```
                    .+ GesVIAlt.SQ004.+ GesVIAlt.SQ005.+
                    GesVIAlt.SQ007.)/4+ (GesVIIAlt.SQ002.+
                    GesVIIAlt.SQ003.+ GesVIIAlt.SQ005.)/3,
49                  ScoreMCAltProz = ScoreMCAlt/15*100,
50                  ScoreMCSV = (Refl03bIVBeob.SQ003.+
                    Refl03bIVBeob.SQ004.+ Refl03bIVBeob.SQ006
                    .)/3+ (Refl03IVUmgang.SQ003.+
                    Refl03IVUmgang.SQ004.+ Refl03IVUmgang.
                    SQ005.)/3,
51                  ScoreMCSVProz = ScoreMCSV/2*100,
52                  ScoreMCNo1 = (Refl02b.SQ001.+ Refl02b.SQ002
                    .+ Refl02b.SQ003.+ Refl02b.SQ004.+
                    Refl02b.SQ005.)/5,
53                  ScoreMCNo2 = (Refl03b.SQ001.+ Refl03b.SQ002
                    .+ Refl03b.SQ003.)/3,
54                  ScoreMCNo3 = (Refl06b.SQ001.+ Refl06b.SQ003
                    .+ Refl06b.SQ004.+ Refl06b.SQ005.)/4,
55                  ScoreMCNo4 = (Refl07b.SQ001.+ Refl07b.SQ002
                    .+ Refl07b.SQ004.+ Refl07b.SQ005.)/4,
56                  ScoreMCNo5 = (Refl09b.SQ001.+ Refl09b.SQ002
                    .+ Refl09b.SQ003.+ Refl09b.SQ004.+
                    Refl09b.SQ006.+ Refl09b.SQ007.)/6,
57                  ScoreMCNo6 = (Refl12b.SQ001.+ Refl12b.SQ002
                    .+ Refl12b.SQ003.)/3,
58                  ScoreMCNo7 = (Refl13b.SQ001.+ Refl13b.SQ002
                    .)/2,
59                  ScoreMCNoG = (Gesb.SQ001.+ Gesb.SQ002.+
                    Gesb.SQ003.+ Gesb.SQ006.+ Gesb.SQ007.)/5,
60                  ScoreMCNo = ScoreMCNo1 + ScoreMCNo2 +
                    ScoreMCNo3 + ScoreMCNo4 + ScoreMCNo5 +
                    ScoreMCNo6 + ScoreMCNo7 + ScoreMCNoG,
61                  Motivation = Mot.SOS2b. + Mot.SOS6b. + Mot.
                    SOS10b.,
62                  PC = paste(toupper(results$PersCode.PC1.),
                    toupper(results$PersCode.PC2.), toupper(
                    results$PersCode.PC3.), toupper(results$
                    PersCode.PC4.), sep="")
63                  )
64 results = results[-21,]
65
66 FS <- as.numeric(gsub("A", "", results$FSPhy))
```

S.2 Zur Profilanalyse

S.2.1 Berechnung der Modelle

```
LPA_Aufgaben<- results[results$Phase!="-oth-",] %>%
  subset(select = c("R1Bew1", "R1Alt1", "R2Bew1", "R2Alt1",
     "R3Bew1", "R3Alt1", "R3Bew2", "R3Alt2", "R4Bew1", "
     R4Alt1", "R4Bew2", "R4Alt2", "R5Bew1", "R5Alt1", "R5Bew2
     ", "R5Alt2", "R5Bew3", "R5Alt3", "R6Bew1", "R6Alt1", "
     R6Bew2", "R6Alt2", "R6Bew3", "R6Alt3", "R7Bew1", "R7Alt1
     ", "R7Bew2", "R7Alt2", "RGBew1", "RGAlt1", "RGBew2", "
     RGAlt2")) %>%
  scale(center=TRUE, scale=FALSE) %>%
  estimate_profiles(1:8)
LPA_Aufgaben$model_1_class_2$fit
LPA_Aufgaben$model_1_class_4$fit
compare_solutions(LPA_Aufgaben, statistics = c("AIC", "BIC",
    "SABIC", "Entropy", "BLRT_p"))
```

Plot der Profile (Für Modell mit 4 Profilen)

```
plot_profiles(LPA_Aufgaben[[4]], rawdata=TRUE, add_line =
   TRUE) +
    theme(axis.text.x = element_text(angle=45, vjust=1,
       hjust=1))
```

Ergänzung der Profilzuordnung (Spalten "Profil2"bzw. "Profil4") sowie numerische Angaben zu Fachsemester und Ausbildungsphase (gegenüber der Darstellung in der Arbeit sind hier - sofern nicht anders angegeben - die Profile in aufsteigender Reihenfolge mit 2, 1, 3 und 4 benannt)

```
results_Profil_Aufgaben <- mutate(results[results$Phase!="-
   oth-",],
                            Profil2 = LPA_Aufgaben$model_1_
                               class_2$dff$Class,
                            Profil4 = LPA_Aufgaben$model_1_
                               class_4$dff$Class,
                            FS_num = as.numeric(gsub("A", ""
                               , results[results$Phase!="-oth
                               -",]$FSPhy)),
                            Phase_num = as.numeric(gsub("A",
                                "", results[results$Phase!="-
                               oth-",]$Phase)))

```

```
#write.csv2(results_Profil_Aufgaben, file="results_Profil_
   Aufgaben.csv") # Bei Bedarf Speichern das Dataframes für
   zukünfitge Analysen
```

S.2.2 Beschreibung der Profile

Betrachtung von Unterschieden zwischen den Profilen auf Ebene der Testwerte

```
kruskal.test(results_Profil_Aufgaben$ScoreMC ~ as.factor(
   results_Profil_Aufgaben$Profil4))
pairwise.wilcox.test(results_Profil_Aufgaben$ScoreMC,
   results_Profil_Aufgaben$Profil4,
                     p.adjust="bonferroni")
results_Profil_Aufgaben %>%
wilcox_effsize(ScoreMC ~ Profil4) %>%
  as.data.frame()

kruskal.test(results_Profil_Aufgaben$ScoreMCBew ~ as.factor(
   results_Profil_Aufgaben$Profil4))
pairwise.wilcox.test(results_Profil_Aufgaben$ScoreMCBew,
   results_Profil_Aufgaben$Profil4,
                     p.adjust="bonferroni")
results_Profil_Aufgaben %>%
wilcox_effsize(ScoreMCBew ~ Profil4) %>%
  as.data.frame()

kruskal.test(results_Profil_Aufgaben$ScoreMCAlt ~ as.factor(
   results_Profil_Aufgaben$Profil4))
pairwise.wilcox.test(results_Profil_Aufgaben$ScoreMCAlt,
   results_Profil_Aufgaben$Profil4,
                     p.adjust="bonferroni")
results_Profil_Aufgaben %>%
wilcox_effsize(ScoreMCAlt ~ Profil4) %>%
  as.data.frame()
```

S.2.3 Längsschnittliche Betrachtungen

Aufbereitung der Datensätze

Erstellung der beiden Dataframes zu echten Bachelor-Längsschnitten

Kohorte mit Code 1005/1012

```
# 1. Studierenden Prä-Post Didaktische Rekonstruktion (PB WS 2021/22) (Dataframe: results_1005_1012_sort)
results_1005 <- subset(results_Profil_Aufgaben[results_Profil_Aufgaben$Code=="1005",])
results_1005 <- mutate(results_1005,
                       PP = "Prae") # für späteres Zusammenfügen der verschiedenen Prä-Post-Gruppen
results_1012 <- subset(results_Profil_Aufgaben[results_Profil_Aufgaben$Code=="1012",])
results_1012 <- mutate(results_1012,
                       PP = "Post") # für späteres Zusammenfügen der verschiedenen Prä-Post-Gruppen

colnames(results_1005) <- paste(colnames(results_1005), ".Prae", sep="")
N_1005 <- length(results_1005$id)

colnames(results_1012) <- paste(colnames(results_1012), ".Post", sep="")
N_1012 <- length(results_1012$id)


results_1005_1012_sort <- data.frame()
for (i in 1:N_1005){
  j <- which((results_1005$PC[i]==results_1012$PC)==TRUE)
  if (length(j)>0) {
    results_1005_1012_sort <- rbind(results_1005_1012_sort, c(results_1005[i,],
                                    results_1012[j,]))
    colnames(results_1005_1012_sort) <- c(colnames(results_1005), colnames(results_1012))

    }
  else {}
}
```

Kohorte mit Code 1021/1027

```
results_1021 <- subset(results_Profil_Aufgaben[results_
   Profil_Aufgaben$Code=="1021",])
results_1021 <- mutate(results_1021,
                       PP = "Prae") # für späteres Zusammenf
                          ügen der verschiedenen Prä-Post-
                          Gruppen
results_1027 <- subset(results_Profil_Aufgaben[results_
   Profil_Aufgaben$Code=="1027",])
results_1027 <- mutate(results_1027,
                       PP = "Post") # für späteres Zusammenf
                          ügen der verschiedenen Prä-Post-
                          Gruppen

colnames(results_1021) <- paste(colnames(results_1021), ".
   Prae", sep="")
N_1021 <- length(results_1021$id)

colnames(results_1027) <- paste(colnames(results_1027), ".
   Post", sep="")
N_1027 <- length(results_1027$id)


results_1021_1027_sort <- data.frame()
for (i in 1:N_1021){
  j <- which((results_1021$PC[i]==results_1027$PC)==TRUE)
  if (length(j)>0) {
    results_1021_1027_sort <- rbind(results_1021_1027_sort,
       c(results_1021[i,],
                      results_1027[j,]))
    colnames(results_1021_1027_sort) <- c(colnames(results_
       1021), colnames(results_1027))

    }
  else {}
}
```

Erstellung des Dataframes für die Prä-Post-PS-Studierenden

```
results_nur_PS_Profil <- subset(results_Profil_Aufgaben[(
  results_Profil_Aufgaben$Code=="1009" | results_Profil_
     Aufgaben$Code=="1024" |
    results_Profil_Aufgaben$Code=="1016" | results_Profil_
       Aufgaben$Code=="1025" |
    results_Profil_Aufgaben$Code=="1029" | results_Profil_
       Aufgaben$Code=="1035" |
```

```
5     results_Profil_Aufgaben$Code=="1030" | results_Profil_
        Aufgaben$Code=="1036"),])
6
7  results_nur_PS_Profil <- mutate(results_nur_PS_Profil,
8                                  PP = case_when(results_nur_
                                     PS_Profil$Code=="1009" ~ "
                                     Prae",
9                                   results_nur_PS_Profil$Code
                                     =="1024" ~ "Post",
10                                  results_nur_PS_Profil$Code
                                     =="1016" ~ "Prae",
11                                  results_nur_PS_Profil$Code
                                     =="1025" ~ "Post",
12                                  results_nur_PS_Profil$Code
                                     =="1029" ~ "Prae",
13                                  results_nur_PS_Profil$Code
                                     =="1035" ~ "Post",
14                                  results_nur_PS_Profil$Code
                                     =="1030" ~ "Prae",
15                                  results_nur_PS_Profil$Code
                                     =="1036" ~ "Post",
16                                  ))
17
18 results_nur_PS_Prae_Profil <- subset(results_nur_PS_Profil[
     results_nur_PS_Profil$PP=="Prae",])
19 colnames(results_nur_PS_Prae_Profil) <- paste(colnames(
     results_nur_PS_Prae_Profil), ".Prae", sep="")
20 N_Prae <- length(results_nur_PS_Prae_Profil$id)
21
22 results_nur_PS_Post_Profil <- subset(results_nur_PS_Profil[
     results_nur_PS_Profil$PP=="Post",])
23 colnames(results_nur_PS_Post_Profil) <- paste(colnames(
     results_nur_PS_Post_Profil), ".Post", sep="")
24 N_Post <- length(results_nur_PS_Post_Profil$id)
25
26
27 results_PS_sort_Profil=data.frame()
28 for (i in 1:N_Prae){
29   j <- which((results_nur_PS_Prae_Profil$PC[i]==results_nur_
       PS_Post_Profil$PC)==TRUE)
30   if (length(j)>0) {
31     results_PS_sort_Profil <- rbind(results_PS_sort_Profil,
         c(results_nur_PS_Prae_Profil[i,],
32                       results_nur_PS_Post_Profil[j,]))
33     colnames(results_PS_sort_Profil) <- c(colnames(results_
         nur_PS_Prae_Profil), colnames(results_nur_PS_Post_
         Profil))
```

```
34
35      }
36    else {}
37 }
```

Zusammenfügen aller Prä-Post-Gruppen in einen Dataframe

```
1  results_Laengsschnitte <- rbind(results_1005_1012_sort,
2                                  results_1021_1027_sort,
3                                  results_PS_sort_Profil)
4
5  results_Laengsschnitte_BA <- rbind(results_1005_1012_sort,
6                                  results_1021_1027_sort)
7  # Hier ggf. noch Infos über Personen ergänzen, die den Test
      >2x bearbeitet haben?
8
9  # Nachfolgend Übersicht über Profilwechsel Gesamt und nach
      Teilkohorten
10 xtabs(~ results_Laengsschnitte$Profil4.Prae + results_
      Laengsschnitte$Profil4.Post)
11
12 xtabs(~ results_1005_1012_sort$Profil4.Prae + results_1005_
      1012_sort$Profil4.Post)
13
14 xtabs(~ results_1021_1027_sort$Profil4.Prae + results_1021_
      1027_sort$Profil4.Post)
15
16 xtabs(~ results_PS_sort_Profil$Profil4.Prae + results_PS_
      sort_Profil$Profil4.Post)
```

Erstellung des Alluvial-Plots

Für den Alluvial-Plot wird die Reihenfolge der Profile angepasst (nun in aufsteigender Reihenfolge: 1, 2, 3, 4)

```
1  Profil4.Prae.rekod <- recode(results_Laengsschnitte$Profil4.
      Prae, "2=1;␣1=2")
2  Profil4.Post.rekod <- recode(results_Laengsschnitte$Profil4.
      Post, "2=1;␣1=2")
3
4  ggplot(data = results_Laengsschnitte,
5         aes(axis1 = Profil4.Prae.rekod, axis2 = Profil4.Post.
             rekod)) +
6    geom_alluvium(aes(fill = factor(Profil4.Prae.rekod)), show
       .legend=FALSE) +
```

```
  geom_stratum() +
  scale_fill_manual(values = c("blue", "green", "red", "
     purple")) +
  geom_text(stat = "stratum",
            aes(label = paste0("Profil␣", stratum,
                               ifelse(nchar(as.character(
                                  stratum)) == 1L,
                                      "\n"), "(",
                               "N␣=␣", after_stat(n), ")")))
                                  +
  scale_x_discrete(limits = c("Prä", "Post"))+
  theme_void()+
  theme(axis.text.x = element_text())
```

S.3 Charakterisierung der Fähigkeitsstufen über inhaltliche Merkmale der Aufgaben

Betrachtung von Unterschieden zwischen den Fähigkeitsstufen auf Ebene der Aufgabenscores

```
kruskal.test(results_Profil_Aufgaben$R1Bew1 ~ as.factor(
   results_Profil_Aufgaben$Profil4))
pairwise.wilcox.test(results_Profil_Aufgaben$R1Bew1, results
   _Profil_Aufgaben$Profil4,
                     p.adjust="bonferroni")
results_Profil_Aufgaben %>%
wilcox_effsize(R1Bew1 ~ Profil4) %>%
  as.data.frame()

kruskal.test(results_Profil_Aufgaben$R1Alt1 ~ as.factor(
   results_Profil_Aufgaben$Profil4))
pairwise.wilcox.test(results_Profil_Aufgaben$R1Alt1, results
   _Profil_Aufgaben$Profil4,
                     p.adjust="bonferroni")
results_Profil_Aufgaben %>%
wilcox_effsize(R1Alt1 ~ Profil4) %>%
  as.data.frame()

kruskal.test(results_Profil_Aufgaben$R2Bew1 ~ as.factor(
   results_Profil_Aufgaben$Profil4))
pairwise.wilcox.test(results_Profil_Aufgaben$R2Bew1, results
   _Profil_Aufgaben$Profil4,
                     p.adjust="bonferroni")
results_Profil_Aufgaben %>%
```

```
19 wilcox_effsize(R2Bew1 ~ Profil4) %>%
20   as.data.frame()
21
22 kruskal.test(results_Profil_Aufgaben$R2Alt1 ~ as.factor(
     results_Profil_Aufgaben$Profil4))
23 pairwise.wilcox.test(results_Profil_Aufgaben$R2Alt1, results
     _Profil_Aufgaben$Profil4,
24                      p.adjust="bonferroni")
25 results_Profil_Aufgaben %>%
26 wilcox_effsize(R2Alt1 ~ Profil4) %>%
27   as.data.frame()
28
29 kruskal.test(results_Profil_Aufgaben$R3Bew1 ~ as.factor(
     results_Profil_Aufgaben$Profil4))
30 pairwise.wilcox.test(results_Profil_Aufgaben$R3Bew1, results
     _Profil_Aufgaben$Profil4,
31                      p.adjust="bonferroni")
32 results_Profil_Aufgaben %>%
33 wilcox_effsize(R3Bew1 ~ Profil4) %>%
34   as.data.frame()
35
36 kruskal.test(results_Profil_Aufgaben$R3Alt1 ~ as.factor(
     results_Profil_Aufgaben$Profil4))
37 pairwise.wilcox.test(results_Profil_Aufgaben$R3Alt1, results
     _Profil_Aufgaben$Profil4,
38                      p.adjust="bonferroni")
39 results_Profil_Aufgaben %>%
40 wilcox_effsize(R3Alt1 ~ Profil4) %>%
41   as.data.frame()
42
43 kruskal.test(results_Profil_Aufgaben$R3Bew2 ~ as.factor(
     results_Profil_Aufgaben$Profil4))
44 pairwise.wilcox.test(results_Profil_Aufgaben$R3Bew2, results
     _Profil_Aufgaben$Profil4,
45                      p.adjust="bonferroni")
46 results_Profil_Aufgaben %>%
47 wilcox_effsize(R3Bew2 ~ Profil4) %>%
48   as.data.frame()
49
50 kruskal.test(results_Profil_Aufgaben$R3Alt2 ~ as.factor(
     results_Profil_Aufgaben$Profil4))
51 pairwise.wilcox.test(results_Profil_Aufgaben$R3Alt2, results
     _Profil_Aufgaben$Profil4,
52                      p.adjust="bonferroni")
53 results_Profil_Aufgaben %>%
54 wilcox_effsize(R3Alt2 ~ Profil4) %>%
55   as.data.frame()
```

```
56
57 kruskal.test(results_Profil_Aufgaben$R4Bew1 ~ as.factor(
      results_Profil_Aufgaben$Profil4))
58 pairwise.wilcox.test(results_Profil_Aufgaben$R4Bew1, results
      _Profil_Aufgaben$Profil4,
59                     p.adjust="bonferroni")
60 results_Profil_Aufgaben %>%
61 wilcox_effsize(R4Bew1 ~ Profil4) %>%
62   as.data.frame()
63
64 kruskal.test(results_Profil_Aufgaben$R4Alt1 ~ as.factor(
      results_Profil_Aufgaben$Profil4))
65 pairwise.wilcox.test(results_Profil_Aufgaben$R4Alt1, results
      _Profil_Aufgaben$Profil4,
66                     p.adjust="bonferroni")
67 results_Profil_Aufgaben %>%
68 wilcox_effsize(R4Alt1 ~ Profil4) %>%
69   as.data.frame()
70
71 kruskal.test(results_Profil_Aufgaben$R4Bew2 ~ as.factor(
      results_Profil_Aufgaben$Profil4))
72 pairwise.wilcox.test(results_Profil_Aufgaben$R4Bew2, results
      _Profil_Aufgaben$Profil4,
73                     p.adjust="bonferroni")
74 results_Profil_Aufgaben %>%
75 wilcox_effsize(R4Bew2 ~ Profil4) %>%
76   as.data.frame()
77
78 kruskal.test(results_Profil_Aufgaben$R4Alt2 ~ as.factor(
      results_Profil_Aufgaben$Profil4))
79 pairwise.wilcox.test(results_Profil_Aufgaben$R4Alt2, results
      _Profil_Aufgaben$Profil4,
80                     p.adjust="bonferroni")
81 results_Profil_Aufgaben %>%
82 wilcox_effsize(R4Alt2 ~ Profil4) %>%
83   as.data.frame()
84
85 kruskal.test(results_Profil_Aufgaben$R5Bew1 ~ as.factor(
      results_Profil_Aufgaben$Profil4))
86 pairwise.wilcox.test(results_Profil_Aufgaben$R5Bew1, results
      _Profil_Aufgaben$Profil4,
87                     p.adjust="bonferroni")
88 results_Profil_Aufgaben %>%
89 wilcox_effsize(R5Bew1 ~ Profil4) %>%
90   as.data.frame()
91
```

```
92 kruskal.test(results_Profil_Aufgaben$R5Alt1 ~ as.factor(
     results_Profil_Aufgaben$Profil4))
93 pairwise.wilcox.test(results_Profil_Aufgaben$R5Alt1, results
     _Profil_Aufgaben$Profil4,
94                      p.adjust="bonferroni")
95 results_Profil_Aufgaben %>%
96 wilcox_effsize(R5Alt1 ~ Profil4) %>%
97   as.data.frame()
98
99 kruskal.test(results_Profil_Aufgaben$R5Bew2 ~ as.factor(
     results_Profil_Aufgaben$Profil4))
100 pairwise.wilcox.test(results_Profil_Aufgaben$R5Bew2, results
     _Profil_Aufgaben$Profil4,
101                      p.adjust="bonferroni")
102 results_Profil_Aufgaben %>%
103 wilcox_effsize(R5Bew2 ~ Profil4) %>%
104   as.data.frame()
105
106 kruskal.test(results_Profil_Aufgaben$R5Alt2 ~ as.factor(
     results_Profil_Aufgaben$Profil4))
107 pairwise.wilcox.test(results_Profil_Aufgaben$R5Alt2, results
     _Profil_Aufgaben$Profil4,
108                      p.adjust="bonferroni")
109 results_Profil_Aufgaben %>%
110 wilcox_effsize(R5Alt2 ~ Profil4) %>%
111   as.data.frame()
112
113 kruskal.test(results_Profil_Aufgaben$R5Bew3 ~ as.factor(
     results_Profil_Aufgaben$Profil4))
114 pairwise.wilcox.test(results_Profil_Aufgaben$R5Bew3, results
     _Profil_Aufgaben$Profil4,
115                      p.adjust="bonferroni")
116 results_Profil_Aufgaben %>%
117 wilcox_effsize(R5Bew3 ~ Profil4) %>%
118   as.data.frame()
119
120 kruskal.test(results_Profil_Aufgaben$R5Alt3 ~ as.factor(
     results_Profil_Aufgaben$Profil4))
121 pairwise.wilcox.test(results_Profil_Aufgaben$R5Alt3, results
     _Profil_Aufgaben$Profil4,
122                      p.adjust="bonferroni")
123 results_Profil_Aufgaben %>%
124 wilcox_effsize(R5Alt3 ~ Profil4) %>%
125   as.data.frame()
126
127 kruskal.test(results_Profil_Aufgaben$R6Bew1 ~ as.factor(
     results_Profil_Aufgaben$Profil4))
```

```
128 pairwise.wilcox.test(results_Profil_Aufgaben$R6Bew1, results
    _Profil_Aufgaben$Profil4,
129                      p.adjust="bonferroni")
130 results_Profil_Aufgaben %>%
131 wilcox_effsize(R6Bew1 ~ Profil4) %>%
132   as.data.frame()
133
134 kruskal.test(results_Profil_Aufgaben$R6Alt1 ~ as.factor(
    results_Profil_Aufgaben$Profil4))
135 pairwise.wilcox.test(results_Profil_Aufgaben$R6Alt1, results
    _Profil_Aufgaben$Profil4,
136                      p.adjust="bonferroni")
137 results_Profil_Aufgaben %>%
138 wilcox_effsize(R6Alt1 ~ Profil4) %>%
139   as.data.frame()
140
141 kruskal.test(results_Profil_Aufgaben$R6Bew2 ~ as.factor(
    results_Profil_Aufgaben$Profil4))
142 pairwise.wilcox.test(results_Profil_Aufgaben$R6Bew2, results
    _Profil_Aufgaben$Profil4,
143                      p.adjust="bonferroni")
144 results_Profil_Aufgaben %>%
145 wilcox_effsize(R6Bew2 ~ Profil4) %>%
146   as.data.frame()
147
148 kruskal.test(results_Profil_Aufgaben$R6Alt2 ~ as.factor(
    results_Profil_Aufgaben$Profil4))
149 pairwise.wilcox.test(results_Profil_Aufgaben$R6Alt2, results
    _Profil_Aufgaben$Profil4,
150                      p.adjust="bonferroni")
151 results_Profil_Aufgaben %>%
152 wilcox_effsize(R6Alt2 ~ Profil4) %>%
153   as.data.frame()
154
155 kruskal.test(results_Profil_Aufgaben$R6Bew3 ~ as.factor(
    results_Profil_Aufgaben$Profil4))
156 pairwise.wilcox.test(results_Profil_Aufgaben$R6Bew3, results
    _Profil_Aufgaben$Profil4,
157                      p.adjust="bonferroni")
158 results_Profil_Aufgaben %>%
159 wilcox_effsize(R6Bew3 ~ Profil4) %>%
160   as.data.frame()
161
162 kruskal.test(results_Profil_Aufgaben$R6Alt3 ~ as.factor(
    results_Profil_Aufgaben$Profil4))
163 pairwise.wilcox.test(results_Profil_Aufgaben$R6Alt3, results
    _Profil_Aufgaben$Profil4,
```

```
164                     p.adjust="bonferroni")
165 results_Profil_Aufgaben %>%
166 wilcox_effsize(R6Alt3 ~ Profil4) %>%
167   as.data.frame()
168
169 kruskal.test(results_Profil_Aufgaben$R7Bew1 ~ as.factor(
      results_Profil_Aufgaben$Profil4))
170 pairwise.wilcox.test(results_Profil_Aufgaben$R7Bew1, results
      _Profil_Aufgaben$Profil4,
171                     p.adjust="bonferroni")
172 results_Profil_Aufgaben %>%
173 wilcox_effsize(R7Bew1 ~ Profil4) %>%
174   as.data.frame()
175
176 kruskal.test(results_Profil_Aufgaben$R7Alt1 ~ as.factor(
      results_Profil_Aufgaben$Profil4))
177 pairwise.wilcox.test(results_Profil_Aufgaben$R7Alt1, results
      _Profil_Aufgaben$Profil4,
178                     p.adjust="bonferroni")
179 results_Profil_Aufgaben %>%
180 wilcox_effsize(R7Alt1 ~ Profil4) %>%
181   as.data.frame()
182
183 kruskal.test(results_Profil_Aufgaben$R7Bew2 ~ as.factor(
      results_Profil_Aufgaben$Profil4))
184 pairwise.wilcox.test(results_Profil_Aufgaben$R7Bew2, results
      _Profil_Aufgaben$Profil4,
185                     p.adjust="bonferroni")
186 results_Profil_Aufgaben %>%
187 wilcox_effsize(R7Bew2 ~ Profil4) %>%
188   as.data.frame()
189
190 kruskal.test(results_Profil_Aufgaben$R7Alt2 ~ as.factor(
      results_Profil_Aufgaben$Profil4))
191 pairwise.wilcox.test(results_Profil_Aufgaben$R7Alt2, results
      _Profil_Aufgaben$Profil4,
192                     p.adjust="bonferroni")
193 results_Profil_Aufgaben %>%
194 wilcox_effsize(R7Alt2 ~ Profil4) %>%
195   as.data.frame()
196
197 kruskal.test(results_Profil_Aufgaben$RGBew1 ~ as.factor(
      results_Profil_Aufgaben$Profil4))
198 pairwise.wilcox.test(results_Profil_Aufgaben$RGBew1, results
      _Profil_Aufgaben$Profil4,
199                     p.adjust="bonferroni")
200 results_Profil_Aufgaben %>%
```

```
wilcox_effsize(RGBew1 ~ Profil4) %>%
  as.data.frame()

kruskal.test(results_Profil_Aufgaben$RGAlt1 ~ as.factor(
   results_Profil_Aufgaben$Profil4))
pairwise.wilcox.test(results_Profil_Aufgaben$RGAlt1, results
   _Profil_Aufgaben$Profil4,
                     p.adjust="bonferroni")
results_Profil_Aufgaben %>%
wilcox_effsize(RGAlt1 ~ Profil4) %>%
  as.data.frame()

kruskal.test(results_Profil_Aufgaben$RGBew2 ~ as.factor(
   results_Profil_Aufgaben$Profil4))
pairwise.wilcox.test(results_Profil_Aufgaben$RGBew2, results
   _Profil_Aufgaben$Profil4,
                     p.adjust="bonferroni")
results_Profil_Aufgaben %>%
wilcox_effsize(RGBew2 ~ Profil4) %>%
  as.data.frame()

kruskal.test(results_Profil_Aufgaben$RGAlt2 ~ as.factor(
   results_Profil_Aufgaben$Profil4))
pairwise.wilcox.test(results_Profil_Aufgaben$RGAlt2, results
   _Profil_Aufgaben$Profil4,
                     p.adjust="bonferroni")
results_Profil_Aufgaben %>%
wilcox_effsize(RGAlt2 ~ Profil4) %>%
  as.data.frame()
```

S.4 Charakterisierung der Fähigkeitsstufen über demografische Variablen

S.4.1 Ausbildungsphase

```
kruskal.test(results_Profil_Aufgaben$Phase_num ~ as.factor(
   results_Profil_Aufgaben$Profil4))
pairwise.wilcox.test(results_Profil_Aufgaben$Phase_num,
   results_Profil_Aufgaben$Profil4,
                     p.adjust="bonferroni")
results_Profil_Aufgaben %>%
wilcox_effsize(Phase_num ~ Profil4) %>%
  as.data.frame()
```

S.4.2 Fachsemester

```
kruskal.test(results_Profil_Aufgaben$FS_num ~ as.factor(
   results_Profil_Aufgaben$Profil4))
```

S.4.3 Umfang der Erfahrungen im Pflichtpraktikum

```
kruskal.test(results_Profil_Aufgaben$WoPflichtpraktika ~ as.
   factor(results_Profil_Aufgaben$Profil4))
pairwise.wilcox.test(results_Profil_Aufgaben$
   WoPflichtpraktika, results_Profil_Aufgaben$Profil4,
                     p.adjust="bonferroni")
results_Profil_Aufgaben %>%
wilcox_effsize(WoPflichtpraktika ~ Profil4) %>%
  as.data.frame()
```

S.4.4 Testmotivation

```
kruskal.test(results_Profil_Aufgaben$Motivation ~ as.factor(
   results_Profil_Aufgaben$Profil4))
pairwise.wilcox.test(results_Profil_Aufgaben$Motivation,
   results_Profil_Aufgaben$Profil4,
                     p.adjust="bonferroni")

results_Profil_Aufgaben %>%
wilcox_effsize(Motivation ~ Profil4) %>%
  as.data.frame()
```

S.4.5 Bearbeitungsdauer

```
kruskal.test(results_Profil_Aufgaben$interviewtime ~ as.
   factor(results_Profil_Aufgaben$Profil4))
pairwise.wilcox.test(results_Profil_Aufgaben$interviewtime,
   results_Profil_Aufgaben$Profil4,
                     p.adjust="bonferroni")

results_Profil_Aufgaben %>%
wilcox_effsize(interviewtime ~ Profil4) %>%
  as.data.frame()
```

T *R*-Skript für die statistischen Auswertungen zum Zusammenhang von Reflexion und Noticing

T.1 Vorbetrachtungen zum Testwert zur Noticingfähigkeit

```
psych::alpha(subset(results_Profil_Aufgaben, select = c("
    ScoreMCNo1","ScoreMCNo2","ScoreMCNo3", "ScoreMCNo4", "
    ScoreMCNo5", "ScoreMCNo6", "ScoreMCNo7", "ScoreMCNoG")),
    check.keys =TRUE)
LH <- c(mean(results_Profil_Aufgaben$ScoreMCNo1), mean(
    results_Profil_Aufgaben$ScoreMCNo2), mean(results_Profil_
    Aufgaben$ScoreMCNo3), mean(results_Profil_Aufgaben$
    ScoreMCNo4), mean(results_Profil_Aufgaben$ScoreMCNo5),
    mean(results_Profil_Aufgaben$ScoreMCNo6), mean(results_
    Profil_Aufgaben$ScoreMCNo7), mean(results_Profil_Aufgaben$
    ScoreMCNoG))
```

T.2 Hypothese 1: Korrelation zwischen den Testwerten der Fähigkeiten zum Noticing und zur Reflexion

```
ks.test(results_Profil_Aufgaben$ScoreMC, "pnorm", mean=mean(
    results_Profil_Aufgaben$ScoreMC), sd=sd(results_Profil_
    Aufgaben$ScoreMC))
ks.test(results_Profil_Aufgaben$ScoreMCNo, "pnorm", mean=
    mean(results_Profil_Aufgaben$ScoreMCNo), sd=sd(results_
    Profil_Aufgaben$ScoreMCNo))
```

```
3 cor.test(results_Profil_Aufgaben$ScoreMC, results_Profil_
    Aufgaben$ScoreMCNo, method = "pearson")
```

T.3 Hypothese 2: Unterschiede in den Testwerten der Fähigkeiten zum Noticing zwischen den Profilen

```
1  ks.test(results_Profil_Aufgaben[results_Profil_Aufgaben$
     Profil4==2,]$ScoreMCNo, "pnorm", mean=mean(results_Profil_
     Aufgaben[results_Profil_Aufgaben$Profil4==2,]$ScoreMCNo),
     sd=sd(results_Profil_Aufgaben[results_Profil_Aufgaben$
     Profil4==2,]$ScoreMCNo))
2
3  ks.test(results_Profil_Aufgaben[results_Profil_Aufgaben$
     Profil4==1,]$ScoreMCNo, "pnorm", mean=mean(results_Profil_
     Aufgaben[results_Profil_Aufgaben$Profil4==1,]$ScoreMCNo),
     sd=sd(results_Profil_Aufgaben[results_Profil_Aufgaben$
     Profil4==1,]$ScoreMCNo))
4
5  ks.test(results_Profil_Aufgaben[results_Profil_Aufgaben$
     Profil4==3,]$ScoreMCNo, "pnorm", mean=mean(results_Profil_
     Aufgaben[results_Profil_Aufgaben$Profil4==3,]$ScoreMCNo),
     sd=sd(results_Profil_Aufgaben[results_Profil_Aufgaben$
     Profil4==3,]$ScoreMCNo))
6
7  ks.test(results_Profil_Aufgaben[results_Profil_Aufgaben$
     Profil4==4,]$ScoreMCNo, "pnorm", mean=mean(results_Profil_
     Aufgaben[results_Profil_Aufgaben$Profil4==4,]$ScoreMCNo),
     sd=sd(results_Profil_Aufgaben[results_Profil_Aufgaben$
     Profil4==4,]$ScoreMCNo))
8
9  leveneTest(results_Profil_Aufgaben$ScoreMCNo, as.factor(
     results_Profil_Aufgaben$Profil4))
10
11 Anova(aov(results_Profil_Aufgaben$ScoreMCNo ~ as.factor(
     results_Profil_Aufgaben$Profil4)), type="3")
12 pairwise.t.test(results_Profil_Aufgaben$ScoreMCNo, as.factor
     (results_Profil_Aufgaben$Profil4), p.adjust="bonferroni")
13 results_Profil_Aufgaben %>%
14   cohens_d(ScoreMCNo ~ Profil4) %>%
15   as.data.frame
16
```

```
mean(results_Profil_Aufgaben[results_Profil_Aufgaben$Profil4
   ==2,]$ScoreMCNo)
mean(results_Profil_Aufgaben[results_Profil_Aufgaben$Profil4
   ==1,]$ScoreMCNo)
mean(results_Profil_Aufgaben[results_Profil_Aufgaben$Profil4
   ==3,]$ScoreMCNo)
mean(results_Profil_Aufgaben[results_Profil_Aufgaben$Profil4
   ==4,]$ScoreMCNo)
```

T.4 Hypothese 3: Betrachtung der Abhängigkeit der Aufgabenscores abhängig von der Anwahl in den Noticing-Aufgaben in den Profilen

T.4.1 Bestimmung der zusätzlichen Reflexionsscores $\text{Reflexionsscore}_{\text{angewählt im Noticing}}$ und $\text{Reflexionsscore}_{\textit{nicht}\text{ angewählt im Noticing}}$

```
results_Profil_Aufgaben_No <- mutate(results_Profil_Aufgaben
   ,
                  ScoreBew_No = (ifelse(Refl02b.SQ003.==1,
                     R1Bew1,0) +
                    ifelse(Refl03b.SQ003.==1,R2Bew1,0) +
                    ifelse(Refl06b.SQ004.==1,R3Bew1,0) +
                       ifelse(Refl06b.SQ005.==1,R3Bew2,0) +
                    ifelse(Refl07b.SQ004.==1,R4Bew1,0) +
                       ifelse(Refl07b.SQ005.==1,R4Bew2,0) +
                    ifelse(Refl09b.SQ001.==1,R5Bew1,0) +
                       ifelse(Refl09b.SQ004.==1,R5Bew2,0) +
                       ifelse(Refl09b.SQ006.==1,R5Bew3,0) +
                    ifelse(Refl12b.SQ001.==1,R6Bew1,0) +
                       ifelse(Refl12b.SQ002.==1,R6Bew2,0) +
                       ifelse(Refl12b.SQ003.==1,R6Bew3,0) +
                    ifelse(Refl13b.SQ001.==1,R7Bew1,0) +
                       ifelse(Refl13b.SQ002.==1,R7Bew2,0) +
                    ifelse(Gesb.SQ006.==1,RGBew1,0) + ifelse
                       (Gesb.SQ007.==1,RGBew2,0))/
                    (Refl02b.SQ003. + Refl03b.SQ003. +
                       Refl06b.SQ004. + Refl06b.SQ005. +
                       Refl07b.SQ004. + Refl07b.SQ005. +
                       Refl09b.SQ001. + Refl09b.SQ004. +
                       Refl09b.SQ006. + Refl12b.SQ001. +
                       Refl12b.SQ002. + Refl12b.SQ003. +
```

```
                    Refl13b.SQ001. + Refl13b.SQ002. + Gesb
                    .SQ006. + Gesb.SQ007.),
11                  ScoreAlt_No = (ifelse(Refl02b.SQ003.==1,
                    R1Alt1,0) +
12                  ifelse(Refl03b.SQ003.==1,R2Alt1,0) +
13                  ifelse(Refl06b.SQ004.==1,R3Alt1,0) +
                    ifelse(Refl06b.SQ005.==1,R3Alt2,0) +
14                  ifelse(Refl07b.SQ004.==1,R4Alt1,0) +
                    ifelse(Refl07b.SQ005.==1,R4Alt2,0) +
15                  ifelse(Refl09b.SQ001.==1,R5Alt1,0) +
                    ifelse(Refl09b.SQ004.==1,R5Alt2,0) +
                    ifelse(Refl09b.SQ006.==1,R5Alt3,0) +
16                  ifelse(Refl12b.SQ001.==1,R6Alt1,0) +
                    ifelse(Refl12b.SQ002.==1,R6Alt2,0) +
                    ifelse(Refl12b.SQ003.==1,R6Alt3,0) +
17                  ifelse(Refl13b.SQ001.==1,R7Alt1,0) +
                    ifelse(Refl13b.SQ002.==1,R7Alt2,0) +
18                  ifelse(Gesb.SQ006.==1,RGAlt1,0) + ifelse
                    (Gesb.SQ007.==1,RGAlt2,0))/
19                  (Refl02b.SQ003. + Refl03b.SQ003. +
                    Refl06b.SQ004. + Refl06b.SQ005. +
                    Refl07b.SQ004. + Refl07b.SQ005. +
                    Refl09b.SQ001. + Refl09b.SQ004. +
                    Refl09b.SQ006. + Refl12b.SQ001. +
                    Refl12b.SQ002. + Refl12b.SQ003. +
                    Refl13b.SQ001. + Refl13b.SQ002. + Gesb
                    .SQ006. + Gesb.SQ007.),
20                  Anzahl_No = (Refl02b.SQ003. + Refl03b.
                    SQ003. + Refl06b.SQ004. + Refl06b.SQ005.
                     + Refl07b.SQ004. + Refl07b.SQ005. +
                    Refl09b.SQ001. + Refl09b.SQ004. +
                    Refl09b.SQ006. + Refl12b.SQ001. +
                    Refl12b.SQ002. + Refl12b.SQ003. +
                    Refl13b.SQ001. + Refl13b.SQ002. + Gesb.
                    SQ006. + Gesb.SQ007.),
21                  Score_No = ScoreBew_No + ScoreAlt_No,
22
23                  ScoreBew_o_No = (ifelse(Refl02b.SQ003.==0,
                    R1Bew1,0) +
24                  ifelse(Refl03b.SQ003.==0,R2Bew1,0) +
25                  ifelse(Refl06b.SQ004.==0,R3Bew1,0) +
                    ifelse(Refl06b.SQ005.==0,R3Bew2,0) +
26                  ifelse(Refl07b.SQ004.==0,R4Bew1,0) +
                    ifelse(Refl07b.SQ005.==0,R4Bew2,0) +
27                  ifelse(Refl09b.SQ001.==0,R5Bew1,0) +
                    ifelse(Refl09b.SQ004.==0,R5Bew2,0) +
                    ifelse(Refl09b.SQ006.==0,R5Bew3,0) +
```

```
28                    ifelse(Refl12b.SQ001.==0,R6Bew1,0) +
                         ifelse(Refl12b.SQ002.==0,R6Bew2,0) +
                         ifelse(Refl12b.SQ003.==0,R6Bew3,0) +
29                    ifelse(Refl13b.SQ001.==0,R7Bew1,0) +
                         ifelse(Refl13b.SQ002.==0,R7Bew2,0) +
30                    ifelse(Gesb.SQ006.==0,RGBew1,0) + ifelse
                         (Gesb.SQ007.==0,RGBew2,0))/
31                    (16 - (Refl02b.SQ003. + Refl03b.SQ003. +
                          Refl06b.SQ004. + Refl06b.SQ005. +
                         Refl07b.SQ004. + Refl07b.SQ005. +
                         Refl09b.SQ001. + Refl09b.SQ004. +
                         Refl09b.SQ006. + Refl12b.SQ001. +
                         Refl12b.SQ002. + Refl12b.SQ003. +
                         Refl13b.SQ001. + Refl13b.SQ002. + Gesb
                         .SQ006. + Gesb.SQ007.)),
32                  ScoreAlt_o_No = (ifelse(Refl02b.SQ003.==0,
                       R1Alt1,0) +
33                    ifelse(Refl03b.SQ003.==0,R2Alt1,0) +
34                    ifelse(Refl06b.SQ004.==0,R3Alt1,0) +
                         ifelse(Refl06b.SQ005.==0,R3Alt2,0) +
35                    ifelse(Refl07b.SQ004.==0,R4Alt1,0) +
                         ifelse(Refl07b.SQ005.==0,R4Alt2,0) +
36                    ifelse(Refl09b.SQ001.==0,R5Alt1,0) +
                         ifelse(Refl09b.SQ004.==0,R5Alt2,0) +
                         ifelse(Refl09b.SQ006.==0,R5Alt3,0) +
37                    ifelse(Refl12b.SQ001.==0,R6Alt1,0) +
                         ifelse(Refl12b.SQ002.==0,R6Alt2,0) +
                         ifelse(Refl12b.SQ003.==0,R6Alt3,0) +
38                    ifelse(Refl13b.SQ001.==0,R7Alt1,0) +
                         ifelse(Refl13b.SQ002.==0,R7Alt2,0) +
39                    ifelse(Gesb.SQ006.==0,RGAlt1,0) + ifelse
                         (Gesb.SQ007.==0,RGAlt2,0))/
40                    (16 - (Refl02b.SQ003. + Refl03b.SQ003. +
                          Refl06b.SQ004. + Refl06b.SQ005. +
                         Refl07b.SQ004. + Refl07b.SQ005. +
                         Refl09b.SQ001. + Refl09b.SQ004. +
                         Refl09b.SQ006. + Refl12b.SQ001. +
                         Refl12b.SQ002. + Refl12b.SQ003. +
                         Refl13b.SQ001. + Refl13b.SQ002. + Gesb
                         .SQ006. + Gesb.SQ007.)),
41                  Anzahl_o_No = 16 - (Refl02b.SQ003. +
                       Refl03b.SQ003. + Refl06b.SQ004. +
                       Refl06b.SQ005. + Refl07b.SQ004. +
                       Refl07b.SQ005. + Refl09b.SQ001. +
                       Refl09b.SQ004. + Refl09b.SQ006. +
                       Refl12b.SQ001. + Refl12b.SQ002. +
                       Refl12b.SQ003. + Refl13b.SQ001. +
```

```
                    Refl13b.SQ002. + Gesb.SQ006. + Gesb.
                    SQ007.),
42                  Score_o_No = ScoreBew_o_No + ScoreAlt_o_No
                    ,
43                  )
44
45 results_Profil_Aufgaben_No_red <- subset(results_Profil_
   Aufgaben_No, !is.na(results_Profil_Aufgaben_No$Score_o_No)
   )
```

T.4.2 Histogramme zu beiden Reflexionsscores

```
1  ggplot(results_Profil_Aufgaben_No_red, aes(x=Score_No/2*100)
   )+
2    geom_histogram(binwidth=5, colour = "black", fill = "white
     ")+
3    #xlim(0,100)+
4    #stat_function(fun=dnorm, args=list(mean=mean(results$
     ScoreMCProz), sd=sd(results$ScoreMCProz)), color="black
     ", size=1)+
5    xlab("Testwerte␣in␣%␣(zusammengefasst␣in␣Intervalle␣mit␣
     einer␣Breite␣von␣5%)")+
6    theme_bw()+
7    theme(axis.text=element_text(size=18)) +
8    scale_y_continuous("Häufigkeit")#, breaks = round(ybreaks/
     (5*192), 3), labels = ybreaks)
9
10 ggplot(results_Profil_Aufgaben_No_red, aes(x=Score_o_No/2*
    100))+
11   geom_histogram(binwidth=5, colour = "black", fill = "white
     ")+
12   #xlim(0,100)+
13   #stat_function(fun=dnorm, args=list(mean=mean(results$
     ScoreMCProz), sd=sd(results$ScoreMCProz)), color="black
     ", size=1)+
14   xlab("Testwerte␣in␣%␣(zusammengefasst␣in␣Intervalle␣mit␣
     einer␣Breite␣von␣5%)")+
15   theme_bw()+
16   theme(axis.text=element_text(size=18)) +
17   scale_y_continuous("Häufigkeit")#, breaks = round(ybreaks/
     (5*192), 3), labels = ybreaks)
```

T.4.3 Ergänzende statistische Auswertung

Prüfung der Normalverteilung

```
ks.test(results_Profil_Aufgaben_No_red$Score_No, "pnorm",
   mean=mean(results_Profil_Aufgaben_No_red$Score_No), sd=sd(
   results_Profil_Aufgaben_No_red$Score_No))
ks.test(results_Profil_Aufgaben_No_red$Score_o_No, "pnorm",
   mean=mean(results_Profil_Aufgaben_No_red$Score_o_No), sd=
   sd(results_Profil_Aufgaben_No_red$Score_o_No))
```

Berechnung gepaarter t-Tests

jeweils mit Voraussetzungstest (Normalverteilung der Testwertdifferenz) und Bestimmung der Effektstärke für die Gesamtstichprobe sowie die Fähigkeitsstufen einzeln

```
ks.test((results_Profil_Aufgaben_No_red$Score_No-results_
   Profil_Aufgaben_No_red$Score_o_No), "pnorm", mean=mean((
   results_Profil_Aufgaben_No_red$Score_No-results_Profil_
   Aufgaben_No_red$Score_o_No)), sd=sd((results_Profil_
   Aufgaben_No_red$Score_No-results_Profil_Aufgaben_No_red$
   Score_o_No)))
t.test(results_Profil_Aufgaben_No$Score_No, results_Profil_
   Aufgaben_No$Score_o_No, paired=TRUE)
lsr::cohensD(results_Profil_Aufgaben_No$Score_No, results_
   Profil_Aufgaben_No$Score_o_No, method="paired")

ks.test((results_Profil_Aufgaben_No_red[results_Profil_
   Aufgaben_No_red$Profil4==2,]$Score_No-results_Profil_
   Aufgaben_No_red[results_Profil_Aufgaben_No_red$Profil4
   ==2,]$Score_o_No), "pnorm", mean=mean((results_Profil_
   Aufgaben_No_red[results_Profil_Aufgaben_No_red$Profil4
   ==2,]$Score_No-results_Profil_Aufgaben_No_red[results_
   Profil_Aufgaben_No_red$Profil4==2,]$Score_o_No)), sd=sd((
   results_Profil_Aufgaben_No_red[results_Profil_Aufgaben_No_
   red$Profil4==2,]$Score_No-results_Profil_Aufgaben_No_red[
   results_Profil_Aufgaben_No_red$Profil4==2,]$Score_o_No)))
t.test(results_Profil_Aufgaben_No[results_Profil_Aufgaben_No
   $Profil4==2,]$Score_No, results_Profil_Aufgaben_No[results
   _Profil_Aufgaben_No$Profil4==2,]$Score_o_No, paired=TRUE)

ks.test((results_Profil_Aufgaben_No_red[results_Profil_
   Aufgaben_No_red$Profil4==1,]$Score_No-results_Profil_
```

```
   Aufgaben_No_red[results_Profil_Aufgaben_No_red$Profil4
   ==1,]$Score_o_No), "pnorm", mean=mean((results_Profil_
   Aufgaben_No_red[results_Profil_Aufgaben_No_red$Profil4
   ==1,]$Score_No-results_Profil_Aufgaben_No_red[results_
   Profil_Aufgaben_No_red$Profil4==1,]$Score_o_No)), sd=sd((
   results_Profil_Aufgaben_No_red[results_Profil_Aufgaben_No_
   red$Profil4==1,]$Score_No-results_Profil_Aufgaben_No_red[
   results_Profil_Aufgaben_No_red$Profil4==1,]$Score_o_No)))
t.test(results_Profil_Aufgaben_No[results_Profil_Aufgaben_No
   $Profil4==1,]$Score_No, results_Profil_Aufgaben_No[results
   _Profil_Aufgaben_No$Profil4==1,]$Score_o_No, paired=TRUE)
lsr::cohensD(results_Profil_Aufgaben_No[results_Profil_
   Aufgaben_No$Profil4==1,]$Score_No, results_Profil_Aufgaben
   _No[results_Profil_Aufgaben_No$Profil4==1,]$Score_o_No,
   method="paired")

ks.test((results_Profil_Aufgaben_No_red[results_Profil_
   Aufgaben_No_red$Profil4==3,]$Score_No-results_Profil_
   Aufgaben_No_red[results_Profil_Aufgaben_No_red$Profil4
   ==3,]$Score_o_No), "pnorm", mean=mean((results_Profil_
   Aufgaben_No_red[results_Profil_Aufgaben_No_red$Profil4
   ==3,]$Score_No-results_Profil_Aufgaben_No_red[results_
   Profil_Aufgaben_No_red$Profil4==3,]$Score_o_No)), sd=sd((
   results_Profil_Aufgaben_No_red[results_Profil_Aufgaben_No_
   red$Profil4==3,]$Score_No-results_Profil_Aufgaben_No_red[
   results_Profil_Aufgaben_No_red$Profil4==3,]$Score_o_No)))
t.test(results_Profil_Aufgaben_No[results_Profil_Aufgaben_No
   $Profil4==3,]$Score_No, results_Profil_Aufgaben_No[results
   _Profil_Aufgaben_No$Profil4==3,]$Score_o_No, paired=TRUE)
lsr::cohensD(results_Profil_Aufgaben_No[results_Profil_
   Aufgaben_No$Profil4==3,]$Score_No, results_Profil_Aufgaben
   _No[results_Profil_Aufgaben_No$Profil4==3,]$Score_o_No,
   method="paired")

ks.test((results_Profil_Aufgaben_No_red[results_Profil_
   Aufgaben_No_red$Profil4==4,]$Score_No-results_Profil_
   Aufgaben_No_red[results_Profil_Aufgaben_No_red$Profil4
   ==4,]$Score_o_No), "pnorm", mean=mean((results_Profil_
   Aufgaben_No_red[results_Profil_Aufgaben_No_red$Profil4
   ==4,]$Score_No-results_Profil_Aufgaben_No_red[results_
   Profil_Aufgaben_No_red$Profil4==4,]$Score_o_No)), sd=sd((
   results_Profil_Aufgaben_No_red[results_Profil_Aufgaben_No_
   red$Profil4==4,]$Score_No-results_Profil_Aufgaben_No_red[
   results_Profil_Aufgaben_No_red$Profil4==4,]$Score_o_No)))
t.test(results_Profil_Aufgaben_No[results_Profil_Aufgaben_No
   $Profil4==4,]$Score_No, results_Profil_Aufgaben_No[results
   _Profil_Aufgaben_No$Profil4==4,]$Score_o_No, paired=TRUE)
```

```
18 lsr::cohensD(results_Profil_Aufgaben_No[results_Profil_
     Aufgaben_No$Profil4==4,]$Score_No, results_Profil_Aufgaben
     _No[results_Profil_Aufgaben_No$Profil4==4,]$Score_o_No,
     method="paired")
```

U *R*-Skript für die statistischen Auswertungen zum Entwicklungszusammenhang von Reflexion und Noticing (CLP-Design)

U.1 PLS-SEM

U.1.1 Gesamtstichprobe

```
1  library(seminr)
2  # specifying measurement model
3  mm <- constructs(
4    composite("Refl.Prae", single_item("ScoreMC.Prae")),
5    composite("Refl.Post", single_item("ScoreMC.Post")),
6    composite("No.Prae", single_item("ScoreMCNo.Prae")),
7    composite("No.Post", single_item("ScoreMCNo.Post"))
8              )
9  # specifying structural model
10 sm <- relationships(
11 paths(from = c("Refl.Prae", "No.Prae"), to = c("Refl.Post",
     "No.Post")))
12
13 # estimating model
14 CLP_PLS <- estimate_pls(data = results_Laengsschnitte,
15 measurement_model = mm,
16 structural_model = sm)
17
18 # Summarize the model results
19 summary_CLP_PLS <- summary(CLP_PLS)
20 summary_CLP_PLS$paths
21 summary_CLP_PLS$fSquare
22
```

```
# Bootstrap the model
boot_CLP_PLS <- bootstrap_model(seminr_model = CLP_PLS,
nboot = 1000,
cores = NULL,
seed = 123
,alpha = 0.05
)
sum_boot_CLP_PLS <- summary(boot_CLP_PLS, alpha = 0.05)
sum_boot_CLP_PLS$bootstrapped_paths

plot(boot_CLP_PLS, title = "Bootstrapped Model")
```

U.1.2 Obere drei Fähigkeitsstufen

Modell für die oberen drei Fähigkeitsstufen

```
# specifying measurement model
mm <- constructs(
  composite("Refl.Prae", single_item("ScoreMC.Prae")),
  composite("Refl.Post", single_item("ScoreMC.Post")),
  composite("No.Prae", single_item("ScoreMCNo.Prae")),
  composite("No.Post", single_item("ScoreMCNo.Post"))
            )
# specifying structural model
sm <- relationships(
paths(from = c("Refl.Prae", "No.Prae"), to = c("Refl.Post",
   "No.Post")))

# estimating model
CLP_PLS_o <- estimate_pls(data = subset(results_
   Laengsschnitte, Profil4.Prae==1 | Profil4.Prae==3 |
   Profil4.Prae==4),
measurement_model = mm,
structural_model = sm)

# Summarize the model results
summary_CLP_PLS_o <- summary(CLP_PLS_o)
summary_CLP_PLS_o$paths
summary_CLP_PLS_o$fSquare

# Bootstrap the model
boot_CLP_PLS_o <- bootstrap_model(seminr_model = CLP_PLS_o,
nboot = 1000,
cores = NULL,
seed = 123,
```

```
alpha = 0.05
)
sum_boot_CLP_PLS_o <- summary(boot_CLP_PLS_o, alpha = 0.05)
sum_boot_CLP_PLS_o$bootstrapped_paths

plot(boot_CLP_PLS_o, title = "Bootstrapped Model Profile 1 u 3 u u 4")
```

U.2 (Partielle) Korrelationen

U.2.1 Gesamtstichprobe

Prüfung der Normalverteilungen

```
ks.test(results_Laengsschnitte$ScoreMC.Prae, "pnorm", mean=mean(results_Laengsschnitte$ScoreMC.Prae), sd=sd(results_Laengsschnitte$ScoreMC.Prae))
ks.test(results_Laengsschnitte$ScoreMC.Post, "pnorm", mean=mean(results_Laengsschnitte$ScoreMC.Post), sd=sd(results_Laengsschnitte$ScoreMC.Post))
ks.test(results_Laengsschnitte$ScoreMCNo.Prae, "pnorm", mean=mean(results_Laengsschnitte$ScoreMCNo.Prae), sd=sd(results_Laengsschnitte$ScoreMCNo.Prae))
ks.test(results_Laengsschnitte$ScoreMCNo.Post, "pnorm", mean=mean(results_Laengsschnitte$ScoreMCNo.Post), sd=sd(results_Laengsschnitte$ScoreMCNo.Post))
```

synchrone Korrelationen zwischen den Konstrukten

```
cor.test(results_Laengsschnitte$ScoreMC.Prae, results_Laengsschnitte$ScoreMCNo.Prae)
cor.test(results_Laengsschnitte$ScoreMC.Post, results_Laengsschnitte$ScoreMCNo.Post)
```

bivariate Autokorrelationen

```
cor.test(results_Laengsschnitte$ScoreMC.Prae, results_Laengsschnitte$ScoreMC.Post)
cor.test(results_Laengsschnitte$ScoreMCNo.Prae, results_Laengsschnitte$ScoreMCNo.Post)
```

Kreuzkorrelationen

```
1 ppcor::pcor.test(results_Laengsschnitte$ScoreMC.Prae,
    results_Laengsschnitte$ScoreMCNo.Post, results_
    Laengsschnitte[,c("ScoreMCNo.Prae")], method="pearson")
2 ppcor::pcor.test(results_Laengsschnitte$ScoreMCNo.Prae,
    results_Laengsschnitte$ScoreMC.Post, results_
    Laengsschnitte[,c("ScoreMC.Prae")], method="pearson")
```

U.2.2 Obere drei Fähigkeitsstufen

Erstellung des reduzierten Dataframes

```
1 results_Laengsschnitte_Prof134 <- results_Laengsschnitte[
    results_Laengsschnitte$Profil4.Prae!=2,]
2 ppcor::pcor.test(results_Laengsschnitte_Prof134$ScoreMCNo.
    Prae, results_Laengsschnitte_Prof134$ScoreMC.Post, results
    _Laengsschnitte_Prof134[,c("ScoreMC.Prae")], method="
    pearson")
```

Prüfung der Normalverteilungen

```
1 ks.test(results_Laengsschnitte_Prof134$ScoreMC.Prae, "pnorm"
    , mean=mean(results_Laengsschnitte_Prof134$ScoreMC.Prae),
    sd=sd(results_Laengsschnitte_Prof134$ScoreMC.Prae))
2 ks.test(results_Laengsschnitte_Prof134$ScoreMC.Post, "pnorm"
    , mean=mean(results_Laengsschnitte_Prof134$ScoreMC.Post),
    sd=sd(results_Laengsschnitte_Prof134$ScoreMC.Post))
3 ks.test(results_Laengsschnitte_Prof134$ScoreMCNo.Prae, "
    pnorm", mean=mean(results_Laengsschnitte_Prof134$ScoreMCNo
    .Prae), sd=sd(results_Laengsschnitte_Prof134$ScoreMCNo.
    Prae))
4 ks.test(results_Laengsschnitte_Prof134$ScoreMCNo.Post, "
    pnorm", mean=mean(results_Laengsschnitte_Prof134$ScoreMCNo
    .Post), sd=sd(results_Laengsschnitte_Prof134$ScoreMCNo.
    Post))
```

synchrone Korrelationen zwischen den Konstrukten

```
cor.test(results_Laengsschnitte_Prof134$ScoreMC.Prae,
   results_Laengsschnitte_Prof134$ScoreMCNo.Prae)
cor.test(results_Laengsschnitte_Prof134$ScoreMC.Post,
   results_Laengsschnitte_Prof134$ScoreMCNo.Post)
```

bivariate Autokorrelationen

```
cor.test(results_Laengsschnitte_Prof134$ScoreMC.Prae,
   results_Laengsschnitte_Prof134$ScoreMC.Post)
cor.test(results_Laengsschnitte_Prof134$ScoreMCNo.Prae,
   results_Laengsschnitte_Prof134$ScoreMCNo.Post)
```

Kreuzkorrelationen

```
ppcor::pcor.test(results_Laengsschnitte_Prof134$ScoreMC.Prae
   , results_Laengsschnitte_Prof134$ScoreMCNo.Post, results_
   Laengsschnitte_Prof134[,c("ScoreMCNo.Prae")], method="
   pearson")
ppcor::pcor.test(results_Laengsschnitte_Prof134$ScoreMCNo.
   Prae, results_Laengsschnitte_Prof134$ScoreMC.Post, results
   _Laengsschnitte_Prof134[,c("ScoreMC.Prae")], method="
   pearson")
```

Danksagung

Zuletzt möchte ich mich bei den Personen herzlich bedanken, die mir die Promotionszeit ermöglicht, diese bereichert und begleitet haben.

Besonderer Dank gilt meinem Betreuer Christoph Kulgemeyer: Du hast gemeinsam mit Horst Schecker - ebenfalls vielen Dank! - im Studium mein Interesse für die Physikdidaktik geweckt und mir zusammen mit Maren Kempin - vielen Dank auch Dir! - erste konkrete Einblicke in die physikdidaktische Forschung und daran anschließend die Promotion ermöglicht. Vielen Dank für die umfassende und wertvolle Unterstützung in den letzten Jahren, die insgesamt wunderbare Betreuung und diverse inhaltliche und methodische Anregungen, die diese Arbeit geformt und bereichert haben. Danke auch besonders dafür, dass Du nicht nur eine Frage mehrfach beantwortet und Diskussion mehrfach geführt hast, bis ich meine Gedanken dazu sortiert hatte.

Ebenfalls bedanken möchte ich mich bei den übrigen Mitgliedern der Projektrunde *ProfiLe-P* für die konstruktive Beratung insbesondere zu Beginn dieses Projekts: Christoph Vogelsang, Andreas Borowski, Tanja Mutschler, David Buschhüter, Peter Reinhold und im Speziellen Melanie Jordans für die schöne und unkomplizierte Zusammenarbeit, z. B. in der Koordination gemeinsamer Erhebungen, und Josef Riese zusätzlich für die Übernahme des Zweitgutachtens.

Für viele tolle fachliche Gespräche, aber auch genau die richtige Ablenkung in der (digitalen) Kaffeepause zwischendurch, danke ich außerdem besonders Madeleine, Stefan, Melissa, Kasim und Ines, aber natürlich auch allen weiteren (ehemaligen) Kolleg*innen in den Arbeitsgruppen in Paderborn und Bremen. Ich habe mich in beiden Gruppen sehr wohl gefühlt!

Außerdem möchte ich mich bei allen Studierenden und Referendar*innen bedanken, die mit mir Interviews geführt oder ihre Testbearbeitungen zur Auswertung zur Verfügung gestellt haben. Ohne diese Bereitschaft hätte ich dieses Projekt nicht durchführen können.

Abschließend danke ich meiner Familie und meinen Freunden für das Zuhören, Mitdenken, Korrekturlesen und generell die vielfältige Unterstützung von musikalischer Ablenkung über Care-Pakete bis hin zum Programmiersupport.

Bisher erschienene Bände der Reihe
Studien zum Physik- und Chemielernen

ISSN 1614-8967

Vollständige Übersicht auf unserer Website

https://www.logos-verlag.de/spcl

Aktuelle Bände

310 Wolfgang Becker (2021): Auswirkungen unterschiedlicher experimenteller Repräsentationen auf den Kenntnisstand bei Grundschulkindern
ISBN 978-3-8325-5255-8 50.00 EUR

311 Marvin Rost (2021): Modelle als Mittel der Erkenntnisgewinnung im Chemieunterricht der Sekundarstufe I. Entwicklung und quantitative Dimensionalitätsanalyse eines Testinstruments aus epistemologischer Perspektive
ISBN 978-3-8325-5256-5 44.00 EUR (open access)

312 Christina Kobl (2021): Förderung und Erfassung der Reflexionskompetenz im Fach Chemie
ISBN 978-3-8325-5259-6 41.00 EUR (open access)

313 Ann-Kathrin Beretz (2021): Diagnostische Prozesse von Studierenden des Lehramts. eine Videostudie in den Fächern Physik und Mathematik
ISBN 978-3-8325-5288-6 45.00 EUR (open access)

314 Judith Breuer (2021): Implementierung fachdidaktischer Innovationen durch das Angebot materialgestützter Unterrichtskonzeptionen. Fallanalysen zum Nutzungsverhalten von Lehrkräften am Beispiel des Münchener Lehrgangs zur Quantenmechanik
ISBN 978-3-8325-5293-0 50.50 EUR (open access)

315 Michaela Oettle (2021): Modellierung des Fachwissens von Lehrkräften in der Teilchenphysik. Eine Delphi-Studie
ISBN 978-3-8325-5305-0 57.50 EUR (open access)

316 Volker Brüggemann (2021): Entwicklung und Pilotierung eines adaptiven Multistage-Tests zur Kompetenzerfassung im Bereich naturwissenschaftlichen Denkens
ISBN 978-3-8325-5331-9 40.00 EUR (open access)

317 Stefan Müller (2021): Die Vorläufigkeit und soziokulturelle Eingebundenheit naturwissenschaftlicher Erkenntnisse. Kritische Reflexion, empirische Befunde und fachdidaktische Konsequenzen für die Chemielehrer*innenbildung
ISBN 978-3-8325-5343-2 63.00 EUR

318 Laurence Müller (2021): Alltagsentscheidungen für den Chemieunterricht erkennen und Entscheidungsprozesse explorativ begleiten
ISBN 978-3-8325-5379-1 59.00 EUR

319 Lars Ehlert (2021): Entwicklung und Evaluation einer Lehrkräftefortbildung zur Planung von selbstgesteuerten Experimenten
ISBN 978-3-8325-5393-7 41.50 EUR (open access)

320 Florian Seiler (2021): Entwicklung und Evaluation eines Seminarkonzepts zur Förderung der experimentellen Planungskompetenz von Lehramtsstudierenden im Fach Chemie
ISBN 978-3-8325-5397-5 47.50 EUR (open access)

321 Nadine Boele (2021): Entwicklung eines Messinstruments zur Erfassung der professionellen Unterrichtswahrnehmung von (angehenden) Chemielehrkräften hinsichtlich der Lernunterstützung
ISBN 978-3-8325-5402-6 46.50 EUR

322 Franziska Zimmermann (2022): Entwicklung und Evaluation digitalisierungsbezogener Kompetenzen von angehenden Chemielehrkräften
ISBN 978-3-8325-5410-1 49.50 EUR

323 Lars-Frederik Weiß (2021): Der Flipped Classroom in der Physik-Lehre. Empirische Untersuchungen in Schule und Hochschule
ISBN 978-3-8325-5418-7 51.00 EUR

324 Tilmann Steinmetz (2021): Kumulatives Lehren und Lernen im Lehramtsstudium Physik. Theorie und Evaluation eines Lehrkonzepts
ISBN 978-3-8325-5421-7 51.50 EUR

325 Kübra Nur Celik (2022): Entwicklung von chemischem Fachwissen in der Sekundarstufe I. Validierung einer Learning Progression für die Basiskonzepte „Struktur der Materie“, „Chemische Reaktion“ und „Energie“ im Kompetenzbereich „Umgang mit Fachwissen“
ISBN 978-3-8325-5431-6 55.00 EUR

326 Matthias Ungermann (2022): Förderung des Verständnisses von Nature of Science und der experimentellen Kompetenz im Schüler*innen-Labor Physik in Abgrenzung zum Regelunterricht
ISBN 978-3-8325-5442-2 55.50 EUR

327 Christoph Hoyer (2022): Multimedial unterstütztes Experimentieren im webbasierten Labor zur Messung, Visualisierung und Analyse des Feldes eines Permanentmagneten
ISBN 978-3-8325-5453-8 45.00 EUR

328 Tobias Schüttler (2022): Schülerlabore als interessefördernde authentische Lernorte für den naturwissenschaftlichen Unterricht nutzen
ISBN 978-3-8325-5454-5 50.50 EUR

329 Christopher Kurth (2022): Die Kompetenz von Studierenden, Schülerschwierigkeiten beim eigenständigen Experimentieren zu diagnostizieren
ISBN 978-3-8325-5457-6 58.50 EUR

330 Dagmar Michna (2022): Inklusiver Anfangsunterricht Chemie. Entwicklung und Evaluation einer Unterrichtseinheit zur Einführung der chemischen Reaktion
ISBN 978-3-8325-5463-7 49.50 EUR

331 Marco Seiter (2022): Die Bedeutung der Elementarisierung für den Erfolg von Mechanikunterricht in der Sekundarstufe I
ISBN 978-3-8325-5471-2 66.00 EUR

332 Jörn Hägele (2022): Kompetenzaufbau zum experimentbezogenen Denken und Arbeiten. Videobasierte Analysen zu Aktivitäten und Vorstellungen von Schülerinnen und Schülern der gymnasialen Oberstufe bei der Bearbeitung von fachmethodischer Instruktion
ISBN 978-3-8325-5476-7 56.50 EUR (open access)

333 Erik Heine (2022): Wissenschaftliche Kontroversen im Physikunterricht. Explorationsstudie zum Umgang von Physiklehrkräften und Physiklehramtsstudierenden mit einer wissenschaftlichen Kontroverse am Beispiel der Masse in der Speziellen Relativitätstheorie
ISBN 978-3-8325-5478-1 48.50 EUR (open access)

334 Simon Goertz (2022): Module und Lernzirkel der Plattform FLexKom zur Förderung experimenteller Kompetenzen in der Schulpraxis. Verlauf und Ergebnisse einer Design-Based Research Studie
ISBN 978-3-8325-5494-1 66.50 EUR

335 Christina Toschka (2022): Lernen mit Modellexperimenten. Empirische Untersuchung der Wahrnehmung und des Denkens in Analogien beim Umgang mit Modellexperimenten
ISBN 978-3-8325-5495-8 50.00 EUR (open access)

336 Alina Behrendt (2022): Chemiebezogene Kompetenzen in der Übergangsphase zwischen dem Sachunterricht der Primarstufe und dem Chemieunterricht der Sekundarstufe I
ISBN 978-3-8325-5498-9 40.50 EUR (open access)

337 Manuel Daiber (2022): Entwicklung eines Lehrkonzepts für eine elementare Quantenmechanik. Formuliert mit In-Out Symbolen
ISBN 978-3-8325-5507-8 48.50 EUR

338 Felix Pawlak (2022): Das Gemeinsame Experimentieren (an-)leiten. Eine qualitative Studie zum chemiespezifischen Classroom-Management
ISBN 978-3-8325-5508-5 46.50 EUR

339 Liza Dopatka (2022): Konzeption und Evaluation eines kontextstrukturierten Unterrichtskonzeptes für den Anfangs-Elektrizitätslehreunterricht
ISBN 978-3-8325-5514-6 69.50 EUR

340 Arne Bewersdorff (2022): Untersuchung der Effektivität zweier Fortbildungsformate zum Experimentieren mit dem Fokus auf das Unterrichtshandeln
ISBN 978-3-8325-5522-1 39.00 EUR (open access)

341 Thomas Christoph Münster (2022): Wie diagnostizieren Studierende des Lehramtes physikbezogene Lernprozesse von Schüler*innen?. Eine Videostudie zur Mechanik
ISBN 978-3-8325-5534-4 44.50 EUR (open access)

342 Ines Komor (2022): Förderung des symbolisch-mathematischen Modellverständnisses in der Physikalischen Chemie
ISBN 978-3-8325-5546-7 46.50 EUR

343 Verena Petermann (2022): Überzeugungen von Lehrkräften zum Lehren und Lernen von Fachinhalten und Fachmethoden und deren Beziehung zu unterrichtsnahem Handeln
ISBN 978-3-8325-5545-0 47.00 EUR (open access)

344 Jana Heinze (2022): Einfluss der sprachlichen Konzeption auf die Einschätzung der Qualität instruktionaler Unterrichtserklärungen im Fach Physik
ISBN 978-3-8325-5553-5 42.00 EUR (open access)

345 Jannis Weber (2022): Mathematische Modellbildung und Videoanalyse zum Lernen der Newtonschen Dynamik im Vergleich
ISBN 978-3-8325-5566-5 68.00 EUR (open access)

346 Fabian Sterzing (2022): Zur Lernwirksamkeit von Erklärvideos in der Physik. Eine Untersuchung in Abhängigkeit von ihrer fachdidaktischen Qualität und ihrem Einbettungsformat
ISBN 978-3-8325-5576-4 52.00 EUR (open access)

347 Lars Greitemann (2022): Wirkung des Tablet-Einsatzes im Chemieunterricht der Sekundarstufe I unter besonderer Berücksichtigung von Wissensvermittlung und Wissenssicherung
ISBN 978-3-8325-5580-1 50.00 EUR

348 Fabian Poensgen (2022): Diagnose experimenteller Kompetenzen in der laborpraktischen Chemielehrer*innenbildung
ISBN 978-3-8325-5587-0 48.00 EUR

349 William Lindlahr (2023): Virtual-Reality-Experimente. Entwicklung und Evaluation eines Konzepts für den forschend-entwickelnden Physikunterricht mit digitalen Medien
ISBN 978-3-8325-5595-5 49.00 EUR

350 Bert Schlüter (2023): Teilnahmemotivation und situationales Interesse von Kindern und Eltern im experimentellen Lernsetting KEMIE
ISBN 978-3-8325-5598-6 43.00 EUR

351 Katharina Nave (2023): Charakterisierung situativer mentaler Modellkomponenten in der Chemie und die Bildung von Hypothesen. Eine qualitative Studie zur Operationalisierung mentaler Modell-komponenten für den Fachbereich Chemie
ISBN 978-3-8325-5599-3 43.00 EUR

352 Anna B. Bauer (2023): Experimentelle Kompetenz Physikstudierender. Entwicklung und erste Erprobung eines performanzorientierten Kompetenzstrukturmodells unter Nutzung qualitativer Methoden
ISBN 978-3-8325-5625-9 47.00 EUR (open access)

353 Jan Schröder (2023): Entwicklung eines Performanztests zur Messung der Fähigkeit zur Unterrichtsplanung bei Lehramtsstudierenden im Fach Physik
ISBN 978-3-8325-5655-6 46.50 EUR

354 Susanne Gerlach (2023): Aspekte einer Fachdidaktik Körperpflege. Ein Beitrag zur Standardentwicklung
ISBN 978-3-8325-5659-4 45.00 EUR

355 Livia Murer (2023): Diagnose experimenteller Kompetenzen beim praktisch-naturwissenschaftlichen Arbeiten. Vergleich verschiedener Methoden und kognitive Validierung eines Testverfahrens
ISBN 978-3-8325-5657-0 41.50 EUR (open access)

356 Andrea Maria Schmid (2023): Authentische Kontexte für MINT-Lernumgebungen. Eine zweiteilige Interventionsstudie in den Fachdidaktiken Physik und Technik
ISBN 978-3-8325-5605-1 57.00 EUR (open access)

357 Julia Ortmann (2023): Bedeutung und Förderung von Kompetenzen zum naturwissenschaftlichen Denken und Arbeiten in universitären Praktika
ISBN 978-3-8325-5670-9 37.00 EUR (open access)

358 Axel-Thilo Prokop (2023): Entwicklung eines Lehr-Lern-Labors zum Thema Radioaktivität. Eine didaktische Rekonstruktion
ISBN 978-3-8325-5671-6 49.50 EUR

359 Timo Hackemann (2023): Textverständlichkeit sprachlich variierter physikbezogener Sachtexte
ISBN 978-3-8325-5675-4 41.50 EUR (open access)

360 Dennis Dietz (2023): Vernetztes Lernen im fächerdifferenzierten und integrierten naturwissenschaftlichen Unterricht aufgezeigt am Basiskonzept Energie. Eine Studie zur Analyse der Wirksamkeit der Konzeption und Implementation eines schulinternen Curriculums für das Unterrichtsfach „Integrierte Naturwissenschaften 7/8“
ISBN 978-3-8325-5676-1 49.50 EUR

361 Ann-Katrin Krebs (2023): Vielfalt im Physikunterricht. Zur Wirkung von Lehrkräftefortbildungen unter Diversitätsaspekten
ISBN 978-3-8325-5672-3 65.50 EUR (open access)

362 Simon Kaulhausen (2023): Strukturelle Ursachen für Klausurmisserfolg in Allgemeiner Chemie an der Universität
ISBN 978-3-8325-5699-0 37.50 EUR (open access)

363 Julia Eckoldt (2023): Den (Sach-)Unterricht öffnen. Selbstkompetenzen und motivationale Orientierungen von Lehrkräften bei der Implementation einer Innovation untersucht am Beispiel des Freien Explorierens und Experimentierens
ISBN 978-3-8325-5663-1 48.50 EUR (open access)

364 Albert Teichrew (2023): Physikalische Modellbildung mit dynamischen Modellen
ISBN 978-3-8325-5710-2 58.50 EUR (open access)

365 Sascha Neff (2023): Transfer digitaler Innovationen in die Schulpraxis. Eine explorative Untersuchung zur Förderung der Implementation
ISBN 978-3-8325-5687-7 59.00 EUR (open access)

366 Rahel Schmid (2023): Verständnis von Nature of Science-Aspekten und Umgang mit Fehlern von Schüler*innen der Sekundarstufe I. Am Beispiel von digital-basierten Lernprozessen im informellen Lernsetting Smartfeld
ISBN 978-3-8325-5722-5 53.50 EUR (open access)

367 Dennis Kirstein (2023): Individuelle Bedingungs- und Risikofaktoren für erfolgreiche Lernprozesse mit kooperativen Experimentieraufgaben im Chemieunterricht. Eine Untersuchung zum Zusammenhang von Lernvoraussetzungen, Lerntätigkeiten, Schwierigkeiten und Lernerfolg beim Experimentieren in Kleingruppen der Sekundarstufe I
ISBN 978-3-8325-5729-4 50.50 EUR (open access)

368 Frauke Düwel (2024): Argumentationslinien in Lehr-Lernkontexten. Potenziale englischer Fachtexte zur Chromatografie und deren hochschuldidaktische Einbindung
ISBN 978-3-8325-5731-7 63.00 EUR (open access)

369 Fabien Güth (2023): Interessenbasierte Differenzierung mithilfe systematisch variierter Kontextaufgaben im Fach Chemie
ISBN 978-3-8325-5737-9 48.00 EUR (open access)

370 Oliver Grewe (2023): Förderung der professionellen Unterrichtswahrnehmung und Selbstwirksamkeitsüberzeugungen hinsichtlich sprachsensibler Maßnahmen im naturwissenschaftlichen Sachunterricht. Konzeption und Evaluation einer video- und praxisbasierten Lehrveranstaltung im Masterstudium
ISBN 978-3-8325-5738-6 44.50 EUR (open access)

371 Anna Nowak (2023): Untersuchung der Qualität von Selbstreflexionstexten zum Physikunterricht. Entwicklung des Reflexionsmodells REIZ
ISBN 978-3-8325-5739-3 59.00 EUR (open access)

372 Dominique Angela Holland (2023): Bildung für nachhaltige Entwicklung (BNE) kooperativ gestalten. Vergleich monodisziplinärer und interdisziplinärer Kooperation von Lehramtsstudierenden bei der Planung, Durchführung und Reflexion von Online-BNE-Unterricht
ISBN 978-3-8325-5760-7 47.00 EUR (open access)

373 Renan Marcello Vairo Nunes (2024): MINT-Personal an Schulen. Eine Untersuchung der Arbeitssituation und professionellen Kompetenzen von MINT-Lehrkräften verschiedener Ausbildungswege
ISBN 978-3-8325-5778-2 51.00 EUR (open access)

374 Mats Kieserling (2024): Digitalisierung im Chemieunterricht. Entwicklung und Evaluation einer experimentellen digitalen Lernumgebung mit universeller Zugänglichkeit
ISBN 978-3-8325-5786-7 45.50 EUR

375 Cem Aydin Salim (2024): Die Untersuchung adaptiver Lernsettings im Themenbereich „Schwimmen und Sinken“ im naturwissenschaftlichen Unterricht
ISBN 978-3-8325-5787-4 49.00 EUR (open access)

376 Novid Ghassemi (2024): Evaluation eines Lehramtsmasterstudiengangs mit dem Profil Quereinstieg im Fach Physik. Erkenntnisse zu Eingangsbedingungen, professionellen Kompetenzen und Aspekten individueller Angebotsnutzung
ISBN 978-3-8325-5789-8 41.50 EUR (open access)

377 Martina Flurina Cavelti (2024): Entwicklung und Validierung eines Messinstruments zur Erfassung der Schülerkompetenzen im Bereich des wissenschaftlichen Skizzierens im Fach Chemie in der Sekundarstufe I
ISBN 978-3-8325-5829-1 45.00 EUR (open access)

378 Tom Bleckmann (2024): Formatives Assessment auf Basis von maschinellem Lernen. Eine Studie über automatisiertes Feedback zu Concept Maps aus dem Bereich Mechanik
ISBN 978-3-8325-5842-0 46.50 EUR (open access)

379 Jana Marlies Rehberg (2024): Das physikspezifische Mindset zum Studienbeginn. Fragebogenentwicklung und Aufbau einer Online-Intervention
ISBN 978-3-8325-5850-5 59.50 EUR (open access)

380 Florian Trauten (2024): Entwicklung und Evaluation von automatisierten Feedbackschleifen in Online-Aufgaben im Fach Chemie
ISBN 978-3-8325-5859-8 46.00 EUR (open access)

381 Johanna Dejanovikj (2024): Reflexion als Lerngegenstand im Unterricht zur Förderung von Bewertungskompetenz
ISBN 978-3-8325-5860-4 41.00 EUR (open access)

382 Katharina Flieser (2024): Verständlichkeit physikalischer Sachtexte. Untersuchungen zum Wirkungsgefüge zwischen sprachlicher Textgestaltung und der Behaltensleistung sowie der Textwahrnehmung im Schulfach Physik
ISBN 978-3-8325-5858-1 44.00 EUR (open access)

383 Stephanie Neppl (2024): Perspektivenübernahme im Physikunterricht. Explorative Interviewstudie zu einer Seminarkonzeption mit dem Schwerpunkt Perspektivenübernahme bei der Planung von Physikunterricht
ISBN 978-3-8325-5865-9 48.00 EUR (open access)

384 Katja Plicht (2024): Ein Physikübungskonzept zur Förderung der Problemlösekompetenz. Entwicklung und empirische Evaluation eines Strategietrainings auf der Basis von Expertisemerkmalen
ISBN 978-3-8325-5875-8 45.00 EUR (open access)

385 Svenja Boegel (2024): Feedback beim Experimentieren: Zur Rolle von Cognitive Load und Motivation
ISBN 978-3-8325-5911-3 45.50 EUR (open access)

386 Laura Pannullo (2025): Wahlmöglichkeiten beim Experimentieren: Entwicklung und Erprobung eines Konzeptes für Experimente in inklusiven Lerngruppen im Physikunterricht
ISBN 978-3-8325-5916-8 46.50 EUR (open access)

387 Matthias Schweinberger (2025): „Stumme Videos“. Aufmerksamkeitslenkende Moderation von Demonstrationsexperimenten
ISBN 978-3-8325-5930-4 50.50 EUR

388 Rike Große-Heilmann (2025): Entwicklung fachdidaktischen Wissens zum Einsatz digitaler Medien im Fach Physik
ISBN 978-3-8325-5943-4 51.00 EUR (open access)

389 Lion Cornelius Glatz (2025): Eine vergleichende Betrachtung der Überzeugungskraft von Experimenten zum Teilchenmodell
ISBN 978-3-8325-5954-0 42.00 EUR (open access)

390 Rasmus Viefers (2025): Förderung der Variablenkontrollstrategie im Sachunterricht
ISBN 978-3-8325-5968-7 44.50 EUR (open access)

391 Luc Albrecht (2025): Kritisches Denken in der Chemielehrer*innenbildung. Eine konzeptionell-empirische Studie zur Integration und Operationalisierung
ISBN 978-3-8325-5970-0 57.00 EUR

392 Anna Weißbach (2025): Reflexion von Physikunterricht. Entwicklung und Validierung eines Testinstruments mit Assessment-Feedback zur Reflexionsfähigkeit Physik-Lehramtsstudierender
ISBN 978-3-8325-5977-9 51.50 EUR

393 Robert Gieske (2025): Förderung fachwissenschaftlicher und kommunikativer Kompetenzen von Schüler*innen durch Disaggregate Instruction im Fach Chemie. Konzeption und Durchführung einer vergleichenden Interventionsstudie zur Analyse der Wirksamkeit sprachsensibler Unterrichtsansätze in der Sekundarstufe I
ISBN 978-3-8325-5984-7 50.00 EUR

Vollständige Übersicht unter: https://www.logos-verlag.de/spcl

Alle erschienenen Bücher können unter der angegebenen ISBN direkt online (http://www.logos-verlag.de) oder telefonisch (030 - 42 85 10 90) beim Logos Verlag Berlin bestellt werden.

Studien zum Physik- und Chemielernen

Herausgegeben von Martin Hopf und Mathias Ropohl

Die Reihe umfasst inzwischen eine große Zahl von wissenschaftlichen Arbeiten aus vielen Arbeitsgruppen der Physik- und Chemiedidaktik und zeichnet damit ein gültiges Bild der empirischen physik- und chemiedidaktischen Forschung im deutschsprachigen Raum.
Die Herausgeber laden daher Interessenten zu neuen Beiträgen ein und bitten sie, sich im Bedarfsfall an den Logos-Verlag oder an ein Mitglied des Herausgeberteams zu wenden.

Kontaktadressen:

Univ.-Prof. Dr. Martin Hopf
Universität Wien,
Österreichisches Kompetenzzentrum
für Didaktik der Physik,
Porzellangasse 4, Stiege 2,
1090 Wien, Österreich,
Tel. +43-1-4277-60330,
e-mail: martin.hopf@univie.ac.at

Prof. Dr. Mathias Ropohl
Didaktik der Chemie,
Fakultät für Chemie,
Universität Duisburg-Essen,
Schützenbahn 70, 45127 Essen,
Tel. 0201-183 2704,
e-mail: mathias.ropohl@uni-due.de

Die Fähigkeit zur Unterrichtsreflexion stellt eine für Lehrkräfte zentrale Fähigkeit zur Professionalisierung und Verbesserung der Unterrichtsqualität dar. Es liegen Befunde vor, dass diese Fähigkeit bei angehenden Lehrkräften häufig nur gering ausgeprägt ist, sodass Bedarf zu ihrer Förderung besteht. Wie sich die Reflexionsfähigkeit entwickelt und im Lehramtsstudium gefördert werden kann, ist allerdings weitgehend unklar.

In dieser Arbeit wird ein Test zur validen Erfassung der Reflexionsfähigkeit angehender Physiklehrkräfte mit anschließendem Feedback für Teilnehmende entwickelt und evaluiert. Die Evaluation erfolgt auf Grundlage von $N = 192$ Testbearbeitungen sowie von Interviews, die die kognitiven Prozesse während der Testbearbeitung und die Interpretation des Feedbacks durch Teilnehmende erheben. Weiterhin werden eine latente Profilanalyse durchgeführt und Zusammenhänge zur Noticingfähigkeit u. a. im Cross-Lagged-Panel-Design analysiert, um weiterführende Erkenntnisse zur Entwicklung der Reflexionsfähigkeit zu gewinnen.

Mit dem positiv evaluierten Test kann die Reflexionsfähigkeit valide erhoben und rückgemeldet werden. Als Ergebnis der Profilanalyse werden vier aufsteigende Fähigkeitsstufen identifiziert, die u. a. Zusammenhänge zum Umfang absolvierter Schulpraktika aufweisen. Ein weiteres Ergebnis ist, dass die Noticingfähigkeit zur Entwicklung der Reflexionsfähigkeit beiträgt und daher künftig im Rahmen ihrer Förderung berücksichtigt werden sollte.

Logos Verlag Berlin

ISBN 978-3-8325-5977-9