

Studien zum Physik- und Chemielernen

H. Niedderer, H. Fischler, E. Sumfleth [Hrsg.]

236

Anja Schödl

FALKO-Physik – Fachspezifische Lehrerkompetenzen im Fach Physik

Entwicklung und Validierung
eines Testinstruments zur Erfassung
des fachspezifischen Professionswissens
von Physiklehrkräften

λογος

Studien zum Physik- und Chemielernen

Herausgegeben von Hans Niedderer, Helmut Fischler und Elke Sumfleth

Diese Reihe im Logos-Verlag bietet ein Forum zur Veröffentlichung von wissenschaftlichen Studien zum Physik- und Chemielernen. In ihr werden Ergebnisse empirischer Untersuchungen zum Physik- und Chemielernen dargestellt, z. B. über Schülervorstellungen, Lehr-/Lernprozesse in Schule und Hochschule oder Evaluationsstudien. Von Bedeutung sind auch Arbeiten über Motivation und Einstellungen sowie Interessensgebiete im Physik- und Chemieunterricht. Die Reihe fühlt sich damit der Tradition der empirisch orientierten Forschung in den Fachdidaktiken verpflichtet. Die Herausgeber hoffen, durch die Herausgabe von Studien hoher Qualität einen Beitrag zur weiteren Stabilisierung der physik- und chemiedidaktischen Forschung und zur Förderung eines an den Ergebnissen fachdidaktischer Forschung orientierten Unterrichts in den beiden Fächern zu leisten.

Hans Niedderer

Helmut Fischler

Elke Sumfleth

Studien zum Physik- und Chemielernen

Band 236

Anja Schödl

FALKO-Physik

Fachspezifische Lehrerkompetenzen
im Fach Physik

Entwicklung und Validierung eines Testinstruments zur
Erfassung des fachspezifischen Professionswissens
von Physiklehrkräften

Logos Verlag Berlin



Studien zum Physik- und Chemielernen

Hans Niedderer, Helmut Fischler, Elke Sumfleth [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.

Zugl.: Dissertation, Universität Regensburg, Fachbereich Physik, 2017

© Copyright Logos Verlag Berlin GmbH 2017

Alle Rechte vorbehalten.

ISBN 978-3-8325-4553-6

Logos Verlag Berlin GmbH
Comeniushof, Gubener Str. 47,
D-10243 Berlin

Tel.: +49 (0)30 / 42 85 10 90

Fax: +49 (0)30 / 42 85 10 92

<http://www.logos-verlag.de>

FALKO-Physik

- Fachspezifische Lehrerkompetenzen im Fach Physik –
Entwicklung und Validierung eines Testinstruments
zur Erfassung des fachspezifischen Professionswissens
von Physiklehrkräften



DISSERTATION ZUR ERLANGUNG DES AKADEMISCHEN GRADES EINER
DOKTORIN
DER DIDAKTIK DER NATURWISSENSCHAFTEN
„DR. PHIL. NAT.“ (DOCTOR PHILOSOPHIAE NATURALIS)
IM PROMOTIONSFACH NATURWISSENSCHAFT UND TECHNIK (NWT)
DER FAKULTÄT PHYSIK
DER UNIVERSITÄT REGENSBURG

vorgelegt von
Anja Schödl (geb. Altmann)

aus
Passau

im Jahr
2017

Das vorliegende Forschungsprojekt wurde im Zeitraum von September 2012 bis Mai 2016 in der Arbeitsgruppe von Frau Dr. Anja Göhring und von Juni 2016 bis Mai 2017 in der Arbeitsgruppe von Herrn Prof. Dr. Karsten Rincke durchgeführt.

Gutachter/innen:

Prof. Dr. Karsten Rincke

Prof. Dr. Oliver Tepner

Prüfungsausschuss:

Prof. Dr. Astrid Rank (Vorsitzende)

Prof. Dr. Karsten Rincke (Erstbetreuer/-gutachter)

Prof. Dr. Oliver Tepner (Zweitbetreuer/-gutachter)

Prof. Dr. Stefan Krauss (weiterer Prüfer)

Kontakt:

anja.schoedl@physik.uni-regensburg.de

anja.schoedl@me.com

Im Vergleich zur Prüfungsversion ist das Kodiermanual in dieser Dissertationsschrift nicht enthalten, kann aber bei der Autorin erfragt werden.

„I'll take my chance against the law. You'll take yours against the sea.“
Lt. Fletcher Christian, Mutiny on the Bounty (1935 film)

Danke...

„Soll ich? Oder soll ich nicht?“ An diese Frage werden sich zahlreiche Personen in meinem Umfeld erinnern, als das Promotionsangebot 2012 plötzlich im Raum stand. Leicht war es nicht, nach einigen Jahren Schule die gewohnten Gefilde zu verlassen und noch einmal neu anzufangen, zumal in meinem Fall auch noch die Promotionsberechtigung erworben werden musste (als Realschullehrkraft besitzt man aufgrund der zu kurzen Regelstudienzeit in Bayern per se keine Promotionsberechtigung). Vielen Dank daher an meine Eltern und meine Oma, meine LehrerkollegInnen, an Freunde und Bekannte, welche mich in meiner Entscheidung bestärkten und nicht zuletzt an meinen Mann, der mich während der ganzen Promotionszeit ertrug und mich unterstützte, vor allem dann, wenn die Selbstzweifel mal wieder Überhand nehmen wollten. Danke, dass ihr an mich geglaubt habt.

Bedanken möchte ich mich auch bei Frau Dr. Anja Göhring, welche mich als Initiatorin und Betreuerin im Anfangsstadium meiner Arbeit unterstützte und es mir ermöglichte, diese auf zahlreichen Tagungen vorzustellen. Vielen Dank auch an die Forschungsgruppe FALKO für die freundliche Aufnahme und die produktiven Arbeitssitzungen. In diesem Zusammenhang auch vielen Dank an Alfred für die statistische Unterstützung sowie an Prof. Dr. Sven Hilbert für die statistische Beratung.

Weiterhin ein herzliches „Danke!“ an Karsten Rincke, welcher mich in der Endphase meiner Promotion in seine Arbeitsgruppe aufnahm und mich als Erstbetreuer und -gutachter und „Adoptiv“-Doktorvater bei der Fertigstellung meiner Arbeit unterstützte. Von Anfang an wurde ich von meinem Zweitbetreuer und -gutachter Oliver Tepner begleitet. Sowohl bei Fragen statistischer und konzeptioneller Art als auch bei persönlichen Problemen stand er mir immer beratend zur Seite - vielen, vielen herzlichen Dank! Auch möchte ich meinen Dank richten an Stefan Krauss, der mir in zahlreichen intensiven Arbeitssitzungen dabei half, dem FALKO-Physik-Buchkapitel in der FALKO-Monographie, welches im engeren Sinne ja eine Auskopplung aus meiner Dissertationsschrift darstellt, die nötige Qualität zu verleihen.

Weiterhin erhielt ich Beratung im Rahmen des ein oder anderen Vortrags in diversen Forschungskolloquien an der Universität Regensburg (z. B. im Naturwissenschafts- und mathematikdidaktischen Kolloquium) und im Doktorierendenkolloquium der GDGP: Vielen Dank an die Teilnehmenden für die Unterstützung!

Als nächstes möchte ich mich bei den Mitgliedern meiner „alten“ Arbeitsgruppe (Naturwissenschaft und Technik, NWT) bedanken. Vorne weg dir, liebe Julia als meine Doktorierenden- und Bürokollegin, die immer wieder mal meiner Impulsivität ausgesetzt war, vor allem dann, wenn der Arbeitsberg mal wieder ins Unermessliche wuchs. Danke fürs Zuhören, Korrekturlesen, Fragen beantworten und die ein oder andere Leckerei (Eiskaffee, ...). Danke an Werner, meinen anfänglichen Bürokollegen und langjährigen Freund, für den ganzen Spaß und die nötige Ablenkung! Auch Astrid, Inken, Susanne, Barbara S., Maria, Helmut, Schnuppi, Andreas und Didi (erweiterte Arbeitsgruppe): Danke für eure Unterstützung! Euch ist es geschuldet, dass ich mich hier immer sehr wohl gefühlt habe, sei es bei gemeinsamen Mittags- oder Kaffeepausen oder auch nur bei kurzen Gesprächen zwischen den Seminaren. Besondere Erwähnung soll hierbei noch das Ritual des freitäglichen Frühstückens (Julia, Maria, Barbara S.) im SoSe 16 erwähnt werden.

Auch die Mitglieder meiner „neuen“ Arbeitsgruppe (Physik-Didaktik) trugen zum Gelingen der Arbeit bei: Danke an Josef, Andi, Jana, Barbara M., Sandra, Magda und Claudia fürs „Ertragen“ der Bürobesuche, wenn mal wieder eine Pause oder Hilfe nötig war. Danke besonders an Hans für die „materielle“ Unterstützung (Versuchsmaterialien,

Kaffee, Geschirrspülerservice). Ein ganz besonders herzliches „Danke!“ richte ich an Chris, der, ähnlich wie Julia, des öfteren Opfer meiner Impulsivität wurde und mich diesbezüglich immer wieder einbremsste. Danke auch für die wichtigen Ablenkungen in Form von Kaffeepausen und Lauftrainingseinheiten.

Letzten Endes dürfen auch meine Forschungs-SHKs hier nicht fehlen: Sylvia, Matthias und Josef, vielen Dank für euren unermüdlichen Einsatz. Ohne euch wäre ich so manches Mal fast im Datensumpf versunken. Josef, vor allem dir auch danke für die vielen Gelegenheiten der kritischen Reflexion und Diskussion, immer dann, wenn die „Betriebsblindheit“ mal wieder zu arg zu Tage trat. Auch den lieben Doktorierenden Susan Fried sowie Florian Treisch (Universität Würzburg), Martin Draude sowie Jan-Henrik Kechel (Universität Kassel) und Peter Mayer (LMU München) herzlichen Dank für die Unterstützung bei der Gewinnung von studentischen Studienteilnehmern! Und ebenfalls ein Dankeschön an Yvonne Gramzow und David Woitkowski, die mir ihr Testinstrument für die Validierungsstudien zur Verfügung stellten.

Auch bei Marei möchte ich mich ganz herzlich bedanken. Zum einen für die vielen gegenseitigen aufbauenden Bürobesuche und zum anderen für die massive Unterstützung bei LaTeX-Problemen im Rahmen der Erstellung meiner Dissertationsschrift.

Zum Schluss geht ein großes Dankeschön an die Experten, Lehrkräfte und Studierenden, welche an meiner Studie teilnahmen. Besonders die ersten KollegInnen opferten enorm viel Zeit bei der Bearbeitung bzw. Beurteilung der Testitems. Ohne Sie/euch wäre diese Forschungsarbeit selbstverständlich nicht möglich gewesen. In diesem Zusammenhang auch vielen Dank an Daniela Dietl, die mich bei der Formulierung des Anschreibens zur Gewinnung von Studienteilnehmenden und sonstigen formalen Fragen unterstützte.

Einen lieben Gruß möchte ich noch abschließend an meine Klasse 10a (Schuljahr 2016/2017) der Schönwerth-Realschule Amberg senden, die mir in der akuten Schreibphase viel Verständnis hinsichtlich einer leichten Verplantheit in einigen Unterrichtsstunden entgegenbrachten.

Danke, danke, danke ...

Teile dieser Dissertationsschrift sind bereits erschienen in:

1. Schödl, A., Göhring, A. (2017). FALKO-P: Fachspezifische Lehrerkompetenzen im Fach Physik - Entwicklung und Validierung eines Testinstruments zur Erfassung des fachspezifischen Professionswissens von Physiklehrkräften. In S. Krauss, A. Lindl, A. Schilcher, M. Fricke, A. Göhring, B. Hofmann, P. Kirchhoff & R. H. Mulder (Hrsg.), FALKO: Fachspezifische Lehrerkompetenzen. Konzeption von Professionswissenstests in den Fächern Deutsch, Englisch, Latein, Physik, Musik, Evangelische Religion und Pädagogik (S. 201-244). Münster: Waxmann.
2. Schödl, A., Göhring, A. (2016). Domain-Specific Teachers' Competencies (FALKO) - Sub-Project Physics. In J. Lavonen, K. Juuti, J. Lampiselkä, A. Uitto & K. Hahl (Eds.), *Electronic Proceedings of the ESERA 2015 Conference. Science education research: Engaging learners for a sustainable future, Part 13/13* (co-ed. M. Evagorou & M. Michelini), (pp. 2208-2215). Helsinki, Finland: University of Helsinki.
3. Schödl, A., Göhring, A. (2015). Fachspezifische Lehrerkompetenzen (FALKO) – Teilprojekt Physik . PhyDid B - Didaktik der Physik - Beiträge zur DPG-Frühjahrstagung in Wuppertal.
4. Göhring, A., Schödl, A. (2014). Fachspezifische Lehrerkompetenzen (FALKO) - Teilprojekt Physik. In S. Bernhold (Hrsg.), *Heterogenität und Diversität - Vielfalt der Voraussetzungen im naturwissenschaftlichen Unterricht. Gesellschaft für Didaktik der Chemie und Physik. Jahrestagung in Bremen 2014*, Bd. 35, S. 364-367.

Zusammenfassung

Zusammenfassend war es das Vorhaben des hier vorgestellten Projekts (FALKO-Physik, Teilprojekt der FALKO-Studie), ein Testinstrument zur objektiven, reliablen und validen Messung des fachdidaktischen Wissens und des Fachwissens von Physiklehrkräften der Sekundarstufe zu entwickeln. FALKO-Physik folgte bei der Testkonstruktion einem projektübergreifenden Rahmenmodell, das sich an der COACTIV-Studie (Kunter et al., 2011; Krauss et al., 2011) anlehnt sowie die Wissenstaxonomie Shulmans (1986) berücksichtigt. Die Tests der einzelnen Fachdidaktiken erfassen das fachdidaktische Wissen (pedagogical content knowledge; PCK; FDW) und das Fachwissen (content knowledge; CK; FW) im jeweiligen (Unterrichts-)Fach. Das Testinstrument der Pädagogik misst ergänzend hierzu fächerübergreifend das pädagogische Wissen (pedagogical knowledge; PK; PW) mittels Vignetten. FDW, FW und PW stellen dabei die theoretisch abgeleiteten Kompetenzbereiche des Professionswissens dar (z. B. Bromme, 1997). Innerhalb des FDW werden projektübergreifend die Subwissensfacetten „Wissen über Erklären und Repräsentieren von Inhalten“, „Wissen über typische Schülerschwierigkeiten und -fehler“ und „Wissen über das Potential von Lernmaterialien“ (fachspezifisch) konzeptualisiert. Im Rahmen von FALKO-Physik erfolgte diese Konzeptualisierung über das „Wissen über Messen und Experimentieren“ („M & E“). Das FW wurde jeweils (hauptsächlich) als vertieftes Hintergrundwissen über das Schulcurriculum operationalisiert.

Bei FALKO-Physik wurden während der Validierungsphase als Hauptstichprobe 65 Physiklehrkräfte und 118 angehende Physiklehrkräfte (Studierende) untersucht. Bei der Testkonstruktion wurde besonders darauf Wert gelegt, zentrale physikalische Themengebiete (Mechanik, Elektrizitätslehre, Optik und Wärmelehre) sowohl im Fachwissens- als auch im fachdidaktischen Testteil zu implementieren.

Alle Items konnten objektiv kodiert werden und wurden von Experten und erfahrenen Lehrkräften als berufsrelevant eingestuft (vgl. Abschn. 6.4.3 bzw. 5.3.2). Die Reliabilität des FW-Testteils war zufriedenstellend ($\alpha = 0,82$), während das FDW nur eingeschränkt reliabel gemessen werden konnte ($\alpha = 0,65$). Unter zusätzlicher Berücksichtigung der „M & E“-Subfacette des FDW (vgl. Abschn. 6.2) würde die Reliabilität auf $\alpha = 0,70$ steigen. Sowohl eine konfirmatorische Faktorenanalyse als auch eine Expertenbefragung legten jedoch nahe, diese Facette aus dem FDW-Test zu entfernen, da sie eher FW als FDW abzubilden scheint.

Die manifeste Korrelation ($r = 0,59^{**}$, vgl. Abschn. 6.4.1) zwischen beiden Wissenskategorien (nach Entfernung der „M & E“-Facette) ist ähnlich der entsprechenden Korrelation beim Physik-Test von Kirschner (Kirschner et al., 2017). Als weiteren Beleg für die Unterschiedlichkeit und somit für die differentielle Validität der beiden gemessenen Wissenskategorien können auch die Ergebnisse der durchgeführten Konstruktvalidierungsstudien herangezogen werden (vgl. Abschn. 6.6 und 6.7). Während beispielsweise Gymnasiallehrkräfte und Fachphysiker im FW aufgrund der ähnlichen fachlichen universitären Ausbildung erwartungsgemäß vergleichbar abschnitten ($d = 0,10$; $p = ,69$), war der Vorteil der Lehrkräfte im FDW (ohne „M & E“) dagegen erheblich größer ($d = 0,72^{**}$; $p < ,01$).

Einer der auffälligsten Befunde in Bezug auf zur weiteren Analyse der Hauptstichprobe durchgeführten Regressionsanalysen ist der wesentlich stärkere Einfluss der Statusvariable (Studierende vs. Lehrkräfte) gegenüber der Schulformvariable (Nicht-Gymnasium vs. Gymnasium) bei der Vorhersage des Abschneidens in Bezug auf beide Wissenskategorien (vgl. Abschn. 6.4.3 bzw. Kap. 7). Die vergleichsweise schwache Vorhersagekraft der Schulform, die nicht signifikant ist, kann vermutlich auch darauf zurückgeführt werden, dass beide

Testteile überwiegend schulnah operationalisiert wurden.

Auf diese schulnahe Operationalisierung dürfte wohl auch der einzige signifikante Zusammenhang des Professionswissens mit der Berufserfahrung zurückzuführen sein, der sich in Bezug auf alle FALKO-Fächer beziehungsweise auf COACTIV ergab (vgl. Lindl & Krauss, 2017): Die Korrelation von FW mit der Anzahl der Berufsjahre betrug bei den Physiklehrkräften $r = 0,28^*$. In allen anderen Fächern war dieser Zusammenhang entweder nicht signifikant oder sogar (leicht) negativ. Dieser allgemein auftretende und auf den ersten Blick erstaunlich erscheinende fehlende Zusammenhang, der nicht nur bei FALKO und COACTIV, sondern beispielsweise auch von der ProwiN-Arbeitsgruppe gefunden wurde (und zwar für alle drei Disziplinen Biologie, Chemie und Physik), wurde mittlerweile vielfach diskutiert (Kirschner et al., 2017; Besser & Krauss, 2016). Beispielsweise könnte er auf das Fehlen der Deliberate-Practice-Bedingungen (Ericsson, Krampe & Tesch-Römer, 1993) während der sogenannten dritten Ausbildungsphase zurückzuführen sein. Die in Anhang B enthaltenen FALKO-P-FW-Items (für die Kodieranweisungen zu den FDW- und den FW-Items bitte bei der Autorin nachfragen) können hier als Diskursanregung für die Fachgemeinschaft verstanden werden, um in weiterführenden theoretischen oder empirischen Analysen zu untersuchen, ob in diesen Items eventuell (erstmalig) Charakteristika implementiert wurden, die sensibel auf zunehmende Berufserfahrung reagieren.

Summary

It was the aim of FALKO-physics to develop an objective, reliable and valid test instrument for the measurement of professional knowledge of physics teachers at secondary schools. FALKO-physics is part of the FALKO-study with seven subjects and pedagogy participating. Concerning test development, FALKO-physics took up core elements of the COACTIV-study (Kunter et al., 2011; Krauss et al., 2011), as well as Shulman's (1986) taxonomies. The FALKO-tests capture the pedagogical content knowledge (PCK) and the content knowledge (CK) in the related subjects. In addition to this, the pedagogy test instrument measures interdisciplinary pedagogical knowledge (PK) using vignettes. PCK, CK and PK represent the theoretically derived competencies of professional knowledge (e. g. Bromme, 1997). Within the PCK, the sub-facets „knowledge of instruction“, „knowledge of students' cognitions“ and „knowledge of the potential of learning materials“ were conceptualized specific for each subject throughout the project. In relation to FALKO-physics, this specific sub-facet deals with the „knowledge of measuring and experimenting“ („M & E“). The CK was (mainly) operationalized as enhanced knowledge of the school curriculum.

In FALKO physics, 65 physics teachers and 118 students of physics in the initial phase of teacher training at university were examined as the main test sample during the validation phase. During test construction, it was particularly important to focus on a broad range of central physical topics (mechanics, electricity, optics and thermodynamics) in both the PCK and the CK test section.

The coding of all items fulfilled quantitative test criteria. On top of that, the items were classified as relevant for the teaching profession by experts and expert teachers alike (see 6.4.3 and 5.3.2). The reliability of the CK test section was satisfactory ($\alpha = 0.82$), whereas the PCK could only be measured to an acceptable level ($\alpha = 0.65$). With additional consideration of the „M & E“-sub-facet of the PCK (see 6.2), the reliability would increase to $\alpha = 0.70$. Both a confirmatory factor analysis and an expert survey suggested, however, that this sub-facet should be removed from the PCK test section as it appeared to belong to CK rather than to PCK.

The manifest correlation ($r = 0.59^{**}$, see 6.4.1) between the two categories of knowledge (after removing the „M & E“-sub-facet) is similar to the corresponding correlation in the physics test developed by Kirschner (Kirschner et al., 2017). The results of the construct validation study can also be interpreted as further evidence for the difference and thus for the differential validity of the two categories of measured knowledge (see 6.6 and 6.7). As an example, on the one hand, teachers at higher level secondary schools and physicists performed equally in the CK test section due to their similar university studies ($d = 0.10$; $p = .69$). On the other hand, not surprisingly, teachers have an advantage in the PCK test section (without „M & E“, $d = 0.72^{**}$, $p < .01$).

The regression analyses carried out on the main test sample yielded the most striking finding of the study: The status variable (students vs. teachers) has a much stronger predictive force than the school form variable (non-gymnasium¹ vs. gymnasium; see section 6.4.3 and chapter 7). The comparatively weak predictive power of the type of school, which is not significant, can presumably also be attributed to the fact that both test parts were mainly operationalized on school level rather than on enhanced knowledge of the school curriculum.

Only one significant correlation between the professional knowledge and the professional experience could be found in all FALKO subjects and in COACTIV (see Lindl & Krauss,

¹ „Gymnasium“ is the german term for high-level secondary school respectively grammar school

2017): The correlation between CK and the number of working years was $r = 0.28^*$ for FALKO-physics; in all other subjects this correlation was either not significant or even (slightly) negative. This lack of connection, which seems at first sight a surprising result, was also reported by the ProwiN-group for all of the three participating disciplines: biology, chemistry, and physics. This seemingly absent effect of increasing professional experience on the professional knowledge could be due to the lack of deliberate practice conditions (Ericsson et al., 1993) during the so-called third training phase of teachers. The CK items, which can be found in appendix B (please ask the author for information about the codebook), will ideally serve as a starting ground for further investigation into the effects of increasing professional experience in teaching on teachers' professional knowledge.

Inhaltsverzeichnis

1. Einleitung	1
2. Forschungshintergrund	7
2.1. Professionelle Kompetenz von Lehrkräften	7
2.2. Professionswissen von (Physik-)Lehrkräften	10
3. Anlage und Ziele dieser Arbeit	21
3.1. Forschungsfragen und Hypothesen	21
3.2. Methodisches und statistisches Vorgehen	24
4. Modellierung des Professionswissens	27
4.1. Hintergrundmodell und Vorgehensweise im Überblick	27
4.2. Konzeptualisierung und Operationalisierung: Fachdidaktisches Wissen, Fachwissen	28
4.2.1. Fachdidaktisches Wissen (FDW)	28
4.2.2. Fachwissen (FW)	31
5. Testkonstruktion	35
5.1. Itemkonstruktion	35
5.2. Kodiermanual	38
5.3. Prä-Pilot- und Pilotstudie	40
5.3.1. Stichprobe und Verlauf	44
5.3.2. Psychometrische Gütekriterien	47
5.4. Zusammenfassung (Kriterien der Itemselektion)	50
6. Testvalidierung	53
6.1. Stichprobe und Durchführung	54
6.2. Dimensionalität des fachbezogenen Professionswissens (CFA; faktorielle Validität)	55
6.3. Psychometrische Gütekriterien	56
6.3.1. Interraterreliabilität (Auswertungsobjektivität)	57
6.3.2. Itemschwierigkeit	57
6.3.3. Interne Konsistenz der Konstrukte (Skalenreliabilitäten und Trennschärfen)	58
6.3.4. Augenscheinvalidität: Inhaltliche Validität	58
6.4. Deskriptive Ergebnisse	58
6.4.1. Skalenmittelwerte und Interkorrelationen	59
6.4.2. Gruppenunterschiede zwischen Studierenden und Lehrkräften	59
6.4.3. Schulformunterschiede (Ausbildungssensitivität)	60
6.4.4. Zusammenhänge des fachbezogenen Professionswissens mit Indikatoren des Schul- und Studienerfolgs sowie der Berufs- und Studierendauer: Kriteriale Validität	63

6.5. Repräsentativitätshypothese	64
6.6. Konstruktvalidität: Diskriminante Validität	64
6.7. Konstruktvalidität: Konvergente Validität	70
7. Zusammenfassung, Interpretation und Diskussion der Ergebnisse	77
8. Möglichkeiten der Testweiterentwicklung	87
8.1. Itemergänzung bzw. -neuentwicklung	87
8.2. Überarbeitung der Modellierung	88
8.3. Skalierung	88
9. Ausblick	91
Literatur	93
Abbildungsverzeichnis	107
Tabellenverzeichnis	109
A. Tabellen	111
B. Testheft	201

1. Einleitung

Der „PISA-Schock“ und COACTIV

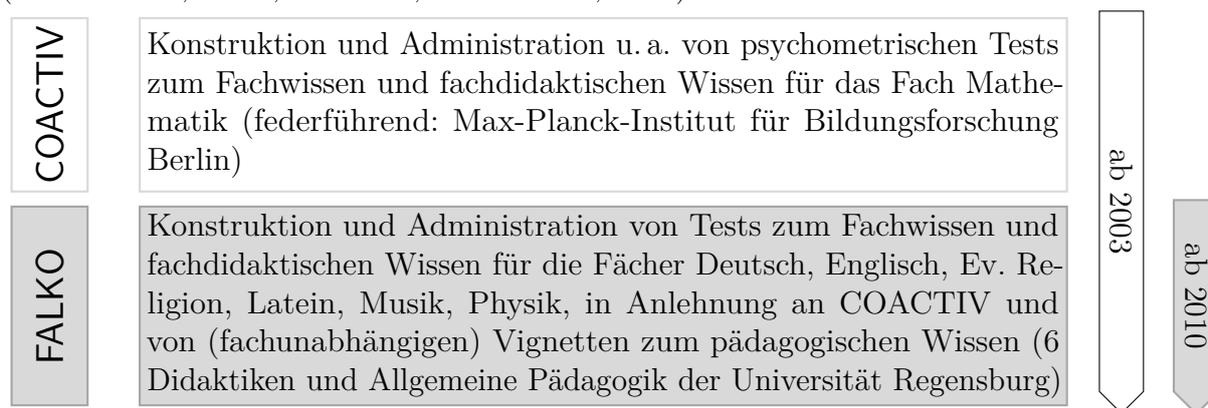
Mit dem unterdurchschnittlichen Abschneiden der deutschen Schüler¹ in der ersten PISA-Studie „PISA 2000“ („Programme for International Student Assessment“; Baumert et al., 2001; untersucht wurden Neuntklässler in den Disziplinen Mathematik, Naturwissenschaften und Lesen) entfachten heftige Diskussionen über möglicherweise zugrunde liegende Ursachen für dieses Resultat, welche in den darauffolgenden Jahren in rege empirische Forschungsanstrengungen, unter anderem in Bezug auf die Lehrkraft als Protagonist im Rahmen der Gestaltung von Unterricht, mündeten.

Während bei PISA 2000 nur die Schüler im Fokus standen, weitete man in der Folge bei PISA 2003/2004 (Prenzel et al., 2004) die Untersuchungen zusätzlich auf die Lehrkräfte der untersuchten Schüler aus. In einer breit angelegten Studie - der COACTIV-Studie („Cognitive Activation in the Classroom“, z. B. Kunter et al., 2011; vgl. Abb. 1.1) - wurden die Mathematiklehrkräfte der untersuchten Klassen einer genaueren Betrachtung unterzogen. Neben biografischen Angaben erfragten die COACTIV-Testinstrumente Daten zu Motivation, zu Aspekten des Unterrichts sowie des Berufserlebens der Lehrkräfte. Daneben bearbeiteten die Studienteilnehmer Items zur Erhebung leistungsrelevanter Unterrichtsaspekte wie zum Beispiel Klarheit und Strukturiertheit von Unterrichtsinhalten oder Aspekte der Klassenführung.

Einen großen Stellenwert nahm weiterhin die Messung des professionellen Wissens der Mathematiklehrkräfte im Sinne des sogenannten „Expertenparadigmas“ (Näheres hierzu siehe Abschnitt 2.2) ein. Dem Professionswissen einer Lehrkraft wird eine besondere Rolle für effektives Handeln im Unterricht und, daraus resultierend, den Lernerfolg von Schülern zugeschrieben (Bromme, 1997). Die enge Verzahnung der COACTIV- mit der PISA-Studie hatte dabei den großen Vorteil, dass Personenmerkmale von Lehrkräften direkt mit dem Lernzuwachs der Schüler und auch mit anderen Unterrichtskriterien über Kausalstrukturen zueinander in Beziehung gesetzt werden konnten. In der Auswertungsphase der Studie wurden folglich zahlreiche Strukturgleichungsmodelle gerechnet, welche auf die im COACTIV-Modell der Lehrerkompetenz (Abb. 2.1) angeführten Kompetenzaspekte und auf ausgewählte Kriterien der Unterrichtsqualität sowie Unterrichtsziele Bezug nehmen (Krauss, Lindl, Schilcher & Tepner, 2017). In Hinblick auf das Forschungsprojekt FALKO („Fachspezifische LehrerKOMPetenzen“; vgl. Abb. 1.1; in diesen Forschungsverbund ist das in dieser Dissertationsschrift abgehandelte Projekt FALKO-Physik eingebettet) mit seiner thematischen Schwerpunktsetzung (Untersuchung des Professionswissens von Lehrkräften) sind vor allem die COACTIV-Ergebnisse in Bezug auf die Kompetenzbereiche Fachwissen (FW), fachdidaktisches Wissen (FDW) und pädagogisch-psychologisches Wissen (PW) interessant: Es konnte gezeigt werden, dass das FDW einer Lehrkraft Merkmale der Unterrichtsqualität (kognitive Aktivierung, Lernunterstützung) und den Lernzuwachs von Schülern maßgeblich erklären kann (Baumert & Kunter, 2011b; Baumert et al., 2010). Für

¹ Personenbezeichnungen beziehen sich aus Gründen der besseren Lesbarkeit lediglich auf ein Geschlecht. Selbstverständlich sind dabei jeweils aber beide Geschlechter gemeint.

Abbildung 1.1.: Die Projekte COACTIV und FALKO mit ihren inhaltlichen Zielsetzungen (nach Krauss, Lindl, Schilcher, Fricke et al., 2017)



das FW einer Lehrkraft zeigten sich keine vergleichbaren signifikanten Effekte in Bezug auf die genannten Kriterien. Allerdings können die Daten dahingehend interpretiert werden, dass das FW als Prädiktor für das FDW angenommen werden kann, da eine ausgeprägte wissenschaftliche Basis einen wichtigen Ausgangspunkt für die Entwicklung von FDW darstellt und sich das FW somit indirekt ebenfalls auf die untersuchten Kriterien auswirkt. Denn Fachwissen bildet die „Grundlage, auf der fachdidaktische Beweglichkeit entstehen kann“ (Baumert & Kunter, 2006, S. 496). In den Folgejahren konnten einige Studien vor allem für den mathematisch-naturwissenschaftlichen Bereich die COACTIV-Ergebnisse zu großen Teilen replizieren (z. B. KiL, „Messung professioneller Kompetenzen in mathematischen und naturwissenschaftlichen Lehramtsstudiengängen“, Kleickmann et al., 2014; ProwiN, „Professionswissen in den Naturwissenschaften, Universität Duisburg-Essen, 2013).

Ermutigt durch COACTIV und weitere (inter)nationale Studien (z. B. MT21, Blömeke, Kaiser und Lehmann, 2008; TEDS-M bzw. TEDS-LT, Blömeke et al., 2009, Blömeke et al., 2011, Blömeke et al., 2013) und in Anbetracht des Fehlens an zu COACTIV vergleichbaren Studien vor allem in Bezug auf Unterrichtsfächer wie beispielsweise Latein, Musik oder Religion, gründete sich 2009 die Forschergruppe FALKO, welche in einem ersten Schritt der projektübergreifenden Fragestellung nachgeht, ob beziehungsweise inwiefern sich professionelle Kompetenzen von Lehrkräften in den teilnehmenden Fächern vergleichbar reliabel, valide und objektiv messen lassen (Krauss, Lindl, Schilcher, Fricke et al., 2017; Lindl & Krauss, 2017).

FALKO und das Teilprojekt FALKO-Physik

Die Projekt-Gruppe FALKO ist ein interdisziplinärer Forschungsverbund, in welchen das Teilprojekt FALKO-Physik eingebettet ist. Am Gesamtprojekt beteiligt sind Fachdidaktiken der Universitäten Regensburg und Augsburg sowie die Allgemeine Pädagogik der Universität Regensburg. Im einzelnen gehören dazu: Deutschdidaktik, Englischdidaktik, Ev. Religionslehre, Lateindidaktik, Mathematikdidaktik, Musikpädagogik, Naturwissenschaft und Technik, allgemeine Pädagogik und Psychologie. Testneukonstruktionen zur Messung des Professionswissens von Lehrkräften der Sekundarstufe I wurden dabei in den Fächern Deutsch (FALKO-D), Englisch (FALKO-E), Evangelische Religionslehre (FALKO-R), Latein (FALKO-L), Musik (FALKO-M), Physik (FALKO-P) und Pädagogik (FALKO-PA) vorgenommen. Weiterführende Informationen über FALKO und die Testkonstruktionen in

den beteiligten Fächer finden sich in Krauss, Lindl, Schilcher und Tepner (2017).

Dem Forschungsprojekt liegt ein gemeinsames konzeptuelles Rahmenmodell zugrunde, das sich an der COACTIV-Studie, in welcher ein Modell professioneller Kompetenz für Mathematik-Lehrkräfte entwickelt wurde (Kunter et al., 2011; Krauss et al., 2011), anlehnt und der Wissenstaxonomie Shulmans (1986) folgt. Dieses gemeinsame Rahmenmodell ermöglicht eine simultane und aufeinander abgestimmte Forschungsstrategie, z. B. vergleichbare Itemformate, gemeinsame Auswertungen (vgl. Lindl & Krauss, 2017). Die Testinstrumente der einzelnen Fachdidaktiken erfassen jeweils das fachdidaktische Wissen (pedagogical content knowledge; PCK) und das Fachwissen (content knowledge; CK) im jeweiligen (Unterrichts-)Fach und das Testinstrument der Pädagogik mittels Vignetten fächerübergreifend das pädagogische Wissen (pedagogical knowledge; PK). PCK, CK und PK stellen dabei die theoretisch abgeleiteten Kompetenzbereiche des Professionswissens dar (vgl. z. B. Bromme, 1997 bzw. Abb. 2.1). Für die beiden Wissenskategorien PCK und CK sollen im Folgenden die entsprechenden deutschen Begriffe fachdidaktisches Wissen (FDW) und Fachwissen (FW) verwendet werden. Zur unterschiedlichen Verwendung der Begriffe siehe Gramzow, Riese und Reinhold (2013). Für internationale Modellierungen des Professionswissens und speziell des PCK vergleiche zum Beispiel Gess-Newsome (1999), Magnusson, Krajcik und Borko (1999) oder Berry, Friedrichsen und Loughran (2015).

Innerhalb des FDW werden projektübergreifend die Subwissensfacetten „Wissen über Erklären und Repräsentieren von Inhalten“, „Wissen über typische Schülerschwierigkeiten und -fehler“ und „Wissen über das Potential von Lernmaterialien“ (fachspezifisch) konzeptualisiert. Das FW wurde jeweils (hauptsächlich) als vertieftes Hintergrundwissen über das Schulcurriculum operationalisiert. Komplettiert wird die Testbatterie durch das Testinstrument FALKO-PA. Es überprüft die pädagogische Kompetenz (PW) in Bezug auf die folgenden fünf professionellen Lehrerrollen (z. B. Sauer, 2016): Die Kompetenz als „Instructor“, als „Coach“, als „Developer“, als „Researcher“ und als „Lifelong Learner“. Jede Rolle wird dabei in Bezug auf die folgenden vier Themenfelder operationalisiert: Kognitive Heterogenität, kulturelle Heterogenität, Medien und Evaluation. Die Erfassung pädagogisch-psychologischer Aspekte geschieht durch einen Fragebogen zur Selbstwirksamkeit, zu Zielorientierungen und zu den impliziten Persönlichkeitstheorien von Lehrkräften. Überprüft wird unter anderem, ob Lehrkräfte, je nachdem, ob sie Persönlichkeitsmerkmale (Begabung, Kreativität) als stabil oder veränderbar einschätzen, sich in ihren didaktischen oder fachwissenschaftlichen Kompetenzen unterscheiden.

FALKO knüpft mit seinen Forschungsfragen an die Arbeit von COACTIV an. Untersucht werden soll unter anderem die Frage, inwieweit sich das Modell und die Befunde von COACTIV auf andere Unterrichtsfächer übertragen lassen. Weiterhin interessieren beispielsweise auch interdisziplinäre Fragestellungen: Gibt es unter den Lehrkräften „den guten Didaktiker“, unabhängig von fachlichen Aspekten? Sind fachspezifische Unterschiede in verschiedenen Kompetenzfacetten messbar? Wie entwickelt sich das Professionswissen im Verlauf der Ausbildung und während des Berufslebens? Um die unterschiedlich gelagerten Forschungsfragen innerhalb von Feldstudien bearbeiten zu können, wurde bei FALKO in einem ersten Schritt die gemeinsame Konzeptualisierung und Operationalisierung des FDW und des FW (jeweils getrennt nach Deutsch, Englisch, Evangelische Religionslehre, Latein Musik und Physik) sowie des PW innerhalb von reliablen, validen und objektiven Papier-und-Bleistift-Tests beziehungsweise (fachunabhängigen) Vignetten (im Fall von FALKO-PA) angestrebt. Ergebnisse hierzu finden sich in Krauss, Lindl, Schilcher, Fricke et al. (2017).

Mit dem Testinstrument FALKO-P soll das fachspezifische Professionswissen (FDW

und FW) von Lehrkräften, welche das Fach Physik in der Sekundarstufe I unterrichten, gemessen werden. Neben Physiklehrkräften an den Schularten Gymnasium und Realschule stehen explizit auch Lehrkräfte, welche an Mittelschulen² unterrichten, im Fokus des Teilprojekts FALKO-Physik.

Das FDW wird, der FALKO-Rahmenkonzeption folgend, als „Wissen über Erklären und Repräsentieren von Inhalten“ („E & R“), „Wissen über typische Schülerschwierigkeiten und -fehler“ („Schk“) und als dritte physikspezifische Subfacette als „Wissen über Messen und Experimentieren“ („M & E“) im Sinne der dritten fachspezifischen FDW-Facette bei FALKO konzeptualisiert und operationalisiert. Studierende der genannten Lehrämter, fachfremde Lehrkräfte an Mittelschulen, fachfremde Lehrkräfte an Realschulen und Gymnasien (Biologie, Chemie, Mathematik) sowie Fachphysiker wurden zu Validierungszwecken ebenfalls in die Gesamtstichprobe aufgenommen.

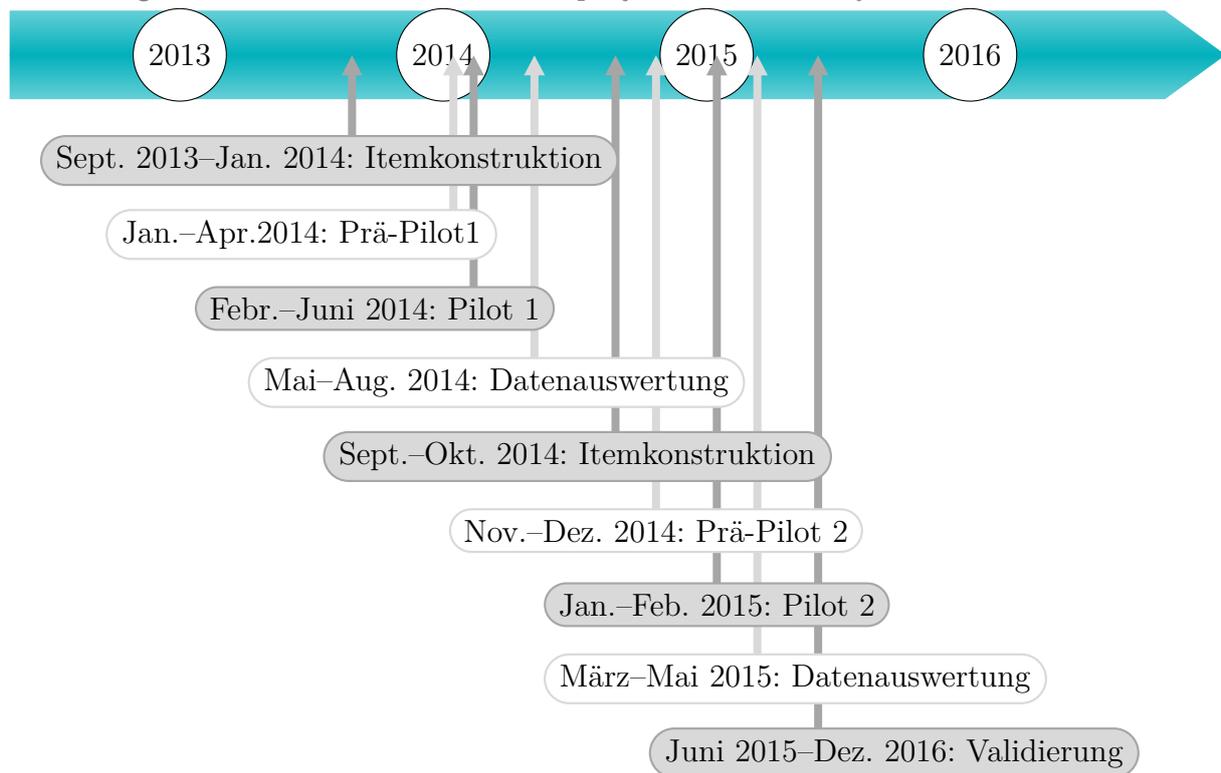
Bevor die Konstruktion des Testinstruments FALKO-P Gegenstand dieser Dissertation wurde, entwickelte Anja Göhring (Staatliches Seminar für Didaktik und Lehrerbildung Rottweil) gemeinsam mit Josef Reisinger (ehemals Physikdidaktik, Universität Regensburg, im Ruhestand) während der Anfangsphase des Teilprojekts FALKO-Physik eine erste Version des Testinstruments. Diese wurde in einer Pilotierung mit gymnasialen Seminarlehrkräften eingesetzt. Die Inhalte einiger Items dienten als Ausgangspunkte für die Itemkonstruktion innerhalb dieses Dissertationsvorhabens und werden in der finalen Testversion auch als solche ausgewiesen.

Struktur dieser Dissertationsschrift

Die vorliegende Disserationsschrift beleuchtet und diskutiert die psychometrische Testkonstruktion (September 2013 bis Dezember 2016; vgl. Abb. 1.2 zum fachdidaktischen Wissen und zum Fachwissen von Sekundarschullehrkräften im Fach Physik. In Kapitel 2 wird detailliert der Forschungshintergrund in Bezug auf professionelle Kompetenz (2.1) und Professionswissen von Lehrkräften skizziert und ausgewählte Projekte zur Professionswissensforschung aus dem mathematisch-naturwissenschaftlichen Bereich auf nationaler Ebene vorgestellt (2.2). Kapitel 3 widmet sich anschließend den Forschungsfragen und den Hypothesen in Bezug auf die in diesem Dissertationsvorhaben durchgeführte Testkonstruktion (3.1) und gibt einen Überblick über das methodische und statistische Vorgehen bei der Testentwicklung (3.2). Bevor konkret im 5. Kapitel auf das Verfahren der Testkonstruktion (Entwicklung von Items und Kodiermanual in 5.1 bzw. 5.2) und auf in diesem Zusammenhang durchgeführte Teilstudien (5.3), welche in die Phase der Itemselektion (5.4) für das final zu validierende Testheft mündeten, eingegangen wird, erfolgt in Kapitel 4 eine explizite Vorstellung des bei FALKO-Physik zugrunde gelegten Hintergrundmodells zum Professionswissen von Physiklehrkräften.

²Die bayerischen Hauptschulen wurden ab dem Schuljahr 2011/2012 sukzessive in Mittelschulen umgewandelt, so dass im weiteren Verlauf der Dissertationsschrift nur dieser Begriff verwendet wird. Während Lehrkräfte an Realschulen und Gymnasien Physik als Unterrichts-/Hauptfach studierten und als singuläres Schulfach unterrichten, werden physikalische Inhalte an der Mittelschule im Regelfall innerhalb eines integrierten Schulfaches, bestehend aus den Teildisziplinen Biologie, Chemie und Physik (z. B. in Bayern: „Physik, Chemie, Biologie“, „PCB“ bzw. „Natur und Technik“, „NT“, LehrplanPLUS), gelehrt. In den meisten Fällen ist es dabei so, dass die Mittelschullehrkräfte keine oder höchstens eine der drei Teildisziplinen als Haupt-/Unterrichts- bzw. Neben-/Didaktikfach studierten, trotzdem aber alle drei Fachdisziplinen im Fächerverbund unterrichten. Folglich müssen die zu unterrichtenden Fachinhalte von den Lehrkräften eigenständig erworben werden, was bedeutet, dass sie, prinzipiell betrachtet, fachfremd unterrichten.

Abbildung 1.2.: Zeitlicher Ablauf des Teilprojekts FALKO-Physik



Im weiteren Verlauf widmet sich das 6. Kapitel der Testvalidierung. Nach einem Überblick über die Zusammensetzung der Gesamtstichprobe und der Durchführung der Testungen (6.1) sowie einer Betrachtung der Dimensionalität des fachbezogenen Professionswissens (faktorielle Validität; Abschn. 6.2) sowie psychometrischen Gütekriterien (Interraterreliabilität, interne Konsistenz der Konstrukte, Augenscheinvalidität; Abschn. 6.3), wird die Stichprobe für die durchzuführenden statistischen Analysen nach beruflichem Status und Schulform aufgetrennt. Im Anschluss an den Bericht deskriptiver Ergebnisse in Form von Gruppenunterschieden und Zusammenhängen des fachbezogenen Professionswissens mit Indikatoren des Schul- und Studienerfolgs sowie der Berufs- und Studiendauer (kriteriale Validität; Abschn. 6.4) beleuchtet Abschnitt 6.5 die in 3.1 aufgeworfene Forschungsfrage, inwiefern sich Wissen in einzelnen Teilgebieten der Physik (z. B. Mechanik) als repräsentativ für fachbezogenes Professionswissen im Fach Physik als Ganzes darstellt. Die Abschnitte 6.6 und 6.7 berichten über weitere Validierungsmaßnahmen (Konstruktvalidität). Bevor in Kapitel 8 über Möglichkeiten der Testweiterentwicklung reflektiert und in Kapitel 9 ein Ausblick auf weitere Forschungsdesiderate gegeben wird, erfolgt in Kapitel 7 die Gesamtzusammenstellung, Interpretation und Diskussion der in Kapitel 6 berichteten Ergebnisse.

Um zum einen die Grundlage für eine vielfältige Weiternutzung, zum anderen aber auch die Basis für eine Weiterentwicklung des Testinstruments FALKO-P zu schaffen, finden sich alle finalen Items (Anhang B) sowie ausgewähltes (statistisches) Material (Anhang A) im Anhang dieser Dissertationsschrift (für die jeweiligen Kodieranweisungen bitte bei der Autorin nachfragen).

2. Forschungshintergrund

Im folgenden Kapitel wird der dieser Dissertationsschrift zugrunde liegende Forschungshintergrund dargelegt. Ausgehend von professionellen Kompetenzen von Lehrkräften (Abschn. 2.1) wird der Begriff des Professionswissens, vor allem in Bezug auf das Fach Physik, genauer beleuchtet (Abschn. 2.2). Weiterhin erfolgt in 2.2 eine zusammenfassende Darstellung von Modellierungen professioneller Kompetenz beziehungsweise von Professionswissen in ausgewählten nationalen Forschungsprojekten und Studien für das Fach Physik.

2.1. Professionelle Kompetenz von Lehrkräften

Spätestens seit den enttäuschenden Ergebnissen der deutschen Schüler bei internationalen Leistungsuntersuchungen wie TIMSS oder PISA (z. B. Baumert et al., 2001) geriet das deutsche Bildungssystem und mit ihm der Unterricht an deutschen Schulen immer wieder in die Diskussion, auch wenn sich in darauffolgenden PISA-Studien eine substanzielle Verbesserung der Gesamtestleistung deutscher Schüler insbesondere in den Naturwissenschaften feststellen ließ (Reiss, Sälzer, Schiepe-Tiska, Klieme & Köller, 2016). Die physikdidaktische Forschung unternimmt in diesem Zusammenhang seit einigen Jahren große Anstrengungen zur Identifizierung und Verbesserung von Faktoren, welche einen (positiven) Einfluss auf die Ausbildung im Fach Physik an deutschen Schulen und Hochschulen haben. Diese Faktoren sind von vielschichtiger Natur und betreffen beispielsweise die Strukturiertheit beziehungsweise Klarheit von Unterricht, naturwissenschaftliche Experimente im Unterricht, Schülervorstellungen, Schülerinteresse an Physik, aber auch die Person der Lehrkraft bezüglich deren professionellen Kompetenzen. Letztere sollen in der Folge genauer betrachtet werden.

Aus pädagogischer Perspektive findet man für den Begriff „Kompetenz“ eine Vielzahl von Definitionsansätzen, welche im Kern meist auf ein Fähigkeits- und Bereitschaftskonzept abzielen. So subsummiert beispielsweise Weinert (2001a, S. 27f.) unter Kompetenzen

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

Dieses Kompetenzverständnis stellt das innerhalb der Lehrerbildungsforschung vorherrschende Konzept dar. In Ergänzung dazu versteht Frey (2006, S. 31) Kompetenz

„als ein Bündel von körperlichen und geistigen Fähigkeiten [...], die jemand benötigt, um anstehende Aufgaben oder Probleme zielorientiert und verantwortungsvoll zu lösen, die Lösungen zu reflektieren und zu bewerten und das eigene Repertoire an Handlungsmustern weiterzuentwickeln.“

Nach Vogelsang (2014) können unter Berücksichtigung obiger und weiterer Perspektiven unter anderem folgende Merkmale von Kompetenz festgehalten werden: Kompetenz ist zum

einen auf einen spezifischen situativen Kontext (Handlungsfeld einer inhaltlichen Domäne) bezogen. Zum anderen gilt Kompetenz als kognitive Disposition einer Person, welche aber auch nicht-kognitive Anteile enthält. Ebenfalls stellt Kompetenz eine Handlungsressource dar und ist somit funktionale Basis für das Ausführen von Handlungen. Weiterhin ist Kompetenz erlernbar und kann sich bei verschiedenen Personen in unterschiedlichen Ausprägungen zeigen.

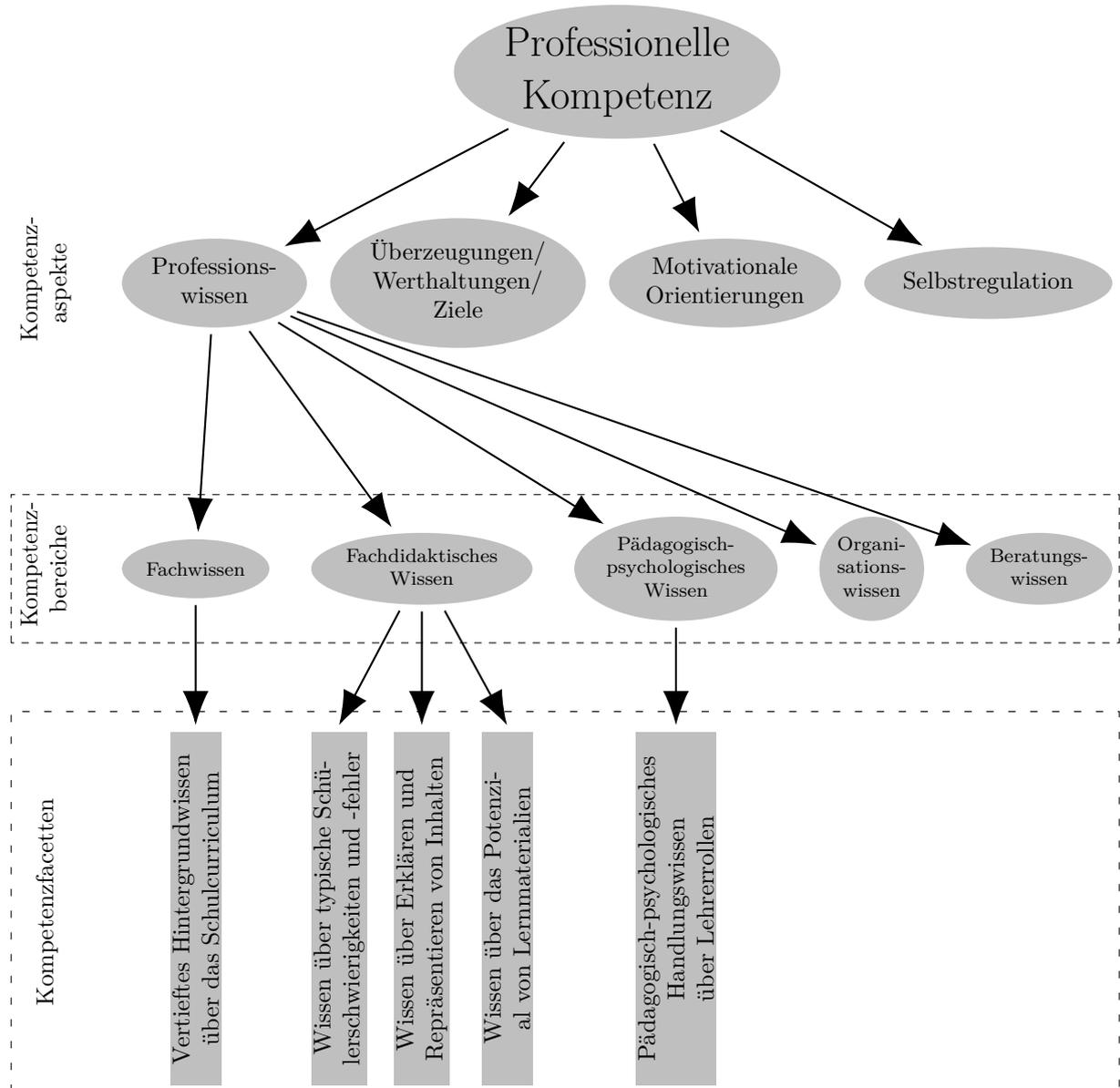
Kompetenzen können demnach als Voraussetzungen gesehen werden, dass eine Person, je nach Anforderungssituation, bestimmte Handlungen in einer festgelegten Form und in einer gewissen Qualität, ausführen kann (Oser, 2002). Geht man davon aus, dass im Falle einer Lehrkraft solche Handlungen das Unterrichten betreffen - nach der Überzeugung vieler Autoren stellt der Akt des Unterrichtens das Kerngeschäft einer Lehrkraft dar (z. B. Tenorth, 2006; Kunter et al., 2011 oder Lipowsky, 2006) - kommt man schnell zu dem Schluss, dass sich erfolgreiches Lehrerhandeln schwer umschreiben lässt, schon allein deshalb, weil erfolgreiches Unterrichten nicht allein von der Lehrkraft (z. B. Vogelsang, 2014) abhängt, sondern im Sinne des Angebots-Nutzungs-Modells nach Helmke (2009) und Lipowsky (2006) ein Zusammenwirken zwischen Schülern (Lernaktivitäten) und deren Lehrkräften sowie weiteren Faktoren wie beispielsweise Kontext (Schulform, Klassenzusammensetzung, ...) oder Familie (Schicht, Sprache, ...) darstellt. Tatsächliches (unterrichtliches) Lehrerhandeln wird in der Regel als „Performanz“ bezeichnet und wird in Forscherkreisen auch als solches analysiert und diskutiert (für eine ausführliche Diskussion zum Zusammenhang zwischen Lehrerkompetenz und „Performanz“, vgl. z. B. Vogelsang, 2014; für eine genauere Betrachtung der Beziehungen zwischen professionellem Wissen, Können und Handeln, vgl. z. B. Krauss, Lindl, Schilcher und Tepner, 2017).

Ein maßgebendes Forschungsdesiderat stellte in den letzten Jahren folglich der Wunsch dar, unter anderem im Sinne einer Möglichkeit der Optimierung von Unterricht, Lehrerkompetenzen detailliert zu untersuchen respektive empirisch erfassbar zu machen. Hierzu wurden in einem ersten Schritt Modellierungen des Konstrukts in Form sogenannter Kompetenzmodelle vorgenommen. Innerhalb solcher Kompetenzmodelle erfolgte eine exakte Zusammenstellung von Teilkompetenzen, welche eine Beschreibung des Gesamtkonstrukts ermöglichen sollten. Es können „standardorientierte“, welche sich an normativen Setzungen orientieren, von „professionswissensorientierten“ Kompetenzmodellen unterschieden werden (Vogelsang, 2014). Als Beispiel für ein standardorientiertes Kompetenzmodell kann das von E. Terhart (2002) genannt werden, welches der Kultusministerkonferenz zur Unterstützung bei der Formulierung der Bildungsstandards (KMK, 2004) vorgelegt wurde.

Das wohl am weitesten verbreitete professionswissensorientierte Kompetenzmodell ist das von Baumert und Kunter (2006), welches im Rahmen der COACTIV-Studie (Kunter et al., 2011) entwickelt wurde. Da dieses Modell (und explizit der enthaltene Kompetenzaspekt des Professionswissens) die Grundlage für das Forschungsprojekt FALKO mit seinem in dieser Dissertationsschrift betrachteten Teilprojekt FALKO-Physik darstellt, soll es in der Folge im Überblick vorgestellt werden (vgl. auch Abb. 2.1 bzw. Baumert & Kunter, 2011), bevor im folgenden Abschnitt detailliert auf das Professionswissen als Teil professioneller Lehrerkompetenz eingegangen wird.

Zusammenfassend werden im Kompetenzmodell von COACTIV Prämissen für ein erfolgreiches Erfüllen von Anforderungen des Lehrerberufs aus unterschiedlichen Perspektiven betrachtet. Folglich bezieht sich professionelle Kompetenz bei COACTIV auf vier übergeordnete Aspekte - Professionswissen, Überzeugungen/Werthaltungen/Ziele, motivationale Orientierungen und selbstregulative Fähigkeiten - welche prinzipiell lern- und vermittelbar sind und aus deren Zusammenspiel professionelles Handeln von Lehrkräften entsteht.

Abbildung 2.1.: Modell professioneller Lehrerkompetenz von COACTIV (nach Baumert & Kunter, 2006) und FALKO (Krauss, Lindl, Schilcher, Fricke et al., 2017)



Durch die Berücksichtigung der nicht-kognitiven Merkmale (motivationale Orientierungen, Selbstregulationsfähigkeit) geht das Kompetenzmodell über das übliche Verständnis von Expertise hinaus. Die im Modell enthaltenen Kompetenzaspekte setzen sich aus Kompetenzbereichen (z. B. für das Professionswissen aus Fachwissen, fachdidaktischem Wissen, pädagogisch-psychologischem Wissen, Orientierungswissen, Beratungswissen) zusammen, welche sich weiter in Kompetenzfacetten (der Kompetenzbereich fachdidaktisches Wissen berücksichtigt z. B. Erklärungswissen, Wissen über das mathematische Denken von Schülern und Wissen über mathematische Aufgaben) aufgliedern (Baumert & Kunter, 2011a).

Innerhalb des Komplexes Überzeugungen/Werthaltungen/Ziele unterscheidet COACTIV zwischen Wertbindungen („value commitments“), epistemologischen Überzeugungen („epistemological beliefs, world views“) subjektiven Theorien über Lehren und Lernen sowie Zielsystemen, wobei der Fokus innerhalb der Studie auf die Untersuchung epistemologischer Überzeugungen, subjektiver Theorien über Lehren und Lernen sowie Unterrichtszielen im Bereich Mathematik gelegt wurde. Die Konzentration auf diese drei Kompetenzbereiche wurde damit begründet, dass diese unmittelbare Handlungsrelevanz für den Unterricht besäßen (Baumert & Kunter, 2011a).

Motivationale Orientierungen und selbstregulative Fähigkeiten als die beiden weiteren Kompetenzaspekte des COACTIV-Modells stellen Merkmale der psychologischen Funktionsfähigkeit von handelnden Personen dar. Beide bilden unter anderem die Basis für die Überwachung und Regulation des beruflichen Handelns. Innerhalb der COACTIV-Studie lag der Fokus des Interesses auf der Untersuchung von intrinsischen motivationalen Orientierungen von Lehrkräften, während berufliche Selbstwirksamkeitsüberzeugungen bezogen auf Unterricht und weitere schulische Anforderungen angesichts eines breiten Forschungshintergrunds nur am Rande erfasst wurden (Baumert & Kunter, 2011a).

2.2. Professionswissen von (Physik-)Lehrkräften

Beim Projekt FALKO und somit auch bei FALKO-Physik liegt der Forschungsfokus auf der Erfassung und Untersuchung des Professionswissens von (Physik-)Lehrkräften als Teilaspekt professioneller Lehrkompetenz. Dem Professionswissen einer Lehrkraft wird eine besondere Rolle für effektives Handeln im Unterricht und, daraus resultierend, den Lernerfolg von Schülern zugeschrieben (z. B. Baumert & Kunter, 2011b; Bromme, 1997; Helmke, 2009; Abell, 2007).

Historisch gesehen lässt sich der Begriff des Professionswissens aus der Entwicklung unterschiedlicher Forschungsparadigmen (Persönlichkeitsparadigma, Prozess-Produkt-Paradigma, Experten-Novizen-Paradigma) ableiten. Ab dem Beginn des 20. Jahrhunderts etablierte sich der Begriff des Persönlichkeitsparadigmas. In den 20er Jahren wurden hierzu auch empirische Studien durchgeführt (Bromme & Haag, 2004). Demnach verweisen „(B)erufszufriedene und erfolgreiche Lehrer [...] auf ihre Lehrerpersönlichkeit.“ (Herrmann, 1999, S. 42). „(D)er Lehrer wirke mehr durch das, was er sei, als durch das, was er wisse“ (Lange, 1895, S. 17). Somit wird der Persönlichkeit der Lehrkraft die entscheidende Rolle bei der Erfüllung der Unterrichts- und Erziehungsaufgaben zugeschrieben (Pause, 1970). Ab den 60er Jahren vollzieht sich unter dem Einfluss des Behaviorismus (Krauss et al., 2008) der Übergang zum Prozess-Produkt-Paradigma. Nicht mehr die Persönlichkeit der Lehrkraft steht nun im Vordergrund, sondern das konkrete Lehrerhandeln im Unterrichtsverlauf sowie dessen Effekte auf die Schüler (Anderson, Evertson & Brophy, 1979). Die Wirksamkeit einer Lehrmethode wird als unabhängig von der Lehrperson, dem Lernenden sowie dem

jeweiligen situationellen Kontext gesehen. Des Weiteren werden Charakteristika einer Lehrkraft, u. a. auch das Wissen dieser, als statisch und nur anteilig das Handeln und die Schülerleistung beeinflussend, angesehen (Abell, 2007; Northfield & Fraser, 1977; Weinert, 1996). Seit dem Beginn der 90er Jahre gelten die Aussagen, die das Experten-Novizen-Paradigma trifft, als zentrale Faktoren für erfolgreichen Unterricht. Ähnlich zum Persönlichkeitsparadigma treten die Gegebenheiten einer Lehrkraft in den Vorder- und die sich daraus ergebenden (Unterrichts)Handlungen in den Hintergrund. Allerdings sind „nicht mehr vage definierte Charakterzüge, sondern [...] das Wissen und Können für die Gestaltung von Lerngelegenheiten“ (Bromme, 1997, S. 186) grundlegend. Lernwirksames Unterrichten ist demnach weniger auf Charaktereigenschaften beziehungsweise individuelle Talente, sondern auf erlernbares Wissen und Können zurückzuführen (Helmke, 2007; Oelkers & Reusser, 2008). Für weitere Informationen zum Experten-Paradigma siehe z. B. Krauss und Bruckmaier (2014).

Im Gegensatz zu anderen FALKO-Teilprojekten (z. B. für Ev. Religion, Latein oder Musik) gibt es für das Fach Physik – vorrangig im deutschsprachigen Raum – mittlerweile vielfältige Ansätze zur Konzeptualisierung und Operationalisierung des fachbezogenen Professionswissens von Lehrkräften (vgl. Tabelle 2.1). Die meisten dieser psychometrischen Testkonstruktionen folgen der taxonomischen Strukturierung von Shulman (1986), der als Professionswissenskategorien vor allem fachdidaktisches Wissen (Pedagogical Content Knowledge, PCK), Fachwissen (Content Knowledge, CK) und pädagogisches Wissen (Pedagogical Knowledge, PK) identifiziert (für weitere Bereiche siehe z. B. Shulman, 1986 oder Kunter et al., 2011), die in der empirischen Bildungsforschung und in der Fachdidaktik als Kernbereiche des Professionswissens angesehen werden (Blömeke, Felbrich & Müller, 2008; Lipowsky, 2006).

Shulman charakterisiert dabei das FDW beziehungsweise das PCK folgendermaßen:

Within the category of pedagogical content knowledge I include, for the most regularly taught topics in one's subject area, the most useful forms of representation of those ideas, the most powerful analogies, illustrations, examples, explanations, and demonstrations – in a word, the ways of representing and formulating the subject that make it comprehensible to others. Since there are no single most powerful forms of representation, the teacher must have at hand a veritable armamentarium of alternative forms of representation, some of which derive from research whereas others originate in the wisdom of practice. Pedagogical content knowledge also includes an understanding of what makes the learning of specific topics easy or difficult: the conceptions and preconceptions that students of different ages and backgrounds bring with them to the learning of those most frequently taught topics and lessons. If those preconceptions are misconceptions, which they so often are, teachers need knowledge of the strategies most likely to be fruitful in reorganizing the understanding of learners, because those learners are unlikely to appear before them as blank slates. (Shulman, 1986, S. 9-10)

Hieraus leiten sich die beiden Subfacetten des fachdidaktischen Wissens „Wissen über Erklären und Repräsentieren von Inhalten“ („E & R“) und „Wissen über typische Schüler-schwierigkeiten und -fehler“ („Schk“) beim FALKO-Modell ab (vgl. Kap. 4).

Für das FW beziehungsweise das CK subsummiert Shulman:

To think properly about content knowledge requires going beyond knowledge of the facts or concepts of a domain. It requires understanding the structures

Tabelle 2.1.: Übersicht über Testinstrumente bzw. Studien auf nationaler Ebene bezogen auf das Fach Physik (es sind nur Instrumente zur Messung des FDW und des FW aufgeführt)

Studie bzw. Testinstrument

PLUS

Modell: FDW, FW als Facetten des professionellen Lehrerwissens

- FDW-Subfacetten: Wissen über Bedingungen des Lernens, Wissen über instruktionale Aktivitäten
- FW-Subfacetten: Komplexität (Fakt, Zusammenhang, Konzept), Inhalt (Schul-/universitäres Wissen)

Themenbereich: Aggregatzustände und ihre Übergänge am Beispiel Wasser

Zielgruppe: Lehrkräfte (Primar-, Sekundarstufe)

- FDW: 14 Items (überwiegend offen)
- FW: 42 Items (geschlossen)

Referenzen: Lange, Kleickmann, Tröbst und Möller (2012); Lange (2010); Ohle, Kauertz und Fischer (2010); Ohle (2010)

Paderborner Instrument

Modell: FW, FDW, PW als Facetten des Professionswissens; beliefs, motivationale Orientierungen)

FW: Dreidimensionales Modell (Achsen: Inhaltsbereich, kogn. Aktivität, Niveaustufen)

- Subfacetten Inhaltsbereich: Kinematik, Kraft, Energie, Impuls
- Subfacetten kog. Aktivität: Reproduzieren, Verstehen, Beurteilen und Analysieren
- Subfacetten Niveaustufen: Schulwissen, vertieftes Wissen, universitäres Wissen

FDW: Eindimensionales Modell

Subfacetten: Wissen über (allgemeine) Aspekte physikalischer Lernprozesse (deklarativ), Wissen über den Einsatz von Experimenten (deklarativ), Gestaltung und Planung von Lernprozessen (prozedural), Beurteilung, Analyse und Reflexion von Lernprozessen (prozedural), adäquate Reaktion in kritischen Unterrichtssituationen (prozedural)

Themenbereich: Mechanik

Zielgruppe: Lehramtsstudierende

- FW: 10 Aufgaben mit 28 Teilitems (offen und geschlossen)
- FDW: 8 Aufgaben mit 39 Teilitems (offen und geschlossen)

Referenzen: Riese (2009); Vogelsang (2014)

QuIP

Modell:

FDW-Facetten: Wissen über Schülerfehlvorstellungen, Wissen über das Curriculum, Wissen über Schwierigkeiten von Aufgaben und Inhalten

Themenbereich: Elektrizitätslehre

Zielgruppe: Lehrkräfte in der Bundesrepublik Deutschland, Schweiz und Finnland

- Theoretisches FDW: 29 (bzw. 25) Items (offen)
- Lehrerhandeln: Videografie von zwei Unterrichtsstunden (Videokodiersystem)

Referenzen: Olszewski (2010)

Fortsetzung auf der nächsten Seite

Tabelle 2.1.: (Fortsetzung von S. 12)

ProwiN

Modell: FW, FDW und PW als Facetten des Professionswissens

FDW: Wissensbereiche (deklarativ, prozedural, konditional), Subfacetten (Wissen über Konzepte, Wissen über Experimente, Wissen über Schülervorstellungen)

FW: Inhaltsbereich (Mechanik), Wissensstufen (Schulwissen, vertieftes Schulwissen)

Themenbereich: Mechanik; zusätzlich Elektrizitätslehre/Optik, themenunabhängige FDW-Items

Zielgruppe: Lehramtsstudierende, Lehrkräfte und Lehrkräfte im Vorbereitungsdienst

– FW: 15 (geschlossen)

– FDW: 17 (überwiegend halboffen)

Referenzen: Tepner et al. (2012); Kirschner (2013); Cauet, Liepertz, Kirschner, Borowski und Fischer (2014); Cauet, Borowski und Fischer (2015); Cauet (2016)

KiL, KeiLa

Modell: FW, FDW und PW als Facetten des Professionswissens.

FDW (dreidimensionales Modell):

– Wissensart: deklaratives Wissen, prozedurales Wissen, schematisches/strategisches Wissen

– fachdidaktischer Inhalt: Schülerkognition, Instruktionsstrategien, Curriculum, Assessment

– Fachinhalt: Mehrere physikalische Inhaltsbereiche

FW (zweidimensionales Modell):

– Wissensart: deklaratives Wissen, prozedurales Wissen, schematisches/strategisches Wissen

– Fachinhalt: Mehrere Inhaltsbereiche

Themenbereich: Mehrere

Zielgruppe: Lehramtsstudierende

– FW: 59 Items (überwiegend geschlossen, z. T. halboffen)

– FDW: 39 Items (überwiegend geschlossen, z. T. halboffen)

Referenzen: Kröger, Neumann und Petersen (2014); Sorge, Kröger, Petersen und Neumann, manuscript submitted for publication; Hohenstein, Kleickmann, Zimmermann, Köller und Möller (in press)

Fortsetzung auf der nächsten Seite

Tabelle 2.1.: (Fortsetzung von S. 12)

ProfiLe-P

Modell: Universitär erwerbbares FW, universitär erwerbbares deklarative/analytische/prozedurale Aspekte des FDW als Teile des Professionswissens angehender Lehrkräfte

FW: Fachstufen (Schulwissen, vertieftes Schulwissen, universitäres Wissen)

FDW: Facetten (Instruktionsstrategien, Schülervorstellungen, Experimente/Vermittlung eines angemessenen Wissenschaftsverständnisses, fachdidaktische Konzepte), kognitive Anforderungen (Reproduzieren, Analysieren, Anwenden)

EW (Erklärungswissen): Unterrichtsnahe, dialogische Erklärungssituationen von Physik

Themenbereich: Mechanik **Zielgruppe:** Lehramtsstudierende *ProfiLe-P*, *DaWis*: FDW: 2 Testhefte mit insg. 91 Items in 50 Aufgaben (offen, geschlossen)

ProfiLe-P, *FaWis*: FW: 40 Items (geschlossen)

ProfiLe-P, *EWis*: EW (Erklärungswissen): Praxisnahe Rollenspiele (videografierte, standardisierte, interaktive Testsituationen; verschiedene physikalische Themenbereiche)

ProfiLe-P: Erweiterungsprojekt FW (verfeinerte Modellierung): Items (offen, geschlossen) auf je vier Kompetenzstufen in Blöcken auf 10 Testhefte verteilt

FW-Subfacette „Nutzung von Fachwissen“: 87 Items

FW-Subfacette „Nutzung von Experimenten und Modellen“: 49 Items

Referenzen: Riese et al. (2015); Gramzow, Riese und Reinhold (2014); Gramzow (2014);

Walzer, Fischer und Borowski (2014); Tomczyszyn, Kulgemeyer und Schecker (2014);

Woitkowski, Riese und Reinhold (2011); Woitkowski (2015)

Ko-WADiS

Kompetenzmodell: Naturwissenschaftliche Untersuchungen (Experimente/Beobachtungen; vier Subdimensionen: Fragestellung, Hypothese, Planung/Durchführung, Auswertung/Reflexion), Modelle nutzen (Zweck, Testen, Ändern von Modellen)

Themenbereich: Naturwissenschaftliche Erkenntnisgewinnung, Modelle

Zielgruppe: Lehramtsstudierende, Bachelor- und Masterstudierende

FW: 70 Items pro Fach (geschlossen)

Referenzen: Straube, Stiller, Tiemann und Nordmeier (2014); Straube, Tiemann, Upmeier zu Belzen, Krüger und Nordmeier (2014); Stiller, Straube, Hartmann, Nordmeier und Tiemann (2015)

FuN-EKoL

Modell: Test- und Videovignetten zur Erfassung handlungsrelevanten Professionswissens

EKoL 8: FDW im Bereich Experimentieren

EKoL 9: FDW (Wissen und Können) im Bereich Modellkompetenz

Referenz: Wilhelm et al. (2016)

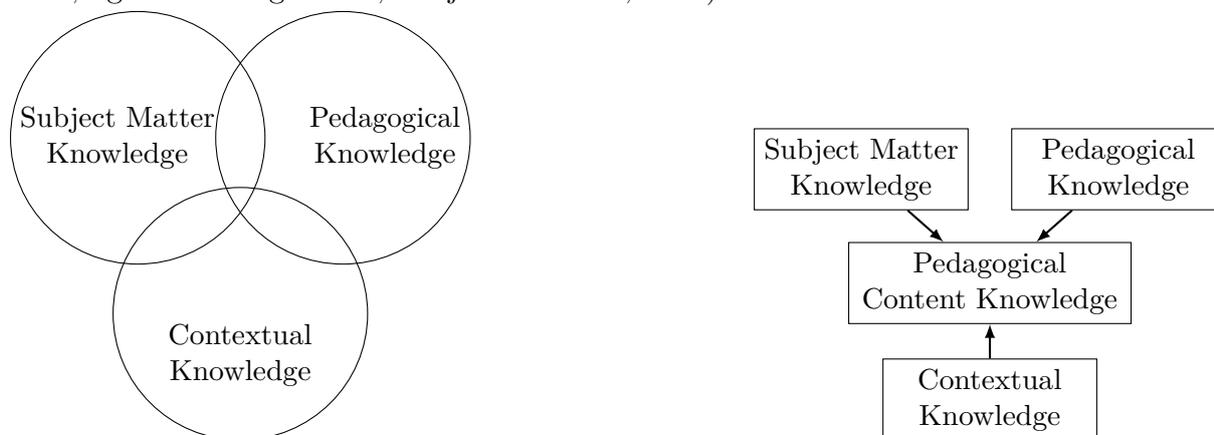
of the subject matter [...]. [...] [T]he structures of a subject include both the substantive and syntactic structure. The substantive structures are the variety of ways in which the basic concepts and principles of the discipline are organized to incorporate its facts. The syntactic structure of a discipline is the set of ways in which truth or falsehood, validity or invalidity, are established. [...]. The teacher need not only to understand that something is so, the teacher must further understand why it is so, on what grounds its warrant can be asserted, and under what circumstances our belief in its justification can be weakened and even denied. (Shulman, 1986, S. 9)

Im Vergleich zum überfachlichen pädagogischen Wissen (PW) werden fachdidaktisches Wissen und Fachwissen generell als fach- beziehungsweise sogar als themenspezifisch angesehen (Park & Chen, 2012), das heißt, dass beispielsweise von einer hohen didaktischen Kompetenz im Bereich der Atomphysik nicht notwendigerweise auch auf vergleichbares didaktisches Wissen im Bereich der Mechanik geschlossen werden kann (Kirschner, 2013). Dennoch wird im Rahmen vieler Testkonstruktionen bislang angenommen, dass eine umfangreiche Wissensbasis in der Mechanik einen guten Prädiktor für Wissen in anderen physikalischen Teilgebieten darstellt (z. B. Friege & Lind, 2004) und demzufolge operationalisieren mehrere Instrumente fachspezifisches Professionswissen in Physik (vornehmlich) über das Wissen im Bereich Mechanik (Tab. 2.1). Da diese artifizielle Einschränkung der inhaltsreichen Physikdomäne als einer „Komplexdisziplin“ testtheoretisch kritisch gesehen werden kann, werden im Teilprojekt FALKO-Physik sowohl Fachwissen als auch fachdidaktisches Wissen in größerer fachlicher Breite abgebildet (für eine Diskussion zur „Repräsentativitätshypothese“ siehe Abschn. 6.5).

Physikalisches Fachwissen wird in verschiedenen Testkonzeptionen auf unterschiedlichen Wissensstufen konzeptualisiert, die man grob in fachliches Alltagswissen, Schulwissen, vertieftes Schulwissen (d. h. fachliches Hintergrundwissen zum Schulcurriculum) und universitäres Wissen trennen kann (Riese, 2009; Baumert et al., 2010). Während es sich beim Schulwissen um Wissensbestände handelt, über die Schüler verfügen sollten, sehen beispielsweise Woitkowski, Riese und Reinhold (2011) vertieftes Schulwissen als „Verknüpfungswissen“ zwischen Schulwissen und universitärem Wissen. Beim universitären Wissen handelt es sich im Gegensatz dazu um Wissen, das genuin während des Studiums an der Universität erworben wird, aber in der Regel losgelöst vom Schulstoff ist. Fachliches Alltagswissen (welches bei Professionswissenstests nicht operationalisiert wird) schließlich umfasst all jenes Wissen, über das Personen unabhängig von schulischen oder universitären Lerngelegenheiten allein durch Begegnungen mit Alltagsphänomenen verfügen. In bisherigen Testkonstruktionen zum Fach Physik wird FW meist als Schulwissen oder vertieftes Schulwissen, gelegentlich auch als universitäres Wissen, operationalisiert. In den Projekten KiL oder KeiLa beispielsweise wird physikalisches Fachwissen auf universitärem Niveau erhoben (in diesen beiden Projekten werden über die Mechanik hinaus z. B. auch Wissen über Festkörperphysik oder spezielle Relativitätstheorie erfasst).

Auch für das FDW erfolgen Testkonstruktionen oft unter ausschließlicher Berücksichtigung des Themengebiets Mechanik (Tab. 2.1). Weitere Inhaltsbereiche, wie zum Beispiel Elektrizitätslehre oder Optik, werden oft nicht (z. B. Paderborner Instrument, Riese, 2009) oder nur mit einer geringen Itemzahl berücksichtigt (z. B. ProwiN, Kirschner, 2013). Es gibt auch Ansätze, in denen das FDW nur über weitere Einzelgebiete der Physik, beispielsweise über die Elektrizitätslehre (QuiP-Studie, Olszewski, 2010) erfasst wird, auf der anderen Seite aber auch sehr breite und sogar naturwissenschaftsübergreifende Instrumente (FuN-EKoL). Häufig operationalisierte Wissensfacetten sind „Instruktion“, „Schülerver-

Abbildung 2.2.: Integratives Modell von PCK (linke Seite; nach Gess-Newsome, 1999) und transformatives Modell des PCK (rechte Seite; nach Gess-Newsome, 1999; Grossman, 1990; vgl. auch Magnusson, Krajcik & Borko, 1999)



ständnis“, „Modelle/Konzepte“, „Experimente“ oder „Curriculum“, wobei vor allem die beiden erstgenannten Subfacetten in beinahe allen Projekten zu finden sind. Ebenso werden bei den meisten Testkonstruktionen implizit oder explizit sowohl deklarative, prozedurale sowie (möglichst) auch konditionale Wissensaspekte berücksichtigt (Anderson et al., 2001; Gruber & Renkl, 2000; Paris, Lipson & Wixson, 1983). Eine explizite Unterscheidung dieser Wissensarten bei der Itemkonstruktion zum FDW findet sich beispielsweise bei Kirschner (2013).

Neben dem Fokus auf Modellierung und Konzeptualisierung von Lehrerkompetenzen, richtet sich das Forschungsinteresse beispielsweise auch auf die Frage nach Zusammenhängen zwischen den Professionswissenskategorien FDW, FW und PW. Hierbei stechen aus der Literatur insbesondere zwei grundlegend unterschiedliche Annahmen in Bezug auf das fachdidaktische Wissen ins Auge. Auf der einen Seite wird FDW als integriertes Wissen angesehen, welches aus der Überschneidung von FW, PW und Kontextwissen resultiert (Gess-Newsome, 1999; vgl. Abb. 2.2, linke Seite; nach Grossman, 1990 umfasst Kontextwissen Kenntnisse des jeweiligen Schulumfelds, der Schule selbst sowie der jeweiligen Schülerschaft). Hieraus ergäbe sich der Schluss, dass FDW nicht von den anderen beiden Wissenskategorien trennbar wäre. Auch Bass und Ball (2004) betonen die Untrennbarkeit von FDW und FW. Auf der anderen Seite kann fachdidaktisches Wissen auch als eigene transformierte Wissenskategorie betrachtet werden, welche sich aus Fachwissen, pädagogischem Wissen und Kontextwissen speist (Gess-Newsome, 1999; Baumert & Kunter, 2006; Grossman, 1990; Magnusson et al., 1999); vgl. Abb. 2.2, rechte Seite). In diesem Fall wäre es anzunehmen, dass sich FDW von FW (empirisch) trennen ließe, was auch Shulmans (1986; 1987) Postulat von zwei sich zwar überlappenden, aber dennoch disjunkten Professionswissenskategorien stützen würde. Trotz mittlerweile zahlreicher Untersuchungen zum fachspezifischen Professionswissen, vor allem im Bereich der Naturwissenschaften, ist aufgrund der uneinheitlichen Befundlage immer noch nicht geklärt, welche der beiden obigen Annahmen final zu stützen ist, wobei sich aber immer wieder die enge Verknüpfung von FDW mit FW zu bestätigen scheint (z. B. Hill, Schilling & Ball, 2004; Kirschner, 2013; Lindl & Krauss, 2017).

Für die COACTIV-Studie ergab sich beispielsweise, dass sich zwar FDW und FW bei den untersuchten nicht-gymnasialen Lehrkräften trennen lassen ($r = 0,64$), dies für die gymnasialen Lehrkräfte (zumindest auf latenter Ebene; $r = 0,94$) aber nicht mehr der Fall

zu sein scheint (Krauss et al., 2008). Die Gymnasiallehrkräfte bei COACTIV verfügen – neben deutlich mehr Fachwissen – durchschnittlich auch über mehr fachdidaktisches Wissen als ihre Kollegen aus den anderen Sekundarschularten. Weiterhin gilt aber, dass ihr fachdidaktisches Wissensniveau stärker von ihrem Fachwissen determiniert ist. Diese stärkere Vernetzung der unterschiedlichen Wissensarten kann in Übereinstimmung mit der Expertisetheorie (z. B. Ericsson et al., 1993) gesehen werden.

In weiteren Studien zeigten sich, ähnlich wie bei COACTIV, mittlere bis hohe positive Korrelationen zwischen FDW und FW (Baumert et al., 2010; Krauss et al., 2008; Olszewski, Neumann & Fischer, 2010; Riese & Reinhold, 2008). Dies zeigt einerseits den engen Zusammenhang zwischen den beiden Professionswissenskategorien, verdeutlicht zum anderen aber, dass die beiden Konstrukte statistisch voneinander zu trennen sind, was Shulmans (1986) Postulat stützen würde. Zu erklären versucht werden diese Befunde durch die Hypothese, dass FW eine wichtige Basis für die Ausbildung von FDW, welches zur erfolgreichen Gestaltung von Unterricht benötigt wird, bildet (Baumert et al., 2010; Hashweh, 1987; Loewenberg Ball, Theule Lubienski & Spangler Mewborn, 2001). Im Umkehrschluss soll das FDW wiederum positiv auf das FW rückwirken (Magnusson et al., 1999). Beispiele für Korrelationswerte verschiedener Studien finden sich beispielsweise bei Kirschner (2013).

Beim Projekt FALKO stützen konfirmatorische Faktorenanalysen in allen beteiligten Fächern Shulmans Hypothese, wobei aber die erhaltenen Korrelationswerte eine hohe Schwankungsbreite aufweisen. Lindl und Krauss (2017) folgern daraus:

Die zwischen beiden Wissensbereichen auftretenden Korrelationen spiegeln indes deren reziprok-kohäsive, konfligierende Relation adäquat wieder (H. G. Neuweg, 2014, S. 591–592). Eine transfachliche, metaanalytische Verallgemeinerung der Größe dieses Zusammenhangs erscheint jedoch angesichts der interdisziplinären Abweichungen nicht möglich, sodass die Differenzen durch die facheigenen Operationalisierungen bedingt und damit als Domänenspezifika anzusehen sind, [...]. (Lindl & Krauss, 2017, S. 422)

Beim Vergleich der FALKO-Resultate mit weiteren Studien ergibt sich folgendes Bild: Die für die Fächer Deutsch, Englisch und Mathematik erhaltenen Korrelationswerte ähneln denen von MT21 (Blömeke, Kaiser & Lehmann, 2008, S. 69), TEDS-LT (Bremerich-Vos, Dämmer, Willenberg & Schwippert, 2011, S. 62; Buchholtz, Kaiser & Stancel-Piątak, 2011, S. 128; Roters, Nold, Haudeck, Keßler & Stancel-Piątak, 2011, S. 95; Jansing, Haudeck, Keßler, Nold & Stancel-Piątak, 2013, S. 94-95; Buchholtz & Kaiser, 2013, S. 133-134) sowie PKE (König et al., 2016, S. 328). Gleiches zeigt sich in Bezug auf die Fächer Biologie, Chemie, Physik und Mathematik im Rahmen von KiL. Bezogen auf die ProwiN-Studie (Biologie, Chemie und Physik) liegen die bei FALKO erhaltenen Werte etwas höher, gerade auch für das Fach Physik sowohl bei KiL ($r = 0,40^{**}$) als auch bei ProwiN ($r = 0,45^{**}$) (Jüttner, Boone, Park & Neuhaus, 2013, S. 60, 63; Universität Duisburg-Essen, 2013; Großschedl, Harms, Glowinski & Waldmann, 2014; Kleickmann et al., 2014; Tepner & Dollny, 2014, S. 249).

Aufgrund der weiterhin heterogenen Forschungslage und der überwiegend schwierigen Vergleichbarkeit aufgrund der unterschiedlichen Modellierungen und Konzeptualisierungen des Professionswissens wird von den Autoren daher angeregt, diesbezügliche Forschungsbestreben domänenspezifisch zu vertiefen.

Ein weiterer prominenter Diskussionspunkt innerhalb der Professionswissensforschung betrifft die Art der Erfassung des Professionswissens. Da Erhebungen in Form von Papier-

und-Bleistift-Tests, wie sie bei den meisten der in Tabelle 2.1 genannten Forschungsvorhaben durchgeführt wurden, aufgrund des bei solchen Studien üblichen Laborsettings kritisiert werden können, nutzen einzelne Studien zur Erhöhung der ökologischen Validität weitere Instrumente wie zum Beispiel die Analyse von Videoaufzeichnungen aus dem Unterricht der Lehrkräfte (QuiP) oder Rollenspiele (ProfiLe-P, Teilprojekt EWis; Tomczyszyn, Kulgemeyer und Schecker, 2014). Eine Diskussion über die Natur des im Rahmen von FALKO gemessenen Wissens findet sich in Krauss, Lindl, Schilcher, Fricke et al. (2017). Zur Diskussion über das Spannungsfeld zwischen theoretischem Wissen, situationsspezifischen Kompetenzen und tatsächlichem Handeln gibt es zahlreiche Literatur mit explizitem Bezug zum Professionswissen von Lehrkräften (vgl. die Unterscheidung von „cognitive perspective“ und „situated perspective“, Depaepe, Verschaffel & Kelchtermans, 2013; Blömeke, Gustafson & Shavelson, 2015; Krauss, Lindl, Schilcher, Fricke et al., 2017).

In den Naturwissenschaften - hier sind diese Diskussion derzeit besonders virulent - gibt es mittlerweile auch Ansätze, in denen sowohl das Wissen (operationalisiert über Papier- und Bleistift-Tests) als auch das Handeln im Unterricht (über Videoaufzeichnungen) von Lehrkräften erhoben werden und die gemessenen (beziehungsweise kodierten) Konstrukte mit Schülermerkmalen in Beziehung gesetzt werden (im Projekt ProwiN für das Fach Physik z. B. Cauet, Liepertz, Kirschner, Borowski und Fischer, 2014, Liepertz und Borowski, 2015; für das Fach Chemie z. B. Dollny, 2011, Strübe, Tepner und Sumfleth, 2015, Tröger, Sumfleth und Tepner, 2016; für das Fach Biologie z. B. Jüttner und Neuhaus, 2013; sowie im Projekt ProfiLe-P z. B. Vogelsang, 2014; Gramzow, 2014; Woitkowski, 2015). Hierbei ist die Stärke von ProwiN, dass innerhalb des Projekts für die Naturwissenschaften Biologie, Chemie und Physik, ähnlich wie bei FALKO, zur Erhöhung der Vergleichbarkeit in den Tests zur Erfassung des FDW jeweils konzeptuell identische Facetten (Experimente, Modelle, Schülervorstellungen; Tepner et al., 2012, Kirschner et al., 2017) domänenspezifisch operationalisiert wurden (Krauss, Lindl, Schilcher & Tepner, 2017). Dabei konnten in der Biologie keine direkten Effekte von FW und FDW auf den Lernerfolg der Schüler nachgewiesen werden, in Bezug auf das FDW zeigte sich jedoch ein indirekter positiver Effekt, mediiert über die kognitive Aktivierung im Unterricht, auf den Lernerfolg (Förtsch, Werner, von Kotzebue & Neuhaus, 2016). Im Fach Chemie konnten direkte Zusammenhänge sowohl des FW als auch des FDW mit der Schülerleistung nachgewiesen werden (Strübe, Tepner & Sumfleth, 2016; Tröger et al., 2016). Für das Fach Physik (vgl. Tab. 2.1) wurde ein Effekt des FW, aber nicht des FDW auf eine kognitiv aktivierende Unterrichtsgestaltung gefunden, ein Einfluss auf die Schülerleistung konnte jedoch weder ausgehend von FW noch von FDW nachgewiesen werden (Cauet, 2016). Möglicherweise wurde also durch den FDW-Test bezüglich Physik bei ProwiN entweder kein handlungsrelevantes Wissen erfasst oder es wurden durch die gewählte Kodierung der Unterrichtsqualität keine von diesem Wissen beeinflussten Faktoren identifiziert (vgl. z. B. Cauet, Borowski & Fischer, 2015).

In der PLUS-Studie (Möller, Kleickmann & Lange, 2013, bzw. Tab. 2.1) konnten dagegen für das fachdidaktische Wissen einer Physiklehrkraft Zusammenhänge mit der Schülerleistung, aber auch mit dem Kompetenzerleben beziehungsweise dem Fachinteresse der Schüler gefunden werden (Lange, Kleickmann, Tröbst & Möller, 2012). Zwischen dem physikalischen Fachwissen einer Lehrkraft und dem Lernerfolg der Schüler konnte in PLUS dagegen kein direkter Zusammenhang nachgewiesen werden, allerdings scheint die Wirkung des Fachwissens auf den Lernerfolg durch die Sequenzierung des Unterrichts (erhoben über Videografie) moderiert zu werden (Ohle, Fischer & Kauertz, 2011).

An den teilweise widersprüchlichen Befunden zur prädiktiven Validität von physika-

lischem FDW setzen die Projekte Ko-WADiS und FuN-EKoL an, in denen das Wissenschaftsverständnis hinsichtlich der naturwissenschaftlichen Arbeitsweisen Beobachten, Experimentieren und des Nutzens von Modellen operationalisiert wurde (die Fachdisziplinen Biologie, Chemie und Physik werden dort – dem PISA-Verständnis folgend – gemeinsam als Naturwissenschaften gefasst). Im Gegensatz zur üblichen Anlehnung an Shulman (1986; 1987) erfolgte die Modellierung des Professionswissens im Projekt Ko-WADiS durch eine Transformation der Inhalte der Bildungsstandards für den mittleren Schulabschluss in Lehrerkompetenzen. In FuN-EKoL wiederum werden anstelle von Papier-und-Bleistift-Tests Videovignetten als Stimuli für die Testteilnehmer eingesetzt. Für das Fach Mathematik siehe hierzu z. B. Lindmeier, Heinze und Reiss (2013), Bruckmaier, Krauss, Blum und Leiss (2016). Aktuell liegen noch keine Veröffentlichungen zu Ergebnissen der Projekte Ko-WADiS und FuN-EKoL vor.

3. Anlage und Ziele dieser Arbeit

Nachdem in der Einleitung und im vorangegangenen Kapitel eine Darstellung des dieser Dissertation zugrunde liegenden Forschungshintergrunds erfolgte, werden im Folgenden Anlage und Ziele dieser Arbeit mit ihren Forschungsfragen und Hypothesen beschrieben (nähere Informationen zu Anlage und Zielen des Gesamtprojekts FALKO sowie zum dortigen methodischen und statistischen Vorgehen bei den durchgeführten Untersuchungen finden sich in Krauss, Lindl, Schilcher und Tepner, 2017 sowie Lindl und Krauss, 2017).

3.1. Forschungsfragen und Hypothesen

Ziel des Forschungsprojekts FALKO war die Konstruktion von objektiven, reliablen und validen Testinstrumenten zur Erhebung des fachspezifischen Professionswissens in den beteiligten Disziplinen Deutsch, Englisch, evangelische Religion, Latein, Musik, Physik sowie Pädagogik. Innerhalb des gemeinsamen Rahmenmodells, welches sich an der Taxonomie von Shulman (1986; 1987) sowie am COACTIV-Modell der Lehrerkompetenzen orientiert (vgl. Krauss, Lindl, Schilcher und Tepner, 2017, bzw. Kapitel 1), wurde dabei eine möglichst breite Erfassung der Konstrukte FDW und FW angestrebt. Diese Zielsetzung hatte im Fall des Testinstruments FALKO-P zur Folge, dass sowohl FDW als auch FW nicht nur über ein (vermeintlich repräsentatives) Teilthema der Physik, zum Beispiel Mechanik, operationalisiert wurde, sondern über die vier Themenbereiche Mechanik, Elektrizitätslehre, Optik und Wärmelehre, welche in ihrer Gesamtheit den größten Anteil an den Physik-Curricula der untersuchten Schularten einnehmen. Dies steht im Gegensatz zu den meisten anderen Testkonstruktionen im Fach Physik, welche das Professionswissen im Regelfall nur über ein einziges Teilgebiet der Physik (z. B. Paderborner Instrument; Riese, 2009) oder aber auf einem anderen fachlichen Niveau (KiL, KeiLa; z. B. Kröger, Neumann & Petersen, 2014) als bei FALKO angestrebt, operationalisieren (vgl. Tab. 2.1 in Abschnitt 2.2). Aus diesen übergeordneten Zielsetzungen leiten sich die Forschungsfragen und Hypothesen des FALKO-Projekts (vgl. Krauss, Lindl, Schilcher & Tepner, 2017; Lindl & Krauss, 2017) und diejenigen des in diesem Band beschriebenen Dissertationsvorhabens ab, welche im Folgenden vorgestellt werden.

Insgesamt sollen durch den Einsatz des Testinstruments FALKO-P in einem ersten Schritt die folgenden Forschungsfragen beantwortet werden können:

- FF1** Lässt sich das fachspezifische Professionswissen (FDW, FW) von Physiklehrkräften objektiv, reliabel und valide messen?
- FF2** Lässt sich die zweidimensionale Struktur des Konstrukts (transformatives Verständnis des PCK nach Gess-Newsome, 1999; Baumert & Kunter, 2006; Grossman, 1990; Magnusson et al., 1999) stützen?
- FF3** Wie unterscheiden sich die untersuchten Teilgruppen (Studierende und Lehrkräfte an Mittelschulen, Realschulen, Gymnasien; fachfremde Mittelschullehrkräfte, vgl.

Kapitel 1, Fußnote 2; Biologie-, Chemie- und Mathematiklehrkräfte; Fachphysiker) in Bezug auf ihr Professionswissen?

- FF4** Wie unterscheiden sich Studienteilnehmer in Bezug auf ihr Professionswissen beim Vergleich ihrer Testleistungen bei FALKO-P und in den ProfiLe-P-Testinstrumenten zum FDW (Gramzow, 2014) und FW (Woitkowski, 2015)?
- FF5** Wie beeinflussen die Prädiktoren Status (Lehrkräfte vs. Studierende) und Schulform (Gymnasium vs. Nicht-Gymnasium) – bei Kontrolle von Geschlecht und kognitiver Leistungsfähigkeit – das FDW und das FW?
- FF6** Wie hängt das fachbezogene Professionswissen mit Indikatoren des Schul- und Studienerfolgs sowie der Berufs- und Studiendauer zusammen?
- FF7** Ist Mechanik (oder eines der anderen operationalisierten Teilgebiete der Physik) repräsentativ für das Gesamtkonstrukt physikalisches FW und FDW? („Repräsentativitätshypothese“)

Da im Ergebnisteil neben deskriptiven und inferenzstatistischen Analysen sowie der Dimensionalitätsprüfung hauptsächlich Ergebnisse zu den Prädiktoren Status (Lehrkräfte vs. Studierende vs. Kontrastpopulationen) und Gruppen (Gymnasium vs. Nicht-Gymnasium) berichtet werden sollen, leiten sich diesbezüglich aus den Forschungsfragen folgende Hypothesen ab, welche anhand der während der Hauptstudie erhobenen Daten geprüft werden sollen (in Anlehnung an Lindl & Krauss, 2017).

Bezüglich des fachdidaktischen Wissens lauten die Hypothesen:

- Hypothese 1a** Lehrkräfte verfügen über mehr fachdidaktisches Wissen als Studierende.
- Hypothese 2a** Die fachdidaktischen Leistungen von Studierenden und Lehrkräften unterscheiden sich hinsichtlich ihrer Schulformzugehörigkeit.
- Hypothese 3a** Bei gleicher Schulformzugehörigkeit sind die Lehrkräfte den Studierenden im fachdidaktischen Wissen überlegen.
- Hypothese 4a** Physiklehrkräfte verfügen über mehr fachdidaktisches Wissen als Mittelschullehrkräfte, welche Physik fachfremd im Fächerverbund „PCB“ beziehungsweise „NT“ unterrichten.
- Hypothese 5a** Physiklehrkräfte verfügen über mehr fachdidaktisches Wissen als Lehrkräfte (RS, Gy), die Biologie, Chemie oder Mathematik unterrichten.
- Hypothese 6a** Gymnasiale und nicht-gymnasiale Physiklehrkräfte verfügen über mehr fachdidaktisches Wissen als Fachphysiker.
- Hypothese 7a** Eine höhere kognitive Leistung (approximiert über die Abiturnote) führt zu einem höheren fachdidaktischen Wissen.
- Hypothese 8a** Männliche Studienteilnehmer verfügen über ein höheres fachdidaktisches Wissen als weibliche.
- Hypothese 9a** Die Schulformzugehörigkeit (Gy vs. nGy) trägt im Vergleich zur Statuszugehörigkeit (Studierende vs. Lehrkräfte) zu einer höheren Varianzaufklärung hinsichtlich des FDW bei.

Hypothese 10a Ein gutes Abschneiden im ersten beziehungsweise zweiten Staatsexamen zieht ein höheres fachdidaktisches Wissen nach sich.

Hypothese 11a Eine höhere Studien-/Berufserfahrung zieht ein höheres fachdidaktisches Wissen nach sich.

Hypothese 12a Physiklehrkräfte erzielen im Fachdidaktikteil des Testinstruments FALKO-P ähnliche Resultate wie im Fachdidaktik-Test aus dem Projekt ProfiLe-P (Gramzow, 2014).

Das Fachwissen betreffend können folgende Hypothesen aufgestellt werden:

Hypothese 1b Lehrkräfte verfügen über mehr Fachwissen als Studierende.

Hypothese 2b Die fachwissenschaftlichen Leistungen von Studierenden und Lehrkräften unterscheiden sich hinsichtlich ihrer Schulformzugehörigkeit.

Hypothese 3b Bei gleicher Schulformzugehörigkeit sind die Lehrkräfte den Studierenden im Fachwissen überlegen.

Hypothese 4b Physiklehrkräfte verfügen über mehr Fachwissen als Mittelschullehrkräfte, welche im Physik fachfremd im Fächerverbund „PCB“ beziehungsweise „NT“ unterrichten.

Hypothese 5b Physiklehrkräfte verfügen über mehr Fachwissen als Lehrkräfte (RS, Gy), die Biologie, Chemie oder Mathematik unterrichten.

Hypothese 6b Gymnasiale Physiklehrkräfte verfügen über ein ähnliches, nicht-gymnasiale Physiklehrkräfte über weniger Fachwissen als Fachphysiker.

Hypothese 7b Eine höhere kognitive Leistung (approximiert über die Abiturnote) führt zu einem höheren Fachwissen.

Hypothese 8b Männliche Studienteilnehmer verfügen über ein höheres Fachwissen als weibliche.

Hypothese 9b Die Schulformzugehörigkeit (Gy vs. nGy) trägt im Vergleich zur Statuszugehörigkeit (Studierende vs. Lehrkräfte) zu einer höheren Varianzaufklärung hinsichtlich des FW bei.

Hypothese 10b Ein gutes Abschneiden im ersten beziehungsweise zweiten Staatsexamen zieht ein höheres Fachwissen nach sich.

Hypothese 11b Eine höhere Studien-/Berufserfahrung zieht ein höheres Fachwissen nach sich.

Hypothese 12b Physiklehrkräfte erzielen im Fachwissensteil des Testinstruments FALKO-P ähnliche Resultate wie im Fachwissens-Test aus dem Projekt ProfiLe-P (Woitkowski, 2015).

Zur Frage der Repräsentativität eines einzelnen Themengebiets der Physik (z. B. Mechanik) für das Professionswissen im Fach Physik kann, der in Forscherkreisen vorherrschenden Überzeugung entsprechend, folgende Hypothese aufgestellt werden:

Hypothese 13 Mechanik (oder ein anderes der untersuchten Teilgebiete der Physik) ist repräsentativ für das Professionswissen im Fach Physik.

3.2. Methodisches und statistisches Vorgehen

Um innerhalb des Gesamtprojekts FALKO ein möglichst gleichschrittiges Vorgehen zu erzielen, gleichzeitig aber den Spezifika des Fachs Physik vor allem in Bezug auf das FDW gerecht zu werden, wurde in einem ersten Schritt das FALKO-Modell zum Professionswissen (FDW und FW) für das Fach Physik adaptiert (vgl. Kap. 4). Das Testinstrument FALKO-P sollte als Papier-und-Bleistift-Test (Powertest; z. B. Bühner, 2011), genau wie diejenigen der anderen FALKO-Teilprojekte, für (vergleichende) Analysen auf Populationsebene einsetzbar sein und nicht der Individualdiagnostik dienen. Um der von FALKO geforderten Breite des Konstrukts gerecht zu werden, wurden sowohl beim FDW als auch beim FW die Themengebiete Mechanik, Elektrizitätslehre, Optik und Wärmelehre in die Itemkonstruktion einbezogen. Das adaptierte Modell, welches weiterhin im FDW zentrale Aspekte des fachdidaktischen Physikwissens (vgl. Kap. 1 und 4) und im FW sowohl vertieftes Hintergrundwissen zu Themengebieten der Physikcurricula als auch Schulwissen abbilden sollte, stellte die heuristische Grundlage zur Itementwicklung dar. Nach ausführlicher Literaturrecherche (vgl. Abschn. 5.1) und unter Berücksichtigung allgemeiner Test- und Aufgabenmerkmale (z. B. Bühner, 2011) erfolgte die Konstruktion von insgesamt 116 offenen und geschlossenen Items mit zugehörigen Kodieranweisungen. Zur Sicherstellung der inhaltlichen Validität (Augenscheinvalidität) wurden die Items innerhalb einer Präpilottierungsphase (vgl. Abschn. 5.3) Fachdidaktikern und erfahrenen Lehrkräften unterschiedlicher Schularten zur Begutachtung (Bearbeitung und Rating der Relevanz, vgl. Tab. 5.6, Abschn. 5.3) vorgelegt (teilweise wurden die Präpilottierungsteilnehmer bei der Bearbeitung der Items zum Lauten Denken aufgefordert und dabei videografiert). In der anschließenden Pilotierungsphase (vgl. Abschn. 5.3) bearbeitete eine Stichprobe von Lehrkräften und Studierenden unterschiedlicher Schularten den als Folge der Präpilottierungsphase reduzierten Itempool. Die Augenscheinvalidität wurde hierbei erneut erhoben. Nach Durchführung statistischer Analysen wurden unter Berücksichtigung der klassischen Testtheorie (inhaltliche Validität, Auswertungsobjektivität, Schwierigkeit, Reliabilität und Trennschärfe; vgl. Tab. 3.1) die Itemauswahl (vgl. Abschn. 5.4) für das finale Testheft (14 offene Items zum FDW, 18 offene und geschlossene Items zum FW) getroffen. Die zugehörigen Kodieranweisungen wurden fortwährend mit Hilfe der Antworten der Studienteilnehmer aus der Präpilottierungs- und der Pilotierungsphase optimiert.

Während der Hauptstudie (vgl. Kap. 6) sollten Validitätshinweise vor allem in Bezug auf Konstruktvalidität (diskriminant mittels Kontrastvalidierung, konvergent mittels Kreuzvalidierung), kriteriale Validität und faktorielle Validität erhalten werden. Hierzu wurde das Testinstrument Studienteilnehmern einer gemischten Stichprobe aus Lehrkräften und Studierenden der Schularten Mittelschule, Realschule und Gymnasium zur Bearbeitung vorgelegt. Als Kontrastgruppen kamen Lehrkräfte an bayerischen Mittelschulen, welche zwar Physik im Fächerverbund „PCB“ (Physik-Chemie-Biologie) beziehungsweise „NT“ (Natur und Technik; LehrplanPLUS, Bayern) unterrichten, das Fach aber nicht im Rahmen ihrer universitären Lehramtsausbildung studierten, fachfremde Lehrkräfte (Schularten Realschule und Gymnasium) der Fächer Biologie, Chemie und Mathematik sowie Fachphysiker hinzu. Weiterhin erfolgte innerhalb einer Teilstudie die Prüfung konvergenter Validität, indem eine kleine Stichprobe an Lehrkräften und Studierenden ($n = 28$) sowohl das Testinstrument FALKO-P als auch die ProfiLe-P-Testinstrumente zum FDW (Gramzow, 2014) und zum FW Woitkowski, 2015) bearbeiteten.

Neben der Betrachtung gängiger psychometrischer Gütekriterien wie beispielsweise Auswertungsobjektivität und Reliabilität liegt der Fokus des Ergebnisberichts zusätzlich auf

Tabelle 3.1.: Richtwerte zur Einschätzung der Internen Konsistenz (Cronbachs α) und der Trennschärpen (r_{it}) sowie der Interraterreliabilität (Spearman's ρ) (Zusammenfassung nach Krauss, Lindl, Schilcher & Tepner, 2017)

<i>Gütekriterium</i>	<i>Bedeutung</i>	<i>Statistischer Indikator</i> (Wertebereich)	<i>Interpretation</i> (Literatur)
Interne Konsistenz (einer Skala)	Genauigkeit, Zuverlässigkeit	Cronbachs α (0 bis 1)	– $\alpha < 0,80$: niedrig – $0,80 \leq \alpha \leq 0,90$: mittel – $\alpha > 0,90$: hoch (Field, 2013; Fisseni, 1997)
Trennschärfe (eines Items)	Repräsentativität für die Skala	Part-whole-korrigierte Korrelation mit der Skala (-1 bis +1)	– $r_{it} < 0,20$: niedrig – $0,20 \leq r_{it} \leq 0,50$: mittel – $r_{it} > 0,50$: hoch (Field, 2013; Fisseni, 1997)
Interraterreliabilität (eines Items)	Auswertungsobjektivität (Übereinstimmung unab- hängiger Kodierer)	Spearman's ρ (0 bis 1)	– $\rho < 0,60$: niedrig – $0,60 \leq \rho \leq 0,90$: mittel – $\rho > 0,90$: hoch (Shavelson & Webb, 1991)
Augenschein- validität (eines Items)	Inhaltliche Gültigkeit (erhoben durch Ein- schätzung der Berufsre- levanz)	Likert-Skala (1 bis 4)	Item ist ... – 1 = nicht, – 2 = kaum , – 3 = eher, – 4 = genau ... berufsrelevant

deskriptiven Ergebnissen bezüglich Skalenmittelwerten und Interkorrelationen, Gruppen- und Schulformunterschieden (Ausbildungssensitivität) sowie Zusammenhängen des fachbezogenen Professionswissens mit Indikatoren des Schul-/Studienerfolgs und der Berufs- und Studiendauer (kriteriale Validität). Mittels linearen Regressionsanalysen soll Auskunft über den Einfluss von Status und Schulform auf beide Konstrukte FDW und FW unter Kontrolle der Eingangsselektivität (Approximierung über die Abiturnote Baron-Boldt, Schuler & Funke, 1988; vgl. Lindl & Krauss, 2017) erhalten werden. Eine Prüfung der Passung (Dimensionalität des fachbezogenen Professionswissens) des FALKO-Modells wird mittels einer konfirmatorischen Faktorenanalyse vorgenommen (faktorielle Validität).

Wie bereits erwähnt, war es ein wesentliches Ziel der Konstruktion von FALKO-P, sowohl das FDW als auch das FW anhand der laut Prüfung der Curricula zentralen physikalischen Themenbereiche Mechanik, Elektrizitätslehre, Optik und Wärmelehre möglichst breit zu operationalisieren, um so auch bisherige Annahmen zum Zusammenhang dieser Wissensbereiche sowie zur (vermeintlichen) Repräsentativität der Mechanik überprüfen zu können („Repräsentativitätshypothese“, vgl. Abschn. 6.5). Hierzu werden im Rahmen der statistischen Auswertung separate Skalen für die vier Themenbereiche gebildet und deren Zusammenhänge mittels Korrelationsmatrizen untersucht.

Da, wie bei Lindl und Krauss (2017, S. 384) angemahnt, die FALKO-einheitliche Konzeptualisierung mit ihren fachbezogenen Operationalisierungen des Professionswissens vorwiegend als Grundlagenforschung zu verstehen ist und zum inner- sowie zum überfachlichen Diskurs anregen soll, werden über die bereits im Rahmen dieser Dissertation durchgeführten Validierungsstudien hinaus Möglichkeiten der Weiterentwicklung des Testinstruments FALKO-P vorgeschlagen (Kap. 8).

4. Modellierung des Professionswissens

Die Modellierung und Operationalisierung des FDW und des FW bei FALKO-Physik erfolgt, wie bereits genauer in Kapitel 2 dargelegt, in enger Anlehnung an das Modell professioneller Lehrerkompetenz, wie es in der COACTIV-Studie verwendet wurde (Krauss et al., 2011) und basiert als Folge auf einem dreidimensionalen Hintergrundmodell. Im Folgenden wird dieses Modell zunächst erläutert sowie ein kurzer Überblick über das Vorgehen bei der Testkonstruktion gegeben (Abschn. 4.1; Näheres hierzu in Kap. 5). Anschließend wird die Konzeptualisierung und Operationalisierung des FDW und des FW detailliert dargestellt (Abschn. 4.2).

4.1. Hintergrundmodell und Vorgehensweise im Überblick

Die erste Dimension des Modells betrifft die Kompetenzbereiche des fachbezogenen Professionswissens selbst. Basierend auf der Wissenstaxonomie von Shulman (1986) erfolgt eine Unterteilung in die beiden Wissenskategorien FDW und FW. Innerhalb des FDW fokussiert FALKO-Physik auf die beiden Facetten „Wissen über Erklären und Repräsentieren von Inhalten“ („E & R“) und „Wissen über Schülerfehler“ („Schk“) sowie als dritte zusätzliche physikspezifische FDW-Facette auf „Wissen über Messen und Experimentieren“ („M & E“) (Abb. 2.1). Die Items zur Messung des FW werden auf den Wissensstufen „Schulwissen“ und „vertieftes Schulwissen“ operationalisiert, jedoch erfolgt keine Postulierung von Subfacetten.

Dimension 2 erstreckt sich auf die physikalischen Inhaltsbereiche. Zwar zeigte sich im Hinblick auf den Kompetenzerwerb von Schülern im Fach Physik, dass Wissen bezüglich Mechanik ein relativ verlässlicher Prädiktor für physikalisches Wissen generell ist (Friege & Lind, 2004), für das Professionswissen von Lehrkräften gibt es jedoch noch keine vergleichbaren empirischen Befunde. Bei FALKO-Physik wird deshalb sowohl FDW als auch FW nicht nur über das Wissen im Teilbereich Mechanik, sondern breiter über das Wissen in den vier Themengebieten Mechanik, Elektrizitätslehre, Optik und Wärmelehre erfasst. Diese vier Inhaltsbereiche nehmen in ihrer Gesamtheit den größten Anteil am Physikcurriculum der Sekundarstufe I der verschiedenen Schularten ein (KMK, 2004; Lehrpläne bayerische Sekundarschulen¹). Die Elektrizitätslehre beansprucht dabei in etwa einen doppelt so hohen Anteil wie die Mechanik und diese wiederum in etwa einen doppelt so hohen wie die Optik oder die Wärmelehre (im FW-Test wurde versucht, diesem Verhältnis explizit Rechnung zu tragen). Auch bei den fachdidaktischen Items sind alle vier Themengebiete vertreten, um so die vielfach postulierten Zusammenhänge zwischen den physikalischen Inhaltsbereichen datenbasiert überprüfen zu können („Repräsentativitätshypothese“; vgl. Abschn. 6.5).

¹<http://www.isb.bayern.de/schulartspezifisches/lehrplan/>

Die dritte Dimension bezieht sich auf die Wissensarten. Die einzelnen Items des Testinstruments, vor allem im Bereich des FDW, sind so konzipiert, dass sie sowohl deklaratives als auch prozedurales sowie konditionales Wissen abbilden, da die Berücksichtigung aller drei Wissensarten vor allem zur Erfassung des FDW in den Naturwissenschaften mittlerweile weit verbreitet ist (in ProwiN für den Bereich Physik: z. B. Kirschner, 2013; für den Bereich Chemie: z. B. Dollny, 2011; für den Bereich Biologie: z. B. Jüttner & Neuhaus, 2013). In den Testitems werden die Studienteilnehmer dabei oft aufgefordert, ihre Antworten ausführlich zu erläutern beziehungsweise zu begründen. Da sowohl das prozedurale als auch das konditionale Wissen situations- und ablauforientiert organisiert ist („Handlungsschemata“, vgl. Blömeke, Felbrich & Müller, 2008; Krauss, Lindl, Schilcher, Fricke et al., 2017), sind viele der FDW-Items so formuliert, dass sie für den Physikunterricht typische und didaktisch „kritische“ Unterrichtssituationen beschreiben. Diese Wissensarten (Dimension 3) wurden jedoch – genauso wie auch die Inhalte (Dimension 2) – nicht systematisch variiert, sondern dienten lediglich als heuristisches Hintergrundmodell zur Itemkonstruktion. Im Ergebnisteil werden folglich vor allem Resultate in Bezug auf die Wissenskategorien FDW und FW sowie auf die Facetten des FDW (Dimension 1) berichtet (zur „Repräsentativitätshypothese“ bzgl. Dimension 2 siehe Abschn. 6.5).

Zur Erfassung des FDW werden im Testinstrument FALKO-P bei allen Items offene Antwortformate gewählt, zum einen, um nicht nur das „Wiedererkennen“ von Erklärungsstrategien beziehungsweise Schülerfehlern zu testen und zum anderen, um keine Lenkung durch vorgegebene Antwortalternativen zu verursachen (zur Verwendung verschiedener Antwortformate siehe z. B. Lienert & Raatz, 1998; Hill, Loewenberg Ball & Schilling, 2008; Bühner, 2011; Moosbrugger & Kelava, 2012). Aus testökonomischen Gründen – die Bearbeitung des gesamten Tests sollte unter Berücksichtigung der Zumutbarkeit für die Studienteilnehmer maximal 90 Minuten dauern (FDW: 60 Minuten, FW: 30 Minuten) – erfolgte die Erfassung des FW mit Hilfe einer Mischung aus offenen und geschlossenen Items. Hier war es oftmals möglich, geeignete Distraktoren zu formulieren, die zwar normativ falsch waren, sich aber als angemessen persuasiv für die Studienteilnehmer herausstellten. Diese Distraktoren wurden zum einen aus der Literatur, zum anderen aus Falschantworten der Studienteilnehmer in der Pilotstudie gewonnen. Im Verlauf einer Prä-Pilot- und einer Pilotstudie wurden aus einem Pool von insgesamt 116 konstruierten Items 32 (14 FDW, 18 FW) geeignet erscheinende identifiziert (für die Kriterien der Itemselektion siehe Abschn. 5.4), welche Eingang in die zu validierende Testversion fanden (Tab. 4.1). Alle 32 finalen Items finden sich in Anhang B (für die Kodieranweisungen bitte bei der Autorin nachfragen).

4.2. Konzeptualisierung und Operationalisierung: Fachdidaktisches Wissen, Fachwissen

4.2.1. Fachdidaktisches Wissen (FDW)

Das fachdidaktische Wissen wird bei FALKO-Physik in die beiden Standard-Subfacetten „Wissen über Erklären und Repräsentieren“ („E & R“), „Wissen über Schülerfehler“ („Schk“) sowie in eine zusätzliche fachspezifische Komponente, dem „Wissen über Messen und Experimentieren“ („M & E“), unterteilt. Dabei werden die vier Inhaltsbereiche Elektrizitätslehre, Mechanik, Optik und Wärmelehre abgebildet, wobei aus Gründen der Testökonomie nicht jede fachdidaktische Subfacette Items aller vier Bereiche enthält. Für

Tabelle 4.1.: Übersicht über die Anzahl und Verteilung der Items (sowie Antwortformat)

Facetten	Item- anzahl	(offen/ geschlossen)	Mechanik	Elektrizitätslehre	Optik	Wärmelehre
FW	18	(12/6)	4	8	3	3
FDW	14	(14/0)	2	7	2	3
<i>E & R</i>	5	(5/0)	1	4	0	0
<i>Schk</i>	5	(5/0)	1	1	1	2
<i>M & E</i>	4	(4/0)	0	2	1	1
Σ	32	(26/6)	6	15	5	6

einen Überblick über Itemzahl pro didaktischer Subfacette sowie deren Verteilung über die physikalischen Inhaltsbereiche, siehe Tabelle 4.1. Dabei dienen die drei konzeptualisierten Subfacetten lediglich als heuristische Grundlage für die Itemkonstruktion. Analysen auf Subfacetten-Ebene sind vorerst nicht angedacht.

Wissen über Erklären und Repräsentieren („E & R“)

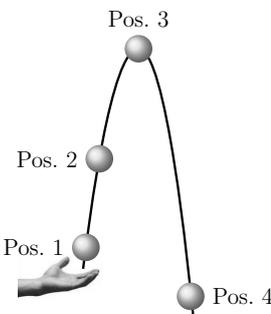
Lehrkräfte benötigen Erklärungswissen, um fachliche Inhalte in Bezug auf ihre Schülerschaft angemessen aufzubereiten und darzustellen (Kulgemeyer & Schecker, 2013; Schilcher, Krauss, Rincke & Hilbert, 2017). Eine Lehrkraft muss demnach Instruktionsstrategien adressatenorientiert auswählen können. Dabei macht es einen Unterschied, ob sie in ihrem Unterricht beispielsweise mit leistungsstärkeren oder leistungsschwächeren Schülern konfrontiert ist oder ob ein bestimmtes physikalisches Themengebiet in der 7. oder in der 9. Jahrgangsstufe unterrichtet werden soll. Da eine Lehrkraft in diesem Sinne über ein Repertoire an Erklärungsansätzen für bestimmte Inhaltsbereiche verfügen sollte, wurden die Studienteilnehmer bei einigen Items dazu aufgefordert, möglichst viele ihnen bekannte Instruktionsstrategien anzugeben. Zum Wissen über Erklären und Repräsentieren von physikalischen Inhalten wurden fünf offene Items entwickelt (für ein Beispielim und ein Kodierungsbeispiel siehe Tab. 4.2 bzw. 4.3).

Wissen über Schülerkognitionen („Schk“)

Unter das Wissen über Schülerkognitionen fallen Kenntnisse zu typischen Schüler(Fehl)vorstellungen wie zum Beispiel zum „Stromverbrauch“ im Bereich der Elektrizitätslehre (z. B. Rhöneck, 1986b) sowie zu kennzeichnenden Argumentationsweisen und Lernprozessen von Schülern. Diesbezügliche Kenntnisse sind eine Voraussetzung dafür, um typische Schülerfehler prognostizieren zu können (Olszewski, 2010) und um im Unterricht entsprechend präventiv und reaktiv handeln zu können. Weiterhin ist Wissen über Schüler(Fehl)vorstellungen notwendig zur Planung von Physikunterricht (Korneck, Lamprecht, Wodzinski & Schecker, 2010). Insgesamt wurden fünf offene Items zu typischen Schülerfehlern konstruiert (für ein Beispielim und ein Kodierungsbeispiel siehe Tab. 4.2 bzw. 4.3).

Tabelle 4.2.: Finale Beispielitems aus dem Testinstrument FALKO-P (skizzierte Lösungsmuster finden sich in Tabelle 4.3)

FDW = Fachdidaktisches Wissen; FW = Fachwissen;

Kategorie	Beispielitems
<p>FDW „Wissen über Instruktion“ Kurz: „E & R“</p>	<p>„Stromverbrauch“ In Ihrer einführenden Unterrichtsstunde zum Themengebiet Elektrizitätslehre erzählt ein Schüler von einem Ferienerlebnis: „Im Zeltlager habe ich nachts ein spannendes Comic-Heft gelesen. Mit der Zeit hat meine Taschenlampe immer schwächer geleuchtet. Nach zwei Stunden hat sie dann gar nicht mehr funktioniert, denn sie hat den ganzen Strom verbraucht!“ <i>Beschreiben Sie jeweils kurz und prägnant möglichst viele grundlegend unterschiedliche Möglichkeiten, um dieser „Stromverbrauchsvorstellung“ im weiteren Verlauf des Unterrichts entgegenzuwirken. Bitte nummerieren Sie Ihre Lösungsvorschläge fortlaufend durch.</i></p>
<p>FDW „Wissen über Schülerkognition“ Kurz: „Schk“</p>	<p>„Wärmeausbreitung“ Fragt man Schüler nach der Bedeutung des Begriffs „Wärme“, erhält man oftmals Aussagen wie die folgende: „Der gesamte Stab wird heiß, weil die Wärme sich in einem Teil staut, bis er sie nicht mehr halten kann und dann bewegt sie sich im Stab entlang.“ <i>Welche Schülervorstellung zum Wärmebegriff könnte hinter dieser Aussage stecken?</i></p>
<p>FDW „Wissen über Messen und Experimentieren“ Kurz: „M & E“</p>	<p>„Spiegelbild“ Viele Schüler haben die Vorstellung, dass das Spiegelbild eines Gegenstands auf dem Spiegel und nicht, wie es sich mit Hilfe des Verständnisses von Lichtausbreitung und -reflexion erklären lässt, hinter dem Spiegel entsteht, und zwar genau in demselben Abstand hinter der Spiegelebene, in dem sich ein Gegenstand vor der Spiegelebene befindet. <i>Beschreiben Sie kurz einen anschaulichen Schülerversuch (Material, Durchführung, Auswertung/Ergebnisdarstellung), mit dem Sie Ihren Schülern den Ort des Spiegelbilds verdeutlichen könnten. Bei Bedarf können Sie zusätzlich erläuternde Skizzen anfertigen.</i></p>
<p>FW</p>	<p>„Senkrechter Wurf“ Ein Ball wird mit der Hand aus der Ruhe senkrecht nach oben beschleunigt. In Position 1 hat der Ball gerade die Hand verlassen und mit Position 3 erreicht er seinen höchsten Punkt. Luftwiderstandskräfte spielen bei dem Vorgang keine Rolle (idealer Vorgang!). <i>Zeichnen Sie in das Bild für die Positionen 1 bis 4 jeweils alle auf den Ball wirkenden Kräfte ein!</i> <small>Die Verzerrung in horizontaler Richtung dient lediglich der besseren Übersicht!</small></p> 

Wissen über Messen und Experimentieren („M & E“)

Experimente sind zentral für den naturwissenschaftlichen Erkenntnisprozess (Scientific Inquiry „SI“, Nature of Science „NOS“, z. B. Kock, Taconis, Bolhuis & Gravemeijer, 2013). Wichtige Aspekte dieser fachdidaktischen Kompetenzfacette sind eine angemessene Auswahl von Experimenten, deren zielgerichtete Vor- und Nachbereitung sowie die didaktisch reflektierte Aufteilung der einzelnen Arbeitsschritte (Hofstein & Lunetta, 2004; Rincke, 2011). Nicht zuletzt findet sich die hervorgehobene Rolle von Experimenten im naturwissenschaftlichen Unterricht explizit auch in den Bildungsstandards zum Fach Physik für den Mittleren Schulabschluss (KMK, 2004) sowie in den Lehrplänen der einzelnen Schularten wieder (ISB: LehrplanPLUS Bayern²). Insgesamt wurden vier offene Items zu dieser Facette konstruiert (für ein Beispielitem und ein Kodierungsbeispiel siehe Tab. 4.2 bzw. 4.3).

Bei FALKO-Physik liegt der Fokus dabei auf der Auswahl eines geeigneten Experiments zur Darstellung eines bestimmten physikalischen Sachverhalts, auf der Identifikation von „Fehlern“ bei Aufbau und Durchführung von (Schüler-)Experimenten sowie der Korrektur dieser Fehler. Im Gegensatz dazu fasst Riese (2009) unter seine Facette „Experimente“ das Wissen über Arten, Funktionen und die Bedeutung von Experimenten im naturwissenschaftlichen Erkenntnisprozesses sowie das Wissen über das Vorgehen beim Experimentieren in der Unterrichtssituation. Das Wissen über Modelle (sowie deren Rolle für den Erkenntnisprozess und für den Unterricht) ist in FALKO-Physik in allen Facetten des FDW enthalten, während zum Beispiel bei ProwiN das Wissen über Modelle und Konzepte in einer eigenen Facette abgebildet wird (vgl. Tepner et al., 2012).

Zur Facette „M & E“ gilt es an dieser Stelle bereits zu erwähnen, dass eine konfirmatorische Faktorenanalyse indizierte, dass diese Facette – zumindest in der vorliegenden Operationalisierung – letztlich eher dem Fachwissen als dem fachdidaktischen Wissen zuzuordnen ist (für Details siehe Abschn. 6.2). Da aber auch auftretende Probleme im Laufe von Testkonstruktionen wertvolle Informationen darstellen, sind diese vier Items – auch wenn sie im Ergebnisteil zum Teil nicht in die Berechnungen zum FDW eingehen – ebenfalls in Anhang B abgebildet.

4.2.2. Fachwissen (FW)

Zur fachlich adäquaten Vorbereitung und Durchführung von Unterricht ist physikalisches Fachwissen erforderlich, das über das Niveau der zu unterrichtenden Schulstufen hinausgeht. Kognitiv aktivierender Unterricht kann nur dann gelingen, wenn Lehrkräfte ein fundiertes Verständnis fachlicher Konzepte und Zusammenhänge aufweisen (Woitkowski et al., 2011). Fachwissen ist weiterhin die Grundlage, um sicher und fachlich korrekt auf Fragen der Schüler einzugehen oder um Schülervorstellungen adäquat zu diagnostizieren. Auch das Erklärungsrepertoire ist augenscheinlich von der Breite und der Tiefe des verfügbaren Fachwissens abhängig (in diesem Sinne ist Fachwissen nicht nur eine Voraussetzung für didaktische Flexibilität, sondern für unterrichtliches Erklären überhaupt; siehe auch Baumert & Kunter, 2006; Braaten & Windschitl, 2011).

In Bezug auf gängige Unterscheidungen von Niveaustufen des Fachwissens wie zum Beispiel „Alltagswissen“, „Schulwissen“, „vertieftes Schulwissen“ oder „universitäres Wissen“ (z. B. Krauss et al., 2011) orientierte sich FALKO-Physik an Schulwissen und vertieftem Schulwissen. Sechs der insgesamt 18 Items des FW-Tests erfordern dabei vertieftes

²<https://www.isb.bayern.de/>

Tabelle 4.3.: Exemplarische Musterlösungen (Auswahl) zu den finalen Beispielitems aus dem Testinstrument FALKO-P (vgl. Tab. 4.2)

FDW = Fachdidaktisches Wissen; FW = Fachwissen;

Kategorie	Ausgewählte Musterlsg. zu den Bsp.-Items aus Tab. 4.2
FDW „Wissen über Instruktion“ Kurz: „E & R“	„Stromverbrauch“ <i>Experimentell, quantitativ: z. B.</i> – Betrag der elektrischen Stromstärke in einem Stromkreis „vor“ und „hinter“ einem Lämpchen messen → Betrag ist immer gleich. – Betrag der elektrischen Stromstärke an verschiedenen Stellen eines unverzweigten ($I_1 = I_2 = I_3 = \dots$) oder verzweigten Stromkreises ($I_{\text{ges}} = I_1 + I_2 + I_3 + \dots$) messen.
FDW „Wissen über Schülerekognition“ Kurz: „Schk“	„Wärmeausbreitung“ <i>Antworten, die der folgenden Vorstellung entsprechen:</i> „Wärmestoffvorstellung“: SuS vergleichen Wärme z. B. mit Stoffen wie Rauch, Dampf und häufig auch mit Luft; evtl. zusätzlich Argumentation mit „Kältestoff“; evtl. auch Flüssigkeit, also Substanzcharakter.
FDW „Wissen über Messen und Experimentieren“ Kurz: „M & E“	„Spiegelbild“ <i>Versuche, die eindeutig beschrieben sind mit „Bild wird tatsächlich „handelnd“ HINTER dem Spiegel verifiziert (z. B. Messen)“</i> Beispiele: – Kerze im Wasserglas – Wasser schütten – Entfernungsmesser (Spiegelreflexkamera, opt. Aufbau diesbzgl.)
FW	„Senkrechter Wurf“ Es wird bei allen vier Positionen lediglich F_g eingezeichnet; die Pfeile sind auch gleich lang; wird der Angriffspunkt nicht im Körperschwerpunkt gewählt, aber AM Körper, ist dies auch richtig zu werten.

Hintergrundwissen zu Themengebieten des Fachlehrplans der Sekundarstufe (wie beispielsweise auch bei MT21, TEDS-M, ProfiLe-P) und zwölf FW-Items wurden auf Schulniveau formuliert (teilweise am Übergang zum vertieften Schulwissen). Letztere Items könnten, zumindest prinzipiell, auch von sehr guten Schülern gelöst werden. Wie schon beim FDW-Testteil wurden auch zur Erfassung des FW Items zu allen vier Inhaltsbereichen Mechanik, Elektrizitätslehre, Optik und Wärmelehre (insgesamt 10 offene, 8 geschlossene; Tab. 4.1) konstruiert. Ein Beispielitem sowie ein Kodierungsbeispiel findet sich in Tabelle 4.2 bzw. Tabelle 4.3.

5. Testkonstruktion

Im Folgenden wird die Entwicklung des Testinstruments FALKO-P beschrieben. Dabei bildet die klassische Testtheorie (KTT) die Grundlage für die Konstruktion der Items und der zugehörigen Kodieranweisungen. Um dem Leser einen Einblick in den Ablauf der Konstruktionsarbeit zu vermitteln, werden in 5.1 der Ablauf der Testentwicklung sowie unterschiedliche Items und deren zugehörige Kodieranweisungen vorgestellt. Abschnitt 5.2 widmet sich im Überblick der Entwicklung des Kodiermanuals. Im Anschluss wird die Durchführung der Prä-Pilot- und der Pilotstudie detailliert beschrieben (Abschn. 5.3; aus konzeptuellen Gründen wurden die beiden Studien zeitlich und inhaltlich aufgesplittet; Näheres hierzu in Abschn. 5.3.1) sowie auf Kriterien der Itemselektion für das finale Testinstrument eingegangen (Abschn. 5.4).

Neben der Umsetzung des FALKO-einheitlichen Rahmenmodells hinsichtlich Konzeptualisierung und Operationalisierung der beiden Professionswissenskategorien FDW und FW war ein wesentliches Ziel der Testkonstruktion bei FALKO-P, sowohl das FDW als auch das FW anhand der zentralen physikalischen Themenbereiche Mechanik, Elektrizitätslehre, Optik und Wärmelehre abzubilden (vgl. hierzu auch Abschn. 4.1), um so auch bisherige Annahmen zum Zusammenhang dieser Wissensbereiche sowie zur (vermeintlichen) Repräsentativität der Mechanik überprüfen zu können (im Folgenden: „Repräsentativitätshypothese“; vgl. Abschn. 6.5).

Zur Erhebung des FDW wurden ausschließlich Items mit offenem Antwortformat konstruiert. Aus testökonomischen Gründen – die Bearbeitung des gesamten Tests sollte unter Berücksichtigung der Belastbarkeit der Studienteilnehmer maximal 90 Minuten dauern (FDW: 60 Minuten, FW: 30 Minuten) – erfolgte die Erfassung des FW mit Hilfe einer Mischung aus offenen und geschlossenen Items. Hier war es oftmals möglich, gute Distraktoren zu formulieren, welche zwar eindeutig falsch waren, sich aber als angemessen persuasiv für die Studienteilnehmer herausstellten (Distraktoren wurden aus der Literatur sowie durch Falschantworten der Studienteilnehmer in der Pilotierung gewonnen). Zusammenfassend gesehen kamen im Laufe der Testkonstruktionsphase insgesamt neun Testheftversionen zum Einsatz, wobei vier Versionen die erste Prä-Pilotierungs- und Pilotierungsrunde (Abschn. 5.3), zwei die zweite (Abschn. 5.3) und drei die Validierungsstudie (Kap. 6) betrafen. In der letzten (aktuellsten) Version wurde lediglich das Layout an manchen Stellen noch für den zukünftigen Einsatz optimiert.

5.1. Itemkonstruktion

Der Entwicklung der Items beider Wissensbereiche wurde neben den deutschen Bildungsstandards (KMK, 2004) und den (bayerischen) Schulcurricula¹ gängige fachdidaktische (z. B. Muckenfuß, 1995; Müller, Wodzinski & Hopf, 2011; Labudde, 2010) und fachwissenschaftliche Literatur (z. B. Tipler, 1994; Dorn & Bader, 2002; Grehn & Krause, 2003; Grehn & Krause, 2005; Epstein & Lessing, 2011) zugrunde gelegt. Weiterhin

¹<http://www.isb.bayern.de/schulartspezifisches/lehrplan/>

Tabelle 5.1.: Übersicht über die unterschiedlichen Testheftversionen inkl. Testitems

<i>Version</i>	<i>Themengebiete</i>	<i>Itemanzahl und -format</i>	
1.0 (Herbst 2013)	Elektrizitätslehre (e)	<i>FDW</i> : 22 (offen) – 6 „E & R“ – 10 „Schk“ – 6 „M & E“	<i>FW</i> : 46, davon – 7 geschlossen – 12 Rechnungen – 27 Frei(text)ergänzungen/offen
1.1 (Winter 2013, Frühjahr 2014)	E-Lehre (e)	<i>FDW</i> : 15 (offen) – 7 „E & R“ – 3 „Schk“ – 5 „M & E“	<i>FW</i> : 42, davon – 7 geschlossen – 8 Rechnungen – 27 Frei(text)erg./off.
2.2 (Frühjahr 2014)	E-Lehre (e)	<i>FDW</i> : 13 (offen) – 6 „E & R“ – 3 „Schk“ – 4 „M & E“	<i>FW</i> : 26, davon – 5 geschlossen – 3 Rechnungen – 18 Frei(text)erg./off.
2.3 (Sommer 2014)	E-Lehre (e)	<i>FDW</i> : 15 (offen) – 6 „E & R“ – 4 „Schk“ – 5 „M & E“	<i>FW</i> : 19, davon – 4 geschlossen – 3 Rechnungen – 12 Frei(text)erg./off.
3.0 (Herbst 2014)	Mechanik (m) Optik (o) Wärmelehre (w)	<i>FDW</i> : 15 (offen) (6 m, 4 o, 5 w) – 8 „E & R“ – 2 „Schk“ – 5 „M & E“	<i>FW</i> : 31, davon – 13 geschlossen (5 m, 4 o, 4 w) – 18 Frei(text)erg./off. (8 m, 3 o, 7 w)
3.2 (Winter 2014, Frühjahr 2015)	Mechanik (m) Optik (o) Wärmelehre (w)	<i>FDW</i> : 13 (offen) (4 m, 4 o, 5 w) – 6 „E & R“ – 4 „Schk“ – 3 „M & E“	<i>FW</i> : 19, davon – 13 geschlossen (5 m, 4 o, 4 w) – 6 Frei(text)erg./off. (2 m, 1 o, 3 w)
4.0a/b (Sommer 2015)	Mechanik (m) E-lehre (e) Optik (o) Wärmelehre (w)	<i>FDW</i> : 17 (offen) (4 m, 8 e, 3 o, 3 w) – 7 „E & R“ – 5 „Schk“ – 5 „M & E“	<i>FW</i> : 20, davon – 7 geschlossen (2 m, 1 e, 2 o, 2 w) – 13 Frei(text)erg./off. (3 m, 7 e, 1 o, 2 w)
4.2 (Herbst 2015 bis Sommer 2016)	Mechanik (m) E-lehre (e) Optik (o) Wärmelehre (w)	<i>FDW</i> : 14 (offen) (2 m, 7 e, 3 o, 3 w) – 5 „E & R“ – 5 „Schk“ – 4 „M & E“	<i>FW</i> : 18, davon – 6 geschlossen (2 m, 1 e, 2 o, 1 w) – 12 Frei(text)erg./off. (2 m, 7 e, 1 o, 2 w)
4.3 (Herbst und Winter 2016) Layoutänderungen	Mechanik (m) E-lehre (e) Optik (o) Wärmelehre (w)	<i>FDW</i> : 14 (offen) (2 m, 7 e, 3 o, 3 w) – 5 „E & R“ – 5 „Schk“ – 4 „M & E“	<i>FW</i> : 18, davon – 6 geschlossen (2 m, 1 e, 2 o, 1 w) – 12 Frei(text)erg./off. (2 m, 7 e, 1 o, 2 w)

Tabelle 5.2.: Formulierung des Items „Stromverbrauch“

In Ihrer einführenden Unterrichtsstunde zum Themengebiet Elektrizitätslehre erzählt ein Schüler von einem Ferienerlebnis:

„Im Zeltlager habe ich nachts ein spannendes Comic-Heft gelesen. Mit der Zeit hat meine Taschenlampe immer schwächer geleuchtet. Nach zwei Stunden hat sie dann gar nicht mehr funktioniert, denn sie hat den ganzen Strom verbraucht!“

Beschreiben Sie jeweils kurz und prägnant möglichst viele grundlegend unterschiedliche Möglichkeiten, um dieser „Stromverbrauchsvorstellung“ im weiteren Verlauf des Unterrichts entgegenzuwirken.

Bitte nummerieren Sie Ihre Lösungsvorschläge fortlaufend durch.

konnte auf einige Items aus der Projektvorbereitung von Anja Göhring und Josef Reisinger zurückgegriffen werden (diese Items wurden im finalen Testheft mit einer entsprechenden Quellenangabe versehen). Final fanden die aus der Prä-Pilotstudie (siehe Abschn. 5.3) gewonnenen Erkenntnisse (z. B. Expertenfeedback) Berücksichtigung.

Im Testinstrument FALKO-P werden zur Erfassung des Professionswissens im Fach Physik überwiegend offene Antwortformate gewählt. Im Gegensatz zu offenen Antwortformaten, ist die Beantwortung geschlossener Items zwar zeitökonomischer, wodurch prinzipiell ein „Mehr“ an Wissen erhoben werden kann. Jedoch beeinflussen vorgegebene Lösungsalternativen das Antwortverhalten, da nicht selbstständig aktiv nach möglichen Lösungen gesucht, sondern lediglich aus vorgegebenen Optionen gewählt werden muss. Offene Antwortformate bieten dann vor allem den Vorteil, dass das den Studienteilnehmern „eigene“ Wissen erfasst werden kann, ohne ihr Antwortverhalten gewissermaßen zu „lenken“ („Aha-Effekt“, vgl. z. B. Hill et al., 2008; vgl. auch Abschn. 4.1). Allerdings ergeben sich womöglich Einschränkungen hinsichtlich der Auswertungsobjektivität, da die Lösungen der Studienteilnehmer einem entwickelten Kategoriensystem folgend interpretiert und bewertet werden müssen (für eine Diskussion von Vor- und Nachteilen von Items mit offenem beziehungsweise geschlossenem Antwortformat siehe z. B. Moosbrugger & Kelava, 2012; Bühner, 2011).

Der Vorteil eines offenen Antwortformats, vor allem bei der Erfassung des FDW, soll anhand des folgenden Itemkontextes (Item „Stromverbrauch“, Tab. 5.2) illustriert werden.

Itembeispiel „Stromverbrauch“

Im Item wird berichtet, dass ein Schüler von der „Stromverbrauchsvorstellung“ überzeugt ist. Die Studienteilnehmer werden in einem offenen Antwortformat aufgefordert, substantiell unterschiedliche Strategien zu nennen und zu beschreiben, die genutzt werden könnten, um bei diesem Schüler einen Conceptual Change anzubahnen.

Dies könnte beispielsweise mit Hilfe eines Experiments zum Nachweis der Stromstärkeerhaltung vor und nach einem Lämpchen innerhalb eines Stromkreises geschehen oder durch Verwendung verdeutlichender Analogien zur Erklärung der Stromstärkeerhaltung insgesamt (z. B. das Fahrradketten- oder Wasserkreislaufmodell). Im Regelfall reicht eine einzelne Strategie aber nicht aus, um eine Änderung von Schülervorstellungen hervorzurufen, unter anderem da das Wort „Stromverbrauch“ im Alltag natürlich häufig verwendet wird und man ja offensichtlich auch „Stromrechnungen“ für diesen „Verbrauch“ bezahlen muss - korrekterweise müsste man in diesem Fall von „Energierrechnung“ sprechen (für Strategien zum Umgang mit Schülervorstellungen allgemein und zum Themenbereich Elektrizitäts-

lehre siehe z. B. Jung, 1986; Rhöneck, 1986a; Rhöneck, 1986b). Eine Lehrkraft muss vielmehr über mehrere dieser Strategien verfügen (auch schon, um adressatengerecht auf einzelne Schüler eingehen zu können). Hier wäre es sicher kontraproduktiv, diese unterschiedlichen Strategien vorzugeben, denn dadurch käme es wohl eher zu einer Messung der „Bewertungskompetenz“ der Studienteilnehmer. Durch die Instruktion „[...] Beschreiben Sie jeweils kurz und prägnant möglichst viele grundlegend unterschiedliche Möglichkeiten, um dieser „Stromverbrauchsvorstellung“ im weiteren Verlauf des Unterrichts entgegenzuwirken [...]“ wurde stattdessen versucht, die Breite und Tiefe des Wissens der Studienteilnehmer in angemessener Weise abzubilden, ohne bereits Optionen vorzugeben. Der zugehörigen Kodieranweisung (Tab. 5.3), welche während der Entwicklungsphase fortlaufend optimiert wurde, kann der finale Bewertungshorizont für dieses Item entnommen werden.

An einem weiteren Beispiel (Item „Steckbrett“; siehe Tab. 5.4) wird nun exemplarisch der Entwicklungsprozess im Verlauf der vier Testheftversionen (Prä-Pilot 1 und Pilot 1) beschrieben:

Itembeispiel „Steckbrett“

In einem ersten Schritt wurden die Studienteilnehmer in einem offen Antwortformat aufgefordert, welche (allgemeinen) Aspekte eine Lehrkraft vor dem ersten Experimentieren mit Steckbrettern mit der Schülerschaft abklären müsse (Version 1.0, Tab. 5.4): „In einer Unterrichtsstunde Ihrer Einsatzreferendarin zur Elektrizitätslehre sollen die Schüler das erste Mal selbst mit Hilfe von Netzwürfeln ...“. Die Experten und erfahrenen Lehrkräfte aus der Prä-Pilotstudie waren sich einig, dass hier eine unüberschaubar große Anzahl an Antworten denkbar und somit eine adäquate Bearbeitung des Items schwer möglich wäre. Es folgte der Versuch einer Eingrenzung, bezogen auf den konkreten Inhalt der Unterrichtsstunde (Version 1.1, Tab. 5.4): „In einer Ihrer Unterrichtsstunden zu den unterschiedlichen Arten von Schaltungen sollen Ihre Schüler das erste Mal selbst mit Hilfe von ...“. Diese Änderung befanden die Experten als noch nicht hinreichend und mahnten eine weitere Konkretisierung hinsichtlich des Bearbeitungsauftrags an. Letztendlich wurde der Inhalt des Items derart eingegrenzt, dass die Studienteilnehmer Fehler in einem aufgebauten Schülerversuch identifizieren müssen (Versionen 2.2, 2.3, Tab. 5.4): „[...] Beurteilen Sie den von den Schülern entworfenen Schaltplan vergleichend mit der aufgebauten Schaltung und notieren Sie enthaltene Fehler.“ Das Erkennen und Korrigieren fehlerhafter Versuchsaufbauten innerhalb handlungsorientierten Unterrichts stellt eine alltägliche Aufgabe einer Physiklehrkraft dar, weshalb dieses Item einer realen Unterrichtssituation sehr nahe kommt. Gleichzeitig wurden erklärende Abbildungen eingefügt, welche in einem letzten Schritt (Versionen 4.0a/b, 4.2, 4.3, Tab. 5.4) noch vergrößert wurden, um die Studienteilnehmer bei der Beantwortung des Items adäquat unterstützen zu können. Parallel zur Aufgabenmodifikation erfolgten entsprechende Optimierungen beziehungsweise Präzisierungen im Kodiermanual (die finale Kodieranweisung für dieses Item ist in Tab. 5.5 ersichtlich).

5.2. Kodiermanual

Zur Kodierung der offenen Items beider Testteile wurden parallel zur Itemkonstruktion Kodieranweisungen entwickelt sowie Rater zur Auswertung der Antworten der Studienteilnehmer geschult. Die Kategorien sowie Kategoriengrenzen wurden zum einen in Anlehnung an fachdidaktische Literatur (vgl. Abschn. 5.1), zum anderen mit Hilfe des Feedbacks von Fachdidaktikern sowie erfahrenen Lehrkräften festgelegt. Während die Antworten bei

Tabelle 5.3.: Exemplarische Kodieranweisung (gekürzter Auszug aus dem Kodiermanual) zum Beispiel-Item „Stromverbrauch“

Codierung	Gesamtscore: 4	
Hier werden die Scores der einzelnen Antworten aufsummiert → hier also max. 4 Scores möglich. Pro Möglichkeit gibt es eine Variable, so dass bei jeder Code 0 oder 1 vergeben werden kann.		
Antworten	Antwort-	Score
	code	
Missing (gar nichts, Strich, „?“ , ...)	999	0
Aufgabe in Testheftversion nicht vorhanden.	88	0
Nicht zuzuordnen („keine Lust“, „keine Zeit“, „nicht Schulstoff“, ...)	97	0
Falsche/unverständliche/unvollständige Lösung, Ausführungen nicht lesbar, Aufgabe missverstanden (Lehrer, die die Frage missverstehen, z. B. aufschreiben, welche Fehler ihre Schüler machen würden, ...)	0	0
Code 1: Experimentell, qualitativ, z. B.	1	1
Reihenschaltung aus mehreren baugl. Lämpchen → Helligkeit aller Lämpchen ist identisch. <i>Ergänzungen aus Testheften:</i> „volle und leere Batterien wiegen“; „Ladung“ von Batterien „messen“		
Code 2: Experimentell, quantitativ, z. B.	1	1
Betrag der el. Stromstärke in einem Stromkreis „vor“ und „hinter“ einem Lämpchen messen → Betrag ist immer gleich. <i>Ergänzungen aus Testheften:</i> Betrag der el. Str.strk. an versch. Stellen eines unverzw. ($I_1 = I_2 = I_3 = \dots$) oder verzw. Stromkreises ($I_{\text{ges}} = I_1 + I_2 + I_3 + \dots$) messen.		
Code 3: Analogien, Modelle, z. B.	1	1
– Fahrradketten-, Riementransmissions-, Wasser(kreislauf)modell → Kettenglieder (e^-) bzw. Wasserteilchen (e^-) verschwinden nicht. – Förderbandmodell → e^- (z. B. Pinguine, Kugeln, Skifahrer, ...) erhalten Energie beim „Hochtransport“ und geben diese beim „Herunterrollen“ ab (an einen Energiewandler), wandern aber selbst „im Kreis“.		

Fortsetzung auf der nächsten Seite

Tabelle 5.3.: (Fortsetzung von S. 39)

-
- Schülermodell, z. B. Turnhalle → Schüler als „Elektronen“ durchlaufen einen Parcours, in dem sie Energie, z. B. in Form von Bällen an einer Stelle des „Stromkreises“ bekommen (Energiewandler 1) und an einer anderen Stelle wieder (an einen zweiten Energiewandler abgeben) → auch Bienenmodell, Energiehutmodell, Kraftprotzmodell.
 - „Nicht endliche“ Spannungsquellen einsetzen, z. B. Dynamot, so dass ein direkter Zusammenhang zwischen investierter mechanischer Energie und der umgewandelten el. Energie deutlich wird → Weg der Energie nachvollziehen; el. Strom als „Transportmittel“ el. Energie.
-

Code 4: Allgemeine Erklärungen, z. B. „Strategie Umdeuten“, z. B. 1 1

- anhand der Abgrenzung der Begriffe „el. Strom“ und „el. Energie“.
 - anhand Erklärungen zu Alltagsbeispielen.
-

den fachwissenschaftlichen Items relativ einfach als „richtig“ oder „falsch“ kodiert werden können (sollten), gibt es bei den fachdidaktischen Items oftmals keine eindeutig „richtigen“ oder „falschen“ Lösungen (sondern eher didaktisch „geeignete“ und „ungeeignete“ Antworten). „Korrekte“ Lösungen beim Fachdidaktik-Test zeichnen sich vielmehr durch eine hohe Akzeptanz beziehungsweise Übereinstimmung seitens der Testentwickler und der befragten Experten aus. Durch zahlreiche Präzisierungen des Kodiermanuals konnten für den Fachdidaktik-Test jedoch zum Fachwissenstest vergleichbare Interraterübereinstimmungen erzielt werden (vgl. Tab. A.2 in Anhang A; dies ist auch bei den anderen FALKO-Teilprojekten der Fall, vgl. Krauss, Lindl, Schilcher, Fricke et al., 2017). Durch die Antworten der Studienteilnehmer in der Pilotierungsphase wurden die Kategorien noch mit Anker- und Negativbeispielen ergänzt, so dass eine Zuordnung von neuen Antworten vereinfacht wurde. Weiterhin fand ein regelmäßiger Austausch zwischen den Kodierern statt, der zum Ziel hatte, Redundanzen und Überspezifizierungen aus dem Kodiermanual zu entfernen sowie die Kategoriengrenzen noch weiter auszuscharfen. Items, bei denen sich trotz einer Präzisierung der Kodieranweisung keine angemessene Übereinstimmung der Rater zeigte, wurden entfernt. Tabelle 5.3 gibt anhand des Beispielitems „Stromverbrauch“ einen Einblick in das finale Kodiermanual.

Insgesamt wurden bei der endgültigen Kodierung (Validierungsphase) 60 der insgesamt 270 Testhefte von zwei unabhängigen geschulten Ratern doppelt kodiert. Bei abweichendem Urteil einigten sich beide Rater im Diskurs auf einen Score (Rater-Agreement; Werte für die Interraterreliabilität siehe A.2 in Anhang A). Letztendlich bleibt noch festzuhalten, dass die Entwicklung des finalen Kodiermanuals stetigen Wachstums- und Änderungsprozessen unterworfen war.

5.3. Prä-Pilot- und Pilotstudie

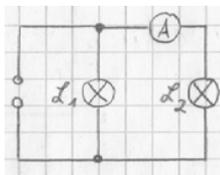
Die gesamte Prä-Pilotierungs und Pilotierungsphase, welche von Januar 2014 bis Februar 2015 andauerte (vgl. Abb. 5.1), teilte sich themenspezifisch in je eine Prä-Pilot- und eine Pilotstudie auf. Während innerhalb der Prä-Pilot- und Pilotstudie 1 Items zum Inhaltsbereich Elektrizitätslehre entwickelt und getestet wurden, geschah dies innerhalb

Tabelle 5.4.: Entwicklung des Testitems „Steckbrett“ über die einzelnen Testheftversionen hinweg

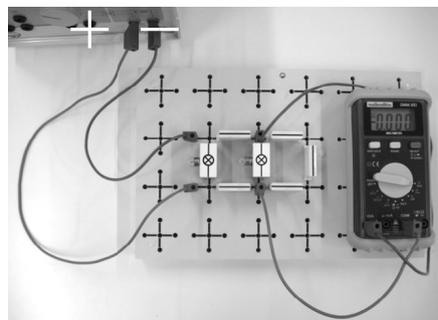
Vers.	Itemformulierungen
1.0	In einer Unterrichtsstunde Ihrer Einsatzreferendarin zur Elektrizitätslehre sollen die Schüler das erste Mal selbst mit Hilfe von Netzwürfeln, Steckbrettern und entsprechendem Kabelmaterial elektrische Schaltungen aufbauen und mit einem Vielfachmessgerät die elektrische Stromstärke messen. Welche Aspekte muss die Referendarin mit den Schülern hinsichtlich Aufbau der Schaltungen und Durchführung der Messungen thematisieren und klären, bevor sie anfangen zu experimentieren?
1.1	In einer Ihrer Unterrichtsstunden zu den unterschiedlichen Arten von Schaltungen sollen Ihre Schüler das erste Mal selbst mit Hilfe von <i>Netzgeräten, Steckbrettern, Glühlämpchen</i> und entsprechendem Kabelmaterial einfache elektrische Schaltungen aufbauen und damit einem <i>analogen Vielfachmessgerät</i> die Beträge der <i>elektrischen Stromstärke</i> messen. Welche Aspekte muss die Referendarin mit den Schülern hinsichtlich Aufbau der Schaltungen und Durchführung der Messungen thematisieren und klären, bevor sie anfangen zu experimentieren?
2.2	Sie führen mit Ihrer Klasse eine Schülerübung zum Bau von einfachen Stromkreisen mit Hilfe von Netzgeräten, Steckbrettern, Lämpchen, Kabelmaterial und Vielfachmessgeräten durch. Unter anderem soll in einer Parallelschaltung aus zwei Lämpchen die elektrische Stromstärke in den einzelnen Zweigen gemessen werden. Hierzu sollen die Schüler einen geeigneten Schaltplan erstellen, nach diesem die Schaltung aufbauen, diese in Betrieb nehmen und die Messungen durchführen. Sie als Lehrkraft kontrollieren die Schaltpläne und Aufbauten der Schüler vor deren Inbetriebnahme und können bei einer der Schülergruppen Folgendes beobachten:
2.3	Ausschnitt aus dem Versuchsprotokoll

Versuch 2
Messung der Stromstärke durch L_2 .

Schaltplan



Aufbau im Überblick



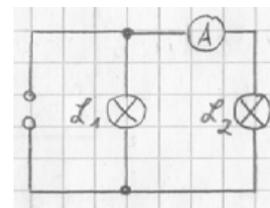
Vergrößertes Messgerät



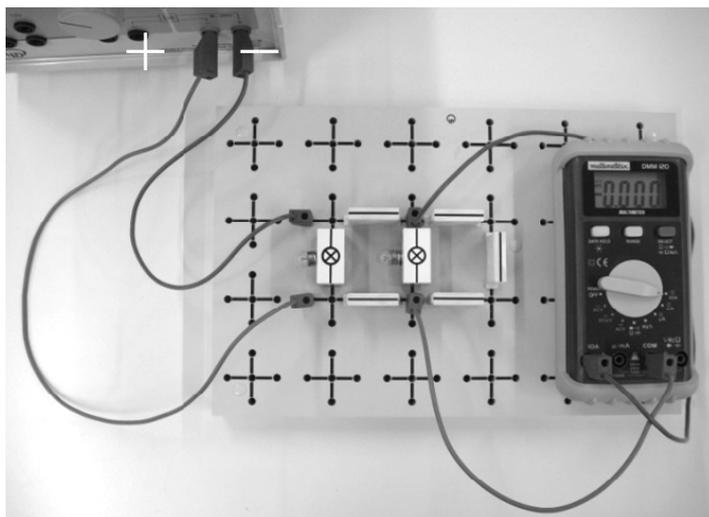
Tabelle 5.4.: (Fortsetzung von S. 41)

	<i>Beurteilen Sie den von den Schülern entworfenen Schaltplan vergleichend mit der aufgebauten Schaltung und notieren Sie enthaltene Fehler.</i>
4.0a/b	Sie führen mit Ihrer Klasse eine Schülerübung zum Bau von einfachen Stromkreisen mit Hilfe von Netzgeräten, Steckbrettern, Lämpchen, Kabelmaterial und
4.2	Vielfachmessgeräten durch. Unter anderem soll in einer Parallelschaltung aus zwei Lämpchen die elektrische Stromstärke in den einzelnen Zweigen gemessen werden. Hierzu sollen die Schüler einen geeigneten Schaltplan erstellen, nach diesem die Schaltung aufbauen, sie in Betrieb nehmen und die Messungen durchführen. Sie als Lehrkraft kontrollieren die Schaltpläne und Aufbauten vor deren Inbetriebnahme und können bei einer der Schülergruppen Folgendes beobachten:
4.3	Ausschnitt aus dem Versuchsprotokoll

*Versuch 2.
Messung der Stromstärke durch L₂.*



Aufbau im Überblick



Vergrößertes Messgerät



Beurteilen Sie den von den Schülern entworfenen (korrekten!) Schaltplan vergleichend mit der aufgebauten Schaltung und notieren Sie enthaltene Fehler.

Abbildung 5.1.: Zeitlicher Ablauf der Pilotstudien 1 und 2



Tabelle 5.5.: Finale Kodieranweisung zum Item „Steckbrett“

Codierung	Gesamtscore: 4	
<p>Hier werden die Scores der einzelnen Antworten aufsummiert → hier also max. 3 Scores möglich. Hier gibt es pro eingebautem Fehler je eine Variable, so dass bei jeder Variable die „1“ oder die „0“ oder ... verteilt werden kann, wenn der Fehler „entdeckt“ wurde oder eben nicht</p>		
Antworten	Antwort-	Score
	code	
Missing (gar nichts, Strich, „?“ , ...)	999	0
Aufgabe in Testheftversion nicht vorhanden.	88	0
Nicht zuzuordnen („keine Lust“, „keine Zeit“, „nicht Schulstoff“, ...)	97	0
Falsche/unverständliche/unvollständige Lösung, Ausführungen nicht lesbar, Aufgabe missverstanden (Lehrer, die die Frage missverstehen, z. B. aufschreiben, welche Fehler ihre Schüler machen würden, ...)	0	0
Code 1	1	1
<p>Fehler 1: Die Parallelschaltung enthält einen Kurzschluss ganz rechts; beide Lämpchen wären demnach überbrückt. Ebenfalls müsste man einen Kurzschluss über das Strommessgerät annehmen → mind. einer der beiden Kurzschlüsse muss genannt sein → es muss über Kurzschluss und nicht nur über „unnötige Schlaufe“ argumentiert werden</p>		
Code 2	1	1
<p>Fehler 2: Das Messgerät ist nicht in Reihe (zur Messung der Stromstärke), sondern parallel zum 2. Lämpchen geschaltet, so dass man eigentlich die Spannung messen würde.</p>		
Code 3	1	1
<p>Fehler 3: Die Kabel sind am Messgerät bei „A“ und nicht bei „mA“ eingesteckt und/oder Messbereich zu hoch gewählt.</p>		

der Prä-Pilot- und Pilotstudie 2 zu den Inhaltsbereichen Mechanik, Optik und Wärmelehre.

Die Trennung in die beiden Prä-Pilotstudien und Pilotstudien ist konzeptionellen Gründen geschuldet: Zu Projektbeginn war angedacht, ein Testinstrument mit Fokus auf die Elektrizitätslehre als alleinigen Inhaltsbereich zu entwickeln, da dieses Themengebiet mit einer überproportional hohen Präsenz in den Curricula der (bayerischen) Sekundarschularten vertreten ist und somit seitens der Testkonstrukteure eine hohe prädiktive Validität dieses Inhaltsbereichs für Physikwissen allgemein vermutet wurde. Der gemeinsamen Projektkonzeption von FALKO folgend, welche vorsieht, das Professionswissen in der jeweiligen Disziplin möglichst breit abzubilden, wurden aber die anderen drei genannten Inhaltsbereiche, welche ebenfalls einen bemerkenswerten Anteil am Curriculum einnehmen (vgl. Abschn. 4.1), ergänzt, mit dem Ziel, fachspezifisches und nicht themenspezifisches Physikwissen abzubilden sowie bisherige Annahmen zum Zusammenhang der operationalisierten Wissensbereiche sowie zur (vermeintlichen) Repräsentativität der Mechanik überprüfen zu können („Repräsentativitätshypothese“, vgl. Abschn. 6.5). Während der Präpilotierungs- und Pilotierungsphase wurden der überwiegende Anteil der Testhefte von zwei geschulten unabhängigen Ratern doppelt kodiert.

5.3.1. Stichprobe und Verlauf

Prä-Pilot-Studie 1 und 2

Im Verlauf zweier Prä-Pilotierungsrunden (Prä-Pilot 1: Elektrizitätslehre, Prä-Pilot 2: Mechanik, Optik, Wärmelehre) wurden die Items - nach deren kritischer inhaltlicher Analyse und Diskussion innerhalb der Projektgruppe FALKO und FALKO-Physik (z. T. wurden Items bereits zu diesem Zeitpunkt für ungeeignet befunden und vor der eigentlichen Prä-Pilotierung entfernt) - Fachdidaktikern (Experten) verschiedener Universitäten (Prä-Pilot 1: $N = 3$, Prä-Pilot 2: $N = 10$) sowie erfahrenen Lehrkräften der Schularten Realschule (Prä-Pilot 1: $N = 6$, Prä-Pilot 2: $N = 1$) und Mittelschule (Prä-Pilot 1: $N = 2$, Prä-Pilot 2: $N = 2$) zur Bearbeitung und zur Bestimmung der inhaltlichen Validität (Augenscheinvalidität) mittels Ratingskalen vorgelegt (die Geschlechtsverteilung, Altersstruktur sowie die Anzahl der Berufsjahre sind den Tab. A.13 und A.14 in Anhang A zu entnehmen; der größte Teil dieser Stichprobe nahm im weiteren Verlauf an der Pilotierungsphase teil, so dass die konkreten Werte für die erhaltenen Augenscheinvaliditäten erst im Zusammenhang mit der Pilotstudie berichtet werden).

Die Augenscheinvalidität wurde mit Hilfe von Tabelle 5.6 erhoben. Auf einer vierstufigen Skala (1 = „trifft nicht zu“ bis 4 = „trifft genau zu“) schätzten die Studienteilnehmer zum einen die Eindeutigkeit des jeweiligen Items ein. Als zweites wurde nach der Berufsrelevanz der Iteminhalte gefragt. In einem dritten Schritt sollte noch angegeben werden, inwieweit die Studienteilnehmer es für sinnvoll hielten, den jeweiligen Iteminhalt bereits während des Studiums beziehungsweise Referendariats zu thematisieren. Items, deren Eindeutigkeit und Berufsrelevanz im Mittel sowie nach mündlicher Auskunft („lautes Denken“, siehe unten) am niedrigsten eingeschätzt wurde, wurden überarbeitet oder fanden keinen Eingang in die nächste Testheftversion (die Mittelwerte der erhobenen Augenscheinvaliditäten sind den beiden Tab. A.4 und A.5 in Anhang A zu entnehmen).

Einige Studienteilnehmer sollten bei der Testbearbeitung auch „laut denken“ und wurden dabei videografiert. Das erhaltene Feedback und die schriftlichen beziehungsweise zum Teil auch mündlichen Antwortvorschläge wurden dazu genutzt, die Itemformulierungen zu überarbeiten, ungeeignet erscheinende Items zu identifizieren, um diese zu modifizieren oder zu entfernen, und, zusätzlich zu den in der Literatur recherchierten Lösungsvorschlägen,

Tabelle 5.6.: Tabelle zur Erhebung der Augenscheinvalidität: 1 „Trifft nicht zu“, 2 „Trifft kaum zu“, 3 „Trifft eher zu“, 4 „Trifft genau zu“

Bitte kreuzen Sie an: Diese Aufgabe/Inhalte	Trifft nicht zu	Trifft kaum zu	Trifft eher zu	Trifft genau zu
... ist/sind eindeutig gestellt.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
... beinhaltet berufsrelevantes Wissen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
... sollten im Lehramtsstudium/Referendariat berücksichtigt werden.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Tabelle 5.7.: Beispiel-Item „Kochplatte“ zum Themenbereich Elektrizitätslehre (wurde im Zuge der Pilotierung modifiziert)

Itemformat	Beispiel-Item „Kochplatte“
Rechenaufgabe	Eine elektrische Kochplatte wird von zwei Spiralen geheizt, die einzeln, in Reihe oder parallel geschaltet werden können (4 Stufen). Bei einer Spannung von 230 V ist die niedrigste Heizleistung 200 W und die höchste 2000 W. Welche Widerstände haben die Heizspiralen? <i>Ihr Lösungsweg muss klar nachvollziehbar sein!</i>
Qualitative Aufgabe	Eine elektrische Kochplatte mit einem Drehschalter enthält zwei unterschiedliche Heizspiralen. Durch unterschiedliche Schaltungen dieser lassen sich vier verschiedene (Heiz)Stufen realisieren: Jede der Heizspiralen dieser einen Kochplatte im Einzelbetrieb, beide in Reihe oder beide parallel geschaltet. <i>Begründen Sie qualitativ knapp, mit welcher der genannten Schaltungen sich die höchste (Heiz)Stufe realisieren lässt.</i>

das Kodiermanual zu erstellen beziehungsweise zu optimieren.

Im Zuge der ersten Testmodifikation (Version 1.0 auf Version 1.1; vgl. Tab. 5.1) fand auch die erste Itemreduktion (FDW: 22 auf 15, FW: 46 auf 42) statt. Items, deren Inhalte nicht in den Physik-Curricula aller untersuchten Sekundarstufenschularten (z. B. Themengebiet „elektrische Felder“) verankert sind, wurden entfernt. Innerhalb einer zweiten Modifikation (Version 1.1 auf 2.2; vgl. Tab. 5.1) wurden weitere Items entfernt, welche von den Experten als wenig eindeutig und berufsrelevant eingeschätzt wurden beziehungsweise Inhalte aus fachlicher Sicht redundant innerhalb unterschiedlicher Aufgabenstellungen beleuchteten. Einige Berechnungsitems wurden in qualitativ zu beantwortende Formate umgewandelt, da ja mit FALKO-P primär physikalisches und nicht mathematisches Wissen erfasst werden soll. Dies bedeutet, dass bei einem solchen Item (FW: Item „Kochplatte“, Tab. 5.7) beispielsweise nicht mehr die exakten Widerstandsbeträge der Heizspiralen eines Elektroherds berechnet werden müssen, statt dessen aber argumentiert werden soll, mit welcher Art von Schaltung (Reihenschaltung, Parallelschaltung, Einzelbetrieb) die höchste Heizstufe erreicht würde.

Einige der verbliebenen Items wurden zum besseren Verständnis um Abbildungen ergänzt

(z. B. Item „Steckbrett“, Tab. 5.4). Bereits bestehende Abbildungen wurden mit einem kurzen erklärenden Text versehen (z. B. Item „senkrechter Wurf“, Tab. 4.2) oder es wurden zusätzlich zur Gesamtabbildung vergrößerte Abbildungsausschnitte hinzugefügt (z. B. Item „Steckbrett“, Tab. 5.4). Bezogen auf den Pool an Elektrizitätslehreaufgaben (Prä-Pilot 1) erfolgte schließlich eine Verkleinerung der Itemanzahl beim FDW von 22 auf 13 und beim FW von 46 auf 26. Für den Fall der 2. Prä-Pilotstudie sank die Zahl der fachdidaktischen Items von 15 auf 13 und die der Fachwissensitems von 31 auf 19 (vgl. Tab. 5.1).

Pilot-Studie 1 und 2

Um in einem nächsten Schritt Aussagen über Itemschwierigkeiten, -trennschärfen, Interraterreliabilitäten und interne Konsistenzen zu erhalten, wurden die Items innerhalb der Pilotierungsphase an jeweils einer Stichprobe von Lehrkräften unterschiedlicher Sekundar-schularten, Lehramtsstudierenden sowie Universitätsmitarbeitenden pilotiert. Aufgrund anschließender statistischer Analysen geeignet erscheinende Items fanden Eingang in die Testendversion (vgl. Abschn. 5.4 und Anhang B). Insgesamt teilte sich die Pilotstudie entsprechend der Prä-Pilotstudie in zwei Phasen auf (Abb. 5.1). Während der ersten Phase (Pilot 1) wurden Items zum Inhaltsbereich Elektrizitätslehre pilotiert. Die zweite Phase (Pilot 2) verlief analog zur ersten und fokussierte auf die Inhaltsbereiche Mechanik, Optik und Wärmelehre.

Im Verlauf der dritten Testheftmodifikation in Vorbereitung der Pilot 1 (Version 2.2 auf Version 2.3; vgl. Tab. 5.1) wurden innerhalb der FDW-Skala zwei Items (eines zur Subfacette „Schk“ und eines zur Subfacette „M & E“) ergänzt, um jede der drei Subfacetten mit einer angemessenen Itemanzahl abbilden zu können. Im Fall des Fachwissens wurden aus testökonomischen Gründen weitere Items, welche fachlich ähnliche Inhalte aufwiesen, entfernt. Letztendlich enthielt diese Pilotierungsversion 15 Items zum FDW und 19 zum FW. Im Zeitraum von Februar bis Juni 2014 wurden diese Items innerhalb der Pilot 1 einer gemischten Stichprobe ($N = 75$) aus Lehrkräften, Studierenden und Universitätsmitarbeitenden zur Bearbeitung vorgelegt. Die Studienteilnehmer bearbeiteten den Test unentgeltlich von einem Testleiter beaufsichtigt einzeln oder in Gruppen ohne Hilfsmittel, wie z. B. Formelsammlungen oder Taschenrechner (die genaue Zusammensetzung sowie weitere Daten zur Stichprobe sind der Tabelle A.15 in Anhang A zu entnehmen). Das Alter der Studienteilnehmer betrug im Mittel 37,80 Jahre ($SD = 11,34$, range: 22 – 74), die mittlere Berufserfahrung 10,94 Jahre ($SD = 10,12$, range: 0 – 37). 36 % der Stichprobe war weiblichen Geschlechts.

Von Januar 2015 bis Februar 2015 wurden innerhalb der Pilot 2 die Items zu den Inhaltsbereichen Mechanik, Optik und Wärmelehre wiederum einer gemischten Stichprobe ($N = 55$) aus Lehrkräften, Studierenden und Universitätsmitarbeitenden vorgelegt. Innerhalb dieser Stichprobe wurden die Fachwissens-Items von einer kleinen Teilstichprobe fachfremder Akademiker, unter anderem wissenschaftliche Mitarbeiter anderer (naturwissenschaftlicher) Fachdisziplinen ($n = 11$) bearbeitet, um diejenigen Items zu selektieren und entfernen zu können, welche mit Hilfe eines hohen (physikalischen) Alltagswissens beantwortet werden konnten. Bei einem dieser Items (Item „Jahreszeiten“, Tab. 5.8) wurde beispielsweise eine Zuordnung der Erdposition in Bezug auf die Sonne zu den Jahreszeiten auf der Nordhalbkugel verlangt. Dieses Item wies innerhalb der obigen Stichprobe eine Lösungswahrscheinlichkeit nahe 1 auf.

Von den ursprünglich für die Prä-Pilot 2 entwickelten 15 Fachdidaktik-Items und den 31 Fachwissens-Items (Version 3.0; vgl. Tab. 5.1) fanden 13 bzw. 19 Eingang in die Pilot-2-Version (Version 3.2; vgl. Tab. 5.1). Das Vorgehen bei der Itemselektion erfolgte

Tabelle 5.8.: Beispiel-Item „Jahreszeiten“ zum Themenbereich Optik (wurde im Zuge der Pilotierung entfernt)

Beispiel-Item: „Jahreszeiten“	
<p>Ordnen Sie jedem der vier folgenden Ereignisse die Nummer der Position der Erde beim Umlauf um die Sonne – bezogen auf Deutschland (Nordhalbkugel) – zu. Bitte tragen Sie hierzu die jeweilige Nummer in das entsprechende Kästchen ein.</p>	
<input type="radio"/> Winteranfang	<input type="radio"/> Frühlingsanfang
<input type="radio"/> Sommeranfang	<input type="radio"/> Herbstanfang

analog zur ersten Prä-Pilotstudie. Auch innerhalb der Pilot 2 wurde die Bearbeitung unentgeltlich von einem Testleiter beaufsichtigt einzeln oder in Gruppen ohne Hilfsmittel wie beispielsweise Taschenrechner oder Formelsammlungen vorgenommen (die exakte Stichprobenszusammensetzung ist in der Tabelle A.16 in Anhang A dargestellt). In diesem Fall betrug das Alter der Studienteilnehmer im Mittel 40,46 Jahre ($SD = 12,43$, range: 25 – 75), die mittlere Berufserfahrung 11,97 Jahre ($SD = 11,60$, range: 0 – 40). 42 % der Stichprobe war weiblich.

Neben der Bearbeitung der Items hatten die Studienteilnehmer in beiden Pilotierungsrunden, wie schon die Experten und erfahrenen Lehrkräfte in der Prä-Pilotstudie, die Aufgabe, die Items auf einer vierstufigen Skala hinsichtlich Eindeutigkeit, Berufs- und Ausbildungsrelevanz zu beurteilen (Tab. 5.6). Damit sollte die Inhaltsvalidität der Items weiter sichergestellt werden. Nach Ende beider Pilotierungen wurden nach intensiver statistischer Analyse die geeignetsten Items für den finalen Test ausgewählt, welcher im Zeitraum von Juni 2015 bis Dezember 2016 validiert wurde.

5.3.2. Psychometrische Gütekriterien

Im folgenden Abschnitt werden die Testgütekriterien Objektivität und Reliabilität in Bezug auf die Pilotstudien 1 und 2 betrachtet. In diesem Zusammenhang wird auf die Interraterreliabilität (Auswertungsobjektivität), die Itemschwierigkeit, die interne Konsistenz der FDW- und FW-Skalen, die Trennschärfekoeffizienten sowie die Augenscheinvalidität (Inhaltsvalidität) eingegangen.

Interraterreliabilität (Auswertungsobjektivität)

Ein Maß für die Übereinstimmung der Kodierer bei der Beurteilung von Antworten ist Spearmans ρ^2 .

Im Rahmen der Pilotstudie 1 lagen insgesamt 56 Testhefte vor (75 abzüglich der Testhefte der Experten und derjenigen Lehrkräfte, welche oftmals nicht alle Items beantworteten, statt dessen aber Hinweise zur Eindeutigkeit usw. notierten). Von diesen 56 Testheften lagen 44 vollständig ausgefüllt vor. 28 davon wurden doppelt kodiert, so dass deren Scores (Punktwerte) zur Bestimmung der Interraterreliabilität herangezogen werden konnten. Bei abweichendem Urteil einigten sich beide Rater anschließend im Diskurs auf einen Score

²Werte für ρ von $0,50 < \rho \leq 0,70$ bedeuten eine mittlere, Werte von $\rho > 0,70$ eine hohe und Werte von $\rho > 0,90$ eine sehr hohe Übereinstimmung bzw. Auswertungsobjektivität (Wirtz & Caspar, 2002)

(Rater-Agreement bzw. Konsenskodierung), der dann als Berechnungsgrundlage für die weiteren statistischen Analysen diene. Wie in Tabelle 5.9 ersichtlich, lagen die Werte des Spearmans ρ für die FDW-Items im Mittel bei 0,81 ($SD = 0,08$) und für die FW-Items bei 0,77 ($SD = 0,23$).

Im Falle der Pilotstudie 2 lagen 46 (von 55) vollständige Datensätze für das FW (inkl. fachfremde Akademiker, vgl. Abschn. 5.3.1) und 35 für das FDW (ohne fachfremde Akademiker) zur Bestimmung des Spearmans ρ vor. Tabelle 5.9 zeigt die entsprechenden Werte auf Einzelitemebene. Im Mittel lagen diese für die FDW-Skala bei 0,89 ($SD = 0,06$) und für die FW-Skala bei 0,87 ($SD = 0,07$).

Itemschwierigkeit

Eine generelle Beurteilung von Schwierigkeitsindices (bzw. Lösungswahrscheinlichkeiten) stellt sich im Regelfall als schwierig dar, weil dies meistens vom Kontext oder auch der Art des Testinstruments (objektiver Test, Persönlichkeitstest, projektiver Test), von der Stichprobe (homogen, heterogen) sowie von der Art sowie der Breite des Konstrukts abhängt (Bühner, 2011). Nach Fisseni (1997, S. 124) gelten jedoch Itemschwierigkeiten von $P > 0,80$ als niedrig, solche von $0,80 \geq P \geq 0,20$ als mittel und solche mit $P < 0,20$ als hoch. Im Allgemeinen werden Itemschwierigkeiten mit einer guten Streuung im mittleren Bereich, ergänzt durch einige wenige im niedrigen und hohen Bereich, bevorzugt, um eine annähernd gleich gute Differenzierung von Studienteilnehmern mit unterschiedlicher Fähigkeitsausprägung zu ermöglichen (Lienert & Raatz, 1998; Bortz & Döring, 2006). Für die Pilot 1 betrug der mittlere Schwierigkeitsindex für das Professionswissen (FDW, FW) $P = 40,63$ ($SD = 20,02$) und für die Pilot 2 $P = 48,32$ ($SD = 19,89$), was bedeutet, dass die Items der Pilot 2 im Mittel etwas einfacher als die der Pilot 1 anzusehen sind. Betrachtet man die FDW- und die FW-Skala jeweils getrennt, ergeben sich für die Pilot 1 die Werte $P = 37,74$ (FDW: $SD = 17,92$; range: 14,20 – 68,18 und $P = 42,76$; FW: $SD = 21,67$; range: 9,09 – 94,32) sowie $P = 38,97$ (FDW: $SD = 19,70$; range: 10,00 – 68,57) und $P = 54,71$ (FW: $SD = 17,78$; range: 23,91 – 80,43) für die Pilot 2. Es kann festgestellt werden, dass sich der höhere Mittelwert für die Schwierigkeit des Professionswissens (FDW, FW) allein aus den Schwierigkeitsindices der FW-Skala in der Pilot 2 erklären lässt (die Schwierigkeitsindices auf Einzelitemebene finden sich in den Tabellen A.6 und A.7 in Anhang A. Bei der Auswahl der Items für die finale Testversion wurde darauf geachtet, dass die Schwierigkeitsindices, der allgemeinen Forderung folgend, gut um den mittleren Bereich streuen, in den Extrembereichen aber noch genügend Items vorhanden sind, um auch ausreichend zwischen Studienteilnehmern mit hohem Professionswissen bzw. zwischen solchen mit niedrigem Professionswissen differenzieren zu können.

Interne Konsistenz der Konstrukte (Skalenreliabilität und Trennschärfen)

Mit Hilfe des Reliabilitätsindikators Cronbachs α lässt sich die interne Konsistenz einer Skala abschätzen. Um eine statistische Vergleichbarkeit zu gewährleisten, wurden für die Berechnung der Werte die insgesamt 36 Datensätze derjenigen Studienteilnehmer verwendet, die sowohl bei der Pilot 1 als auch bei der Pilot 2 mitgewirkt hatten. Für die Pilot 1 ergab sich für das FDW ein Cronbachs α von 0,85³ und für das FW ein Wert von 0,92 (vgl. Tab. 5.9). Im Falle der Pilot 2 lag das Cronbachs α für das FDW bei 0,70 und

³für die Berechnung dieses Wertes wurden nur 14 der 15 Items miteinbezogen, da das Item „Ladungstrennung“ von kaum einem Studienteilnehmer beantwortet, sondern meistens mit einem nachfragenden Kommentar versehen wurde.

Tabelle 5.9.: Interne Konsistenzen der Gesamtskalen (Cronbachs α) und Trennschärfen (r_{it}), Augenscheinvalidität (Berufsrelevanz) sowie Interraterreliabilität (Spearmans ρ)

Skala (Itemanzahl)	Interne Konsistenz ($N = 56$, Pilot 1; $N = 46$, Pilot 2)	Trennschärfe $r_{it}: M(SD)$	Augenscheinva- lidität (Berufsrelevanz; $n = 24 - 60$)	Interraterrelia- bilität (doppelt kodierte Testhefte; $n = 28$, Pilot 1; $n = 35$, Pilot 2)
	Cronbachs α		$M(SD)$	$\rho: M(SD)$
FDW Pilot 1 (14)	0,85		3,44 (0,26) range 2,58 – 3,79	0,81 (0,08) range 0,65 – 0,91 (14 offene Items)
FW Pilot 1 (19)	0,92		3,21 (0,32) range 2,58-3,85	0,77 (0,23) range 0,23-1,00 (16 offene Items)
FDW Pilot 2 (13)	0,70		3,46 (0,21) range 3,22 – 3,90	0,89 (0,06) range 0,80 – 1,00 (13 offene Items)
FW Pilot 2 (18)	0,77		3,49 (0,15) range 3,24 – 3,78	0,87 (0,07) range 0,76-0,96 (6 offene Items)
FDW P1 + P2 (27)	0,86	0,39 (0,16) range 0,12 – 0,70	3,47 (0,23) range 3,06 – 3,90	0,85 (0,08) range 0,65 – 1,00 (27 offene Items)
FW P1 + P2 (37)	0,92	0,46 (0,22) range –0,07 – 0,82	3,35 (0,28) range 2,58 – 3,85	0,80 (0,08) range 0,65 – 1,00 (22 offene Items)

Bem.: M : arithm. Mittel; SD : Standardabweichung; FW: Fachwissen; FDW: Fachdidaktisches Wissen. Skalierung Augenscheinvalidität (Berufsrelevanz): 1 Trifft nicht zu; 2 Trifft kaum zu; 3 Trifft eher zu; 4 Trifft genau zu.

für das FW bei 0,77⁴ (vgl. Tab. 5.9). Somit sind die Skalen beider Pilotierungen angesichts der Testneukonstruktion als hinreichend reliabel zu bewerten (Field, 2013, S. 709; Brunner & Süß, 2007).

Zur Berechnung der Trennschärfen wurden die Items beider Pilotierungsphasen zu den beiden Skalen FDW und FW zusammengefasst, da für die finale Testversion Items beider Pilotierungen enthalten sein sollten. Demnach lagen die Werte der Trennschärfen r_{it} der Items des FDW zwischen 0,12 und 0,70 (vgl. Tab. 5.9). Für das FW reichten sie von -0,07 bis 0,82 (vgl. Tab. 5.9). Für die Endauswahl wurden Items mit einer negativen Trennschärfe keinesfalls und solche mit kleinen Trennschärfen ($r_{it} < 0,30$) in der Regel nicht berücksichtigt (Fisseni, 1997). Eine ausführliche Übersicht über die Trennschärfen der Einzelitems findet sich in den Tabellen A.8 und A.9 in Anhang A.

Augenscheinvalidität (Inhaltsvalidität)

Zur Überprüfung der Augenscheinvalidität sollten die Experten und erfahrenen Lehrkräfte aus den Prä-Pilotstudien sowie die Studienteilnehmer der Pilotstudien pro Item drei Kriterien auf einer vierstufigen Ratingskala (vgl. Tab. 5.6) bewerten. Pro Skala wurden dann jeweils die arithmetischen Mittelwerte dieser Einschätzungen berechnet. Da der größte Teil der Experten und erfahrenen Lehrkräfte im Verlauf der Testkonstruktionsphase auch an den Pilotierungsrunden teilnahmen, werden erst für diese Phase die konkreten Mittelwerte der Augenscheinvaliditäten berichtet (auf Einzelitemebene finden sich aber alle Werte der Prä-Pilot- und Pilotstudien in den Tab. A.4, A.5, A.10 und A.11 in Anhang A). Für die Pilot 1 lagen die Werte der Augenscheinvaliditäten bezogen auf die Berufsrelevanz (vgl. Tab. 5.9), welche das ausschlaggebendste Kriterium bei der finalen Itemauswahl darstellte, im Mittel für die FDW-Skala bei 3,44 ($SD = 0,26$) und für die FW-Skala bei 3,21 ($SD = 0,32$). Für die Pilot 2 ergaben sich für die FDW-Skala ein Mittelwert von 3,46 ($SD = 0,21$) und für die FW-Skala einer von 3,49 ($SD = 0,15$). Des Weiteren sollten die Teilnehmer die sprachliche Eindeutigkeit der Itemformulierungen beurteilen. Hier ergaben sich innerhalb der Pilot 1 Durchschnittswerte für das FDW von 3,60 ($SD = 0,11$) und für das FW von 3,67 ($SD = 0,13$) und innerhalb der Pilot 2 für das FDW von 3,53 ($SD = 0,26$) und für das FW von 3,69 ($SD = 0,13$). Bezüglich der Einschätzung der Notwendigkeit, die Iteminhalte in der Ausbildung zu berücksichtigen, ergab sich in der Pilot 1 für die FDW-Items ein Mittelwert von 3,50 ($SD = 0,21$) und für die FW-Items ein Mittelwert von 3,29 ($SD = 0,28$) sowie ein Wert von 3,49 ($SD = 0,19$) und 3,56 ($SD = 0,12$) für die FDW- und die FW-Skala in der Pilot 2.

Die Mittelwerte aus beiden Pilotstudien liegen im Hinblick auf die eingeschätzten Punkte jeweils in einem akzeptablen Bereich. Auf Einzelitemebene betrachtet wurden hauptsächlich diejenigen Items mit den höchsten Werten für die Eindeutigkeit und die Berufsrelevanz für die Endauswahl berücksichtigt (vgl. Tab. A.10 und A.11 in Anhang A).

5.4. Zusammenfassung (Kriterien der Itemselektion)

Wie bereits in 4.1 erwähnt, wurden aus ursprünglich 116 Itemvorschlägen im Rahmen einer Präpilotierungs- und einer Pilotierungsphase für die endgültige Testversion 14 Fachdidaktik-

⁴für die Berechnung dieses Wertes wurden nur 18 der 19 Items miteinbezogen, da das Item „Jahreszeiten“ aufgrund der Stichprobe aus fachfremden Akademikern (vgl. Abschnitt 5.3.1) als physikalisches Alltagswissen identifiziert wurde und demnach nicht zum Professionswissen gezählt wird (vgl. Abschnitt 2.2).

und 18 Fachwissensitems ausgewählt (vgl. Tab. A.2 und A.3 in Anhang A). Dabei wurden die soeben in Abschnitt 5.3 berichteten Kriterien Auswertungsobjektivität (Interraterreliabilität), Schwierigkeit, Augenscheinvalidität („Berufsrelevanz“; Reliabilität und Trennschärfe berücksichtigt (diese finale Itemreduktion erfolge im Übergang von Version 2.3 und Version 3.2 über Version 4.0a/b auf Version 4.2; vgl. Tab. 5.1). Die Reihenfolge der Themenbereiche innerhalb beider Testteile ergab sich durch die geäußerten „Vorlieben“ der Studienteilnehmer und aus den Itemschwierigkeiten. Die insgesamt etwas leichteren Mechanik-Items wurden an den Testanfang gestellt, um den Mitwirkenden die Testeingangsphase zu erleichtern. Die schwierigeren Elektrizitätslehre-Items wurden jeweils ungefähr in der Testmitte positioniert. Den jeweiligen Abschluss bildeten die wiederum etwas einfacheren und in einer vergleichsweise geringeren Anzahl vorkommenden Items zur Optik und Wärmelehre. Die Anordnung der Items innerhalb der Themengebiete erfolgte – in Entsprechung der Struktur eines Power-Tests (Bühner, 2011) – nach ansteigender Schwierigkeit.

Im ersten Schritt der Itemselektion wurden die Mittelwerte der Augenscheinvaliditäten aus der Pilotstudie berücksichtigt (vgl. Tab. A.10 und A.11 in Anhang A). Items mit einer geringen Akzeptanz seitens der Studienteilnehmer („Eindeutigkeit“ und „Berufsrelevanz“; Mittelwerte um $M = 3,2$ und niedriger) wurden entfernt oder innerhalb des Teilprojekts FALKO-Physik nochmals einer kritischen Diskussion unterzogen und gegebenenfalls aus normativen Gründen unter Abwägung der Alternativen (inhaltsgleiche Alternativitems) ausgeschlossen oder trotzdem für die Testendversion ausgewählt.

Die Schwierigkeit betreffend gelten Items, die von kaum einem beziehungsweise von annähernd allen Studienteilnehmern gelöst werden können, als testtheoretisch bedenklich, da in Bezug auf die Testleistung Decken- oder Bodeneffekte resultieren können. Daher fanden in der Regel Items mit einem mittleren Schwierigkeitsgrad, ergänzt durch solche mit einer Schwierigkeit von $0,20 \leq P \leq 0,80$, Eingang in die Testendversion (für die Schwierigkeitsindices der Items siehe Tab. A.6 und A.7 in Anhang A).

Bei den Werten für den Trennschärfekoeffizienten – dieser gibt im Prinzip an, wie gut ein Item zu einer Gesamtskala passt – gelten Werte von $< 0,30$ als niedrig, Werte von $0,30 - 0,50$ als mittel und Werte von $> 0,50$ als hoch (Fisseni, 1997). Insgesamt gilt, dass zur Erfassung eines engen Konstrukts die Trennschärfen der verwendeten Items eine gewisse Homogenität aufweisen sollen. Möchte man dagegen ein Konstrukt breiter abbilden, sollte ein Testinstrument Items mit größerer Trennschärfenvarianz beinhalten (vgl. Bühner, 2011). Folglich wurden Items mit negativen oder sehr niedrigen Trennschärfen entfernt sowie ansonsten auf eine ausreichende Streuung der Trennschärfen geachtet (für die Trennschärfen der Items vgl. Tab. A.8 und A.9 in Anhang A).

Konnte eine niedrige Interrater-Übereinstimmung ($\rho \leq 0,80$) (Shavelson & Webb, 1991) bei einem Item weder durch eine Optimierung des Kodiermanuals (z. B. Hinzufügen von Anker- bzw. Negativ-Beispielen), noch durch weitergehende Raterschulungen in akzeptabler Weise verbessert werden, wurde auch dieses Item entfernt, außer es sprachen normative Gründe (z. B. ein hoher Wert für die Augenscheinvalidität) dagegen. Die meisten Testhefte der Pilotstudie wurden von zwei unabhängigen geschulten Ratern doppelt kodiert (für die Werte der Interraterreliabilitäten im Überblick für die vollständigen Datensätze siehe Tab. 5.9; für die Werte auf Einzelitemebene vgl. Tab. A.17 in Anhang A). Schließlich gingen Items, welche zu einer Verschlechterung der internen Konsistenz der Skalen (FDW, FW) führten – ein Wert für das Cronbachs α von $\geq 0,70$ ist dabei erstrebenswert (z. B. Nunally & Bernstein, 1994) – nicht in die Testendversion ein.

Es sei final noch einmal darauf hingewiesen, dass an mancher Stelle die angegebenen Cut-off-Werte nicht exakt eingehalten wurden, da bei der finalen Auswahl der Items auch

normative Gründe Beachtung fanden. Wenige Aufgaben, deren statistische Kennwerte sich grenzwertig zeigten, von den befragten Experten und den Studienteilnehmern aus der Pilotstudie aber als inhaltlich valide eingestuft wurden (hohe Werte bezogen auf die Augenscheinvaliditäten sowie Aussagen der Studienteilnehmer, wie z. B. „Das sollte ich eigentlich wissen.“), wurden trotzdem in die Testendversion übernommen.

6. Testvalidierung

Für die Testvalidierung (Hauptstudie) wurden 183 der insgesamt 269 (Gesamtzusammensetzung der Stichprobe siehe Tab. A.18 in Anhang A) Studienteilnehmer als Hauptstichprobe für die durchgeführten statistischen Analysen herangezogen. Die verbleibenden 86 Datensätze dienten, zusammengefasst zu unterschiedlichen Kontraststichproben, dem Erhalt von Validitätshinweisen in Bezug auf unterschiedliche Aspekte der Konstruktvalidität. Die gesamte Validierungsphase fand im Zeitraum von Juni/Juli 2015 bis Dezember 2016 statt (vgl. Abb. 6.1).

In Abschnitt 6.1 wird zunächst näher auf die Zusammensetzung der Hauptstichprobe sowie die Durchführung der Hauptstudie eingegangen. In Abschnitt 6.2 wird die Dimensionalität des fachbezogenen Professionswissens (Modellpassung) mittels einer konfirmatorischen Faktorenanalyse überprüft (faktorielle Validität). Abschnitt 6.3 berichtet psychometrische Gütekriterien wie Interraterreliabilität (Auswertungsobjektivität), interne Konsistenz der Konstrukte (Skalenreliabilitäten, Trennschärfen) und Augenscheinvalidität (inhaltliche Validität). Abschnitt 6.4 widmet sich nachfolgend deskriptiven Ergebnissen in Bezug auf Skalenmittelwerte und Interkorrelationen, auf Gruppenunterschiede hinsichtlich der untersuchten Teilstichproben, Schulformunterschieden sowie Zusammenhängen des fachbezogenen Professionswissens mit Indikatoren des Schul- und Studienerfolgs sowie der Berufs- und Studiendauer (kriteriale Validität). Während in Abschnitt 6.6 Ergebnisse der Konstruktvalidität (diskriminante Validität) bezüglich der untersuchten Kontraststichproben berichtet werden, beschäftigt sich Abschnitt 6.5 zuvor mit der Untersuchung der bereits in 3.1 aufgestellten Repräsentativitätshypothese. Abgerundet wird das gesamte Kapitel in Abschnitt 6.7 mit den Ergebnissen einer Kreuzvalidierung mit je einem bereits existierenden FDW- und FW-Test (Forschungsprojekt ProfiLe-P, Gramzow, 2014; Woitkowski, 2015; Konstruktvalidität: konvergente Validität).

Abbildung 6.1.: Zeitlicher Ablauf der Validierungsphase

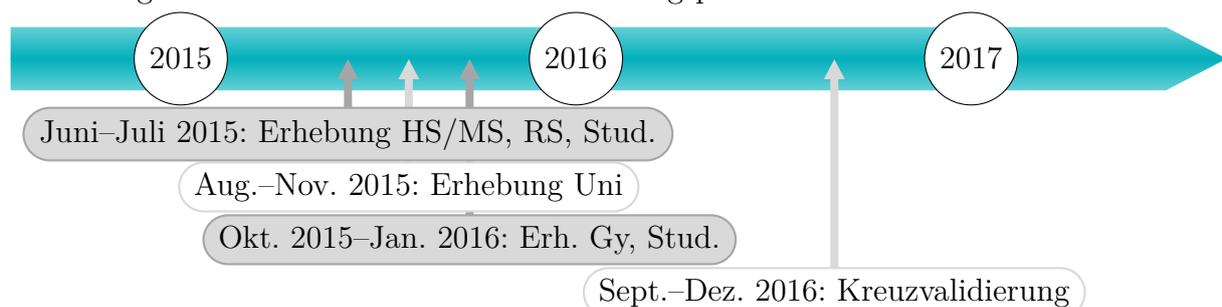


Tabelle 6.1.: Stichprobenzusammensetzung der Hauptstichprobe aus der Hauptstudie

	Nichtgymnasial			Gymnasial		Summe
	GS	MS	RS	Gy	FOS/ BOS	
Studierende			32	86		118
Referendare			4			4
Lehrkräfte		6	20	35		61

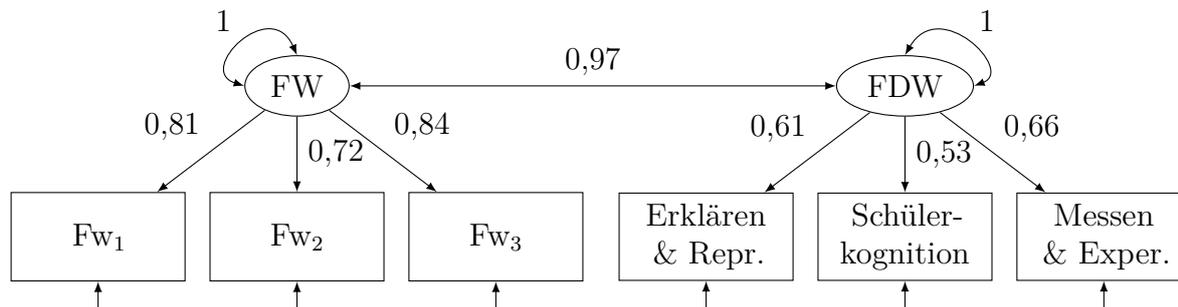
6.1. Stichprobe und Durchführung

Die Hauptstichprobe der Validierungsphase ($N = 183$; vgl. Tab. 6.1) setzte sich aus $n = 65$ bayerischen Physiklehrkräften verschiedener Sekundarschularten und $n = 118$ Lehramtsstudierenden mit dem Unterrichtsfach Physik zusammen.

Zu den Physiklehrkräften wurden auch $n = 4$ Referendare gezählt. Zur statistischen Rechtfertigung der Zuordnung der Referendare zu den Lehrkräften bei allen FALKO-Disziplinen siehe Abschnitt 2.1 in Lindl und Krauss (2017). Von den 65 Lehrkräften (gymnasial, Gy: $n = 35$, nicht-gymnasial, nGy: $n = 30$) waren 32 % weiblich. Im Mittel waren sie 42,58 Jahre alt ($SD = 10,65$; range 26 – 70) und hatten eine Berufserfahrung von 13,54 Jahren ($SD = 9,58$; range 1 – 40). Die Zusammenlegung von $n = 24$ Realschul- und $n = 6$ Mittelschullehrkräften mit Unterrichtsfach (Hauptfach; $n = 2$) oder Didaktikfach (Nebenfach; $n = 4$) Physik zur Gruppe der nicht-gymnasialen Lehrkräfte kann mit der Gleichheit beziehungsweise Ähnlichkeit der universitären Lehramtsausbildung in Bayern (hier entstammen alle bisher untersuchten Lehrkräfte sowie der größte Teil der Studierenden) begründet werden (vgl. hierzu auch Abschnitt 4 in Krauss, Lindl, Schilcher & Tepner, 2017). Trotz des hohen Anteils an bayerischen Studienteilnehmern ist das Testinstrument FALKO-P, wie auch die anderen Testinstrumente des FALKO-Projekts (vgl. Krauss, Lindl, Schilcher & Tepner, 2017), durch die Beachtung der nationalen Bildungsstandards bei der Testkonstruktion prinzipiell bundeslandunabhängig ausgerichtet.

Von den 118 Lehramtsstudierenden waren 28 % weiblich. Ihr Durchschnittsalter lag bei 24,39 Jahren ($SD = 2,95$; range 20 – 37) und die durchschnittliche Studienerfahrung betrug 6,35 Semester ($SD = 2,83$, range 2 – 12). 86 der Studierenden waren für das Lehramt an Gymnasien eingeschrieben (\emptyset -Sem 6,84, $SD = 2,90$) und 32 für das Lehramt an Realschulen (\emptyset -Sem 5,03, $SD = 2,15$). Lehramtsstudierende der Schulart Mittelschule konnten nicht untersucht werden, da diese das Unterrichtsfach Physik im Rahmen ihres Studiums nur vergleichsweise selten wählen.

In der Hauptstudie bearbeiteten alle Studienteilnehmer den Test unter Beaufsichtigung eines Testleiters einzeln oder in Gruppen. Die beteiligten Lehrkräfte konnten in der Regel durch Fortbildungsangebote gewonnen werden, während die Studierenden den Test meist in regulären Lehrveranstaltungen (an den Universitäten Regensburg, Würzburg, Kassel und der LMU München) bearbeiteten. Die Teilnahme am Test erfolgte auf freiwilliger Basis unter Gewährleistung der Anonymität der Testteilnehmer (in der Regel ohne finanzielle Entschädigung). Da das Testinstrument als Power-Test konzipiert wurde, war die Bearbeitungszeit nicht beschränkt, hierfür allerdings durchschnittlich 90 Minuten veranschlagt (FDW: 60 Minuten, FW: 30 Minuten). Studierende, die länger als 90 Minuten benötigten, blieben nach Ende der regulären Veranstaltungszeit bis zur Fertigstellung der Bearbeitung

Abbildung 6.2.: Konfirmatorisches Faktorenmodell des Fach- und fachdidaktischen Wissens (mit „M & E“) (Modellierung auf Basis der Hauptstichprobe $N = 183$)

Bem.: Modellfit: $\chi^2(8, N = 183) = 11,70$; $p = ,17$; CFI = 0,99; RMSEA = 0,05; SRMR = 0,03; Werte von $p \geq ,01$, CFI (Comparative Fit Index) $> 0,95$, RMSEA (Root-Mean-Square Error of Approximation) $< 0,08$ und SRMR (Standardized Root Mean Residual) $< 0,01$ indizieren einen akzeptablen Modellfit (Schermelleh-Engel, Moosbrugger & Müller, 2003; Hu & Bentler, 1999).

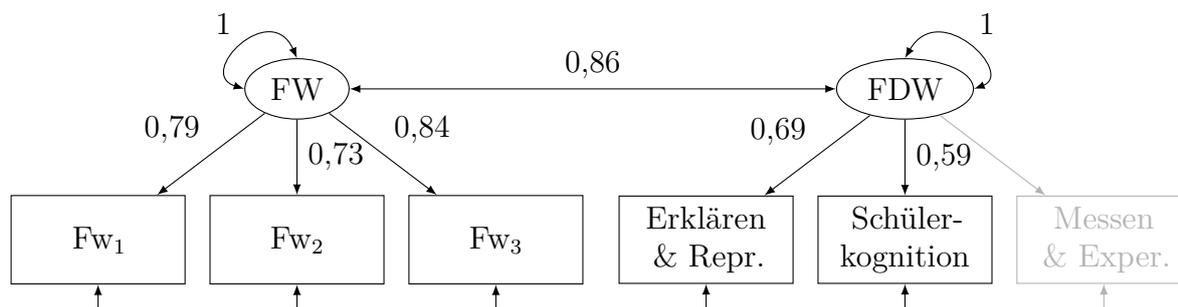
(Studierende, die die Bearbeitung nach eineinhalb Stunden abbrachen, wurden bei den Analysen nicht berücksichtigt).

6.2. Dimensionalität des fachbezogenen Professionswissens (CFA; faktorielle Validität)

Ergebnisse in Bezug auf die Konstrukte FDW und FW lassen sich nur belastbar interpretieren, wenn deren jeweilige Eindimensionalität konstatiert werden kann. Zur Prüfung der Eindimensionalität wurde eine konfirmatorische Faktorenanalyse durchgeführt. Um ein akzeptables Verhältnis zwischen Modellkomplexität und Datengrundlage zu erreichen (es sollte ein ausreichendes Verhältnis von Studienteilnehmern pro zu schätzendem Parameter gewährleistet werden; Little, Cunningham, Shahar & Widaman, 2002, S. 154-155), wurden für die Modellspezifikation die Items des FW-Tests zu drei Parcels und die des FDW-Tests ebenfalls zu drei Parcels (entsprechend den drei Facetten „E & R“, „Schk“ und „M & E“ des FDW) gebündelt, welche als manifeste Indikatoren der latenten Konstrukte Fachwissen und fachdidaktisches Wissen fungieren (zur Begründung dieses Vorgehens vgl. Krauss et al., 2008, S. 721). Die FW-Items wurden zufällig zu drei manifesten Indikatoren gruppiert, jedoch unter Berücksichtigung der Randbedingung, dass die Beträge der arithmetischen Mittelwerte und Standardabweichungen der resultierenden Parcels vergleichbar waren (Little et al., 2002, S. 165-166). Für eine Überprüfung der Unterscheidung der minderungskorrigierten Korrelation (zwischen FDW und FW) von 1 erfolgte eine Modelltestung gegen ein weiteres strukturgleiches Modell, bei welchem die latente Korrelation zwischen FDW und FW auf den Wert 1 fixiert war.

Bei Betrachtung der Ergebnisse erwiesen sich zwar die einzelnen Faktorladungen als hinreichend groß und signifikant, um auf einen guten lokalen Fit zu schließen. Auch die globale Anpassungsgüte des Modells war positiv zu beurteilen. Alle Fit-Indizes genügten den vorgeschlagenen Cutoff-Kriterien für die Annahme eines Modells (Schermelleh-Engel, Moosbrugger & Müller, 2003). Allerdings wurde als Ergebnis eine Korrelation zwischen den beiden Wissenskategorien FDW und FW erhalten (Abb. 6.2, Tab. A.39 in Anhang A), die nicht mehr signifikant von 1 unterscheidbar war (Chi-Quadrat-Differenzentest: $\Delta\chi^2(1, N = 183) = 0,29$; $p < ,59$).

Abbildung 6.3.: Konfirmatorisches Faktorenmodell des Fach- und fachdidaktischen Wissens (ohne „M & E“) (Modellierung auf Basis der Hauptstichprobe $N = 183$)



Bem.: Modellfit: $\chi^2(4, N = 183) = 4,95$; $p = ,29$; CFI = 1,00; RMSEA = 0,04; SRMR = 0,02; Werte von $p \geq ,01$, CFI (Comparative Fit Index) $> 0,95$, RMSEA (Root-Mean-Square Error of Approximation) $< 0,08$ und SRMR (Standardized Root Mean Residual) $< 0,01$ indizieren einen akzeptablen Modellfit (Schermelleh-Engel, Moosbrugger & Müller, 2003; Hu & Bentler, 1999).

Da diese Ergebnisse nicht hypothesenkonform waren (vgl. Kapitel 3.1) und bisherigen Forschungsergebnissen aus dem deutschsprachigen Raum (z. B. Forschungsprojekte ProWiN und ProfiLe-P) widersprachen, wurden zusätzliche explorative Analysen durchgeführt, die ergaben, dass die „M & E“-Items stärker mit der Dimension FW als der Dimension FDW assoziiert waren. Auf Basis dieser Erkenntnis wurden diese Items in einer Expertenbefragung reevaluiert, was den fachwissenschaftlichen Charakter dieser Items bestätigte (eine spezifische Zuordnung zu den Bereichen FW oder FDW wurde im Rahmen der Überprüfung der Augenscheinvalidität (vgl. Abschn. 5.3.2) von den Studienteilnehmern nicht erbeten, wodurch keine entsprechende Hypothese im Vorfeld generiert wurde).

Nach Entfernung der „M & E“-Facette wies das Modell (Abb. 6.3, Tab. A.40 in Anhang A) mit einem Wert von $r = 0,86$ eine zwar hohe, aber von 1 signifikant unterschiedliche Korrelation zwischen beiden Wissenskategorien auf (Chi-Quadrat-Differenzentest: $\Delta\chi^2(1, N = 183) = 3,87$; $p < ,05$). Um eine Vergleichbarkeit zur ersten Modelltestung zu erreichen, wurden für die modifizierte Modellspezifikation die Items des FW-Tests erneut zu drei Parcels (wiederum unter Berücksichtigung der oben beschriebenen Randbedingungen) und des FDW-Tests nun zu zwei Parcels (entsprechend den beiden verbliebenen Facetten) gebündelt, die wiederum als manifeste Indikatoren der latenten Konstrukte Fachwissen und fachdidaktisches Wissen fungierten. Resultierend erwiesen sich die einzelnen Faktorladungen erneut als hinreichend groß und signifikant, um auf einen guten lokalen Fit zu schließen. Auch die globale Anpassungsgüte des Modells war erneut positiv zu beurteilen. Alle Fit-Indizes genügten ebenfalls den vorgeschlagenen Cutoff-Kriterien für die Annahme eines Modells (Schermelleh-Engel et al., 2003). Diesem Ergebnis folgend wurden die im weiteren Verlauf dieses Kapitels durchgeführten statistischen Analysen anhand der beiden unterscheidbaren Skalen FDW und FW durchgeführt.

6.3. Psychometrische Gütekriterien

Im Folgenden werden die Testgütekriterien Objektivität, Reliabilität und Validität betrachtet. Wie bereits erwähnt, wurde für sämtliche Analysen (Gütekriterien, deskriptive und inferenzstatistische Analysen) bezüglich FDW auf die Facette Messen und Experimentieren („M & E“) verzichtet. Die Grundlage für die Entfernung dieser Facette war die in Abschnitt 6.2 durchgeführte konfirmatorische Faktorenanalyse, die indizierte, dass die Items von „M

Tabelle 6.2.: Interne Konsistenzen der Gesamtskalen (Cronbachs α) und Trennschärfen (r_{it}), Augenscheinvalidität (Berufsrelevanz) sowie Interraterreliabilität (Spearman's ρ)

Skala (Itemanzahl)	Interne Konsistenz ($N = 183$)	Trennschärfe (part-whole-korrigiert)	Augenscheinvalidität (Berufsrelevanz; $n = 34$; Experten, Lehrkräfte)	Interraterreliabilität (doppelt kodierte Testhefte; $n = 60$)
	Cronbachs α	r_{it} : M (SD)	M (SD)	ρ : M (SD)
FW (18)	0,82	0,41 (0,13) range 0,16 – 0,57	3,44 (0,28) range 2,86 – 3,89	0,94 (0,06) range 0,85 – 1,00 (12 offene Items)
FDW (10)	0,65	0,31 (0,05) range 0,24 – 0,40	3,58 (0,22) range 3,17 – 3,90	0,89 (0,05) range 0,83 – 0,96 (10 offene Items)

Bem.: M: arithm. Mittel; SD: Standardabweichung; FW: Fachwissen; FDW: Fachdidaktisches Wissen. Skalierung Augenscheinvalidität (Berufsrelevanz): 1 Trifft nicht zu; 2 Trifft kaum zu; 3 Trifft eher zu; 4 Trifft genau zu.

& E“ nicht ausreichend mit dem latenten Konstrukt des FDW assoziiert waren.

6.3.1. Interraterreliabilität (Auswertungsobjektivität)

Für die Bestimmung der Interraterreliabilität während der Hauptstudie wurden 60 der insgesamt 183 Testhefte aus der Hauptstichprobe von zwei geschulten Ratern doppelt kodiert. Bei abweichendem Urteil einigten sich beide Rater im Diskurs auf einen Score (Rater-Agreement bzw. Konsenskodierung), der dann als Berechnungsgrundlage für anschließende Analysen diente. Dabei lagen die pro Wissensbereich gebildeten arithmetischen Mittelwerte für das FDW bei $\rho = 0,89$ ($SD = 0,05$) und für das FW bei $\rho = 0,94$ ($SD = 0,06$). Wie zusätzlich in Tabelle 6.2 illustriert, konnten die Antworten auf die Items des FALKO-P-Tests für beide Wissenskategorien somit im Durchschnitt sehr objektiv beurteilt werden, sodass die Konsenskodierung nur in wenigen Fällen erforderlich war. Aufgrund dieser hohen Werte (eine ausführliche Übersicht über die Interraterreliabilitäten auf Einzelitemebene findet sich in A.2 in Anhang A) wurden die verbleibenden Testhefte nur noch einfach kodiert. Auf eine Skalierung des Testinstruments wurde zu diesem Zeitpunkt verzichtet. Allerdings stellt eine solche und als eventuelle Folge eine Testüberarbeitung eine Möglichkeit der Testweiterentwicklung dar, welche in Zukunft verfolgt werden könnte (vgl. Kap. 8).

6.3.2. Itemschwierigkeit

Wie bereits in Abschnitt 5.3.2 berichtet, gelten nach Fisseni (1997, S. 124) Itemschwierigkeiten von $P > 0,80$ als niedrig, solche von $0,80 \geq P \geq 0,20$ als mittel und solche mit $P < 0,20$ als hoch. Für die Hauptstudie (vgl. auch Abschnitt 6.4.1) betrug der mittlere Schwierigkeitsindex für das Professionswissen (FDW, FW) $P = 42,31$ ($SD = 21,07$). Betrachtet man die FDW- und die FW-Skala getrennt, ergibt sich für erstere $P = 38,73$ ($SD = 19,19$; range: 7,65 – 78,14) und für zweitere $P = 45,09$ ($SD = 22,56$; range: 9,56 – 85,79). Bei der Betrachtung der Schwierigkeitsindices auf Einzelitemebene (vgl. Tab. A.12 in Anhang A)

kann festgestellt werden, dass die Werte, der allgemeinen Forderung folgend, gut um den mittleren Bereich streuen, in den Extrembereichen aber noch genügend Items vorhanden sind, um auch ausreichend zwischen Studienteilnehmern mit hohem Professionswissen bzw. zwischen solchen mit niedrigem Professionswissen differenzieren zu können (Lienert & Raatz, 1998; Bortz & Döring, 2006).

6.3.3. Interne Konsistenz der Konstrukte (Skalenreliabilitäten und Trennschärfen)

Für das FW liegt der Wert der internen Konsistenz $\alpha = 0,82$ und für das FDW bei $\alpha = 0,65$, was angesichts einer Testneukonstruktion (und der aufgrund der Entfernung einer Facette geringeren Itemzahl im FDW) als hinreichend reliabel zu bewerten ist (Field, 2013, S. 709; Brunner & Süß, 2007). Die um Autokorrelationen korrigierten Trennschärfen r_{it} der Items des FDW lagen zwischen 0,24 und 0,40 und für FW von 0,16 bis 0,57 (siehe Tab. 6.2). Die Skala zum FW enthält zwei Items mit einer Trennschärfe von $r_{it} \leq 0,20$. Obwohl Trennschärfen üblicherweise erst ab $r_{it} \geq 0,20$ als akzeptabel erachtet werden, wurden diese beiden Items aus Gründen der didaktischen beziehungsweise fachlichen Relevanz (indiziert durch Expertenrückmeldungen) in der jeweiligen Skala belassen. Eine ausführliche Übersicht über die Trennschärfen der Einzelitems findet sich in A.2 in Anhang A.

6.3.4. Augenscheinvalidität: Inhaltliche Validität

Zur Überprüfung der Augenscheinvalidität sollten 34 (erfahrene) Lehrkräfte (diese Zahl entspricht dem Teil der Studienteilnehmer aus den Pilotstudien, welche alle finalen Items bearbeitet hatten und somit zur Stichprobe der Hauptstudie hinzugezogen werden konnten) pro Item drei Kriterien auf einer vierstufigen Ratingskala (vgl. Tab. 5.6 in Abschn. 5.3.2) bewerten. Pro Skala wurden dann jeweils wiederum die arithmetischen Mittelwerte dieser Einschätzungen berechnet. Diese lagen im Hinblick auf die eingeschätzte Berufsrelevanz für das FDW bei 3,58 und für das FW bei 3,44 (vgl. Tab. 6.2). Des Weiteren sollten die Teilnehmer die sprachliche Eindeutigkeit der Itemformulierungen beurteilen. Hier ergaben sich Durchschnittswerte für das FDW von 3,58 ($SD = 0,22$; range: 3,17 – 3,90) und für das FW von 3,64 ($SD = 0,13$; range: 3,44 – 3,86). Bezüglich der Einschätzung der Notwendigkeit, die Iteminhalte bereits in der Lehramtsausbildung zu berücksichtigen, wurde für die FDW-Items ein Mittelwert von 3,59 ($SD = 0,28$; range: 3,00 – 4,00) und für die FW-Items ein Mittelwert von 3,47 ($SD = 0,28$; range: 2,82 – 3,89) erhalten. Die Werte auf Einzelitemebene in Bezug auf die Berufsrelevanz, welche als das ausschlaggebendste Kriterium angesehen wird, finden sich in Tab. A.2. Tabelle A.23 in Anhang A zeigt die Werte für die anderen beiden Einschätzungen ebenfalls auf Einzelitemebene.

6.4. Deskriptive Ergebnisse

Im Folgenden wird der Fokus, dem theoretischen Professionswissensansatz folgend (vgl. Krauss, Lindl, Schilcher & Tepner, 2017), als erstes auf Vergleiche zwischen Lehrkräften (L) und Studierenden (S; Statusgruppenunterschiede, Abschn. 6.4.2) und in einem zweiten Schritt auf Schulformunterschiede innerhalb beider Gruppen gelegt (Abschn. 6.4.3). Demnach werden aus der Hauptstichprobe zwei Gruppen (Studierende S, Lehrkräfte L) und vier nach Schulart (Gymnasium Gy, nicht-Gymnasium nGy) getrennten Untergruppen (nGy-S,

Tabelle 6.3.: Mittelwerte (M), Standardabweichungen (SD) und empirische Maxima auf Skalenebene sowie Skaleninterkorrelationen (Produkt-Moment-Korrelationen)

	Deskriptive Ergebnisse			Korrelationen		
	$N = 183$	M (SD)	emp. Max.	FW	E & R	Schk
FW (18)	12,34	(5,99)	27	1		
FDW (10)	7,41	(3,27)	16	0,59**		
E & R (5)	3,84	(2,11)	9	0,53**	1	
Schk (5)	3,57	(1,78)	8	0,46**	0,41**	1

Bem.: FW: Fachwissen; FDW: Fachdidaktisches Wissen; E & R: Erklären und Repräsentieren; Schk: Schülerkognition; signifikanter Zusammenhang: * $p \leq ,05$; ** $p \leq ,01$.

nGy-L, Gy-S und Gy-L) gebildet, die verschiedenen vergleichenden Analysen unterzogen werden. Nachfolgend werden die Zusammenhänge des fachbezogenen Professionswissens mit Indikatoren des Schul- und Studienerfolgs beziehungsweise der Studien- oder Berufsdauer berichtet (Abschn. 6.4.4).

Zunächst sollen jedoch deskriptive Ergebnisse in Bezug auf die untersuchte Hauptstichprobe ($N = 183$; Tab. 6.1) vorgestellt werden (Abschn. 6.4.1). Dafür werden jeweils Summenscores zu FDW und FW sowie zu den zwei FDW-Facetten „Erklären und Repräsentieren“ („E & R“) und „Schülerkognitionen“ („Schk“) gebildet.

6.4.1. Skalenmittelwerte und Interkorrelationen

Im Mittel erzielten die 183 Studienteilnehmer beim FW einen Score von 12,34 Punkten ($SD = 5,99$; Max_{emp} : 27; Max_{theo} : 29) und beim FDW von 7,41 Punkten ($SD = 3,27$; Max_{emp} : 16; Max_{theo} : 23). Somit erreichten die Studienteilnehmer in beiden Testteilen im Schnitt knapp 50 % der maximal zu erzielenden Punktzahl (vgl. auch Abschn. 6.3.2), was als Indikator für eine angemessene Schwierigkeit des Gesamttests gesehen werden kann (für die Schwierigkeit der Einzelitems siehe Tab. A.12 und A.2 in Anhang A).

Sowohl die Korrelationen zwischen FW und FDW als auch diejenigen zwischen den beiden verbleibenden Facetten des fachdidaktischen Wissens sind erwartungsgemäß hoch (Tab. 6.3). Die Tatsache, dass sie sich dennoch deutlich von 1 unterscheiden, impliziert aber, dass jeweils unterschiedliche Konstrukte gemessen wurden.

6.4.2. Gruppenunterschiede zwischen Studierenden und Lehrkräften

Die 183 Studienteilnehmer unterteilen sich in 65 Lehrkräfte (darunter vier Referendare) und 118 Studierende. Für einen Überblick über Statusgruppenunterschiede werden die Testleistungen der Gruppe der Lehrkräfte – zunächst noch ohne Berücksichtigung der unterschiedlichen Schulformen – denjenigen der Studierenden gegenübergestellt.

Hinsichtlich der mittleren Testleistung schneiden die Lehrkräfte beim FW wesentlich besser ab als die Studierenden ($d = 1,27^{**}$; zur Interpretation siehe Legende zu Tab. 6.4). Auch beim FDW erzielen die Lehrkräfte einen höheren durchschnittlichen Score ($d = 0,51^{**}$). Aus Tabelle 6.4 wird weiterhin ersichtlich, dass dieser Unterschied vor allem auf die Facette „E & R“ ($d = 0,68^{**}$) zurückzuführen ist. Statusgruppenunterschiede (auf Itemebene)

Tabelle 6.4.: Gruppenunterschiede zwischen Studierenden und Lehrkräften im Fachwissen und im fachdidaktischen Wissen

	Studierende ($n = 118$)		Lehrkräfte ($n = 65$)		Effektstärke d (Leh. vs. Stud.)	p -Wert
	M (SD)	emp. Max.	M (SD)	emp. Max.		
$N = 183$						
FW (18)	10,03 (4,68)	22,5	16,54 (5,85)	27	1,27**	< ,01
FDW (10)	6,84 (3,20)	15	8,45 (3,16)	16	0,51**	< ,01
E & R (5)	3,35 (2,05)	9	4,72 (1,93)	9	0,68**	< ,01
Schk (5)	3,49 (1,77)	7	3,72 (1,81)	8	0,13	< ,40

Bem.: M : arith. Mittel; SD : Standardabweichung; FW: Fachwissen; FDW: Fachdidaktisches Wissen; E & R: Erklären und Repräsentieren; Schk: Schülerkognition. Die Effektstärke d berechnet sich aus der Mittelwertdifferenz beider Gruppen ($M_{\text{Lehrkräfte}} - M_{\text{Studierende}}$) dividiert durch die gepoolte Standardabweichung. Nach Cohen (1992) entspricht $d = 0,20$ einem kleinen, $d = 0,50$ einem mittleren und $d = 0,80$ einem großen Effekt. Signifikanter Unterschied: * $p \leq ,05$; ** $p \leq ,01$.

finden sich in Tabelle A.3 in Anhang A (dort werden auch die Statusgruppenunterschiede der vier entfernten „M & E“-Items berichtet).

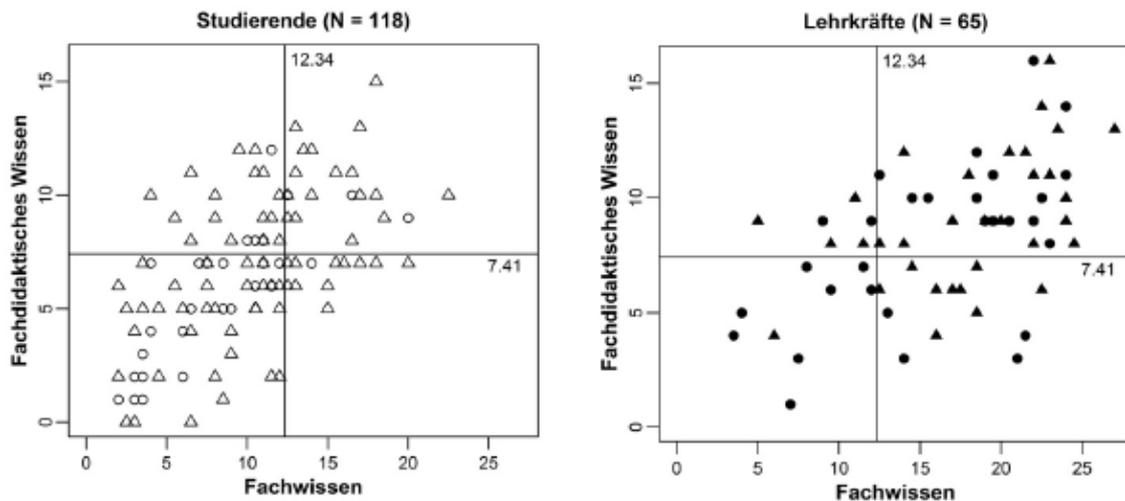
Die beiden Scatterplots in Abbildung 6.4 veranschaulichen noch einmal die Testleistungen der Studierenden (links) und der Lehrkräfte (rechts) sowie den Zusammenhang zwischen FDW und FW in beiden Gruppen. Die angegebenen Korrelationen für beide Teilgruppen (die Produkt-Moment-Korrelation bei den Studierenden beträgt $r = 0,53^{**}$ und bei den Lehrkräften $r = 0,59^{**}$) stellen lediglich einen ersten Anhaltspunkt dar, da sich Korrelationen erst für größere Stichprobenzahlen stabilisieren (Schönbrodt & Perugini, 2013).

6.4.3. Schulformunterschiede (Ausbildungssensitivität)

Bislang wurde die Verteilung der Schulformen in der Lehrer- und in der Studierendengruppe noch nicht berücksichtigt. Potentielle Schulformunterschiede, die in Abbildung 6.4 bereits durch unterschiedliche Symbole graphisch angedeutet sind, sollen nun genauer beleuchtet werden.

In FALKO und somit auch für FALKO-Physik war der Status der Studienteilnehmer (Studierende vs. Lehrkräfte) – gemäß dem Begriff „Professionswissen“ – primäres Strukturierungskriterium bezüglich der Gesamtstichprobe, natürlich sind aber auch jeweils Unterschiede in Bezug auf die Schulform zu erwarten. Die Hypothesen sind dabei (vgl. auch Kapitel 3.1), dass Lehrkräfte in beiden Wissensbereichen besser als Studierende abschneiden und in diesen beiden Gruppen jeweils die gymnasialen Studienteilnehmer wiederum besser als die nicht-gymnasialen Teilnehmer. Entsprechend dieser Hypothesen sind auch die Vorzeichen der Effektstärken d in Tabelle 6.5 gewählt (erwartungskonforme Unterschiede erhalten dabei eine positive Effektstärke).

Abbildung 6.4.: Streudiagramme Fachwissen und fachdidaktisches Wissen (nach Statusgruppen getrennt) mit Skalennittelwerten der Gesamtstichprobe



Bem.: ○/•: Nicht-Gymnasium; Δ/▲: Gymnasium; die Gitterlinien kennzeichnen jeweils die Mittelwerte der Gesamtstichprobe. Produkt-Moment-Korrelationen von FW und FDW bei Studierenden ($n = 118$) und Lehrkräften ($n = 65$): $r = 0,59^{**}$ bzw. $r = 0,53^{**}$; signifikanter Zusammenhang * $p \leq 0,05$; ** $p \leq 0,01$.

Tabelle 6.5.: Paarweise Gruppenunterschiede zwischen Gy-Lehrkräften, nGy-Lehrkräften, Gy-Studierenden und nGy-Studierenden im Fachwissen und fachdidaktischen Wissen

Fachwissen (FW)				Fachdidaktisches Wissen (FDW)			
M (SD)	Studierende		Lehrkräfte	M (SD)	Studierende		Lehrkräfte
	n = 86	d = 1,43**	n = 35		n = 86	d = 0,53**	n = 35
GY	10,54 (4,69)		17,56 (5,43)	GY	7,23 (3,23)		8,89 (2,84)
	d = 0,41*		d = 0,38		d = 0,46*		d = 0,31
	n = 32	d = 1,26**	n = 30		n = 32	d = 0,68*	n = 30
NGY	8,64 (4,42)		15,35 (6,18)	NGY	5,78 (2,89)		7,93 (3,47)

Bem.: M: arithm. Mittel; SD: Standardabweichung; GY: Gymnasium, NGY: Nicht-Gymnasium
Die Effektstärke d berechnet sich aus der Mittelwertdifferenz ($M_{\text{Lehrkräfte}} - M_{\text{Studierende}}$ bzw. $M_{\text{GY}} - M_{\text{NGY}}$) beider Gruppen dividiert durch die gepoolte Standardabweichung. Nach Cohen (1992) entspricht $d = 0,20$ einem kleinen, $d = 0,50$ einem mittleren und $d = 0,80$ einem großen Effekt. Signifikanter Unterschied: * $p \leq 0,05$; ** $p \leq 0,01$.

Von den 65 teilnehmenden Lehrkräften unterrichteten $n = 35$ an Gymnasien (Gy) und $n = 30$ an Real- oder Mittelschulen (nGy; vgl. hierzu Fußnote 2 in Kapitel 1). Der Grund für die Zusammenlegung von 24 Realschul- und sechs Mittelschullehrkräften liegt in der vergleichbaren universitären Ausbildung, falls Lehrkräfte der Mittelschule im Studium Physik als Unterrichts- beziehungsweise Didaktikfach (in diesem Fall zumindest bezüglich der didaktischen Ausbildung) gewählt haben (vgl. Abschnitt 4 in Krauss, Lindl, Schilcher & Tepner, 2017). Mittelschullehrkräfte ohne entsprechendes Physikstudium wurden in der Hauptstichprobe nicht berücksichtigt, da sie in der universitären Ausbildung keine professionellen Lerngelegenheiten für das Fach Physik hatten; sie werden daher als Kontraststichprobe behandelt (vgl. Abschnitt 6.6). Detailliertere Analysen (t-Tests) zeigen zwar, dass sich nur die Mittelschullehrkräfte mit studiertem Unterrichtsfach Physik nicht, wohl aber die Mittelschullehrkräfte mit Didaktikfach Physik signifikant von den Realschullehrkräften unterscheiden. Aufgrund der Unzuverlässigkeit von t-Tests bei sehr kleinen Stichproben (hierbei besteht die Gruppe der Didaktikfach-Mittelschullehrkräfte nur aus vier Testpersonen) werden die vier Datensätze dennoch in der nicht-gymnasialen Stichprobe belassen (vgl. Tab. A.19, A.20 und A.21 in Anhang A). Eventuell müsste die hier angewandte Gruppenzuordnung im Rahmen zukünftiger größerer Stichproben noch einmal überprüft werden.

Von den $n = 118$ Studierenden waren $n = 86$ für ein Lehramt an Gymnasien und $n = 32$ für ein Lehramt an Realschulen immatrikuliert. Mittelschulstudierende mit entsprechendem Studium konnten nicht untersucht werden, da das Unterrichts- beziehungsweise Didaktikfach Physik von dieser Studierendengruppe nur vergleichsweise selten gewählt wird, auch wenn in Bayern alle Lehrkräfte an der Mittelschule prinzipiell Physik im Fächerverbund mit Biologie und Chemie unterrichten müssen (für Details zum bayerischen Lehrerbildungssystem vgl. Abschnitt 4, Krauss, Lindl, Schilcher & Tepner, 2017).

In Tabelle 6.5 wird die Testleistung der vier Subgruppen ersichtlich: Demnach zeigen die Gy-Lehrkräfte sowohl beim FW als auch beim FDW eine höhere Testleistung als die nGy-Lehrkräfte, gefolgt von den Gy-Studierenden und den nGy-Studierenden (eine graphische Darstellung der insgesamt acht Gruppenunterschiede im FW und im FDW für alle FALKO-Fächer und COACTIV im Vergleich findet sich im Anhang zu (Lindl & Krauss, 2017)).

In zwei getrennten klassischen linearen Regressionsmodellen wurde darüber hinaus der Einfluss der unabhängigen Variablen Status (S vs. L) und Schulform (jeweils nGy vs. Gy) auf das FW und das FDW unter gleichzeitiger Kontrolle von Geschlecht und kognitiver Leistungsfähigkeit untersucht, wobei letztere über die Abiturnote der Studienteilnehmer approximiert wurde (Baron-Boldt et al., 1988). Wie Tabelle 6.6 zu entnehmen ist, wird ein Zusammenhang zwischen Abiturnote und den Testleistungen von $-0,20^{**}$ (FW) und $-0,22^{**}$ (FDW)¹ erhalten. Aufgrund der gegenläufigen Skalierung von Schulnoten sind diese Korrelationen negativ. Auch die Kovariate Geschlecht zeigt eine signifikante Korrelation mit beiden Wissenskategorien (FW: $0,29^{**}$; FDW: $0,19^{**}$).

¹Aufgrund fehlender Hypothesen wurden beim Regressionsmodell keine Interaktionen modelliert. Bei Implikation einer Interaktion zwischen Abiturnote und beruflichem Status würde diese bezüglich des FDW signifikant, da FDW und Abiturnote vor allem in der Gruppe der Studierenden assoziiert sind (vgl. auch Tab. 6.7).

Tabelle 6.6.: Regressionen von Fachwissen (FW) und fachdidaktischem Wissen (FDW) auf die Prädiktoren beruflicher Status und Schulform unter Kontrolle von Abiturnote und Geschlecht

AV	β_{Abitur}	$\beta_{\text{Geschlecht}}$	β_{Status}	$\beta_{\text{Schulform}}$	R^2
FW	-0,20**	0,29**	0,51**	0,06	0,40
FDW	-0,22**	0,19**	0,22**	0,08	0,15

Bem.: Geschlecht: 0 = weiblich, 1 = männlich; berufliche Stellung: 0 = Studierende, 1 = Lehrkräfte; Schulformvariable: 0 = Nicht-Gymnasium, 1 = Gymnasium; β : standardisierte Regressionsgewichte; R^2 : aufgeklärte Varianz. Da einige Personen fehlende Werte bei einem oder mehreren Prädiktoren aufweisen, wurde eine Full-Information-Maximum-Likelihood-Prozedur angewandt (Enders, 2010). Signifikantes Regressionsgewicht: * $p \leq ,05$; ** $p \leq ,01$.

Tabelle 6.7.: Produkt-Moment-Korrelationen (nullter Ordnung) zwischen dem Professionswissen und Indikatoren des Schul-/Studienerfolgs bzw. der Berufs-/Studiendauer

$N = 183$	Abiturnote	Fachsemester	Note 1. StEx.	Note 2. StEx.
Stu. ($n = 118$)				
FW (18)	-0,28**	0,58**		
FDW (10)	-0,31**	0,40**		
Leh. ($n = 65$)		Berufsjahre		
FW (18)	-0,18	0,28*	-0,22	0,10
FDW (10)	-0,11	0,01	-0,24	-0,17

Bem.: FW: Fachwissen; FDW: Fachdidaktisches Wissen. Da einige Personen fehlende Werte bei einem oder mehreren Prädiktoren aufweisen, wurde eine Full-Information-Maximum-Likelihood-Prozedur angewandt (Enders, 2010). Signifikanter Zusammenhang: * $p \leq 0,05$; ** $p \leq 0,01$.

6.4.4. Zusammenhänge des fachbezogenen Professionswissens mit Indikatoren des Schul- und Studienerfolgs sowie der Berufs- und Studiendauer: Kriteriale Validität

Interessante Aufschlüsse liefern auch Analysen zu Zusammenhängen des Professionswissens mit Indikatoren des Ausbildungserfolgs oder der Berufs- respektive Studiendauer (Produkt-Moment-Korrelationen nullter Ordnung). Der auffallendste Befund in Tabelle 6.7 ist der hohe Zusammenhang zwischen Studiendauer (Fachsemester) und fachbezogenem Professionswissen in der Gruppe der Studierenden (FW: $r = 0,58^{**}$; FDW: $r = 0,40^{**}$).

Weiterhin zeigt sich eine Korrelation mit der Abiturnote von $r = -0,28^{**}$ für das FW und $r = -0,31^{**}$ für das FDW bei Studierenden, es ist jedoch kein signifikanter Zusammenhang mit den Noten des ersten oder zweiten Staatsexamens feststellbar. In der Gruppe der Lehrkräfte ist dagegen kein signifikanter Zusammenhang zwischen Berufserfahrung und FDW nachzuweisen, jedoch scheinen Lehrkräfte mit mehr Berufserfahrung tendenziell über ein höheres FW ($r = 0,28^*$) zu verfügen. In der Tat ist die ausbleibende Zunahme didaktischen Wissens im Beruf jedoch alles andere als singulär. Dieses Resultat wurde in sämtlichen weiteren FALKO-Disziplinen ebenso gefunden wie auch in COACTIV und in allen drei naturwissenschaftlichen Fächern (Biologie, Chemie, Physik) bei ProwiN. Erstaunlicherweise ist ganz im Gegenteil sogar die signifikante Korrelation von FW mit

Tabelle 6.8.: Manifeste Korrelationen (nach Pearson) der vier separaten Skalen nach Inhalt (FDW + FW)

	Mechanik (6 Items)	E-Lehre (15 Items)	Optik (5 Items)	Wärmelehre (6 Items)
Mechanik (6 Items)	1,00			
E-Lehre (15 Items)	0,57	1,00		
Optik (5 Items)	0,41	0,53	1,00	
Wärmelehre (6 Items)	0,41	0,62	0,48	1,00

der Berufserfahrung singular, die wahrscheinlich auf die bereits oben beschriebene stark unterrichtsnahe Operationalisierung des Fachwissens – inklusive dezidierter Items auf Schulniveau – zurückzuführen ist. Diese Befunde werden in der Diskussion noch einmal aufgegriffen und vertieft.

6.5. Repräsentativitätshypothese

Um die Hypothese der Repräsentativität des Professionswissens in Mechanik für die Gesamtheit der Teilbereiche der Physik zu untersuchen, wurden separate Skalen (FDW + FW; FDW; FW) für die vier Fachbereiche Mechanik, Elektrizitätslehre, Optik und Wärmelehre erstellt (die Reliabilitäten der gebildeten Skalen sind Tabelle A.22 in Anhang A zu entnehmen). Die Korrelationen (ersichtlich in Tab. 6.8) zwischen den einzelnen Inhaltsbereichen liegen mit $r = 0,40$ bis $r = 0,62$ alle lediglich im schwachen bis moderaten Bereich (Cohen, 1988). Gleiches spiegelt sich auch bei getrennten Summenscores für das fachdidaktische und das Fachwissen wider (vgl. Tab. A.24 und A.25 in Anhang A).

6.6. Konstruktvalidität: Diskriminante Validität

Im Folgenden werden im Sinne einer Konstruktvalidierung Gruppenvergleiche zwischen Teilen der Hauptstichprobe (nGy- und Gy-Lehrkräfte, $n = 30$ bzw. $n = 35$) und Kontrastpopulationen berichtet. Die in diesem Fall (diskriminante Validität) untersuchten Kontrastpopulationen wurden anhand theoretischer Überlegungen gewählt. Sofern das FALKO-P-Testinstrument das fachspezifische Professionwissen (FDW, FW) von Physiklehrkräften misst, sollten die für die Gruppenvergleiche mit den Kontrastpopulationen aufgestellten Hypothesen (vgl. Abschnitt 3.1 bzw. siehe unten) bestätigt werden können. Folgende Kontrastpopulationen wurden gewählt: Fachfremde Lehrkräfte an Mittelschulen ($n = 25$ mit Unterricht im Fach „PCB“ bzw. „NT“, aber ohne eine universitäre Ausbildung im Fach Physik), fachfremde Lehrkräfte an Realschulen und Gymnasien ($n_{\text{ges}} = 32$; Biologie: $n = 7$, Chemie: $n = 13$, Mathematik: $n = 12$; keiner der Mathematik-Lehrkräfte hatte Physik als Beifach), sowie Fachphysiker ($n = 29$ Bachelor- und Masterstudierende, Doktorierende, akademische Räte, Universitätsprofessoren). Selbstverständlich könnten die genannten Teilstichproben für exaktere Vergleichsanalysen weiter aufgeteilt werden. Jedoch wird aufgrund der dabei immer kleiner werdenden Stichprobengröße (für Gruppenvergleiche sollten die untersuchten Gruppen idealerweise auch ähnlich groß sein, vgl. Field, 2013) darauf verzichtet. Im Falle des FDW wird weiterhin die Summenskala aus den

nach der Modelltestung verbliebenen zehn Items verwendet. Die Studienteilnehmer aller Kontrastpopulationen bearbeiteten die Testitems unentgeltlich einzeln oder in Kleingruppen unter der Aufsicht eines geschulten Testleiters. Auch ihnen war die Verwendung von Hilfsmitteln wie Taschenrechner und Formelsammlungen untersagt. Als Testzeit wurden erneut ca. 90 Minuten veranschlagt, wobei wenige Minuten Überziehungszeit aufgrund der Testkonzeption als Power-Test (Bühner, 2011) gewährt wurden.

Für alle genannten Teilstichproben wurde mittels konfirmatorischer Faktorenanalyse (Bayes-Modellierung aufgrund der geringen Stichprobenzahl; für weitere Informationen zur bayesianischen Datenanalyse vgl. Gelman et al., 2013) geprüft, ob das FALKO-Modell in seiner eingeschränkten Version (ohne die FDW-Subfacette „M & E“) auch dort Gültigkeit hat. Auch wenn für die Stichproben der fachfremden RS- und Gy-Lehrkräfte und der Fachphysiker eine Modellgültigkeit nicht gezeigt werden konnte, werden trotzdem Gruppenvergleiche vorgenommen, da das Modell in allen anderen Teilstichproben und insbesondere auch innerhalb der Subgruppen der Hauptstudie Gültigkeit hat (die Outputs der Modellprüfung finden sich in den Tabellen A.44, A.45, A.46, A.47, A.48, A.49, A.50, A.51, A.52, A.53, A.54, A.55, A.56, A.57 und A.58 in Anhang A).

Die mit den Gruppenvergleichen zu prüfenden Hypothesen lauten für das FDW wie folgt (vgl. Kap. 3 und 7):

Hypothese 4a Physiklehrkräfte verfügen über mehr fachdidaktisches Wissen als Mittelschullehrkräfte, welche im Physik fachfremd im Fächerverbund „PCB“ beziehungsweise „NT“ unterrichten.

Hypothese 5a Physiklehrkräfte verfügen über mehr fachdidaktisches Wissen als Lehrkräfte (RS, Gy), die Biologie, Chemie oder Mathematik unterrichten.

Hypothese 6a Gymnasiale und nicht-gymnasiale Physiklehrkräfte verfügen über mehr fachdidaktisches Wissen als Fachphysiker.

Für das FW werden folgende Hypothesen aufgestellt (vgl. auch Kap. 3 und 7):

Hypothese 4b Physiklehrkräfte verfügen über mehr Fachwissen als Mittelschullehrkräfte, welche im Physik fachfremd im Fächerverbund „PCB“ beziehungsweise „NT“ unterrichten.

Hypothese 5b Physiklehrkräfte verfügen über mehr Fachwissen als Lehrkräfte (RS, Gy), die Biologie, Chemie oder Mathematik unterrichten.

Hypothese 6b Gymnasiale Physiklehrkräfte verfügen ein ähnliches, nicht-gymnasiale Physiklehrkräfte über weniger Fachwissen als Fachphysiker.

Wie bereits berichtet, bestand die Gruppe der $n = 65$ Physiklehrkräfte aus $n = 35$ gymnasialen Lehrkräften (Gy) sowie $n = 30$ nicht-gymnasialen Lehrkräften (nGy; Realschule und Mittelschule). Eine weitere Unterteilung dieser Gruppe erfolgte für die Gruppenvergleiche zu den fachfremden Mittelschullehrkräften und den fachfremden Realschul- und gymnasialen Lehrkräften nicht, da die Unterschiede in der mittleren Testleistung zwischen den beiden Gruppen sowohl für das FDW als auch für das FW nicht besonders groß ausfielen und die Effektstärken auch nicht signifikant wurden (vgl. Abb. 6.5 in Abschn. 6.4.3) ist. Lediglich für den Gruppenvergleich mit den Fachphysikern wurde eine Aufteilung der Stichprobe der Physiklehrkräfte in Gy und nGy vorgenommen, aufgrund

Tabelle 6.9.: Gruppenunterschiede zwischen Lehrkräften (Gy, nGy) und fachfremden Mittelschullehrkräften im Fachwissen (FW) und fachdidaktischen Wissen (FDW)

	Lehrkräfte Gy + nGy (n = 65)		fachfremde Lehrkräfte MS (n = 25)		Effektstärke d (Gy + nGy Leh. vs. ff MS-Leh.)	p-Wert
	<i>M</i> (<i>SD</i>)	emp. Max.	<i>M</i> (<i>SD</i>)	emp. Max.		
<i>N</i> = 90						
FW (18)	16,54 (5,85)	27	6,64 (5,05)	19,5	1,75 ^{***}	< ,001
FDW (10)	8,45 (3,16)	16	4,64 (3,91)	13	1,13 ^{***}	< ,001
E & R (5)	4,72 (1,93)	9	1,84 (2,10)	7	1,46 ^{***}	< ,001
Schk (5)	3,72 (1,81)	8	2,80 (2,40)	8	0,47 [*]	= ,05

M: arith. Mittel; *SD*: Standardabweichung; FW: Fachwissen; FDW: Fachdidaktisches Wissen; E & R: Erklären und Repräsentieren; Schk: Schülerkognition. Die Effektstärke *d* berechnet sich aus der Mittelwertdifferenz beider Gruppen ($M_{\text{LehGynGy}} - M_{\text{LehffMS}}$) dividiert durch die gepoolte Standardabweichung. Nach Cohen (1992) entspricht $d = 0,20$ einem kleinen, $d = 0,50$ einem mittleren und $d = 0,80$ einem großen Effekt. Signifikanter Unterschied: * $p \leq ,05$; ** $p \leq ,01$; *** $p \leq ,001$.

der Erwartung, dass aufgrund der fachlich ähnlichen Ausbildung Gy-Lehrkräfte eine mit Fachphysikern vergleichbare Leistung im FW erbringen (genauere Informationen zu dieser Stichprobenzusammensetzungen finden sich in Abschn. 6.1 und in den Tabellen A.26, A.27 und A.28 in Anhang A).

Physik-Lehrkräfte (Gy und nGy) versus fachfremde Mittelschullehrkräfte

Die Vergleichsgruppe der fachfremden Mittelschullehrkräfte bestand aus $n = 25$ Studienteilnehmern. Ihr Durchschnittsalter lag bei 35,22 Jahren ($SD = 4,34$; range 29 – 44) und die mittlere Berufserfahrung betrug 7,88 Jahre ($SD = 2,76$, range 4 – 17). 28 % waren weiblich (eine ausführlichere Stichprobenbeschreibung findet sich in der Tabelle A.29 in Anhang A). Für diesen Gruppenvergleich wurde jeweils ein t-Test ($n = 65$ Gy- und nGy-Lehrkräfte vs. $n = 25$ MS-Lehrkräfte) für die gesamte FW- und FDW-Skala sowie für die Skalen der beiden FDW-Subfacetten „E & R“ und „Schk“ durchgeführt.

Wie in Tabelle 6.9 ersichtlich, übertrafen die Gy- und nGy-Physiklehrkräfte die fachfremden Mittelschullehrkräfte im FW ($t(88) = 7,453$, $p \leq ,001$; $d = 1,75^{***}$) und FDW ($t(88) = 4,784$, $p \leq ,001$; $d = 1,13^{***}$) mit einer vergleichsweise hohen Effektstärke. Schlüsselt man weiter nach den FDW-Subfacetten auf, fällt auf, dass im Falle der „Schk“-Skala der Unterschied gerade noch signifikant wird ($t(88) = 6,193$, $p \leq ,001$; $d = 0,47^*$). Im Falle der „E & R“-Skala ist der Vorteil der Physiklehrkräfte dagegen wesentlich größer ($t(88) = 6,193$, $p \leq ,001$; $d = 1,46^{***}$).

Erwähnt sei in diesem Zusammenhang der durchaus interessante Vergleich zwischen den fachfremden Mittelschullehrkräften - diese unterrichten, wie bereits erwähnt, Physik ohne universitäres Studium im Fächerverbund „PCB“ beziehungsweise „NT“ - und den fachfremden Realschul- und Gymnasiallehrkräften. Erstere schneiden nämlich sowohl bei

Tabelle 6.10.: Gruppenunterschiede zwischen fachfremden Lehrkräften (RS, Gy) und fachfremden Mittelschullehrkräften im Fachwissen (FW) und fachdidaktischen Wissen (FDW)

	fachfremde Lehrkräfte MS ($n = 25$)		fachfremde Lehrkräfte RS + Gy ($n = 32$)		Effektstärke d (ff. Leh. RS + Gy vs. ff Leh. MS)	p -Wert
$N = 57$	M (SD)	emp. Max.	M (SD)	emp. Max.		
FW (18)	6,64 (5,05)	19,5	3,89 (2,10)	10	0,68*	< ,05
FDW (10)	4,64 (3,91)	13	1,69 (1,42)	5	0,96**	= ,001
E & R (5)	1,84 (2,10)	7	0,53 (0,80)	2	0,79**	< ,01
Schk (5)	2,80 (2,40)	8	1,16 (1,02)	4	0,86**	< ,01

Bem.: M : arith. Mittel; SD : Standardabweichung; FW: Fachwissen; FDW: Fachdidaktisches Wissen; E & R: Erklären und Repräsentieren; Schk: Schülerkognition. Die Effektstärke d berechnet sich aus der Mittelwertdifferenz beider Gruppen ($M_{\text{ffLehRS Gy}} - M_{\text{ffLehMS}}$) dividiert durch die gepoolte Standardabweichung. Nach Cohen (1992) entspricht $d = 0,20$ einem kleinen, $d = 0,50$ einem mittleren und $d = 0,80$ einem großen Effekt. Signifikanter Unterschied: * $p \leq ,05$; ** $p \leq ,01$; *** $p \leq ,001$.

Betrachtung der FW- ($t(31) = 2,555$, $p \leq ,05$; $d = 0,68^*$) und der FDW-Gesamtskala ($t(30) = 3,590$, $p = ,001$; $d = 0,96^{**}$) als auch bei Betrachtung der beiden FDW-Subskalen „E & R“ ($t(30) = 2,958$, $p \leq ,01$; $d = 0,79^{**}$) und „Schk“ ($t(31) = 3,208$, $p \leq ,01$; $d = 0,86^{**}$) jeweils wesentlich besser ab als die fachfremden RS- und Gy-Lehrkräfte (vgl. Tabelle 6.10), auch wenn der Unterschied für den Fall des FW am geringsten ausfällt.

Physik-Lehrkräfte (Gy und nGy) versus fachfremde Realschul- und Gymnasiallehrkräfte

Für die Vergleichsgruppe der fachfremden Realschul- und Gymnasiallehrkräfte ($n = 32$) ergab sich die folgende Stichprobenstruktur: Unter den 32 Studienteilnehmern, welche allesamt an bayerischen Schulen unterrichten, befanden sich $n = 7$ Biologie-, $n = 13$ Chemie- und $n = 12$ Mathematiklehrkräfte. Ihr Durchschnittsalter lag bei 24,39 Jahren ($SD = 2,95$; range 20 – 37) und die mittlere Berufserfahrung betrug 6,35 Jahre ($SD = 2,83$, range 2 – 12). 66 % waren weiblich (eine ausführlichere Stichprobenbeschreibung findet sich in der Tabelle A.30 in Anhang A). Erneut wurden jeweils ein t-Test ($n = 65$ Gy- und nGy-Lehrkräfte vs. $n = 32$ RS- und Gy-Lehrkräfte) für die gesamte FW- und FDW-Skala sowie für die Skalen der beiden FDW-Subfacetten „E & R“ und „Schk“ durchgeführt.

Sowohl in der FW- ($t(95) = 11,833$, $p \leq ,001$; $d = 2,65^{***}$) und FDW-Gesamtskala ($t(95) = 11,522$, $p \leq ,001$; $d = 2,49^{***}$) als auch in den beiden FDW-Subskalen „E & R“ ($t(95) = 11,756$, $p \leq ,001$; $d = 2,54^{***}$) und „Schk“ ($t(95) = 11,756$, $p \leq ,001$; $d = 1,61^{***}$) zeigte sich sehr deutlich der Vorteil der Physiklehrkräfte gegenüber den fachfremden Lehrkräften (vgl. Tabelle 6.11), der aber wiederum für die „Schk“-Skala am geringsten

Tabelle 6.11.: Gruppenunterschiede zwischen Lehrkräften (Gy, nGy) und fachfremden Realschul- und Gymnasiallehrkräften im Fachwissen (FW) und fachdidaktischen Wissen (FDW)

	Lehrkräfte Gy + nGy ($n = 65$)		fachfremde Lehrkr. RS + Gy ($n = 32$)		Effekt- stärke d (Leh. Gy + nGy vs. ff. Leh. RS + Gy)	p -Wert
$N = 97$	$M (SD)$	emp. Max.	$M (SD)$	emp. Max.		
FW (18)	16,54 (5,85)	27	3,89 (2,10)	10	2,65 ^{***}	< ,001
FDW (10)	8,45 (3,16)	16	1,69 (1,42)	5	2,49 ^{***}	< ,001
E & R (5)	4,72 (1,93)	9	0,53 (0,80)	2	2,45 ^{***}	< ,001
Schk (5)	3,72 (1,81)	8	1,16 (1,02)	4	1,61 ^{***}	< ,001

Bem.: M : arith. Mittel; SD : Standardabweichung; FW: Fachwissen; FDW: Fachdidaktisches Wissen; E & R: Erklären und Repräsentieren; Schk: Schülerkognition. Die Effektstärke d berechnet sich aus der Mittelwertdifferenz beider Gruppen ($M_{\text{LehGynGy}} - M_{\text{ffRSgy}}$) dividiert durch die gepoolte Standardabweichung. Nach Cohen (1992) entspricht $d = 0,20$ einem kleinen, $d = 0,50$ einem mittleren und $d = 0,80$ einem großen Effekt. Signifikanter Unterschied: * $p \leq ,05$; ** $p \leq ,01$; *** $p \leq ,001$.

ausfällt.

Physik-Lehrkräfte (getrennt nach Gy und nGy) versus Fachphysiker

Wie weiter oben bereits konstatiert, wurde die Stichprobe an Physik-Lehrkräften ($n = 65$) für den Vergleich mit der Konstrastgruppe der Fachphysiker ($n = 29$) hypothesenkonform weiter in gymnasiale (Gy) und nicht-gymnasiale (nGy) Lehrkräfte aufgeteilt. Die 35 Gy-Lehrkräfte waren im Mittel 46,11 Jahre alt ($SD = 8,40$; range 33 – 70) und hatten eine mittlere Berufserfahrung von 15,94 Jahren ($SD = 7,81$; range 5 – 40). 20% waren weiblich. Die 30 nGy-Lehrkräfte hatten ein Durchschnittsalter von 38,47 Jahren ($SD = 11,62$; range 26 – 64) und die durchschnittliche Berufserfahrung betrug 10,73 Jahre ($SD = 10,77$; range 1 – 40). 47% waren weiblich. Die Vergleichsgruppe der Fachphysiker bestand aus $n = 29$ Studienteilnehmern der Universität Regensburg. Ihr Durchschnittsalter lag bei 32,48 Jahren ($SD = 12,55$; range 22 – 74). Die mittlere Berufserfahrung wird nicht angegeben, da die Stichprobe sowohl Studierende als auch bereits Berufstätige enthält und diese Angabe somit wenig aussagekräftig wäre. 17% waren weiblich (eine ausführlichere Stichprobenbeschreibung findet sich in der Tabelle A.31 in Anhang A). Da es sich hier nun um drei unterschiedliche Gruppen handelt, wurde auf eine einfaktorielle ANOVA sowie nach Hochberg korrigierte Post-hoc-Tests nach Hochberg (Field, 2013) anstelle eines t-Tests zurückgegriffen.

Die untersuchte Gesamtstichprobe von $n = 94$ wurde wie angekündigt nach gymnasialen und nicht-gymnasialen Lehrkräften sowie nach Fachphysikern aufgeteilt und deren Ergebnisse deskriptiv und inferenzstatistisch mittels einer einfaktoriellen Varianzanalyse mit der abhängigen Variable Wissenskatgorie (FDW, FW) und der unabhängigen Variable Gruppenzugehörigkeit (Gy, nGy, FPh) einander gegenübergestellt. Die Levene-Tests

Tabelle 6.12.: Gruppenunterschiede zwischen Gy-Lehrkräften, nGy-Lehrkräften und Fachphysikern im Fachwissen (FW) und fachdidaktischen Wissen (FDW)

	Gy-Lehrkr. (<i>n</i> = 35)		nGy-Lehrkr. (<i>n</i> = 30)		Fachphysiker (<i>n</i> = 29)		F (2,91)	<i>p</i> -Wert
	<i>M</i> (<i>SD</i>)	emp. Max.	<i>M</i> (<i>SD</i>)	emp. Max.	<i>M</i> (<i>SD</i>)	emp. Max.		
<i>N</i> = 94								
FW (18)	17,56 (5,43)	27	15,35 (6,18)	24	17,03 (4,69)	25	1,40	= ,25
FDW (10)	8,89 (2,84)	16	7,93 (3,47)	16	6,93 (2,53)	12	3,44	≤ ,05
E&R (5)	4,94 (1,76)	8	4,47 (2,11)	9	3,72 (1,53)	7	3,577	≤ ,05
Schk (5)	3,94 (1,68)	8	3,47 (1,94)	7	3,21 (1,73)	6	1,528	= ,22

Bem.: *M*: arith. Mittel; *SD*: Standardabweichung; FW: Fachwissen; FDW: Fachdidaktisches Wissen; E & R: Erklären und Repräsentieren; Schk: Schülerkognition. Signifikanter Unterschied: **p* ≤ ,05.

sind bei keiner der durchgeführten ANOVAs signifikant, sodass von Varianzhomogenität ausgegangen werden kann.

Die Mittelwerte für die Testleistung im FDW ergaben sich wie folgt. Gy: *M* = 8,89 (*SD* = 2,84), nGy: *M* = 7,93 (*SD* = 3,47) und FPh: *M* = 6,93 (*SD* = 2,53). Die ANOVA sowie der Post-hoc-Test nach Hochberg ergab, dass sich die nGy-Lehrkräfte nicht signifikant von den Gy-Lehrkräften und auch nicht signifikant von den Fachphysikern unterscheiden. Lediglich die Gy-Lehrkräfte schnitten signifikant besser als die Fachphysiker ab.

Bei Ausführung von nach den FDW-Subfacette getrennten ANOVAs wurden im Fall der FDW-Subfacette „E & R“ folgende Mittelwerte erhalten. Gy: *M* = 4,94 (*SD* = 1,76), nGy: *M* = 4,47 (*SD* = 2,11) und FPh: *M* = 3,72 (*SD* = 1,53). Die durchgeführte ANOVA und der Post-hoc-Test nach Hochberg zeigte, wie schon bei der FDW-Gesamtskala, dass es keinen signifikanten Unterschied in der Testleistung zwischen den nGy- und den Gy-Lehrkräften, sowie zwischen den nGy-Lehrkräften und den Fachphysikern gab. Lediglich der Unterschied zwischen den Gy-Lehrkräften und Fachphysikern wurde signifikant. Für die FDW-Subfacette „Schk“ ergaben sich folgende Mittelwerte: Gy: *M* = 3,94 (*SD* = 1,68), nGy: *M* = 3,47 (*SD* = 1,94) und FPh: *M* = 3,21 (*SD* = 1,73). Für die mittlere Testleistung im FW wurden folgende Werte erhalten: Gy: *M* = 17,56 (*SD* = 5,43), nGy: *M* = 15,35 (*SD* = 6,18) und FPh: *M* = 17,03 (*SD* = 4,69). Keiner der Mittelwertsunterschiede waren der ANOVA und dem Post-hoc-Test nach Hochberg folgend signifikant.

Im Fall der FW-Skala wurden folgende Mittelwerte erhalten. Gy: *M* = 17,56 (*SD* = 5,43), nGy: *M* = 15,35 (*SD* = 6,18) und FPh: *M* = 17,03 (*SD* = 4,69). Wie bei der FDW-Subskala „Schk“ wurden die Gruppenunterschiede nicht signifikant.

Eine Zusammenstellung der Ergebnisse zu den einzelnen Gruppenvergleichen ist in der Tabelle 6.12 dargestellt.

Um weitergehend zu prüfen, ob sich für die Statusgruppe der Lehrkräfte insgesamt im Vergleich mit den Fachphysikern Unterschiede in der jeweiligen Testleistung ergeben, wurden die obigen Gruppenvergleiche nun noch mit der Gesamtstichprobe an Lehrkräften (*n* = 65) durchgeführt. Hierzu wurde für die beiden Wissenskategorien FDW und FW

sowie für die beiden FDW-Subfacetten „E & R“ und „Schk“ jeweils ein t-Test gerechnet. Für die FDW-Skala ergab sich für die Stichprobe der Lehrkräfte ein Mittelwert von $M = 8,45$ ($SD = 3,16$) und für Stichprobe der Fachphysiker ein Wert von $M = 6,93$ ($SD = 2,53$). Dieser Unterschied zwischen beiden Gruppen ist signifikant ($t(92) = 2,275$, $p < ,05$; $d = 0,51^*$). Im Falle der beiden FDW-Subfacetten „E & R“ und „Schk“ ergaben sich für die Gruppe der Lehrkräfte die Mittelwerte $M = 4,72$ ($SD = 1,93$) und $M = 3,72$ ($SD = 1,80$) sowie die Mittelwerte $M = 3,72$ ($SD = 1,53$) und $M = 3,21$ ($SD = 1,50$) für die Gruppe der Fachphysiker. Der Unterschied in der ersten FDW-Subfacette ist dabei signifikant ($t(92) = 2,458$, $p < ,05$; $d = 0,55^*$), derjenige in der zweiten nicht ($t(92) = 1,345$, $p = ,182$; $d = 0,30$). Für den Unterschied der beiden Gruppen in Bezug auf das FW wurden die folgenden Mittelwerte erhalten: $M = 16,54$ ($SD = 5,85$) für die Lehrkräfte und $M = 17,03$ ($SD = 4,69$) für die Fachphysiker. Der Gruppenunterschied ist nicht signifikant ($t(92) = -0,402$, $p = ,689$; $d = -0,09$).

6.7. Konstruktvalidität: Konvergente Validität

In diesem Abschnitt werden Ergebnisse der Suche nach Hinweisen auf konvergente Validität präsentiert. Hierzu bearbeitete eine Gruppe von Studierenden des Lehramts Physik ($n_{\text{ges}} = 28$) für Realschule ($n = 15$) und Gymnasium ($n = 13$) neben dem Test FALKO-P (FDW und FW) zusätzlich die Tests zum FDW (Gramzow, 2014) und zum FW (Woitkowski, 2015) aus dem Projekt ProfiLe-P. Die gesamte Stichprobe ($n_{\text{ges}} = 28$) war bei FALKO-P-Teilnahme durchschnittlich 23,46 ($SD = 2,13$, range: 20 – 28) Jahre alt und befand sich im Mittel im 7. Fachsemester ($M = 6,57$, $SD = 2,94$, range: 1 – 11). Bei der Teilnahme an den ProfiLe-P-Tests betrug das mittlere Alter 24,14 ($SD = 2,38$, range: 21 - 29) Jahre. Die durchschnittliche Studiendauer lag bei 8,32 Semestern ($SD = 3,16$, range: 2 - 13). 36% der Studienteilnehmer waren weiblich. Für die Teilstichprobe der gymnasialen ($n = 13$) und nicht-gymnasialen ($n = 15$) Studierenden ergaben sich folgende deskriptive Ergebnisse (detailliertere Informationen zur Stichprobenzusammensetzung sind Tabelle A.32 in Anhang A zu entnehmen):

Gymnasiale Studierende: Mittleres Alter von 24,00 Jahren ($SD = 2,58$, range: 20 – 28) bei FALKO-P- und mittleres Alter von 24,85 Jahren ($SD = 2,58$, range: 21 – 29) bei ProfiLe-P-Teilnahme. Durchschnittliche Studiendauer im ersteren Fall von 8,62 ($SD = 1,76$, range: 6 – 11) beziehungsweise 10,46 ($SD = 1,61$, range: 9 – 13) Semestern. 31% dieser Teilstichprobe war weiblich.

Nicht-gymnasiale Studierende: Durchschnittliches Alter von 23,00 Jahren ($SD = 1,60$, range: 21 – 27) bei FALKO-P- und durchschnittliches Alter von 23,53 Jahren ($SD = 2,10$, range: 21 – 29) bei ProfiLe-P-Teilnahme. Mittlere Studiendauer im ersteren Fall von 4,80 ($SD = 2,60$, range: 1 – 11) beziehungsweise 6,47 ($SD = 3,02$, range: 2 – 13) Semestern. 40% der Realschul-Studierenden waren weiblich.

Wie bereits in Abschnitt 6.1 berichtet, fand die erste Erhebung (FALKO-P) innerhalb regulärer universitärer Veranstaltungen unentgeltlich statt. Die 2. Teilnahme (ProfiLe-P) erfolgte außerhalb von regulären Veranstaltungen einzeln oder in Kleingruppen entgeltlich (die Studienteilnehmer erhielten 15,- Euro Aufwandsentschädigung) unter Aufsicht eines geschulten Testleiters und dauerte in etwa 90 Minuten. Die Verwendung von Hilfsmitteln (Taschenrechner, Formelsammlung, ...) war auch in diesem Fall nicht gestattet.

Anhand der deskriptiven Ergebnisse wird ersichtlich, dass zwischen der Bearbeitung der Testinstrumente beider Projekte im Mittel ein Zeitraum von ca. zwei Semestern lag. Dies ist der Tatsache geschuldet, dass die Gruppe der Physik-Lehramtsstudierenden an

der Universität Regensburg durch die Bearbeitung des FALKO-P-Testinstruments nahezu vollständig ausgeschöpft war und somit nur Studierende dieser Teilstichprobe um eine weitere Testteilnahme gebeten werden konnten. Idealerweise sollte eine Kreuzvalidierung möglichst zum selben Messzeitpunkt stattfinden, um eine Verzerrung der Daten vor allem durch zeitliche Faktoren zu vermeiden. Um die sich nun aber ergebende zeitliche Verzerrung innerhalb der statistischen Analysen möglichst zu berücksichtigen - schließlich werden die Studierenden in den zwei weiteren Studiensemestern höchstwahrscheinlich von den zusätzlich wahrgenommenen Lerngelegenheiten profitiert haben - wurden für die Gruppenvergleiche neben t-Tests einfaktorielles Kovarianzanalysen (ANCOVAs) unter Kontrolle des Fachsemesters durchgeführt.

Um zu untersuchen, inwiefern die Skalen der jeweiligen Testinstrumente zusammenhängen, wurden die Summenscores der Skalen (FW und FDW einschließlich der Subskalen) manifest (nach Pearson) und latent (CFA, bayesianische Modellierung) gegeneinander korreliert. Denn Ziel dieser Teil-Untersuchung war es, herauszufinden, ob beide Tests jeweils das gleiche Konstrukt messen. Interessant gestaltet sich diese Frage unter anderem deshalb, da zwar jeweils beide Tests vorgeben, das FDW beziehungsweise das FW im Fach Physik zu messen, dennoch aber zum Teil deutliche Unterschiede in der Modellierung des Professionswissens beziehungsweise in der Operationalisierung der Wissenskategorien und Subfacetten in den Testitems vorliegen (vgl. Kap. 4 und weiter unten im Text). Erwähnt werden muss in diesem Zusammenhang, dass sowohl der FDW- als auch der FW-Test aus konzeptionellen (im FALKO-Modell wurden nicht alle bei ProfiLe-P modellierten FDW-Subfacetten berücksichtigt) und ökonomischen Gründen (die gesamte Bearbeitungszeit sollte auch in diesem Fall 90 Minuten nicht überschreiten) gekürzt wurde, mit der Vorgabe, das ProfiLe-P-Modell trotzdem möglichst genau in den jeweils verbleibenden Testitems abzubilden. Gleichzeitig wurde eine möglichst exakte Passung zum FALKO-Modell (berücksichtigte Subfacetten, Itemzahl, ...) angestrebt.

Bevor die Ergebnisse der Kreuzvalidierung berichtet werden, erfolgt eine kurze Darstellung des ProfiLe-P-Projekts mit seinem Modell des Professionswissens im Fach Physik sowie eine Beschreibung des Vorgehens bei der Zusammenstellung der Items für die beiden Tests zur Erhebung des FDW und des FW.

Ziel des Forschungsverbundes ProfiLe-P („Professionswissen in der Lehramtsausbildung Physik“) war es, ein Rahmenmodell des Professionswissens von angehenden Physiklehrkräften sowie von Übungsleitern der Physik zu entwickeln. Unter anderem sollten folgende Teilziele verwirklicht werden: Modellierung der professionsspezifischen Wissensbereiche, welche Studierende während der universitären Ausbildung erwerben können (inhaltliche Schwerpunktsetzung „Mechanik“ als Teilgebiet der Physik), die Entwicklung von validen Testinstrumenten sowie die Durchführung von Zusammenhangsanalysen in Bezug auf die konzeptualisierten Wissensbereiche (Gramzow, 2014). Das ProfiLe-P-Modell sollte dabei deklarative, analytische und prozedurale Aspekte des FDW sowie das FW im Fach Physik umfassen. Unterteilt war das Projekt in drei Teilprojekte (vgl. auch Tab. 2.1 in Abschn. 2.2): Zum einen das Teilvorhaben „DaWis“, in welchem analytische und deklarative Aspekte des FDW fokussiert wurden, zum anderen „EWis“, das sich mit prozeduralen Komponenten des FDW beschäftigte sowie das Teilprojekt „FaWis“, welches sich bei seinen Untersuchungen dem physikalischen Fachwissen widmete. Die während der Projektlaufzeit in allen drei Teilvorhaben entwickelten Testinstrumente wurden an Physik-Lehramtsstudierenden aller Schulstufen sowie Übungsgruppenleitern validiert. Als weitere Kontraststichproben wurden Vollfachstudierende der Physik sowie Lehramtsstudierende der Mathematik herangezogen (Literatur Gesamt-ProfiLe-P: z. B. Kulgemeyer et al., 2012).

Die beiden für die Kreuzvalidierung eingesetzten Testinstrumente sind den Teilprojekten „DaWis“ und „FaWis“ sowie deren Weiterentwicklungen zuzuordnen. Hauptziele des ersteren Teilprojekts waren die Entwicklung eines möglichst differenzierten Modells zur Beschreibung des FDW (Zielgruppe Lehramtsstudierende der Physik) sowie die Entwicklung eines modellkonformen, validen Testinstruments, welches detaillierte Aussagen zu Zusammenhängen zwischen den Dimensionen des Professionswissens in Bezug auf das Gesamtprojekt, aber auch Aussagen zur inneren Struktur des FDW ermöglichen sollte (Gramzow, 2014). Ursprünglich umfasste das zweidimensionale Modell des FDW acht Subfacetten, welche auf unterschiedliche Themengebiete bezogen werden konnten. Für die Testkonstruktion bei „DaWis“ wurde die Achse der Subfacetten auf die vier FDW-Subfacetten „Instruktionsstrategien“ („ib/id“), „Schülvorstellungen“ („s“), „Experimente und Vermittlung eines angemessenen Wissenschaftsverständnisses“ („ex“) sowie „fachdidaktische Konzepte“ reduziert. Weiterhin erfolgte eine Fokussierung auf den Inhaltsbereich „Mechanik“ sowie die Einführung einer dritten Dimension, welche die kognitiven Anforderungen „reproduzieren“, „anwenden“ und „analysieren“ abbildet (Gramzow, 2014).

Im Fokus des zweiten Teilprojekts stand das FW von Fachphysik- und Physiklehramtstudierenden. Zentrale Zielsetzungen waren zum Beispiel die Modellierung und Konzeptualisierung verschiedener Wissensstufen beziehungsweise -facetten innerhalb des FW, die Erweiterung der bereits im Projekt ProfiLe-P gewonnenen Erkenntnisse auf den Bereich des Fachphysikstudiengangs sowie die differenzierte Diagnose des FW mittels eines im Teilprojekt konstruierten und erprobten Testinstruments (Woitkowski, 2015). Das Modell umfasst fachliches Wissen sowohl in der Schul- als auch in der Hochschulphysik und splittet sich in die beiden Subfacetten „Nutzung von Fachwissen“ („FW“) sowie „Nutzung von Experimenten und Modellen“. Das Teilmodell „Nutzung von Fachwissen“ ist dreidimensional angelegt und umfasst die Achsen „Inhaltsbereich“ („Kinematik“, „Kraft“, „Energie/Impuls“), „Fachstufen“ („Schulwissen“, „vertieftes Wissen“, „universitäres Wissen“) sowie „hierarchische Komplexität“ („Fakten (I)“, „Prozessbeschreibung (II)“, „lineare Kausalität (III)“, „multivariate Interdependenz (IV)“). Auch das Teilmodell „Nutzung von Experimenten und Modellen“ („EM“) ist dreidimensional angelegt und bildet in seinen drei Achsen „Inhaltsbereich“ („Kinematik“, „Kraft“, „Energie/Impuls“), „Suchraum“ („Hypothesen-Raum“, „Experiment-Raum“) und ebenfalls „hierarchische Komplexität“ („Fakten (I)“, „Prozessbeschreibung (II)“, „lineare Kausalität (III)“, „multivariate Interdependenz (IV)“) ab (Woitkowski, 2015).

Wie bereits konstatiert, sollten trotz Item-Reduktion die beiden ProfiLe-P-Modelle so exakt wie möglich abgebildet, gleichzeitig aber zur Durchführung der Gruppenvergleiche eine möglichst gute Passung zum FALKO-Modell erreicht werden. Dies setzte ein genaues Studium der Items sowie ein gezieltes Abwägen unter Berücksichtigung der Modelle voraus. Für den Fall des FDW-Testteils wurden die Items aus Testheft „A“ verwendet, da diese in der Dissertationsschrift von Gramzow (2014) frei zugänglich und dort inklusive der Kodieranweisungen detailliert erläutert sind. Da die Subfacette „fachdidaktische Konzepte“ des DaWis-Modells keine getrennte Berücksichtigung im FALKO-Modell fand und die restlichen drei Sub-Facetten am ehesten der FALKO-P-Konzeption entsprechen, wurden die ihr zugehörigen Items entfernt. Da die Achse „kognitive Anforderungen“ im FALKO-P-Modell ebenfalls keine explizite Berücksichtigung erhielt, fand diesbezüglich keine Itemauswahl statt. Stattdessen wurden alle verbliebenen Items ($N = 20$) der drei Sub-Facetten in das Testinstrument für die Kreuzvalidierung mit einbezogen.

Die Auswahl der FW-Items aus dem Modell von Woitkowski (2015) gestaltete sich komplizierter. Explizit wurden die Items des Teilmodells „Nutzung von Fachwissen“ („FW-

Skala“) als interessant erachtet, da diese konzeptuell am ehesten zum FALKO-P-Modell zu passen schienen. In einem ersten Schritt wurde hierbei entschieden, auf die Items zum „universitären Wissen“ zu verzichten, da keine entsprechende Modellierung im FALKO-P-Modell existiert. Gleiches gilt für die Subfacette „Hypothesen-Raum“ des Teilmodells „Nutzung von Experimenten und Modellen“. Lediglich ein kleiner Teil der Items der Subfacette „Experiment-Raum“ sollte für die Kreuzvalidierung berücksichtigt werden, da diesbezüglich die Korrelation zur FALKO-P-Subfacette „M & E“ untersucht werden sollte, welche ja nach detaillierten Modelltestungen eher zur FW-Skala als zur FDW-Skala zu gehören scheint (vgl. Abschn. 6.2). Da die Gesamtzahl an Testitems bei den Testungen im ProfiLe-P-Projekt, einem Multi-Matrix-Design folgend, auf unterschiedliche Blöcke aufgeteilt und anschließend zu unterschiedlichen Testheften zusammengefasst wurden, musste letztendlich eine Itemauswahl übergreifend aus allen Blöcken vorgenommen werden, um das Modell möglichst adäquat abbilden zu können. Bei der finalen Auswahl der Items für die „FW-Skala“ wurde zum einen das Zahlenverhältnis aller konzeptualisierten Testitems hinsichtlich der „Inhaltsbereiche“ und der „hierarchischen Komplexität“, zum anderen die in der Dissertation von Woitkowski (2015) berichteten Itemfits berücksichtigt. Weiterhin spielte die Zugehörigkeit zu den Subfacetten „Schulwissen“ und „vertieftes Schulwissen“ eine Rolle, denn das Verhältnis aus Items auf dem Niveau des „Schulwissens“ und des „vertieften Schulwissens“ sollte möglichst exakt dem bei FALKO-P entsprechen. Unter Berücksichtigung aller soeben genannten Kriterien fanden letztendlich die Items A3, A7, B2, B4, B7a/b, C4a, C5, D7, E6, F1, F3b, G3a, G4, G5, G6, G7, I1, J3b und J6 für das Teilmodell „Nutzung von Fachwissen“ und die Items 13, 22, 43, 52 und 93 für die Subfacette „Experiment-Raum“ des Teilmodells „Nutzung von Experimenten und Modellen“ (vgl. Woitkowski, 2015) Eingang in das Testheft für die Kreuzvalidierung.

Im Folgenden werden die Ergebnisse der oben genannten Analysen für die FDW- und die FW-(Sub)Skalen präsentiert. Dabei erfolgt zuerst der Bericht zur gesamten Kreuzvalidierungsstichprobe ($n_{\text{ges}} = 28$). Anschließend folgt der nach den Schularten (Gy: $n = 13$, RS bzw. nGy: $n = 15$) getrennte Ergebnisbericht. In die Analysen wurde auch die im FALKO-P-Modell post-hoc entfernte Subfacette „M & E“ mit einbezogen, da diese ja ursprünglich Teil der FALKO-P-Modellierung war und daher in Teilanalysen berücksichtigt werden sollte.

Manifeste Korrelationen zwischen den Skalen:

Für die manifeste Korrelation (nach Pearson) der beiden FDW-Skalen ergab sich für die gesamte Kreuzvalidierungsstichprobe ($n_{\text{ges}} = 28$) ein Wert von $r = 0,69$, für die FW-Skalen einer von $r = 0,67$. Fasst man die beiden Skalen zu einer „Professionswissensskala“ (FDW + FW) zusammen, steigt der Wert auf $r = 0,74$ (vgl. Tab. 6.13). Interessanterweise sind diese Werte bei getrennter Betrachtung der beiden Lehrämter für die Gy-Studienteilnehmer deutlich höher (FDW: $r = 0,76$, FW: $r = 0,76$, FDW + FW: $r = 0,82$) als für die nGy-Studienteilnehmer (FDW: $r = 0,59$, FW: $r = 0,46$, FDW + FW: $r = 0,60$, vgl. A.33 und A.34 in Anhang A). Weiterführend (auch wenn die explizite Betrachtung der FDW-Subfacetten bei FALKO-Physik nicht angestrebt war, vgl. Abschnitt 4.2) wurden auch die Korrelationen zwischen den entsprechenden FDW-Subfacetten und auch zwischen der „M & E“- (FALKO-P) und der „EM“-Subfacette (ProfiLe-P) betrachtet, welche im mittleren bis hohen Bereich liegen (Cohen, 1992, S. 157) vgl. Tab. A.35, A.36 und A.37 in Anhang A). Auch hier ist zu beobachten, dass bis auf die Korrelation zwischen der „M & E“- (FALKO-P) und der „EM“-Subfacette die Werte für die Gy-Studienteilnehmer höher als für die nGy-Studienteilnehmer liegen.

Tabelle 6.13.: Manifeste Korrelationen (nach Pearson) der sechs Professionswissens-Skalen (FDW/FW, FDW, FW), ges. Stichprobe ($n = 28$); $_{FA}$: FALKO, $_{Pro}$: ProfiLe-P

	FDW/- FW $_{FA}$ (28 Items)	FDW/- FW $_{Pro}$ (40 Items)	FDW $_{FA}$ (10 Items)	FDW $_{Pro}$ (20 Items)	FW $_{FA}$ (18 Items)	FW $_{Pro}$ (20 Items)
FDW/- FW $_{FA}$	1,00	0,74				
FDW/- FW $_{Pro}$	0,74	1,00				
FDW $_{FA}$			1,00	0,69		
FDW $_{Pro}$			0,69	1,00		
FW $_{FA}$					1,00	0,67
FW $_{Pro}$					0,67	1,00

Latente Korrelationen zwischen den Skalen:

Für die latente (bayesianische) Modellierung der Zusammenhänge zwischen den Skalen wurden analog zu den Modelltestungen in der Hauptstichprobe (vgl. Abschn. 6.2) im Fall des FDW die jeweiligen Subfacetten und für den Fall des FW jeweils drei gebildete Parcels als manifeste Indikatoren für die latenten Konstrukte verwendet. Im Fall der Gesamtstichprobe ($n_{ges} = 28$) ergab sich dabei für die Korrelation zwischen den beiden FDW-Skalen ein Wert von $r_{latent} = 0,81$ und für die Korrelation zwischen den beiden FW-Skalen $r_{latent} = 0,77$. Bei Betrachtung der beiden Teilstichproben (hier ist das berechnete Modell zum Teil nur eingeschränkt gültig, was eventuell aus der sehr geringen Stichprobenzahl resultieren könnte) lagen die Werte für die FDW-Skalen mit $r_{latent} = 0,84$ (Gy) und $r_{latent} = 0,64$ (nGy) sowie für die FW-Skalen mit $r_{latent} = 0,82$ (Gy) und $r_{latent} = 0,50$ (nGy), wie schon bei den Werten für die manifesten Korrelationen, erneut für die Teilstichprobe der Gymnasialen jeweils deutlich höher (vgl. Tabellen A.53, A.54, A.55, A.56, A.57 und A.58 in Anhang A).

Gruppenvergleiche (t-Tests) ohne Kontrolle des Fachsemesters sowie Gruppenvergleiche (einfaktorielle ANCOVAs) unter Kontrolle des Fachsemesters:

Die Gruppenvergleiche in Bezug auf die betrachteten Skalen (FDW + FW, FDW, FW) und Subskalen (FALKO FDW: „E & R“, „Schk“, „M & E“; ProfiLe-P FDW: „ib/id“, „s“, „ex“; ProfiLe-P FW: „EM“) wurden analog zu den Korrelationsanalysen zuerst unter Betrachtung der gesamten Kreuzvalidierungsstichprobe und anschließend getrennt nach Schulart (Gy vs. nGy) durchgeführt. Die Mittelwerte sowie die geschätzten Mittelwerte der Testleistungen (in Prozent) in den t-Tests und den ANCOVAs sind Tabelle A.38 in Anhang A zu entnehmen. Für die Analysen wurden prozentuale Summenscores der einzelnen Skalen gebildet, da in den jeweiligen Testteilen eine unterschiedliche maximale (theoretische) Anzahl an Punkten erreicht werden konnte. Bezüglich des Professionswissens (FDW + FW) konnten weder bei Betrachtung der Gesamtgruppe ($t(54) = -1,749$, $p = ,09$;

$d = -0,47$), noch der beiden Teilgruppen (Gy: $t(24) = -0,870$, $p = ,39$, $d = -0,34$; nGy: $t(28) = -1,738$, $p = ,09$, $d = -0,64$) ein signifikanter Unterschied zwischen den Leistungen in den beiden Tests festgestellt werden. Die im Anschluss durchgeführten ANCOVAs bestätigten dieses Ergebnis.

Trennt man im nächsten Schritt modellkonform nach FDW und FW, erhält man detailliertere Ergebnisse. Für den Fall des FDW wurden in den t-Tests Unterschiede in der Testleistung, sowohl für die Gesamtstichprobe ($t(54) = -5,798$, $p < ,001$, $d = -1,55^{***}$) als auch für die beiden Teilstichproben (Gy: $t(24) = -4,381$, $p < ,001$, $d = -1,72^{***}$; nGy: $t(28) = -4,143$, $p < ,001$, $d = -1,51^{***}$) erhalten. Hierbei zeigten aber die nachträglichen ANCOVAs eine Abhängigkeit von der Anzahl der studierten Fachsemester. Die Analysen zur FW-Skala zeigten ähnliche Resultate, nur dass hierbei der Unterschied in der Testleistung für die Teilgruppe der gymnasialen Studierenden wohl unabhängig von der Anzahl der studierten Fachsemester ist (Gy + nGy: $t(54) = -3,925$, $p < ,001$, $d = -1,05^{***}$; Gy: $t(24) = -2,654$, $p < ,05$, $d = -1,04^*$; nGy: $t(28) = -3,282$, $p < ,01$, $d = -1,20^{**}$).

Bei der finalen Betrachtung der Ergebnisse zu den FDW-Subskalen sowie zur FW-Subfacette „EM“ zeigte sich, dass bezüglich der „E & R“- und „ib/id“-Skala nur signifikant unterschiedliche Testleistungen für die gesamte ($t(50) = -2,738$, $p < ,01$, $d = -0,73^{**}$) und für die nicht-gymnasiale Stichprobe ($t(28) = -2,651$, $p < ,05$, $d = -0,97^*$) erhalten werden. Diese Unterschiede sind nur bezogen auf die nGy-Studierenden wohl auf die Anzahl der studierten Fachsemester zurückzuführen. Bezüglich der gymnasialen Teilstichprobe zeigte sich kein signifikanter Unterschied ($t(17) = -1,265$, $p = ,22$, $d = -0,50$). Für den Fall der „Schk“- und „s“-Subfacette manifestieren sich in allen drei betrachteten Gruppen (Gy + nGy: $t(54) = -5,000$, $p < ,001$, $d = -1,34^{***}$; Gy: $t(24) = -4,686$, $p < ,001$, $d = -1,84^{***}$; nGy: $t(28) = -2,866$, $p < ,01$, $d = -1,05^{**}$) signifikante Unterschiede in der Testleistung, die alle wohl auf die studierte Fachsemesteranzahl zurückzuführen sind. Bezogen auf die „M & E“- und „ex“-Skala zeigte sich ein ähnliches Bild, nur dass der Unterschied in der Testleistung für die Teilgruppe der nGy-Studierenden nicht signifikant wurde (Gy + nGy: $t(54) = -5,000$, $p < ,001$, $d = -1,34^{***}$; Gy: $t(24) = -4,686$, $p < ,001$, $d = -1,84^{***}$; nGy: $t(28) = -2,866$, $p < ,01$, $d = -1,05^{**}$).

In einem letzten Schritt wurden die Testleistungen aller Gruppen bezogen auf die „M & E“- und die „EM“-Skala untersucht, da sich bei den Korrelationsanalysen zumindest für die nicht-gymnasiale Teilstichprobe ein hoher Zusammenhang zwischen diesen beiden Subfacetten finden ließ (vgl. Tabelle A.37 in Anhang A). Allerdings ergab sich weder für die gesamte ($t(48) = -0,222$, $p = ,83$, $d = -0,06$) noch für die gymnasiale ($t(17) = -0,193$, $p = ,849$, $d = -0,08$) oder nicht-gymnasiale Stichprobe ein signifikanter Unterschied in der Testleistung.

7. Zusammenfassung, Interpretation und Diskussion der Ergebnisse

Insgesamt war es das Vorhaben des Teilprojekts FALKO-Physik, ein Testinstrument (FALKO-P) zur objektiven, reliablen und validen Messung des fachspezifischen Professionswissens (fachdidaktisches und Fachwissen) von Physiklehrkräften der Sekundarstufe zu entwickeln. Physiklehrkräfte waren hierfür per Konvention Lehrkräfte, die das Fach Physik an den Schularten Gymnasium und Realschule unterrichteten. Weiterhin wurden Mittelschullehrkräfte zu dieser Population gezählt, wenn sie a) den Fächerverbund „PCB“ (Physik/Chemie/Biologie) beziehungsweise „NT“ (Natur und Technik, LehrplanPLUS Bayern) unterrichten und b) im Rahmen ihres Studiums Physik entweder als Unterrichts- oder Didaktikfach belegten.

In der vorliegenden Studie wurden als Hauptstichprobe 65 Physiklehrkräfte und 118 angehende Physiklehrkräfte (Studierende) untersucht. Dem potentiellen Einwand mangelnder Messinvarianz, da das für das Professionswissen von Lehrkräften entwickelte Testverfahren auf studentische Stichproben angewendet würde (z. B. Kleickmann et al., 2013, S. 96; Kleickmann et al., 2014, S. 281), kann begegnet werden: In den ersten Multigruppenanalysen ließen sich zwar bei Annahme des reduzierten FALKO-Modells (ohne die FDW-Subfacette „M & E“) die beiden Wissenskategorien (FDW und FW) bei einem guten lokalen und globalen Fit und einer latenten Korrelation von $r = 0,94$ (für den Output vgl. die Tabellen A.42 und A.43 in Anhang A) gerade nicht mehr voneinander trennen (Chi-Quadrat-Differenzentest: $\Delta\chi^2(1, N = 118) = 0,39$; $p < ,53$, Tab. A.43). Allerdings indizierte eine aufgrund der geringen Stichprobengröße durchgeführte bayesianische konfirmatorische Faktorenanalyse, dass sich auch bei alleiniger Betrachtung der Studierendenstichprobe die beiden Wissenskategorien FDW und FW voneinander trennen lassen (für den Output vgl. Tabelle A.44 in Anhang A). Mit einer Korrelation von $r = 0,89$ liegt dieser Wert nur ein wenig höher als derjenige für die Hauptstichprobe ($r = 0,86$). Folglich kann von skalarer Invarianz ausgegangen werden, was heißt, dass sowohl bei Studierenden als auch bei Lehrkräften dieselben Konstrukte gemessen werden (vgl. Lindl & Kloiber, in Vorbereitung).

Weiterhin wurden folgende Studienteilnehmer im Rahmen von Kontraststichproben zur Gewinnung von Validitätshinweisen (Konstruktvalidität, diskriminant und konvergent) herangezogen:

25 fachfremde Mittelschullehrkräfte, welche zwar PCB beziehungsweise „NT“ unterrichten, aber das Fach Physik nicht studierten (3 Referendare, welche aus bereits in Abschnitt 6.1 erläuterten Gründen zu den Lehrkräften gezählt werden, studierten zumindest das Fach Naturwissenschaft und Technik, NWT, an der Universität Regensburg).

32 fachfremde Realschul- und Gymnasiallehrkräfte der Fächer Biologie ($n = 7$), Chemie ($n = 13$) und Mathematik ($n = 12$).

29 Fachphysiker (Bachelor-/Masterstudierende, Doktorierende, akademische Räte, Professoren, welche kein Lehramtsstudium absolvierten, zum Teil aber Lehrveranstaltungen in Form von Vorlesungen, Seminaren und Übungen abhielten).

28 Studierende des Lehramts für Realschule ($n = 15$) und Gymnasium ($n = 13$) mit Unterrichtsfach Physik, welche innerhalb einer Kreuzvalidierung zusätzlich Items des FDW- (Gramzow, 2014) beziehungsweise des FW-Tests (Woitkowski, 2015) aus dem Projekt ProfiLe-P bearbeiteten.

Bei der Testkonstruktion wurde besonders darauf Wert gelegt, zentrale physikalische Themengebiete (Mechanik, Elektrizitätslehre, Optik und Wärmelehre) sowohl im FDW- als auch im FW-Testteil zu implementieren.

Im Folgenden werden die in Kapitel 6 berichteten Ergebnisse zusammengefasst, interpretiert und diskutiert sowie die in Abschnitt 3.1 aufgestellten Forschungsfragen beziehungsweise -Hypothesen beantwortet.

Forschungsfrage 1 (FF1): Lässt sich das fachspezifische Professionswissen (FDW, FW) von Physiklehrkräften objektiv, reliabel und valide messen?

Soweit in Abschnitt 6.3.1 und in Tab. A.2 in Anhang A ersichtlich, konnten alle finalen Items objektiv kodiert werden. Dabei lag für die 60 doppelt kodierten Testhefte der Mittelwert der Interraterübereinstimmung (Spearmans ρ) für die FDW-Skala bei $\rho = 0,89$ und für die FW-Skala bei $\rho = 0,94$.

Die Reliabilität des FW-Testteils war zufriedenstellend ($\alpha = 0,82$), während das FDW nur eingeschränkt reliabel gemessen werden konnte ($\alpha = 0,65$), auch wenn dieser Wert angesichts einer Testneukonstruktion noch als hinreichend reliabel anzusehen ist (Field, 2013, S. 709; Brunner & Süß, 2007). Unter zusätzlicher Berücksichtigung der „M & E“-Facette würde die Reliabilität dieser Skala auf $\alpha = 0,70$ steigen. Sowohl eine konfirmatorische Faktorenanalyse (vgl. Abschn. 6.2) als auch eine Expertenbefragung legten jedoch nahe, diese Facette aus dem FDW-Test zu entfernen (da sie eher FW als FDW abzubilden scheint). An dieser Stelle könnte eine Weiterentwicklung des Testinstruments in Form von beispielsweise Itemneukonstruktionen ansetzen (vgl. Kap. 8). Die um Autokorrelationen korrigierten Trennschärfen r_{it} der Items des FDW liegen mit ihren Werten zwischen 0,24 und 0,40 und für FW von 0,16 bis 0,57 (siehe Tab. 6.2) im gewünschten Bereich, auch wenn die Skala zum FW zwei Items mit einer Trennschärfe von $r_{it} \leq 0,20$ enthält (diese beiden Items wurden aus normativen Gründen in der Skala belassen; vgl. Abschn. 6.3.3).

Zur Überprüfung der inhaltlichen Validität (Augenscheinvalidität) sollte ein Teil der Studienteilnehmer (Experten und erfahrene Lehrkräfte) die Eindeutigkeit ($M_{FDW} : 3,58$, $M_{FW} = 3,64$), die Berufs- ($M_{FDW} : 3,58$, $M_{FW} = 3,44$) sowie die Ausbildungsrelevanz ($M_{FDW} : 3,59$, $M_{FW} = 3,47$) der bearbeiteten Items auf einer vierstufigen Rating-Skala einschätzen (vgl. Abschn. 5.3.1, 5.3.2 und 6.3.4). Alle Mittelwerte zeugen von einer hohen Akzeptanz bei den befragten Studienteilnehmern, so dass bei beiden Testteilen von einer adäquaten inhaltlichen Passung ausgegangen werden kann. Dabei sind nicht nur die hohen Beurteilungen der FDW-Items insgesamt beachtenswert - sie werden außerdem als berufs- und ausbildungsrelevanter als die FW-Items eingeschätzt. Dies ist zum einen als ein Hinweis auf die realitätsnahe Konstruktion und kann zum anderen als besondere Bedeutung der operationalisierten Inhalte für den Unterrichtsalltag gedeutet werden. Weitere Validitätshinweise werden bezüglich der Forschungsfragen 3 und 4 interpretiert und diskutiert.

Forschungsfrage 2 (FF2): Lässt sich die zweidimensionale Struktur des Konstrukts (transformatives Verständnis des PCK nach Gess-Newsome, 1999; Baumert & Kunter, 2006; Grossman, 1990; Magnusson et al., 1999) stützen?

Um zu beurteilen, ob Shulmans (1986; 1987) Hypothese von zwar überlappenden, aber dennoch trennbaren Wissenskategorien FDW und FW gestützt werden kann, wurden konfirmatorische Faktorenanalysen durchgeführt. Bei Annahme des FALKO-übergreifenden Modells mit seinen drei FDW-Subfacetten konnten die beiden Wissenskategorien FDW und FW für das Fach Physik statistisch nicht mehr voneinander getrennt werden. Weitere Analysen (vgl. Abschn. 6.2) ergaben, dass die FDW-Subfacette „M & E“ (vgl. Kapitel 4) eher zum FW als zum FDW zu gehören scheint. Nach Entfernung dieser Subfacette wurde eine adäquate Modellpassung erhalten. Es konnte sowohl ein guter lokaler als auch globaler Fit erzielt werden, die Fit-Indizes entsprachen gängigen Cutoff-Kriterien (Schermelleh-Engel et al., 2003). Auch in den meisten Teilstichproben konnte eine Modellgültigkeit nachgewiesen werden. Somit kann Shulmans Hypothese – nach Entfernung der „M & E“-Facette – für die im Fach Physik erhobenen Daten gestützt werden, wenn auch die latente Korrelation im Gegensatz zu den anderen FALKO-Teilprojekten (vgl. hierzu auch Abschnitt 4 in Lindl & Krauss, 2017) oder zu anderen Professionswissenstests im Fach Physik vergleichsweise hoch ausfällt. Eventuell zeigt sich nach einer Testmodifikation oder der Untersuchung weiterer Stichproben mit FALKO-P eine moderatere Korrelation zwischen den beiden Wissenskategorien. Da aber, in Bezug auf das FALKO-Projekt, auch nicht bei vermeintlich verwandten Fächern, das heißt Geistes-, Sprach- oder eben Naturwissenschaften, ähnliche Korrelationen festzustellen sind, kann dieses Resultat als Anregung für die zukünftige Forschung betrachtet werden, die Frage nach dem Zusammenhang zwischen den beiden theoriebasiert abgeleiteten Wissenskategorien domänenspezifisch zu vertiefen.

Forschungsfrage 3 (FF3): Wie unterscheiden sich die untersuchten Teilgruppen (Studierende und Lehrkräfte an Mittelschulen, Realschulen, Gymnasien; fachfremde Mittelschullehrkräfte, Biologie-, Chemie- und Mathematiklehrkräfte, Fachphysiker) in Bezug auf ihr Professionswissen?

nGy-Studierende und nGy-Lehrkräfte (Real- und Mittelschule) vs. Gy-Studierende und Gy-Lehrkräfte mit Unterrichtsfach Physik:

Beim Blick auf die mittlere Testleistung in den untersuchten Teilstichproben der Hauptstudie (Status: Lehrkräfte vs. Studierende bzw. Schulform: Gy vs. nGy; vgl. z. B. Tab. 6.5) bestätigen sich die im Vorfeld aufgestellten Hypothesen hinsichtlich der Ausprägung des Professionswissens in den verschiedenen Subgruppen: Demnach schneiden Lehrkräfte insgesamt besser sowohl im FDW ($d = 0,51^{**}$) als auch im FW ($d = 1,27^{**}$) ab als Studierende (vgl. Abschn. 6.4.2 bzw. Hypothesen 1a und 1b), was nach dem Experten-Novizen-Paradigma ein zu erwartendes Ergebnis darstellt. Betrachtet man die Scatterplots in Abschnitt 6.4.2 genauer, so ist auffallend, dass bei den Studierenden der rechte untere Quadrant nahezu leer ist, wohingegen bei den Lehrkräften dort vereinzelt hohe Fachwissenswerte im Verbund mit geringen didaktischen Kompetenzen zu finden sind, was – zumindest für die Lehrkräfte – im Kontrast zu dem Befund aus COACTIV steht, dass hohes Fachwissen vor geringem fachdidaktischen Wissen schützt (Krauss et al., 2008). Die manifeste Korrelation ($r = 0,59^{**}$, vgl. Abschn. 6.4.1) zwischen beiden Wissenskategorien (und nach Entfernung der „M & E“-Facette) ist ähnlich der entsprechenden Korrelation

beim Physik-Test von Kirschner (Kirschner et al., 2017).

Weiterhin zeigen die Gy-Lehrkräfte sowohl beim FW als auch beim FDW eine höhere Testleistung als die nGy-Lehrkräfte, gefolgt von den Gy-Studierenden und den nGy-Studierenden (vgl. Abschn. 6.4.3 bzw. Hypothesen 2a und 3a bzw. 2b und 3b). Dies kann als erster Hinweis auf die Kriteriumsvalidität des Tests interpretiert werden (eine graphische Darstellung der insgesamt acht Gruppenunterschiede im FW und im FDW für alle FALKO-Fächer und COACTIV im Vergleich findet sich im Anhang zu Lindl & Krauss, 2017). Erklärt werden kann diese Rangfolge durch das jeweils absolvierte Studium. Während die Gymnasialen innerhalb einer Regelstudienzeit von neun Semestern ihre (zwei) gewählten Disziplinen vertieft und wissenschaftsorientiert, das heißt vergleichbar mit den Lehrveranstaltungen des entsprechenden Bachelor- beziehungsweise Masterstudiengangs belegen, studieren die nicht-gymnasialen Lehrämter mit Unterrichtsfach innerhalb einer Regelstudienzeit von sieben Semestern ihre Disziplin(en) mit einem reduzierten fachlichen Pensum. Zwar ist es mittlerweile so, dass die Gymnasiallehramtsstudierenden ähnlich viele fachdidaktische Lehrveranstaltungen wie die nicht-gymnasialen Lehrämter besuchen, dennoch spielt aber vermutlich auch das Mehr an Fachwissen eine Rolle, denn dieses bildet die „Grundlage, auf der fachdidaktische Beweglichkeit entstehen kann“ (Baumert & Kunter, 2012, S. 308). Dies gilt vor allem für die innerhalb der Studie untersuchten Lehrkräfte, die bereits vor einigen Jahren ihr Studium beendeten, als das gymnasiale Lehramtsstudium noch wesentlich weniger fachdidaktische Inhalte hatte als das nicht-gymnasiale.

nGy- und Gy-Lehrkräfte versus fachfremde Mittelschullehrkräfte:

Die Vergleichsgruppe der fachfremden Mittelschullehrkräfte bestand aus $n = 25$ Studienteilnehmern (eine ausführlichere Stichprobenbeschreibung findet sich in der Tabelle A.29 in Anhang A). Für diesen Gruppenvergleich wurde jeweils ein t-Test ($n = 65$ Gy- und nGy-Lehrkräfte vs. $n = 25$ MS-Lehrkräfte) für die gesamte FW- und FDW-Skala sowie für die Skalen der beiden FDW-Subfacetten „E & R“ und „Schk“ durchgeführt. Erwartungsgemäß schneiden die Gy- und nGy-Lehrkräfte sowohl im FDW als auch im FW deutlich besser als die Mittelschullehrkräfte ab, was selbstverständlich auf die Wirksamkeit der entsprechenden universitären Ausbildung zurückzuführen ist. Allerdings ist es so, dass für den Fall der FDW-Subskala „Schk“ der Unterschied in der Testleistung am geringsten und nur gerade noch signifikant ausfällt, was zeigt, dass sich die Mittelschullehrkräfte im Selbststudium zumindest in Bezug auf die Lernvoraussetzungen von Schülern weitergebildet haben. Die Hypothesen 4a und 4b können demnach gestützt werden.

Diesbezüglich interessant zeigt sich auch der in diesem Zusammenhang durchgeführte Gruppenvergleich zwischen den fachfremden Mittelschul- und den fachfremden Gy- und nGy-Lehrkräften. Immerhin übertrifft die erst- die zweitgenannte Gruppe deutlich in allen betrachteten (Sub)Skalen, was wiederum als Beleg dafür gewertet werden kann, dass sich die Mittelschullehrkräfte im Eigenstudium oder durch den Besuch von Lehrerfortbildungen, welche gerade für den naturwissenschaftlichen Bereich vielfach angeboten werden, Professionswissen im Fach Physik zum Teil selbst auch ohne ein universitäres Studium erarbeiten können, auch wenn der Wert der universitären Ausbildung in keinsten Weise geschmälert werden darf, wenn man den Vergleich der Testleistungen zu den untersuchten Physiklehrkräften zieht.

nGy- und Gy-Lehrkräfte versus fachfremde Realschul- und Gymnasiallehrkräfte:

Für den Gruppenvergleich mit fachfremden Realschul- und Gymnasiallehrkräfte konnten $n = 32$ Studienteilnehmer gewonnen werden (eine ausführlichere Stichprobenbeschreibung findet sich in der Tabelle A.30 in Anhang A). Erneut wurden jeweils ein t-Test ($n = 65$ Gy- und nGy-Lehrkräfte vs. $n = 32$ RS- und Gy-Lehrkräfte) für die gesamte FW- und FDW-Skala sowie für die Skalen der beiden FDW-Subfacetten „E & R“ und „Schk“ durchgeführt. Wie erwartet schneiden die Physiklehrkräfte bezüglich aller betrachteten (Sub)Skalen deutlich besser als die fachfremden RS- und Gy-Lehrkräfte ab. Sowohl in der FW- ($t(95) = 11,833$, $p \leq ,001$; $d = 2,65^{***}$) und FDW-Gesamtskala ($t(95) = 11,522$, $p \leq ,001$; $d = 2,49^{***}$) als auch in den beiden FDW-Subskalen „E & R“ ($t(95) = 11,756$, $p \leq ,001$; $d = 2,54^{***}$) und „Schk“ ($t(95) = 11,756$, $p \leq ,001$; $d = 1,61^{***}$) zeigte sich sehr deutlich der Vorteil der Physiklehrkräfte gegenüber den fachfremden Lehrkräften (vgl. Tabelle 6.11), der aber wiederum für die „Schk“-Skala am geringsten ausfällt, was eventuell als kleiner Hinweis auf die besondere Sensibilität der (naturwissenschaftlichen) Lehrkräfte in Bezug auf Lernvoraussetzungen von Schülern interpretiert werden kann und dies demnach die Zugehörigkeit des Wissens über Schüler(wohl)vorstellungen zum FDW unterstreicht. Somit können die Hypothesen 5a und 5b gestützt werden.

nGy-Lehrkräfte und Gy-Lehrkräfte vs. Fachphysiker:

Für die folgende Untersuchung wurde die Stichprobe an Physik-Lehrkräften ($n = 65$) für den Vergleich mit der Kontrastgruppe der Fachphysiker ($n = 29$) einmal hypothesenkonform weiter in gymnasiale (Gy; $n = 35$) und nicht-gymnasiale (nGy; $n = 30$) Lehrkräfte aufgeteilt (eine ausführlichere Stichprobenbeschreibung findet sich in der Tabelle A.31 in Anhang A) und einmal als Gesamtgruppe berücksichtigt.

Bei den Analysen in Bezug auf die drei Subgruppen schnitten die Gy-Lehrkräfte im FDW signifikant besser als die Fachphysiker ab, während weder zwischen den nGy- und den Gy-Lehrkräften noch zwischen den nGy-Lehrkräften und den Fachphysikern ein signifikanter Unterschied in der Testleistung festgestellt werden kann. Gleiches war bei der FDW-Subskala „E & R“ festzustellen. Für den Fall der FDW-Subskala „Schk“ und des FDW konnten keine signifikanten Unterschiede zwischen den drei untersuchten Gruppen festgestellt werden.

Beim Vergleich der gesamten Lehrkräfte-Gruppe mit den Fachphysikern zeigte sich, dass die Lehrkräfte im FDW signifikant bessere Testleistungen erzielten als die Fachphysiker, wobei dies vor allem auf die FDW-Subskala „E & R“ zurückzuführen ist. Im FW kann erneut kein signifikanter Gruppenunterschied festgestellt werden.

Auch wenn die Hypothesen 6a und 6b somit nur teilweise durch die erhaltenen Ergebnisse gestützt werden können, ist es so, dass plausible Erklärungsansätze für das vermehrt ähnliche Abschneiden der Physiklehrkräfte und der Fachphysiker sowohl im FW als auch in Teilen des FDW gefunden werden können. Zum einen ist es ja so, dass die Fachphysiker ein mit den gymnasialen Lehrkräften vergleichbares Fachstudium absolvieren. Gleichzeitig wurden die FW-Items bei FALKO-P sehr schulnah (vgl. Kap. 4) konstruiert, so dass eigentlich alle betrachteten Gruppen (auch die nGy-Lehrkräfte) über ein ähnliches Wissen verfügen sollten, weshalb die Ergebnisse der Analysen nicht allzu verwunderlich sind. Weiterhin wird ja FW nicht nur als Voraussetzung für didaktische Flexibilität, sondern für unterrichtliches Erklären überhaupt angesehen (vgl. Baumert & Kunter, 2006; Braaten & Windschitl, 2011), so dass prinzipiell auch Fachphysiker zumindest implizit Voraussetzungen für die Entwicklung von FDW innerhalb ihre Studiums erwerben, da sie sich ja deutlich mehr und vertieftere fachwissenschaftliche Inhalte aneignen als beispielsweise die nicht-

gymnasialen Lehramtsstudierenden.

Allerdings schneiden die Fachphysiker in der für die Lehrerprofession entscheidenden Subfacette des FDW, nämlich dem „Erklären und Repräsentieren“ signifikant schlechter als die Gruppe der Physiklehrkräfte ab, denn Lehrkräfte müssen ja über ein breites Repertoire an (unterschiedlichen) Erklärungsstrategien verfügen, um Unterrichtsinhalte adäquat vermitteln zu können (Kulgemeyer & Schecker, 2013; Schilcher et al., 2017).

Als interessant könnten sich hierbei auch noch die Ergebnisse weiterführender Analysen zeigen, nämlich, wenn zukünftig innerhalb der Gruppe der Fachphysiker zwischen Studierenden (Bachelor, Master) und bereits berufstätigen Studienteilnehmern (akademische Räte, Universitätsprofessoren, ...) unterschieden wird (hierzu müsste die Stichprobe noch entsprechend vergrößert werden), denn auch in der Lehre tätige Fachphysiker sollten im Gegensatz zu Studierenden über ein gewisses Erklärungsrepertoire verfügen, um Vorlesungs-/Übungsinhalte adäquat vermitteln zu können.

Insgesamt betrachtet können die erhaltenen Resultate hinsichtlich der Konstrastgruppenvergleiche als weiteren Hinweis für die Unterschiedlichkeit und somit für die differentielle Validität der beiden gemessenen Wissenskategorien herangezogen werden.

Forschungsfrage 4 (FF4): Wie unterscheiden sich Studienteilnehmer in Bezug auf ihr Professionswissen beim Vergleich ihrer Testleistungen bei FALKO-P und in den ProfiLe-P-Testinstrumenten zum FDW (Gramzow, 2014) und FW (Woitkowski, 2015)?

Um Hinweise auf konvergente Validität im Sinne von Konstruktvalidität zu erhalten, bearbeitete eine Gruppe von Studienteilnehmern ($n_{\text{ges}} = 28$) aus 13 Studierenden des Physik-Lehramts für Realschulen und 15 Studierenden des Physik-Lehramts für Gymnasien zusätzlich zum Testinstrument FALKO-P Items des FDW- (Gramzow, 2014) und FW-Tests (Woitkowski, 2015) aus dem Projekt ProfiLe-P. Zur Untersuchung, inwiefern die Skalen der jeweiligen Testinstrumente zusammenhängen, wurden die Summenscores der Skalen (FW und FDW einschließlich der Subskalen) manifest (nach Pearson) und latent (CFA, bayesianische Modellierung) gegeneinander korreliert. Zwischen den beiden Testdurchführungen lagen im Mittel zwei Fachsemester, so dass bei den untersuchten Gruppenvergleichen (n_{ges} , n_{Gy} , n_{nGy}) neben t-Tests auch auf Kovarianzanalysen mit der studierten Fachsemesteranzahl als Kontrollvariable zurückgegriffen wurde.

Sowohl für die manifesten (nach Pearson) als auch für die latenten (bayesianischen) Korrelationen zwischen den untersuchten (Sub)Facetten ergaben sich durchgehend mittlere bis hohe Werte, welche bei Betrachtung der Subgruppen häufig im Falle der gymnasialen Studienteilnehmer deutlich höher ausfielen. Begründungen hierfür können lediglich vermutet und müssten bei weiterführenden Teilstudien genauer untersucht werden. Eventuell könnte es daran liegen, dass die ProfiLe-P-Testinstrumente primär zur Untersuchung von Studierenden sowohl des Physik-Lehramts als auch Studierenden der Fachphysik sowie Übungsbetreuenden konzipiert wurden und die Iteminhalte demnach vorwiegend am universitären Curriculum ausgerichtet sind, während FALKO-P primär zur Untersuchung von Physik-Lehrkräften entwickelt wurde und sich die meisten Testitems sehr am Schulwissen orientieren. Dadurch könnten eben vor allem die Gy-Studierenden mit ihrem an vertieften universitären Inhalten orientierten Studium angesprochen worden sein. Interessanterweise zeigt eine einzige untersuchte Korrelation ein anderes Verhalten. Während sich für die gymnasialen Studierenden kein Zusammenhang zwischen der „M & E“- (FALKO-P) und der „EM“-Facette (ProfiLe-P) zeigt, nimmt dieser für die nicht-gymnasialen Studierenden sogar einen mittleren bis hohen Wert an. In weiterführenden Analysen müsste hierfür

nach einer Begründung gesucht werden. Aufgrund der vor allem für das Professionswissen insgesamt und für die FDW- und FW-Skala erhaltenen hohen Korrelationswerte kann davon ausgegangen werden, dass die jeweils beide Testinstrumente dasselbe Konstrukt messen.

Zusätzlich zu den Korrelationsanalysen wurden Gruppenvergleiche in Bezug auf die gesamte Stichprobe ($n_{\text{ges}} = 28$) und auf die beiden Teilstichproben durchgeführt (Gy: $n = 13$, nGy: $n = 15$). In den meisten Fällen war ein resultierender Gruppenunterschied auf die Anzahl der studierten Fachsemester zurückzuführen, so dass bei gleichem Fachsemester auch von einer ähnlichen Testleistung ausgegangen werden könnte. Trotz dieser (ersten) Ergebnisse wird allerdings angeraten, eine weitere Kreuzvalidierung mit einer größeren Stichprobe zum gleichen Messzeitpunkt durchzuführen, um Verzerrungen der Testleistung aufgrund einer höheren studierten Fachsemesteranzahl zu vermeiden und um die bisherigen Analyseergebnisse zu bestätigen oder zu widerlegen. Ergänzend hierzu wäre wünschenswert, auch eine entsprechende Lehrerstichprobe untersuchen zu können.

Insgesamt gesehen kann trotz der inhaltlich unterschiedlichen Modellierungen des FDW und des FW davon ausgegangen werden, dass FALKO-P und die beiden ProfiLe-P-Testinstrumente auch dieselben Konstrukte messen, wonach die Hypothesen 12a und 12b gestützt werden können.

Forschungsfrage 5 (FF5): Wie beeinflussen die Prädiktoren Status (Lehrkräfte vs. Studierende) und Schulform (Gymnasium vs. Nicht-Gymnasium) – bei Kontrolle von Geschlecht und kognitiver Leistungsfähigkeit – das FDW und das FW?

In zwei getrennten klassischen linearen Regressionsmodellen wurde der Einfluss der unabhängigen Variablen Status und Schulform auf das FW und das FDW unter gleichzeitiger Kontrolle von Geschlecht und kognitiver Leistungsfähigkeit untersucht (vgl. Abschn. 6.4.3). Dabei wurde ersichtlich, dass die Abiturnote erwartungsgemäß (Hypothesen 7a und 7b) mit den Testleistungen in beiden Wissensbereichen zusammenhängt (FW: $-0,20^{**}$; FDW: und $-0,22^{**}$). Auch die Kovariate Geschlecht (Hypothesen 8a und 8b) ist ein signifikanter Prädiktor für die Leistung in beiden Wissensbereichen, wobei der deutliche Vorteil der männlichen Studienteilnehmer Befunde von Riese (2009) sowie Kirschner (2013) repliziert. Demzufolge scheint das Fach Physik (immer noch) eher eine Männerdomäne zu sein.

Entsprechend der theoretischen Strukturierung – zuerst nach Status, dann nach Schulform – ist der erklärungsstärkste Faktor die Unterscheidung zwischen Lehrkräften und Studierenden, wobei der Regressionskoeffizient für das FW im Vergleich zu den anderen FALKO-Disziplinen besonders ausgeprägt ist (FDW: $\beta_{\text{Status}} = 0,22^{**}$; FW: $\beta_{\text{Status}} = 0,51^{**}$). Einen vergleichbar hohen Beitrag liefert die Statusvariable nur noch im Fach Latein, Lindl & Kloiber, 2017). Interessanterweise trägt die Schulform im Fach Physik – entgegen der Erwartungen im Vorfeld (Hypothese 9a und 9b) – kaum zur Varianzaufklärung hinsichtlich beider Wissenskategorien bei. Dies könnte unter anderem daran liegen, dass bei der Konstruktion des FW-Tests kein rein universitäres Wissen implementiert wurde, um den Test auch bei (fachfremden) Mittelschullehrkräften einsetzen zu können. Darüber hinaus unterscheiden sich die schulischen (bayerischen) Curricula sowie mittlerweile auch die fachdidaktischen Anteile im Lehramtsstudium für das Fach Physik zwischen den Schulformen Realschule und Gymnasium (für die Sekundarstufe I) kaum mehr, was das ähnliche Abschneiden auch im FDW-Testteil erklären könnte. In der Tat finden sich nur im Bereich Mechanik und Atomphysik stark differierende Curricula für die unterschiedlichen Schulformen. Diese vertieften gymnasialen Inhalte wurden aber in den Mechanik-Items nicht (und

zwar weder im FW- noch im FDW-Test) thematisiert. Weitere überfachliche Analysen innerhalb des FALKO-Projekts, bezogen auf die untersuchten Prädiktoren finden sich in Lindl und Krauss (2017).

Forschungsfrage 6 (FF6): Wie hängt das fachbezogene Professionswissen mit Indikatoren des Schul- und Studienerfolgs sowie der Berufs- und Studiendauer zusammen?

Interessante Ergebnisse liefern auch die Analysen zu den Zusammenhängen des Professionswissens mit Indikatoren des Ausbildungserfolgs sowie der Berufs- respektive Studiendauer (vgl. Abschn. 6.4.4). Am meisten sticht der hohe Zusammenhang zwischen Studiendauer (Fachsemester) und fachbezogenem Professionswissen in der Gruppe der Studierenden (FW: $r = 0,58^{**}$; FDW: $r = 0,40^{**}$) heraus (Hypothesen 11 a und 11 b). Auch wenn bei der Interpretation querschnittlicher Daten im Hinblick auf Entwicklungsverläufe Vorsicht geboten ist, können diese Zusammenhänge als erste Indikatoren für die Wirksamkeit des universitären Studiums gesehen werden. Um herauszufinden, ob dies auch tatsächlich der Fall ist, stellen längsschnittliche Analysen ein Forschungsdesiderat dar (vgl. Kap. 9).

Wie bereits berichtet, zeigt sich eine Korrelation mit der Abiturnote von $r = -0,28^{**}$ für das FW und $r = -0,31^{**}$ für das FDW bei Studierenden. Jedoch war kein signifikanter Zusammenhang mit den Noten des ersten oder zweiten Staatsexamens feststellbar (Hypothesen 10a und 10b). Für den Fall des ersten Staatsexamens könnte dies als ein Resultat der besonders schulnahen Konstruktion vor allem der fachwissenschaftlichen Items interpretiert werden. Mit derartigen Anforderungen können Studierende während ihrer universitären Ausbildung kaum Erfahrung sammeln, da sie sich in der Regel mit Fachwissen (und auch fachdidaktischem Wissen) auf universitärem Niveau auseinandersetzen. Für den Fall des zweiten Staatsexamens gestaltet sich eine Interpretation komplizierter. Während des zweijährigen (zumindest in Bayern, aus der alle untersuchten Lehrkräfte stammen) Referendariatszeit sollten die Lehrer-Novizen bereits Erfahrungen mit schulnahe FW und FDW gemacht haben, was sich in der Testleistung und auch im Zusammenhang mit der Note des zweiten Staatsexamens niederschlagen sollte. Eine mögliche Erklärung könnte sein, dass während des zweiten Abschnitts der Lehrerbildung anderes Wissen als im FALKO-P-Testinstrument implementiert wurde, erworben wird. Da im Moment die Stichprobe zu wenig Referendare für tiefere Untersuchungen enthält, würden weitere (längsschnittliche) Analysen sicherlich aufschlussreiche Ergebnisse liefern. Auch eine (längsschnittliche) Untersuchung der ersten Berufsjahre nach dem Referendariat führt sicherlich zu interessanten Resultaten hinsichtlich der (Weiter-)Entwicklung des Professionswissens.

In der Gruppe der Lehrkräfte ist dagegen – auf den ersten Blick überraschend (Hypothesen 11a und 11b) – kein signifikanter Zusammenhang zwischen Berufserfahrung und FDW nachzuweisen, jedoch scheinen Lehrkräfte mit mehr Berufserfahrung tendenziell über ein höheres FW ($r = 0,28^*$) zu verfügen. In allen anderen FALKO-Fächern war dieser Zusammenhang entweder nicht signifikant oder sogar (leicht) negativ. Auch bei COACTIV und beispielsweise in allen drei naturwissenschaftlichen Disziplinen (Biologie, Chemie und Physik) bei ProwiN wurde ein ähnliches Ergebnis erhalten. Dieser auf den ersten Blick erstaunlich erscheinende fehlende Zusammenhang in Bezug auf das FW und vor allem auf das FDW, der mittlerweile vielfach diskutiert wurde (Kirschner et al., 2017; Besser & Krauss, 2016), könnte zum Beispiel auf das Fehlen der Deliberate-Practice-Bedingungen (Ericsson et al., 1993) während der sogenannten dritten Ausbildungsphase zurückzuführen sein.

Umso erstaunlicher ist diesbezüglich die innerhalb der bisherigen Forschungsarbeit

zum Professionswissen singuläre signifikante Korrelation zwischen dem FW und der Berufserfahrung der untersuchten Lehrkräfte. Sie ist wahrscheinlich auf die bereits oben beschriebene stark unterrichtsnahe Operationalisierung des Fachwissens – inklusive dezidiert Items auf Schulniveau – zurückzuführen. Die in Anhang B ersichtlichen FW-Items (für die zugehörigen Kodieranweisungen bitte bei der Autorin nachfragen) sollen in diesem Zusammenhang als Diskursanregung für die Fachgemeinschaft verstanden werden, um in weiterführenden theoretischen oder empirischen Analysen zu untersuchen, ob in diesen Items eventuell (erstmalig) Charakteristika implementiert wurden, die sensibel auf zunehmende Berufserfahrung reagieren.

Forschungsfrage 7 (FF7): Ist Mechanik (oder eines der anderen operationalisierten Teilgebiete der Physik) repräsentativ für das Gesamtkonstrukt physikalisches FW und FDW („Repräsentativitätshypothese“)?

Um die Hypothese der Repräsentativität des Professionswissens in Mechanik für die Gesamtheit der Teilbereiche der Physik zu untersuchen, wurden separate Skalen (FDW + FW; FDW; FW) für die vier Fachbereiche Mechanik, Elektrizitätslehre, Optik und Wärmelehre erstellt (vgl. Abschn. 6.5). Auch wenn die Itemanzahl der gebildeten Skalen zum Teil sehr klein ist und die erhaltenen internen Konsistenzen zum Teil auch extrem niedrig sind (was aber als ein Hinweis auf die Breite des jeweiligen Konstrukts interpretiert werden kann), lässt sich für die berechneten Korrelationsmatrizen trotzdem das interessante Ergebnis feststellen, dass sich eine vergleichbare Repräsentativität aller vier Fachbereiche, bei Korrelationen im durchgehend schwachen bis moderaten Bereich (Cohen, 1988) ergab. Es kann somit gefolgert werden, dass keiner der erfassten Inhaltsbereiche des Professionswissens eine ausreichende Repräsentativität für das Gesamtkonstrukt fachspezifisches Professionswissen im Fach Physik aufweist. Hypothese 13, wonach angenommen wird, dass Wissen im Inhaltsbereich Mechanik repräsentativ für das Wissen im Fach Physik ist, kann demnach nach derzeitiger Datenlage nicht gestützt werden. Auch wenn die Zahl der Studienteilnehmer mit $N = 183$ (Hauptstichprobe) noch vergleichsweise klein ist und das Testinstrument für genauere Ergebnisse an einer größeren Stichprobe eingesetzt werden muss, wird angeregt, innerhalb der Forschungsgemeinschaft zu diskutieren, ob im Sinne der Inhaltsvalidität das fachspezifische Professionswissen künftig auf einer breiteren Basis als bei bisherigen Testkonstruktionen geschehen (z. B. ProwiN, QuiP, ProfiLe-P) beziehungsweise über die gesamte Breite des heterogenen Konstrukts (Bühner, 2011) operationalisiert werden sollte.

8. Möglichkeiten der Testweiterentwicklung

Wie bereits in den vorherigen Kapiteln konstatiert, stellt die jetzige Version des Testinstruments FALKO-P mit seinen zehn (bzw. 14) Items zum FDW und 18 Items zum FW (vgl. Testheft in Anhang A) keine in Stein gemeißelte Endversion dar. Vielmehr könnte durch weitere Pilotierungsrunden an Verbesserungen gearbeitet werden, um den zukünftigen Einsatz zur Untersuchung weiterer Forschungsfragen (noch) zuverlässiger gestalten zu können. Im folgenden Abschnitt werden beispielhaft mehrere Möglichkeiten der Testoptimierung beschrieben. Selbstverständlich könnten weitere Optimierungen auch an anderen Stellen erfolgen.

8.1. Itemergänzung bzw. -neuentwicklung

Im Sinne einer Testverbesserung könnte zum einen an der Zusammenstellung des Testinstruments gearbeitet werden. Unter anderem wurden während der Pilotierungsrunden Items verworfen, die nach den durchgeführten Analysen zwar eine ähnliche Eignung zeigten, aber redundante Inhalte abprüften (vgl. Abschn. 5.3.1 bzw. 5.4). Diese könnten innerhalb weiterer Pilotierungsrunden einer erneuten Eignungsüberprüfung unterzogen werden. Dies wäre gerade in Bezug auf die vollständig entfernte FDW-Subfacette „M & E“ eine Möglichkeit, diese Subfacette eventuell doch noch für das FALKO-Modell passend zu konstruieren.

Soll diese dritte FDW-Subfacette weiterhin nicht mehr für das Modell berücksichtigt werden, würde sich anbieten, aus dem verworfenen Item-Pool zumindest weitere Items zu den beiden Subfacetten „E & R“ sowie „Schk“ hinzuzufügen, um die grenzwertige Reliabilität ($\alpha = 0,65$; vgl. Abschn. 6.3.3) der FDW-Skala auf einen akzeptableren Wert zu erhöhen. Eine Fokussierung auf lediglich zwei FDW-Subfacetten erschiene insofern eher wenig problematisch, als dass diese beiden Subfacetten zentrale Bestandteile des theoretisch definierten fachdidaktischen Wissens darstellen (u. a. in den Wissenstaxonomien nach Shulman, 1986; Shulman, 1987; oder der auf dem „PCK Summit“, 2012, entwickelten Definition). Denn auch wenn die verschiedenen FDW-Modelle, welche für die Konstruktion der Testinstrumente anderer nationaler und internationaler Arbeitsgruppen zugrunde gelegt wurden, nicht einheitlich sind und auch keine einheitliche Definition des Konstrukts des fachdidaktischen Wissens existiert, sind die beiden verbliebenen Subfacetten zumindest implizit Bestandteil eines jeden Modells zum fachdidaktischen Wissen, auch wenn sich unterschiedliche Schwerpunktsetzungen bei deren Operationalisierung zeigen (Kapitel 1 sowie Kirschner, 2013). Selbstverständlich müsste dann diesbezüglich das Kodiermanual entsprechend ergänzt beziehungsweise erweitert werden, was ebenfalls für eine eventuelle Überarbeitung der Modellierung gilt (vgl. Abschn. 8.2).

8.2. Überarbeitung der Modellierung

Selbstverständlich könnten aber auch nach weiterer theoretisch fundierter Überlegung neu konstruierte Items, entweder unter einer modifizierten „M & E“- oder in Ergänzung einer weiteren inhaltlich anders gelagerten fachspezifischen Subfacette, wie beispielsweise „Gestaltung von Lernprozessen“ bei Riese (2009), „Wissen über Konzepte“ bei Kirschner (2013) oder „Modelle/Konzepte“ bei Gramzow (2014) zur FDW-Skala hinzugefügt werden. In diesem Sinne würde eine Itemneukonstruktion eigentlich zu einer Überarbeitung der Modellierung des Professionswissens führen. Allerdings sollte bei allen Modifizierungsmöglichkeiten keinesfalls die Struktur des FALKO-Gesamtmodells außer Acht gelassen werden, da ja die Vergleichbarkeit der Modelle in den Einzeldisziplinen und somit der erhaltenen Ergebnisse eine der wesentlichen Stärken des FALKO-Projekts darstellt.

Eine Modellmodifizierung könnte aber auch in Bezug auf die FW-Skala stattfinden. Nach derzeitigem Stand befinden sich unter den 18 FW-Items zwölf, welche dem Niveau des „Schulwissens“ und sechs, welche dem Niveau des „vertieften Hintergrundwissens“ zugeordnet werden können. Zwar sollten, dem FALKO-Modell folgend, alle Items auf letzterem Niveau konstruiert werden, im Fall von FALKO-Physik aber beruhte die zusätzliche Entwicklung von Items auf „Schulniveau“ darauf, dass das Testinstrument auch für Mittelschullehrkräfte einsetzbar sein sollte, welche zwar Physik im Fächerverbund „PCB“ (Physik-Chemie-Biologie) beziehungsweise „NT“ (Natur und Technik; Lehrplan-PLUS Bayern) unterrichten, das Fach aber nicht im Rahmen ihrer Lehramtsausbildung studierten (vgl. Fußnote 2 in Kap. 1). Eventuell ist dieser überwiegend schulnahen Operationalisierung auch die bisher einzig bekannte Korrelation ($r = 0,28^*$) zwischen dem Fachwissen und der Anzahl der Berufsjahre geschuldet (vgl. Kap. 6 und 7). Dennoch bildet die hohe Korrelation zwischen dem fachdidaktischen Wissen und dem Fachwissen (latent: $r = 0,86$) einen Nährboden für etwaige Validitätsdiskussionen. Insofern könnte in weiteren Modelltestungen getrennt nach Fachwissen auf „Schulniveau“ und auf „vertieftem Hintergrundwissen“ geprüft werden, ob sich andere (niedrigere und somit mit den Ergebnissen anderer Arbeitsgruppen vergleichbarere) latente Korrelationen ergäben und die beiden Wissenskategorien somit statistisch „besser“ voneinander trennbar wären. Je nach Ergebnis dieser Analysen müssten vermutlich auch hier Itemergänzungen oder -neukonstruktionen vorgenommen werden, um weiterhin eine ausreichend hohe Reliabilität der beiden FW-Subskalen zu gewährleisten.

8.3. Skalierung

Die Konstruktion der beiden Testteile (FDW, FW) des Testinstruments FALKO-P folgte vornehmlich den Vorgehensweisen der klassischen Testtheorie (KTT). Neben der Ermittlung der Interraterreliabilität bei der Kodierung der Antworten auf die offenen Items, der internen Konsistenz der Skalen sowie der Trennschärfen und Schwierigkeiten, wurde auch die Dimensionalität der Items (Modellpassung) mittels konfirmatorischen Faktorenanalysen geprüft (vgl. Kap. 5, 6, und 7).

Um nun aber einen tieferen Einblick in die Schwierigkeit der entwickelten Items zu erhalten, könnte unter Berücksichtigung der probabilistischen Testtheorie (PTT) eine Skalierung der Items vorgenommen werden, was letztendlich in einen Ersatz beziehungsweise in eine Neukonstruktion solcher münden könnte (siehe oben), wenn einzelne Items sich final als nicht geeignet herausstellten. Die PTT nimmt einen Zusammenhang zwischen Personenfähigkeit und der Itemschwierigkeit an. Konkret bedeutet dies, dass das Ant-

wortverhalten eines Studienteilnehmers im Sinne einer Lösungswahrscheinlichkeit bei der Bearbeitung eines Items nicht einfach null oder eins ist (wie bei der KTT angenommen), sondern probabilistisch für jeden Teilnehmer in einem Wertebereich zwischen null und eins liegt (logistisches Modell). Folglich können Aussagen darüber erhalten werden, welche Items ein Studienteilnehmer mit einer bestimmten Fähigkeitsausprägung wahrscheinlich korrekt lösen kann und welche nicht. Ein weiterer Unterschied zwischen KTT und PTT ist, dass die PTT Antwortmuster betrachtet, so dass Itemantworten manifeste, das heißt beobachtbare, Ausprägungen einer latenten Variable darstellen. Bei Gültigkeit des Modells ist tatsächlich eine Messung von homogenen Eigenschaften oder Fähigkeiten möglich, während im Rahmen der KTT lediglich Itemantworten ohne weitere Prüfung gezählt werden. Folglich sind Korrelationen der Itemantworten bei der PTT auf die eine untersuchte latente Variable zurückzuführen und das Antwortverhalten eines Studienteilnehmers, neben weiteren Einflussgrößen wie die Itemschwierigkeit, kann vorhergesagt werden. Mit Hilfe von probabilistischen Testmodellen (z. B. Rasch-Modell oder Birnbaum-Modell) lassen sich Annahmen formulieren, von welchen Modellparametern, z. B. Personenfähigkeit oder die Zugehörigkeit einer Person zu einer bestimmten Gruppe, die Lösungswahrscheinlichkeit eines Items abhängt. Dabei werden die zur Feststellung der Fähigkeits- oder Eigenschaftsausprägungen nötigen Personenparameter sowie die zur Feststellung der Itemschwierigkeit erforderlichen Itemparameter geschätzt. Verwendete Schätzmethoden sind beispielsweise die verbundene (joint unconditional, uML) oder die konditionale (conditional, cML) Maximum-Likelihood-Schätzung (Bühner, 2011).

Das Rasch-Modell als 1PL-Modell (Ein-Parameter-Logistisches Modell), welches bei der Skalierung des Testinstruments FALKO-P verwendet werden könnte, geht davon aus, dass die Lösungswahrscheinlichkeit für ein bestimmtes Item zum einen von der Fähigkeits- beziehungsweise Eigenschaftsausprägung (Personenparameter) einer Person und zum anderen von der Itemschwierigkeit (Itemparameter) und sonst keinem weiteren Parameter abhängt, wobei der Personenparameter den Wert einer Person auf der zu untersuchenden latenten Variable angibt. Dabei gilt, dass je größer die Differenz zwischen der Fähigkeit einer Person und der Itemschwierigkeit ausfällt, desto größer ist die Wahrscheinlichkeit, dass diese Person das Item löst, also je fähiger eine Person, desto wahrscheinlicher die richtige Lösung des Items. Muss das Rasch-Modell durch die erhobenen Daten nicht verworfen werden, so sind die Itemantworten lokal stochastisch unabhängig voneinander, was als Homogenitätsnachweis für die Items betrachtet werden kann. Weiterhin gelten für die Items innerhalb des Rasch-Modells die gleichen Trennschärfen, weshalb der Summenwert eine umfassende Statistik der Personenfähigkeit darstellt. Muss ein Rasch-Modell verworfen werden, kann es sein, dass innerhalb der untersuchten Stichprobe Personengruppen existieren, welche zur Lösung der Items unterschiedliche Fähigkeiten nutzen. Hierbei lägen in den verschiedenen Gruppen unterschiedliche latente Variablen für die Lösung der Items zugrunde. In diesem Fall könnte mit Hilfe eines Mixed-Rasch-Modells geprüft werden, ob sich innerhalb der Gesamtstichprobe Personengruppen finden lassen, für welche das zuvor angenommene Rasch-Modell dennoch angenommen werden kann. Wenn ein Rasch-Modell angenommen wird, also Gültigkeit hat, treffen folgende Annahmen auf die zu skalierenden Testitems zu: Lokale stochastische Unabhängigkeit, gleiche Trennschärfen, Summenwerte als erschöpfende Statistiken der Personenparameter sowie spezifische Objektivität (Bühner, 2011).

Um eine adäquatere Skalierung für die beiden FALKO-P-Testteile und eine Modellgültigkeit für möglicherweise enthaltene unterschiedliche Personenklassen zu erhalten, müsste eigentlich, wie oben bereits beschrieben, jeweils ein Mixed-Rasch-Modell verwendet werden,

da in der Hauptstichprobe je nach Aufteilung zwei (Studierende vs. Lehrkräfte) oder sogar vier (nGy-Studierende, Gy-Studierende, nGy-Lehrkräfte, Gy-Lehrkräfte) unterschiedliche Gruppen (Klassen) enthalten sind, für die die Gültigkeit des Rasch-Modells getrennt nachgewiesen werden müsste, da ja unterschiedliche Personenklassen voraussetzungsabhängig (Studien-/Berufsjahre, Art der Ausbildung) unterschiedliche Fähigkeiten bei der Lösung der Testitems nutzen könnten (Bühner, 2011). Da aber bereits gezeigt werden konnte (Kap. 6 und 7), dass in beiden Statusgruppen (Studierende vs. Lehrkräfte) das gleiche Konstrukt gemessen wird, könnte für die Skalierung des Testinstruments FALKO-P bei der Betrachtung der Hauptstichprobe ($N = 183$) die Verwendung eines ordinalen Rasch-Modells (Partial-Credit-Modell) in Frage kommen, da die meisten der verwendeten Items nicht dichotom sondern mehrstufig skaliert sind. Ebenfalls aufgrund der bereits durchgeführten Prüfung der Dimensionalität mittels konfirmatorischer Faktorenanalysen, kann auf die Verwendung einer mehrdimensionalen Rasch-Analyse (welche bei Betrachtung des Gesamt-Tests notwendig wäre) verzichtet werden. Statt dessen könnten die beiden Testteile (FDW und FW) getrennt voneinander skaliert werden.

Beim Partial-Credit-Modell erfolgt eine Übertragung der Annahmen des dichotomen Rasch-Modells auf eine mehrstufige Antwortskala (dies betrifft Items, für welche man neben einem Vollscore auch einen Teilscore erhalten kann). Bei Betrachtung der Wahrscheinlichkeiten, mit denen bestimmte (Teil)Antworten gegeben werden, ergeben sich sogenannte Schwellenparameter zwischen den einzelnen Antwortkategorien. Folglich befindet sich ein Studienteilnehmer, je nach Fähigkeitsausprägung, in einer anderen Schwelle (es existieren immer $k - 1$ Schwellen bei k Antwortkategorien bzw. -stufen). Im Idealfall sollten die Schwellen geordnet sein, was mittels einer grafischen Kontrolle bei Annahme des Modells geschehen kann. Die Abstände zwischen den Schwellenparametern müssen nicht unbedingt gleich groß sein, außer es liegen Items vor, die mit Hilfe einer Likert-Skala beantwortet werden sollen (Bühner, 2011).

Für zukünftige Untersuchungen ist angedacht, das Testinstrument FALKO-P einer eingehenden Modifizierung zu unterziehen und gegebenenfalls zu skalieren, um eine (noch) bessere Modellpassung zu erzielen.

9. Ausblick

Insgesamt gesehen bietet sich, den Einsatz des Testinstruments FALKO-P und der weiteren FALKO-Testinstrumente betreffend, eine Vielzahl an Möglichkeiten für weitergehende Untersuchungen und Studien. Dabei kann die Testkonstruktionsphase, über welche hier in dieser Dissertationsschrift und in Krauss, Lindl, Schilcher, Fricke et al. (2017) für alle FALKO-Disziplinen berichtet wird, lediglich als erster Schritt einer Folge möglicher Studien zur Untersuchung einer Vielzahl an Forschungsfragen angesehen werden. Bevor aber eine Vorstellung einiger möglicher Untersuchungsoptionen im Überblick erfolgt, soll betont werden, dass auch, wie in Kapitel 8 dargestellt, zunächst zahlreiche Möglichkeiten der Optimierung des Testinstruments existieren. Solche Weiterentwicklungen können auf eine Item-Ergänzung (Item-Ersatz aus dem bestehenden Pilotierungs-Item-Pool oder Item-Neukonstruktion), vor allem des FDW- aber auch des FW-Teils sowie auf eine Modifizierung des FALKO-Physik-Modells zur Beschreibung des Professionswissens abzielen. Weiterhin ist eine Skalierung des Testinstruments denkbar, was zu einer besseren Beurteilung der Item-Passung und in der Folge vermutlich erneut zu Item-Ersatz oder Item-Neukonstruktionen führen würde.

Betrachtet man beispielsweise die Ergebnisse zum Zusammenhang des fachspezifischen Professionswissens und hier insbesondere des FDW mit der Berufsdauer (vgl. Abschn. 6.4.4 sowie Lindl & Krauss, 2017), sticht einem die im Vorfeld erwartete, aber schlussendlich fehlende positive Korrelation ins Auge. Auch innerhalb der Studien anderer Arbeitsgruppen (z. B. ProwiN) konnte ein solcher Zusammenhang nicht nachgewiesen werden. Dieser fach- und sogar projektübergreifende Befund widerspricht Shulmans (1986) Konzeptualisierung des PCK:

Since there are no single most powerful forms of representation, the teacher must have at hand a veritable armamentarium of alternative forms of representation, some of which derive from research whereas others originate in the wisdom of practice. (Shulman, 1986, S. 9)

Ein diesbezügliches Forschungsdesiderat betrifft die Untersuchung, ob Shulmans Annahme eventuell inkorrekt ist oder ob projektübergreifend alle bisher konstruierten FDW-Items nicht sensitiv auf eine zunehmende Berufserfahrung reagieren beziehungsweise welche Kriterien bei der Konstruktion von Items, die auf Erfahrungswissen reagieren, eine Rolle spielen. Um weiterhin vertieftere Informationen über fachspezifische Kompetenzentwicklungsprozesse bei (angehenden) Lehrkräften zu erhalten, sind längsschnittlich gestaltete Studien anzudenken. Besonders interessant sollten sich hierbei die Übergänge zwischen den drei Phasen der Lehrerbildung gestalten, was eine Längsschnitt-Untersuchung von Studierenden vor dem ersten Staatsexamen über das Referendariat mit dem zweiten Staatsexamen und über die ersten Berufsjahre hinweg erfordern würde (Lindl & Krauss, 2017).

Ein weiteres Beispiel für den Einsatz der FALKO-Testinstrumente stellt eine erste Kooperation zwischen den Teilprojekten FALKO-Physik und FALKO-Pädagogik mittels einer querschnittlich angelegten Untersuchung dar. Hierbei wurde eine Ad-hoc-Stichprobe

bayrischer Physiklehrkräfte ($N = 9$) befragt (vgl. Mulder, Sauer & Kempka, 2017). Um das Persönlichkeitsprofil hinsichtlich des Professionswissens möglichst ganzheitlich zu erfassen, bearbeiteten die Studienteilnehmer sowohl den FALKO-PA- als auch den FALKO-P-Test. Beide Tests wurde von den Lehrkräften im Rahmen von Fortbildungen unter der Aufsicht eines Testleiters auf freiwilliger Basis durchgeführt. Es zeigte sich, dass Studienteilnehmer mit einem geringen Gesamtscore beim PW mittlere bis hohe Gesamtscores beim FW und mittlere Gesamtscores beim FDW erreichen können. Umgekehrt kann festgestellt werden, dass Lehrkräfte mit einem geringeren Gesamtscore beim FW sowie beim FDW einen höheren Gesamtscore beim PW erlangen können. Somit indizieren diese ersten tentativen Ergebnisse, dass das PW in der Tat unabhängig von den beiden fachspezifischen Wissenskategorien sein könnte. Selbstverständlich müsste diese und weitere Hypothesen in größer angelegten quantitativen Studien weiter untersucht werden.

Ferner stellt sich die interessante Frage, ob didaktische Kompetenz „pro Person“ oder „pro Fach“ zu verstehen ist, das heißt, ob hohe didaktische Kompetenz im Unterrichtsfach Physik automatisch auch hohe didaktische Kompetenz im Zweifach, zum Beispiel Mathematik, bedeutet. Mit Hilfe der FALKO-Tests der Einzeldisziplinen könnten solche Zusammenhänge zwischen dem FDW und dem FW (und dem PW) von (angehenden) Lehrkräften, welche oftmals zwei (oder mehrere) Unterrichtsfächer studier(t)en untersucht werden. Auch die Untersuchung von Zusammenhängen mit weiteren Aspekten des Professionswissens (motivationale Orientierungen, selbstregulative Fähigkeiten, etc.) sind denkbar.

Weiterhin ist eine Untersuchung des Einflusses der Kompetenzbereiche des Professionswissens auf Kriterien der Unterrichtsqualität oder auf die Schülerleistung denkbar und wünschenswert (Lindl & Krauss, 2017), zumal gerade für das Fach Physik innerhalb anderer Forschungsprojekte (z. B. ProwiN, ProfiLe-P) diesbezüglich bisher eher enttäuschende Ergebnisse festgestellt wurden (vgl. Kap. 1 sowie Krauss, Lindl, Schilcher, Fricke et al., 2017).

Letzten Endes ist erwähnenswert, dass aus den Forschungsarbeiten der FALKO-Disziplinen heraus ein Folgeprojekt (FALKE - Fachspezifische Lehrerkompetenz im Erklären) hervorging, welches im Rahmen der Qualitätsoffensive Lehrerbildung die im Rahmen von FALKO gewonnenen Erkenntnisse weiterführend untersucht. FALKE plant, die didaktische Qualität unterrichtlichen Erklärens theoretisch und empirisch innerhalb der teilnehmenden Disziplinen (Biologie, Chemie, Deutsch, Englisch, Evangelische Religion, Geschichte, Kunst, Mathematik, Musik, Physik und Grundschuldidaktik; weiterhin beteiligt sind die deutsche Sprachwissenschaft sowie die Sprecherziehung) zu untersuchen. Konkret soll es dabei um unterschiedliche Inszenierungsmöglichkeiten fachlich korrekter Erklärungen gehen. Für jede der beteiligten Disziplinen werden typische unterrichtliche Erklärsituationen in Form kurzer Videovignetten dargestellt. Diese Vignetten sollen anschließend unterschiedlichen Stichproben (z. B. Didaktikern, Seminarlehrern, Lehramtsstudierenden, ...) zur Beurteilung der wahrgenommenen didaktischen Qualität vorgelegt werden. Diese Beurteilung soll mittels eines computergestützten Fragebogens mit geschlossenen Items, welcher einen fachübergreifenden allgemeinen Teil zur Strukturierung, Adressatenorientierung und sprachlichen/sprecherischen Aspekten sowie einen fachspezifischen beziehungsweise themenspezifischen Teil (z. B. Was konkret eine Erklärung im Fach Physik ausmacht) enthält, erfolgen (vgl. Schilcher et al., 2017). Die FALKO-Disziplinen und in diesem Zusammenhang auch das Teilprojekt FALKO-Physik konnten hierbei wertvolle Vorarbeit beispielsweise in Bezug auf die Konzeptualisierung der Professionswissenskategorien oder die Operationalisierung von Testitems leisten.

Literatur

- Abell, S. K. (2007). Research on Science Teacher Knowledge. In S. K. Abell & N. G. Lederman (Hrsg.), *Handbook of Research on Science Education* (S. 1105–1149). Mahwah: Lawrence Erlbaum Associates.
- Abell, S. K. & Lederman, N. G. (Hrsg.). (2007). *Handbook of Research on Science Education*. Mahwah: Lawrence Erlbaum Associates.
- Allemann-Ghionda, C. (Hrsg.). (2006). *Kompetenzen und Kompetenzentwicklung von Lehrerinnen und Lehrern- Ausbildung und Beruf*. Weinheim: Beltz.
- Anderson, L. M., Evertson, C. M. & Brophy, J. E. (1979). An Experimental Study of Effective Teaching in First-Grade Reading Groups. *The Elementary School Journal*, 79(4), 193–223.
- Anderson, L. M., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E. & Pintrich, P. R. (2001). *A Taxonomy for Learning, Teaching and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. New York: Longman.
- Arnold, K.-H. (Hrsg.). (2007). *Unterrichtsqualität und Fachdidaktik*. Bad Heilbrunn: Klinkhardt.
- Baron-Boldt, J., Schuler, H. & Funke, U. (1988). Praedikative Validitaet von Schulabschlussnoten: Eine Metaanalyse. *Zeitschrift für pädagogische Psychologie*, 2(2), 79–90.
- Bass, H. & Ball, D. L. (2004). A practice-based theory of mathematical knowledge for teaching: The case of mathematical reasoning. In J. Wang & B. Xu (Hrsg.), *Trends and challenges in mathematics education* (S. 295–313). Shanghai: East China Normal University Press.
- Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., ... Weiß, M. (2001). *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Opladen: Leske + Budrich.
- Baumert, J. & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 9(4), 469–520.
- Baumert, J. & Kunter, M. (2011a). Das Kompetenzmodell von COACTIV. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV* (S. 29–54). Münster: Waxmann.
- Baumert, J. & Kunter, M. (2011b). Das mathematikspezifische Wissen von Lehrkräften, kognitive Aktivierung im Unterricht und Lernfortschritte von Schülerinnen und Schülern. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV* (S. 163–192). Münster: Waxmann.
- Baumert, J. & Kunter, M. (2012). Professionelle Kompetenz von Lehrkräften. In I. Gogolin, H. Kuper, H. H. Krüger & J. Baumert (Hrsg.), *Stichwort: Zeitschrift für Erziehungswissenschaft* (S. 277–337). Baumert, J. & Kunter, M. (2012). Professionelle Kompetenz von Lehrkräften. In I. Gogolin, Wiesbaden: Springer Fachmedien Wiesbaden.

- Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., ... Tsai, Y.-M. (2010). Teachers' Mathematical Knowledge, Cognitiv Activation in the Classroom, and Student Progress. *American Educational Research Journal*, 47(1), 133–180.
- Baumert, J., Kunter, M., Blum, W., Klusmann, U., Krauss, S. & Neubrand, M. (2011). Professionelle Kompetenz von Lehrkräften, kognitiv aktivierender Unterricht und die mathematische Kompetenz von Schülerinnen und Schülern (COACTIV) - Ein Forschungsprogramm. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV* (S. 7–25). Münster: Waxmann.
- Bernholt, S. (Hrsg.). (2014a). *Naturwissenschaftliche Bildung zwischen Science und Fachunterricht: Jahrestagung in München 2013*. Kiel: IPN. Gesellschaft für Didaktik der Chemie und Physik. Zugriff unter <http://www.gdcp.de/index.php/tagungsbaende/tagungsband-uebersicht/145-tagungsbaende/2013/4220-band33>
- Bernholt, S. (Hrsg.). (2014b). *Naturwissenschaftliche Bildung zwischen Science- und Fachunterricht. Gesellschaft für Didaktik der Chemie und Physik, Jahrestagung in München 2013*. Kiel: IPN.
- Bernholt, S. (Hrsg.). (2014c). *Naturwissenschaftliche Bildung zwischen Science- und Fachunterricht. Gesellschaft für Didaktik der Chemie und Physik, Jahrestagung in München 2013*. Kiel: IPN.
- Bernholt, S. (Hrsg.). (2015). *Heterogenität und Diversität - Vielfalt der Voraussetzungen im naturwissenschaftlichen Unterricht: Jahrestagung in Bremen 2014*. Kiel: IPN. Gesellschaft für Didaktik der Chemie und Physik.
- Berry, A., Friedrichsen, P. & Loughran, J. (Hrsg.). (2015). *Re-Examining Pedagogical Content Knowledge in Science Education*. London: Taylor and Francis.
- Besser, M. & Krauss, S. (2016). Der Lehrer als Experte: Professionswissen verstehen und entwickeln. *Journal für LehrerInnenbildung*, 16(4), 42–47.
- Blömeke, S., Bremerich-Vos, A., Haudeck, H., Kaiser, G., Nold, G., Schwippert, K. & Willenberg, H. (Hrsg.). (2011). *Kompetenzen von Lehramtsstudierenden in gering strukturierten Domänen, Erste Ergebnisse aus TEDS-LT*. Münster: Waxmann.
- Blömeke, S., Bremerich-Vos, A., Kaiser, G., Nold, G., Haudeck, H., Keßler, J.-U. & Schwippert, K. (Hrsg.). (2013). *Professionelle Kompetenzen im Studienverlauf. Weitere Ergebnisse zur Deutsch-, Englisch-, und Mathematiklehrerausbildung aus TEDS-LT*. Münster: Waxmann.
- Blömeke, S., Felbrich, A. & Müller, C. (2008). Messung des erziehungswissenschaftlichen Wissens angehender Lehrkräfte. In S. Blömeke, G. Kaiser & R. Lehmann (Hrsg.), *Professionelle Kompetenz angehender Lehrerinnen und Lehrer. Erste Ergebnisse zur Wirksamkeit der Lehrerausbildung: Wissen, Überzeugungen und Lerngelegenheiten deutscher Mathematikstudierender und -referendare* (S. 171–194). Münster: Waxmann.
- Blömeke, S., Gustafson, J. E. & Shavelson, R. J. (2015). Beyond Dichotomies. Competence viewed as a continuum. *Zeitschrift für Psychologie*, 223(1), 3–13. doi:10.1027/2151-2604/a000194
- Blömeke, S., Kaiser, G. & Lehmann, R. (Hrsg.). (2008). *Professionelle Kompetenz angehender Lehrerinnen und Lehrer. Erste Ergebnisse zur Wirksamkeit der Lehrerausbildung: Wissen, Überzeugungen und Lerngelegenheiten deutscher Mathematikstudierender und -referendare*. Münster: Waxmann.
- Blömeke, S., Kaiser, G., Lehmann, R., König, J., Döhrmann, M., Buchholtz, C. & Hacke, S. (2009). TEDS-M: Messung von Lehrerkompetenzen im internationalen Vergleich. In

- O. Zlatkin-Troitschanskaia, K. Beck, D. Sembill, R. Nickolaus & R. Mulder (Hrsg.), *Lehrerprofessionalität. Bedingungen, Genese, Wirkungen und ihre Messung* (S. 181–209). Weinheim: Beltz.
- Blömeke, S., Seeber, S., Lehmann, R., Kaiser, G., Schwarz, B., Felbrich, A. & Müller, C. (2008). Messung des fachbezogenen Wissens angehender Mathematiklehrkräfte. In S. Blömeke, G. Kaiser & R. Lehmann (Hrsg.), *Professionelle Kompetenz angehender Lehrerinnen und Lehrer. Erste Ergebnisse zur Wirksamkeit der Lehrerausbildung: Wissen, Überzeugungen und Lerngelegenheiten deutscher Mathematikstudierender und -referendare* (S. 49–88). Münster: Waxmann.
- Bortz, J. & Döring, N. (2006). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler* (4. überarbeitete). Heidelberg: Springer.
- Braaten, M. & Windschitl, M. (2011). Working toward a stronger conceptualization of scientific explanation for science education. *Science Education*, 95(4), 639–669.
- Bremerich-Vos, A., Dämmer, J., Willenberg, H. & Schwippert, K. (2011). Professionelles Wissen von Studierenden des Lehramts Deutsch. In S. Blömeke, A. Bremerich-Vos, H. Haudeck, G. Kaiser, G. Nold, K. Schwippert & H. Willenberg (Hrsg.), *Kompetenzen von Lehramtsstudierenden in gering strukturierten Domänen, Erste Ergebnisse aus TEDS-LT* (S. 47–76). Münster: Waxmann.
- Bromme, R. (1997). Kompetenzen, Funktionen und unterrichtliches Handeln des Lehrers. In F. E. Weinert (Hrsg.), *Psychologie des Unterrichts und der Schule: Pädagogische Psychologie (Enzyklopädie der Psychologie)* (Bd. 1 (3), 3, S. 177–212). Göttingen: Hogrefe.
- Bromme, R. & Haag, L. (2004). Forschung zur Lehrerpersönlichkeit. In W. Helsper & J. Böhme (Hrsg.), *Handbuch zur Schulforschung* (1. Aufl., S. 777–793). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Bruckmaier, G., Krauss, S., Blum, W. & Leiss, D. (2016). Measuring mathematics teachers' professional competence by using video clips (COACTIV video). *ZDM – The International Journal on Mathematics Education*, 48(1-2). Heftnummer und Seitenzahl?
- Bruguière, C., Tiberghien, A. & Clément, P. (Hrsg.). (2014). *Topics and Trends in Current Science Education*. Dordrecht u. a.: Springer.
- Brunner, M. & Süß, H.-M. (2007). Wie genau können kognitive Fähigkeiten gemessen werden? *Diagnostica*, 53(4), 184–193. doi:10.1026/0012-1924.53.4.184
- Buchholtz, N. & Kaiser, G. (2013). Professionelles Wissen im Studienverlauf: Lehramt Mathematik. In S. Blömeke, A. Bremerich-Vos, G. Kaiser, G. Nold, H. Haudeck, J.-U. Keßler & K. Schwippert (Hrsg.), *Professionelle Kompetenzen im Studienverlauf. Weitere Ergebnisse zur Deutsch-, Englisch-, und Mathematiklehrausbildung aus TEDS-LT* (S. 107–143). Münster: Waxmann.
- Buchholtz, N., Kaiser, G. & Stancel-Piątak, A. (2011). Professionelles Wissen von Studierenden des Lehramts Mathematik. In S. Blömeke, A. Bremerich-Vos, H. Haudeck, G. Kaiser, G. Nold, K. Schwippert & H. Willenberg (Hrsg.), *Kompetenzen von Lehramtsstudierenden in gering strukturierten Domänen, Erste Ergebnisse aus TEDS-LT* (S. 101–133). Münster: Waxmann.
- Bühner, M. (2011). *Einführung in die Test- und Fragebogenkonstruktion* (3., aktualisierte und erw. Aufl.). Always Learning. München: Pearson Studium.
- Bundesministerium für Bildung und Forschung (Hrsg.). (2008). *Bildungsforschung*. Bonn.
- Carle, U. & Buchen, S. (Hrsg.). (1999). *Jahrbuch für Unterrichtsforschung*. München.

- Cauet, E. (2016). Testen wir relevantes Wissen? Zusammenhang zwischen dem Professionswissen von Physiklehrkräften und gutem und erfolgreichem Unterrichten. In H. Niedderer, H. Fischler & E. Sumfleth (Hrsg.), *Studien zum Physik- und Chemielernen* (Bd. 204). Cauet, Eva (Verfasser). Berlin: Logos.
- Cauet, E., Borowski, A. & Fischer, H. E. (2015). Testen wir relevantes Wissen? Professionswissen von Physiklehrkräften. In S. Bernholt (Hrsg.), *Heterogenität und Diversität - Vielfalt der Voraussetzungen im naturwissenschaftlichen Unterricht: Jahrestagung in Bremen 2014* (35, S. 136–138). Gesellschaft für Didaktik der Chemie und Physik. Kiel: IPN.
- Cauet, E., Liepertz, S., Kirschner, S., Borowski, A. & Fischer, H. E. (2014). Professionswissen von Physiklehrkräften und Schülerleistungen. In S. Bernholt (Hrsg.), *Naturwissenschaftliche Bildung zwischen Science und Fachunterricht: Jahrestagung in München 2013* (34, S. 141–143). Gesellschaft für Didaktik der Chemie und Physik. Kiel: IPN.
- Cohen, J. (1988). *Statistical analysis for the behavioral sciences*. Hillsdale: Lawrence Erlbaum.
- Cohen, J. (1992). A power primer. *Psychological bulletin*, 112(1), 155–159.
- Depaepe, F., Verschaffel, L. & Kelchtermans, G. (2013). Pedagogical content knowledge: A systematic review of the way in which the concept has pervaded mathematics educational research. *Teaching and Teacher Education*, 34, 12–25.
- Dollny, S. (2011). Entwicklung und Evaluation eines Testinstruments zur Erfassung des fachspezifischen Professionswissens von Chemielehrkräften. In H. Niedderer, H. Fischler & E. Sumfleth (Hrsg.), *Studien zum Physik- und Chemielernen* (Bd. 127). Berlin: Logos.
- Dorn, F. & Bader, F. (2002). *Physik in einem Band* (Neubearb. - Druck A 1). Bader, Franz (VerfasserIn). Hannover: Schroedel-Schulbuchverlag.
- Enders, C. K. (2010). *Applied missing data analysis. Methodology in the social sciences*. New York: Guilford Press.
- Epstein, L. C. & Lessing, H. E. (2011). *Denksport-Physik: Fragen und Antworten* (Ungekürzte Ausg., 9. Aufl.). dtv. München: Dt. Taschenbuch-Verl.
- Ericsson, K. A., Krampe, R. T. & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, (100), 363–406.
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics: And sex and drugs and rock 'n' roll* (4. ed.). MobileStudy. Los Angeles u.a.: Sage Publ.
- Fischer, H. E. & Sumfleth, E. (Hrsg.). (2013). *nwu-essen - 10 Jahre Essener Forschung zum naturwissenschaftlichen Unterricht*. Berlin: Logos.
- Fisseni, H. (1997). *Lehrbuch der psychologischen Diagnostik* (2.). Göttingen: Hogrefe.
- Förtsch, C., Werner, S., von Kotzebue, L. & Neuhaus, B. (2016). Effects of biology teachers' professional knowledge and cognitive activation on students' achievement. *International Journal of Science Education*. doi:10.1085/09500693.2016.1257170
- Frey, A. (2006). Methoden und Instrumente zur Diagnose beruflicher Kompetenzen von Lehrkräften - Eine erste Standortbestimmung zu bereits publizierten Instrumenten. In C. Allemann-Ghionda (Hrsg.), *Kompetenzen und Kompetenzentwicklung von Lehrerinnen und Lehrern- Ausbildung und Beruf* (S. 30–45). Weinheim: Beltz.
- Friege, G. & Lind, G. (2004). Leistungsmessung im Leistungskurs. *Der mathematische und naturwissenschaftliche Unterricht - MNU*, 57(5), 259–265.
- Gelman, A., Carlin, J. B., Stern H. S., Dunson, D. B., Vehtari, A. & Rubin, D. B. (2013). *Bayesian Data Analysis* (3. Aufl.). Boca Raton: CRC Press.

- 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). Dordrecht: Kluwer Academic Publishers.
- Gess-Newsome, J. & Lederman, N. G. (Hrsg.). (1999). *Examining pedagogical content knowledge*. Dordrecht: Kluwer Academic Publishers.
- Gogolin, I., Kuper, H., Krüger, H. H. & Baumert, J. (Hrsg.). (2012). *Stichwort: Zeitschrift für Erziehungswissenschaft*. Wiesbaden: Springer Fachmedien Wiesbaden.
- Gogolin, I., Kuper, H., Krüger, H.-H. & Baumert, J. (Hrsg.). (2013). *Stichwort: Zeitschrift für Erziehungswissenschaft*. Wiesbaden: Springer Fachmedien Wiesbaden. doi:10.1007/978-3-658-00908-3
- Gramzow, Y. (2014). Fachdidaktisches Wissen von Lehramtsstudierenden im Fach Physik – Modellierung und Testkonstruktion. In H. Niedderer, H. Fischler & E. Sumfleth (Hrsg.), *Studien zum Physik- und Chemielernen* (Bd. 181). Berlin: Logos.
- Gramzow, Y., Riese, J. & Reinhold, P. (2013). Modellierung fachdidaktischen Wissens angehender Physiklehrkräfte. *Zeitschrift für Didaktik der Naturwissenschaften*, (19), 31–49.
- Gramzow, Y., Riese, J. & Reinhold, P. (2014). DaWiS: Teilprojekt Fachdidaktisches Wissen. In S. Bernholt (Hrsg.), *Naturwissenschaftliche Bildung zwischen Science- und Fachunterricht. Gesellschaft für Didaktik der Chemie und Physik, Jahrestagung in München 2013* (S. 111–113). Kiel: IPN.
- Gräsel, C. & Trempler, K. (Hrsg.). (2017). *Entwicklung von Professionalität pädagogischen Personals. Interdisziplinäre Betrachtungen, Befunde und Perspektiven*. Wiesbaden: Springer.
- Grehn, J. & Krause, J. (Hrsg.). (2003). *Metzler Physik*. Hannover: Schroedel Verlag.
- Grehn, J. & Krause, J. (Hrsg.). (2005). *Metzler Physik Lösungen*. Hannover: Schroedel Verlag.
- Grossman, P. L. (1990). *The making of a teacher: Teacher knowledge and teacher education*. New York: Teachers College Press.
- Großschedl, J., Harms, U., Glowinski, I. & Waldmann, M. (2014). Professionswissen angehender Biologielehrkräfte: Das KiL-Projekt. *Der mathematische und naturwissenschaftliche Unterricht*, 67(8), 457–462.
- Gruber, H. & Renkl, A. (2000). Die Kluft zwischen Wissen und Handeln: Das Problem des trägen Wissens. In G. H. Neuweg (Hrsg.), *Wissen - Können - Reflexion: Ausgewählte Verhältnisbestimmungen* (S. 155–174). Innsbruck: Studienverlag.
- Hashweh, M. Z. (1987). Effects of subject-matter knowledge in the teaching of Biology and Physics. *Teaching and Teacher Education*, 3(2), 109–120.
- Helmke, A. (2007). *Unterrichtsqualität - erfassen, bewerten, verbessern* (5. Aufl.). Seelze-Velber: Erhard Friedrich.
- Helmke, A. (2009). *Unterrichtsqualität und Lehrerprofessionalität: Diagnose, Evaluation und Verbesserung des Unterrichts* (1. Aufl.). Seelze-Velber: Erhard Friedrich.
- Helsper, W. & Böhme, J. (Hrsg.). (2004). *Handbuch zur Schulforschung*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Herrmann, U. (1999). „Lehrer“ - Experte und Autodidakt? Bemerkungen zu den strukturellen Möglichkeiten und Grenzen der „Professionalität“ und der „Professionalisierbarkeit“ des Lehrers und seiner beruflichen Praxen. In U. Carle & S. Buchen (Hrsg.), *Jahrbuch für Unterrichtsforschung* (Bd. 2, 2, S. 33–48). München.

- Hill, H. C., Loewenberg Ball, D. & Schilling, S. G. (2008). Unpacking Pedagogical Content Knowledge: Conceptualizing and Measuring Teachers' Topic-Specific Knowledge of Students. *Journal for Research in Mathematics Education*, 371–400.
- Hill, H. C., Schilling, S. G. & Ball, D. L. (2004). Developing measures of teachers' mathematics knowledge for teaching. *The Elementary School Journal*, 105(1), 11–30.
- Hofstein, A. & Lunetta, V. N. (2004). The Laboratory in Science Education: Foundations for the Twenty-First Century. *Science Education*, 88(1), 28–54.
- Hohenstein, F., Kleickmann, T., Zimmermann, F., Köller, O. & Möller, J. (in press). Erfassung von pädagogischen und psychologischem Wissen in der Lehramtsausbildung: Entwicklung eines Messinstruments. *Zeitschrift für Pädagogik*.
- Höttecke, D. (Hrsg.). (2010). Entwicklung naturwissenschaftlichen Denkens zwischen Phänomen und Systematik, Münster: Lit, 30. Gesellschaft für Didaktik der Chemie und Physik.
- Hu, L. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55.
- Ingenkamp, K. (Hrsg.). (1970). *Handbuch der Unterrichtsforschung, Teil II*. Weinheim: Julius Beltz.
- Jansing, B., Haudeck, H., Keßler, J.-U., Nold, G. & Stancel-Piątak, A. (2013). Professionelles Wissen im Studienverlauf: Lehramt Englisch. In S. Blömeke, A. Bremerich-Vos, G. Kaiser, G. Nold, H. Haudeck, J.-U. Keßler & K. Schwippert (Hrsg.), *Professionelle Kompetenzen im Studienverlauf. Weitere Ergebnisse zur Deutsch-, Englisch-, und Mathematiklehrausbildung aus TEDS-LT* (S. 77–106). Münster: Waxmann.
- Jung, W. (1986). Alltagsvorstellungen und das Lernen von Physik und Chemie. *Naturwissenschaften im Unterricht Chemie/Physik* 34, (13), 2–6.
- Jüttner, M., Boone, W., Park, S. & Neuhaus, B. J. (2013). Development and use of a test instrument to measure biology teachers' content knowledge (CK) and pedagogical content knowledge (PCK). *Educational Assessment, Evaluation and Accountability*, 25(1), 45–67.
- Jüttner, M. & Neuhaus, B. J. (2013). Das Professionswissen von Biologielehrkräften - Ein Vergleich zwischen Biologielehrkräften, Biologen und Pädagogen. *Zeitschrift für Didaktik der Naturwissenschaften*, 19, 31–49.
- Kirschner, S. (2013). Modellierung und Analyse des Professionswissens von Physiklehrkräften. In H. Niedderer, H. Fischler & E. Sumfleth (Hrsg.), *Studien zum Physik- und Chemielernen* (Bd. 161). Berlin: Logos.
- Kirschner, S., Sczudlek, M., Tepner, O., Borowski, A., Fischer, H. E., Lenke, G., ... Wirth, J. (2017). Professionswissen in den Naturwissenschaften (ProwiN). In C. Gräsel & K. Trempler (Hrsg.), *Entwicklung von Professionalität pädagogischen Personals. Interdisziplinäre Betrachtungen, Befunde und Perspektiven* (S. 113–130). Wiesbaden: Springer.
- Kleickmann, T., Großschedl, J., Harms, U., Heinze, A., Herzog, S., Hohenstein, F., ... Zimmermann, F. (2014). Professionswissen von Lehramtsstudierenden der mathematisch-naturwissenschaftlichen Fächer – Testentwicklung im Rahmen des Projektes KIL. *Unterrichtswissenschaft*, 42(3), 280–288.
- Kleickmann, T., Richter, D., Kunter, M., Elsner, J., Besser, M., Krauss, S. & Baumert, J. (2013). Teachers' Content Knowledge and Pedagogical Content Knowledge: The Role of Structural Differences in Teacher Education. *Journal of Teacher Education*, 64(1), 90–106.

- KMK. (2004). Bildungsstandards im Fach Physik für den mittleren Schulabschluss (Jahrgangsstufe 10) - Beschluss der Kultusministerkonferenz vom 16.12.2004.
- Kock, Z.-J., Taconis, R., Bolhuis, S. & Gravemeijer, K. (2013). Some Key Issues in Creating Inquiry-Based Instructional Practices that Aim at the Understanding of Simple Electric Circuits. *Research in Science Education*, 43, 579–597.
- König, J., Lammerding, S., Nold, G., Rohde, A., Strauß, S. & Tachtsoglou, S. (2016). Teachers' Professional Knowledge for Teaching English as a Foreign Language: Assessing the Outcomes of Teacher Education. *Journal of Teacher Education*, 67(4), 320–337.
- Korneck, F., Lamprecht, J., Wodzinski, R. & Schecker, H. (2010). *Quereinsteiger in das Lehramt Physik – Lage und Perspektiven der Physiklehrausbildung in Deutschland*. Deutsche Physikalische Gesellschaft (DPG).
- Krauss, S., Blum, W., Brunner, M., Neubrand, M., Baumert, J., Kunter, M., ... Elsner, J. (2011). Konzeptualisierung und Testkonstruktion zum fachbezogenen Professionswissen von Mathematiklehrkräften. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV* (S. 135–162). Münster: Waxmann.
- Krauss, S. & Bruckmaier, G. (2014). Das Experten-Paradigma in der Forschung zum Lehrerberuf. In R. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrberuf* (2. Aufl., S. 241–261). Münster: Waxmann.
- Krauss, S., Lindl, A., Schilcher, A., Fricke, M., Göhring, A., Hofmann, B., ... Mulder, R. H. (Hrsg.). (2017). *FALKO: Fachspezifische Lehrerkompetenzen. Konzeption von Professionswissenstests in den Fächern Deutsch, Englisch, Latein, Physik, Musik, Evangelische Religion und Pädagogik*. Münster: Waxmann.
- Krauss, S., Lindl, A., Schilcher, A. & Tepner, O. (2017). Das Forschungsprojekt FALKO – ein einleitender Überblick. In S. Krauss, A. Lindl, A. Schilcher, M. Fricke, A. Göhring, B. Hofmann, ... R. H. Mulder (Hrsg.), *FALKO: Fachspezifische Lehrerkompetenzen. Konzeption von Professionswissenstests in den Fächern Deutsch, Englisch, Latein, Physik, Musik, Evangelische Religion und Pädagogik* (S. 9–65). Münster: Waxmann.
- 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 (JMD)*, 29(3/4), 223–258.
- Kröger, J., Neumann, K. & Petersen, S. (2014). Struktur und Entwicklung des Professionswissens angehender Physiklehrkräfte. In S. Bernholt (Hrsg.), *Naturwissenschaftliche Bildung zwischen Science- und Fachunterricht. Gesellschaft für Didaktik der Chemie und Physik, Jahrestagung in München 2013* (S. 117–119). Kiel: IPN.
- Kulgemeyer, C., Borowski, A., Fischer, H. E., Gramzow, Y., Reinhold, P., Riese, J., ... Walzer, M. (2012). ProfiLe-P - Professionswissen in der Lehramtsausbildung Physik: Vorstellung eines Forschungsprojekts. In V. Nordmeier & H. Grötzbauch (Hrsg.), *Didaktik der Physik: Beiträge zur DPG-Frühjahrstagung 2012 in Mainz*. Berlin: FU Berlin.
- Kulgemeyer, C. & Schecker, H. (2013). Students explaining science – assessment of science communication competence. *Research in Science Education*, (43), 2235–2256.
- Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S. & Neubrand, M. (Hrsg.). (2011). *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV*. Münster: Waxmann.

- Labudde, P. (Hrsg.). (2010). *Fachdidaktik Naturwissenschaft: 1.-9. Schuljahr*. Bern: Haupt Verlag.
- Lange, K. (1895). *Lehrmethode und Lehrerpersönlichkeit*. Zwickau.
- Lange, K. (2010). *Zusammenhänge zwischen naturwissenschaftsbezogenem fachspezifisch-pädagogischem Wissen von Grundschullehrkräften und Fortschritten im Verständnis naturwissenschaftlicher Konzepte bei Grundschülerinnen und -schülern*. Berlin: Logos.
- Lange, K., Kleickmann, T., Tröbst, S. & Möller, K. (2012). Fachdidaktisches Wissen von Lehrkräften und multiple Ziele im naturwissenschaftlichen Sachunterricht. *Zeitschrift für Erziehungswissenschaft*, 15(1), 55–75.
- Lavonen, J., Juuti, K., Lampiselkä, J., Uitto, A. & Hahl, K. (Hrsg.). (2016). *Electronic Proceedings of the ESERA 2015 Conference. Science education research: Engaging learners for a sustainable future*. Helsinki.
- Lienert, G. & Raatz, U. (1998). *Testaufbau und Testanalyse* (6. Aufl.). Weinheim: Beltz.
- Liepert, S. & Borowski, A. (2015). Einfluss des Professionswissens von Physiklehrkräften auf Sachstruktur und Schülerleistung. In S. Bernholt (Hrsg.), *Heterogenität und Diversität - Vielfalt der Voraussetzungen im naturwissenschaftlichen Unterricht: Jahrestagung in Bremen 2014* (35, S. 139–141). Gesellschaft für Didaktik der Chemie und Physik. Kiel: IPN.
- Lindl, A. & Krauss, S. (2017). Transdisziplinäre Perspektiven auf domänenspezifische Lehrerkompetenzen. Eine Metaanalyse zentraler Resultate des Forschungsprojektes FALKO. In S. Krauss, A. Lindl, A. Schilcher, M. Fricke, A. Göhring, B. Hofmann, ... R. H. Mulder (Hrsg.), *FALKO: Fachspezifische Lehrerkompetenzen. Konzeption von Professionswissenstests in den Fächern Deutsch, Englisch, Latein, Physik, Musik, Evangelische Religion und Pädagogik* (S. 381–438). Münster: Waxmann.
- Lindmeier, A. M., Heinze, A. & Reiss, K. (2013). Eine Machbarkeitsstudie zur Operationalisierung aktionsbezogener Kompetenz von Mathematiklehrkräften mit videobasierten Maßen. *Journal für Mathematik-Didaktik (JMD)*, 34(1), 99–119. doi:10.1007/513138-012-0046-6
- Lipowsky, F. (2006). Auf den Lehrer kommt es an: Empirische Evidenzen für Zusammenhänge zwischen Lehrerkompetenzen, Lehrerhandeln und dem Lernen der Schüler. *Zeitschrift für Pädagogik*, (Beiheft 51), 47–70. Oder In E. Terhart & C. Allemann-Ghionda (Hrsg.), *Kompetenzen und Kompetenzentwicklung von Lehrerinnen und Lehrern: Ausbildung und Beruf* (Bd. 51, S. 47–70). Weinheim: Beltz Verlag.
- Little, T. D., Cunningham, W. A., Shahar, G. & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighting the merits. *Structural Equation Modeling: A Multidisciplinary Journal*, (9), 151–173.
- Loewenberg Ball, D., Theule Lubienski, S. & Spangler Mewborn, D. (2001). Research in Teaching Mathematics: The Unsolved Problem of Teachers' Mathematical Knowledge. In V. Richardson (Hrsg.), *Handbook of research on teaching* (4. Aufl., S. 433–456). New York: Macmillan.
- 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* (Bd. 6, S. 95–132). Dordrecht: Kluwer Academic Publishers.
- Maurer, C. (Hrsg.). (2016). *Authentizität und Lernen - das Fach in der Fachdidaktik. Gesellschaft für Didaktik der Chemie und Physik, Jahrestagung in Berlin*.
- Möller, K., Kleickmann, T. & Lange, K. (2013). Naturwissenschaftliches Lernen im Übergang von der Grundschule zur Sekundarstufe. In H. E. Fischer & E. Sumfleth (Hrsg.),

- nwu-essen - 10 Jahre Essener Forschung zum naturwissenschaftlichen Unterricht* (S. 57–120). Berlin: Logos.
- Moosbrugger, H. & Kelava, A. (2012). *Testtheorie und Fragebogenkonstruktion*. Heidelberg: Springer.
- Muckenfuß, H. (1995). *Lernen im sinnstiftenden Kontext. Entwurf einer zeitgemäßen Didaktik des Physikunterrichts*. Berlin: Cornelsen.
- Mulder, R. H., Sauer, S. & Kempka, F. (2017). FALKO-PA: Ein Instrument aus flexibel einsetzbaren Vignetten zur Erfassung pädagogischer Kompetenzen. In S. Krauss, A. Lindl, A. Schilcher, M. Fricke, A. Göhring, B. Hofmann, ... R. H. Mulder (Hrsg.), *FALKO: Fachspezifische Lehrerkompetenzen. Konzeption von Professionswissenstests in den Fächern Deutsch, Englisch, Latein, Physik, Musik, Evangelische Religion und Pädagogik* (S. 337–379). Münster: Waxmann.
- Müller, R., Wodzinski, R. & Hopf, M. (Hrsg.). (2011). *Schülervorstellungen in der Physik*. Aulis Verlag in der Stark Verlagsgesellschaft.
- Neuweg, G. H. (Hrsg.). (2000). *Wissen - Können - Reflexion: Ausgewählte Verhältnisbestimmungen*. Innsbruck: Studienverlag.
- Neuweg, H. G. (2014). Das Wissen der Wissensvermittler. Problemstellungen, Befunde und Perspektiven der Forschung zum Lehrwissen. In R. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrberuf* (2. Aufl., S. 583–614). Münster: Waxmann.
- Niedderer, H., Fischler, H. & Sumfleth, E. (Hrsg.). (2009). *Studien zum Physik- und Chemielernen*. Berlin: Logos.
- Niedderer, H., Fischler, H. & Sumfleth, E. (Hrsg.). (2010). *Studien zum Physik- und Chemielernen*. Berlin: Logos.
- Niedderer, H., Fischler, H. & Sumfleth, E. (Hrsg.). (2011). *Studien zum Physik- und Chemielernen*. Berlin: Logos.
- Niedderer, H., Fischler, H. & Sumfleth, E. (Hrsg.). (2013). *Studien zum Physik- und Chemielernen*. Berlin: Logos.
- Niedderer, H., Fischler, H. & Sumfleth, E. (Hrsg.). (2014). *Studien zum Physik- und Chemielernen*. Berlin: Logos.
- Niedderer, H., Fischler, H. & Sumfleth, E. (Hrsg.). (2015). *Studien zum Physik- und Chemielernen*. Berlin: Logos.
- Niedderer, H., Fischler, H. & Sumfleth, E. (Hrsg.). (2016). *Studien zum Physik- und Chemielernen*. Berlin: Logos.
- Nordmeier, V. & Grötzbauch, H. (Hrsg.). (2012). *Didaktik der Physik: Beiträge zur DPG-Frühjahrstagung 2012 in Mainz*. Berlin: FU Berlin.
- Northfield, J. R. & Fraser, B. J. (1977). Teacher Characteristics and Pupil Outcomes in Secondary Science Classroom. *Research in Science Education*, (7), 113–121.
- Nunally, J. C. & Bernstein, I. H. (1994). *Psychometric Theory* (3. Aufl.). New York: McGraw-Hill.
- Oelkers, J. & Reusser, K. (2008). Expertise: Qualität entwickeln - Standards sichern - mit Differenz umgehen. In Bundesministerium für Bildung und Forschung (Hrsg.), *Bildungsforschung* (Bd. 27). Bonn.
- Ohle, A. (2010). Primary School Teachers' Content Knowledge in Physics and its Impact on Teaching and Students' Achievement. In H. Niedderer, H. Fischler & E. Sumfleth (Hrsg.), *Studien zum Physik- und Chemielernen* (Bd. 110). Berlin: Logos.

- 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.
- Ohle, A., Kauertz, A. & Fischer, H. E. (2010). Einfluss des Fachwissens von Lehrkräften auf Unterricht und Leistung. In D. Höttecke (Hrsg.), *Entwicklung naturwissenschaftlichen Denkens zwischen Phänomen und Systematik* (30, S. 380–382). Gesellschaft für Didaktik der Chemie und Physik. Jahrestagung in Dresden 2009. Münster: Lit.
- Olszewski, J. (2010). The impact of physics teachers' Pedagogical content knowledge on teacher action and student outcomes. In H. Niedderer, H. Fischler & E. Sumfleth (Hrsg.), *Studien zum Physik- und Chemielernen* (Bd. 109). Berlin: Logos.
- Olszewski, J., Neumann, K. & Fischer, H. E. (2010). Fachdidaktisches Wissen von Physiklehrkräften und dessen Einfluss auf Unterrichtsqualität. In D. Höttecke (Hrsg.), *Entwicklung naturwissenschaftlichen Denkens zwischen Phänomen und Systematik* (30, S. 374–376). Gesellschaft für Didaktik der Chemie und Physik. Jahrestagung in Dresden 2009. Münster: Lit.
- Oser, F. (2002). Standards in der Lehrerbildung - Entwurf einer Theorie kompetenzbezogener Professionalisierung. *Journal für Lehrerinnen- und Lehrerbildung*, 2(1), 8–19.
- Paris, S. G., Lipson, M. Y. & Wixson, K. K. (1983). Becoming a Strategic Reader. *Contemporary Educational Psychology*, (8), 293–316.
- Park, S. & Chen, Y.-C. (2012). Mapping out the integration of the components of pedagogical content knowledge (PCK): PCK as a conceptual tool to understand teachers as professionals. *Research in Science Education*, 38(3), 261–284.
- Pause, G. (1970). Merkmale der Lehrerpersönlichkeit. In K. Ingenkamp (Hrsg.), *Handbuch der Unterrichtsforschung, Teil II* (S. 1353–1526). Weinheim: Julius Beltz.
- PCK Summit. (2012). Zugriff 15. Februar 2016 unter <http://pcksummit.bsccs.org/>
- Prenzel, M., Baumert, J., Blum, W., Lehmann, R., Leutner, D., Neubrand, M., ... Schiefele, U. (Hrsg.). (2004). *PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland – Ergebnisse des zweiten internationalen Vergleichs*. Münster: Waxmann.
- Reiss, K., Sälzer, C., Schiepe-Tiska, A., Klieme, E. & Köller, O. (2016). *PISA 2015. Eine Studie zwischen Kontinuität und Innovation*. Münster: Waxmann.
- Rhöneck, C. v. (1986a). Alltagsvorstellungen und das Lernen von Physik und Chemie. *Naturwissenschaften im Unterricht Chemie/Physik* 34, (13), 10–14.
- Rhöneck, C. v. (1986b). Vorstellungen vom elektrischen Stromkreis und zu den Begriffen Strom, Spannung und Widerstand. *Naturwissenschaften im Unterricht Chemie/Physik* 34, (13), 10–14.
- Richardson, V. (Hrsg.). (2001). *Handbook of research on teaching*. New York: Macmillan.
- Riese, J. (2009). Professionelles Wissen und professionelle Handlungskompetenz von (angehenden) Physiklehrkräften. In H. Niedderer, H. Fischler & E. Sumfleth (Hrsg.), *Studien zum Physik- und Chemielernen* (Bd. 97). Berlin: Logos.
- Riese, J., Kulgemeyer, C., Zander, S., Borowski, A., Fischer, H. E., Gramzow, Y., ... Tomczyszyn, E. (2015). Modellierung und Messung des Professionswissens in der Lehramtsausbildung Physik. *Zeitschrift für Pädagogik*, (61), 55–79.
- Riese, J. & Reinhold, P. (2008). Entwicklung und Validierung eines Instruments zur Messung professioneller Handlungskompetenz bei (angehenden) Lehrkräften. *Lehrerbildung auf dem Prüfstand*, 1(2), 625–640.

- Riese, J. & Reinhold, P. (2009). Fachbezogene Kompetenzmessung und Kompetenzentwicklung bei Lehramtsstudierenden der Physik im Vergleich verschiedener Studiengänge. *Lehrerbildung auf dem Prüfstand*, 2(1), 104–125.
- Rincke, K. (2011). Verweilen, um voran zukommen. *Praxis der Naturwissenschaften Physik in der Schule*, 4(60), 5–11.
- Roters, B., Nold, G., Haudeck, H., Keßler, J.-U. & Stancel-Piątak, A. (2011). Professionelles Wissen von Studierenden des Lehramts Englisch. In S. Blömeke, A. Bremerich-Vos, H. Haudeck, G. Kaiser, G. Nold, K. Schwippert & H. Willenberg (Hrsg.), *Kompetenzen von Lehramtsstudierenden in gering strukturierten Domänen, Erste Ergebnisse aus TEDS-LT* (S. 77–99). Münster: Waxmann.
- Sauer, S. (2016). *Die Erfassung pädagogischer Professionalität von Lehrenden - Konstruktion und Überprüfung eines Messinstruments. Dissertation. Universität Regensburg: http://uni-regensburg.de/34742/1/Dissertation_Sauer_Druckversion.pdf*.
- Schermelleh-Engel, K., Moosbrugger, H. & Müller, H. (2003). Evaluating the Fit of Structural Equation Models: Tests of Significance and Descriptive Goodness-of-Fit Measures. *Methods of Psychological Research Online*, 8(2), 23–74.
- Schilcher, A., Krauss, S., Rincke, K. & Hilbert, S. (2017). Ausblick - Aus FALKO wird FALKO. Fachspezifische Lehrerkompetenz im Erklären. In S. Krauss, A. Lindl, A. Schilcher, M. Fricke, A. Göhring, B. Hofmann, ... R. H. Mulder (Hrsg.), *FALKO: Fachspezifische Lehrerkompetenzen. Konzeption von Professionswissenstests in den Fächern Deutsch, Englisch, Latein, Physik, Musik, Evangelische Religion und Pädagogik* (S. 439–452). Münster: Waxmann.
- Schönbrodt, F. D. & Perugini, M. (2013). At what sample size do correlations stabilize? *Journal of Research in Personality*, (47), 609–612.
- Shavelson, R. J. & Webb, N. M. (1991). *Generalizability theory. A primer*. Newbury Park: Sage.
- Shulman, L. S. (1986). Those Who Understand: Knowledge Growth in Teaching. *Educational Researcher*, 15(4), 4–14.
- Shulman, L. S. (1987). Knowledge and Teaching: Foundations of the New Reform. *Harvard Educational Review*, 57(1), 1–22.
- Sorge, S., Kröger, J., Petersen, S. & Neumann, K. (manuscript submitted for publication). *Structure and Development of Pre-Service Physics Teachers' Professional Knowledge*.
- Stiller, J. & Laschke, C. (Hrsg.). (2015). *Berlin-Brandenburger Beiträge zur Bildungsforschung 2015. Herausforderungen, Befunde und Perspektiven interdisziplinärer Bildungsforschung*. Frankfurt am Main: Peter Lang.
- Stiller, J., Straube, P., Hartmann, S., Nordmeier, V. & Tiemann, R. (2015). Erkenntnisgewinnungskompetenz Chemie- und Physik-Lehramtsstudierender. Untersuchungen zu Domänenspezifität. In J. Stiller & C. Laschke (Hrsg.), *Berlin-Brandenburger Beiträge zur Bildungsforschung 2015. Herausforderungen, Befunde und Perspektiven interdisziplinärer Bildungsforschung* (S. 179–202). Frankfurt am Main: Peter Lang.
- Straube, P., Stiller, J., Tiemann, R. & Nordmeier, V. (2014). Ko-WADiS – Aspekte der Itemkonstruktion. In S. Bernholt (Hrsg.), *Naturwissenschaftliche Bildung zwischen Science- und Fachunterricht. Gesellschaft für Didaktik der Chemie und Physik, Jahrestagung in München 2013* (S. 162–164). Kiel: IPN.
- Straube, P., Tiemann, R., Upmeyer zu Belzen, A., Krüger, D. & Nordmeier, V. (2014). Ko-WADiS – Theoretische Grundlagen. In S. Bernholt (Hrsg.), *Naturwissenschaftliche Bildung zwischen Science- und Fachunterricht. Gesellschaft für Didaktik der Chemie und Physik, Jahrestagung in München 2013* (S. 159–161). Kiel: IPN.

- Strübe, M., Tepner, O. & Sumfleth, E. (2015). Lehrerprofessions- und Schülerwissen über Modelle und Experimente. In S. Bernholt (Hrsg.), *Heterogenität und Diversität - Vielfalt der Voraussetzungen im naturwissenschaftlichen Unterricht: Jahrestagung in Bremen 2014* (35, S. 612–614). Gesellschaft für Didaktik der Chemie und Physik. Kiel: IPN.
- Strübe, M., Tepner, O. & Sumfleth, E. (2016). Modelle und Experimente: Wissen und Handeln von Lehrkräften und Schülerinnen und Schülern. In C. Maurer (Hrsg.), *Authentizität und Lernen - das Fach in der Fachdidaktik. Gesellschaft für Didaktik der Chemie und Physik, Jahrestagung in Berlin* (S. 140–142).
- Tenorth, H.-E. (2006). Professionalität im Lehrerberuf. Ratlosigkeit der Theorie, gelingende Praxis. *Zeitschrift für Erziehungswissenschaft*, 9(4), 580–597.
- Tepner, O., Borowski, A., Dollny, S., Fischer, H. E., Jüttner, M., Kirschner, S., ... Wirth, J. (2012). Modelle 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). Measuring Chemistry Teachers' Content Knowledge: Is It Correlated to Pedagogical Content Knowledge? In C. Bruguière, A. Tiberghien & P. Clément (Hrsg.), *Topics and Trends in Current Science Education* (S. 243–254). Dordrecht u. a.: Springer.
- Terhart, E. (2002). *Standards für die Lehrerbildung: Eine Expertise für die Kultusministerkonferenz*. ZfL-Text. Münster: Inst. für Schulpädag. und Allgemeine Didaktik Univ. Münster.
- Terhart, R., Bennewitz, H. & Rothland, M. (Hrsg.). (2014). *Handbuch der Forschung zum Lehrberuf*. Münster: Waxmann.
- Tipler, P. A. (Hrsg.). (1994). *Physik*. Heidelberg: Spektrum Akad. Verl. Spektrum-Lehrbuch.
- Tomczyszyn, E., Kulgemeyer, C. & Schecker, H. (2014). Diagnostik des Erklärungswissens von Physikstudierenden. In S. Bernholt (Hrsg.), *Naturwissenschaftliche Bildung zwischen Science- und Fachunterricht. Gesellschaft für Didaktik der Chemie und Physik, Jahrestagung in München 2013* (S. 108–110). Kiel: IPN.
- Tröger, H., Sumfleth, E. & Tepner, O. (2016). Fachdidaktisches Wissen und Lehrerhandeln im Chemieunterricht. In C. Maurer (Hrsg.), *Authentizität und Lernen - das Fach in der Fachdidaktik. Gesellschaft für Didaktik der Chemie und Physik, Jahrestagung in Berlin* (S. 143–145).
- Professionswissen in den Naturwissenschaften – die ProwiN-Studie. Bericht Phase 1. (2013). Zugriff 28. Februar 2017 unter <https://www.uni-due.de/prowin/phase1-phy.shtml>
- Vogelsang, C. (2014). Validierung eines Instruments zur Erfassung der professionellen Handlungskompetenzen von (angehenden) Physiklehrkräften. Zusammenhangsanalysen zwischen Lehrerkompetenzen und Lehrerperformanz. In H. Niedderer, H. Fischler & E. Sumfleth (Hrsg.), *Studien zum Physik- und Chemielernen* (Bd. 174). Berlin: Logos.
- Walzer, M., Fischer, H. E. & Borowski, A. (2014). Fachwissen im Studium zum Lehramt der Physik. In S. Bernholt (Hrsg.), *Naturwissenschaftliche Bildung zwischen Science- und Fachunterricht. Gesellschaft für Didaktik der Chemie und Physik, Jahrestagung in München 2013* (S. 108–110). Kiel: IPN.
- Wang, J. & Xu, B. (Hrsg.). (2004). *Trends and challenges in mathematics education*. Shanghai: East China Normal University Press.

- Weinert, F. E. (1996). „Der gute Lehrer“, „die gute Lehrerin“ im Spiegel der Wissenschaft: Was macht Lehrende wirksam und was führt zu ihrer ‚Wirksamkeit‘? *Beiträge zur Lehrerbildung*, 14 (2), 141–151.
- Weinert, F. E. (Hrsg.). (1997). *Psychologie des Unterrichts und der Schule: Pädagogische Psychologie (Enzyklopädie der Psychologie)*. Göttingen: Hogrefe.
- Weinert, F. E. (Hrsg.). (2001a). *Leistungsmessung in Schulen*. Weinheim: Beltz.
- Weinert, F. E. (2001b). Vergleichende Leistungsmessung in Schulen - Eine umstrittene Selbstverständlichkeit. In F. E. Weinert (Hrsg.), *Leistungsmessung in Schulen* (S. 17–32). Weinheim: Beltz.
- Wilhelm, M., Vollmer, C., Tempel, B. J., Rehm, M., Bölsterli, K. & Brovelli, D. (2016). Empirical Validation of a Competency Model for Science Teaching. In J. Lavonen, K. Juuti, J. Lampiselkä, A. Uitto & K. Hahl (Hrsg.), *Electronic Proceedings of the ESERA 2015 Conference. Science education research: Engaging learners for a sustainable future*. (13, Bde. 13, S. 1923–1930). Helsinki.
- Wirtz, M. & Caspar, F. (2002). *Beurteilerübereinstimmung und Beurteilerreliabilität*. Göttingen: Hogrefe.
- Woitkowski, D. (2015). Fachliches Wissen in der Hochschulausbildung – Konzeptualisierung, Messung, Niveaubildung. In H. Niedderer, H. Fischler & E. Sumfleth (Hrsg.), *Studien zum Physik- und Chemielernen* (Bd. 185). Berlin: Logos.
- Woitkowski, D., Riese, J. & Reinhold, P. (2011). Modellierung fachwissenschaftlicher Kompetenz angehender Physiklehrkräfte. *Zeitschrift für Didaktik der Naturwissenschaften*, (17), 289–313.
- Zlatkin-Troitschanskaia, O., Beck, K., Sembill, D., Nickolaus, R. & Mulder, R. (Hrsg.). (2009). *Lehrerprofessionalität. Bedingungen, Genese, Wirkungen und ihre Messung*. Weinheim: Beltz.

Abbildungsverzeichnis

1.1. Projekte COACTIV und FALKO	2
1.2. Zeitlicher Ablauf FALKO-Physik	5
2.1. Modell professioneller Lehrerkompetenz	9
2.2. Modelle PCK	16
5.1. Zeitlicher Ablauf Pilotstudien 1 und 2	42
6.1. Zeitlicher Ablauf Validierungsphase	53
6.2. CFA komplett (alle drei Subfacetten)	55
6.3. CFA final (ohne Subfacette „M & E“)	56
6.4. Streudiagramme FW und FDW, Hauptstichprobe	61

Tabellenverzeichnis

2.1. Beispielprojekte Professionswissen	12
3.1. Statistische Kennwerte - Richtwerte	25
4.1. Itemverteilung finale Testversion	29
4.2. Finale Beispielitems	30
4.3. Exemplarische Musterlösungen zu finalen Beispielitems	32
5.1. Testheftversionen	36
5.2. Item Stromverbrauch	37
5.3. Kodieranweisung Item Stromverbrauch	39
5.4. Entwicklung Item Steckbrett	41
5.5. Kodieranweisung Item Steckbrett	43
5.6. Abfrage Augenscheinvalidität	45
5.7. Beispielitem Kochplatte	45
5.8. Beispielitem Jahreszeiten	47
5.9. Psychometrische Kennwerte Pilot-Studien 1 + 2	49
6.1. Hauptstichprobe Hauptstudie	54
6.2. Psychometrische Kennwerte Hauptstudie	57
6.3. Deskriptive Daten Hauptstudie	59
6.4. Gruppenunterschiede Hauptstichprobe Hauptstudie	60
6.5. Gruppenunterschiede L-Gy, L-nGy, S-Gy und S-nGy	61
6.6. Kennzahlen Regressionsanalysen	63
6.7. Produkt-Moment-Korrelationen FW und FDW mit Prädiktoren	63
6.8. Manifeste Korrelationen Skalen FDW+FW nach Fachinhalt	64
6.9. t-Test Leh. vs. fachfremde MS-Leh.	66
6.10. t-Test fachfr. Leh. RS/Gy vs. fachfr. Leh. MS	67
6.11. t-Test fachfr. Leh. RS/Gy vs. fachfr. Leh. MS	68
6.12. ANOVA Gy-Lehrkr. vs. nGy-Lehrkr. vs. Fachphysik	69
6.13. Manifeste Korrelationen Skalen FDW+FW Kreuzvalidierung	74
A.1. Itemübersicht, Modellzuordnung finales Testheft	111
A.2. Psychometrische Daten auf Einzelitemebene	113
A.3. Gruppenunterschiede Std. vs. Lehrkr. auf Einzelitemebene	115
A.4. Mittelwerte Augenscheinvaliditäten Prä-Pilotstudie 1	117
A.5. Mittelwerte Augenscheinvalidität Prä-Pilotstudie 2	120
A.6. Schwierigkeitsindices Pilotstudie 1	123
A.7. Schwierigkeitsindices Pilotstudie 2	124
A.8. Trennschärfen auf Einzelitemebene FDW	125
A.9. Trennschärfen auf Einzelitemebene FW	126
A.10. Mittelwerte Augenscheinvaliditäten Pilotstudie 1	127
A.11. Mittelwerte Augenscheinvaliditäten Pilotstudie 2	130

A.12.Schwierigkeitsindices Hauptstudie	133
A.13.Deskriptive Statistik Stichprobe Prä-Pilot 1	134
A.14.Deskriptive Statistik Stichprobe Prä-Pilot 2	135
A.15.Stichprobenzusammensetzung Pilot 1	136
A.16.Stichprobenzusammensetzung Pilot 2	137
A.17.Interraterreliabilitäten Einzelitemebene Pilot 1 + 2	138
A.18.Stichprobenzusammensetzung Hauptstudie	139
A.19.t-Test Zusammenlegung nGy-Lehrkräfte	141
A.20.t-Test Zusammenlegung nGy-Lehrkräfte, U-Fach	143
A.21.t-Test Zusammenlegung nGy-Lehrkräfte, D-Fach	145
A.22.Reliabilitäten (FDW + FW, FDW, FW) nach Fachinhalt	147
A.23.Augenscheinvaliditäten Hauptstudie	148
A.24.Manifeste Korrelationen FDW-Skala nach Fachinhalt	151
A.25.Manifeste Korrelationen FW-Skala nach Fachinhalt	151
A.26.Stichprobenzusammensetzung N65 (Lehrkräfte nGy, Gy)	152
A.27.Stichprobenzusammensetzung N35 (Lehrkräfte Gy) Hauptstudie	153
A.28.Stichprobenzusammensetzung N30 (Lehrkräfte nGy) Hauptstudie	154
A.29.Stichprobenzusammensetzung N25 (fachfr. Leh. MS) Hauptstudie	155
A.30.Stichprobenzusammensetzung N32 (fachfremde Lehrkräfte nGy, Gy)	156
A.31.Stichprobenzusammensetzung N29 (Fachphysik)	157
A.32.Stichprobenzusammensetzung N28 Kreuzvalidierung	158
A.33.Manifeste Korrelationen Skalen FDW+FW Kreuzvalidierung, nur Gy	160
A.34.Manifeste Korrelationen Skalen FDW+FW Kreuzvalidierung, nur nGy	160
A.35.Manifeste Korrelationen Skalen FDW+FW Kreuzvalidierung, Subskalen	161
A.36.Manif. Korr. Skalen FDW+FW Kreuzvalidierung, Subskalen, nur Gy	161
A.37.Manif. Korr. Skalen FDW+FW Kreuzvalidierung, Subskalen, nur nGy	162
A.38.Gruppenunterschiede Kreuzvalidierung	163
A.39.Output CFA komplett	176
A.40.Output CFA FDW/FW N183 (ohne „M & E“)	178
A.41.Output bayesianische CFA (N183, Hauptstichprobe)	180
A.42.Output CFA FDW/FW N118 einfakt. (ohne „M & E“)	181
A.43.Output CFA FDW/FW N118 zweifakt. (ohne „M & E“)	183
A.44.Output bayesianische CFA (N118, Gy+nGy Studierende)	185
A.45.Output bayesianische CFA (N86, Gy Studierende)	186
A.46.Output bayesianische CFA (N32, nGy Studierende)	187
A.47.Output bayesianische CFA (N65, Gy+nGy Lehrkräfte)	188
A.48.Output bayesianische CFA (N35, Gy Lehrkräfte)	189
A.49.Output bayesianische CFA (N30, nGy Lehrkräfte)	190
A.50.Output bayesianische CFA (N32, ff. RS-Gy-Lehrkräfte)	191
A.51.Output bayesianische CFA (N29, Fachphysiker)	192
A.52.Output bayesianische CFA (N25, ff. MS-Lehrkräfte)	193
A.53.Output bayesianische CFA (N28, Kreuzvalidierung, FDW)	194
A.54.Output bayesianische CFA (N13, Kreuzvalidierung, FDW)	195
A.55.Output bayesianische CFA (N15, Kreuzvalidierung, FDW)	196
A.56.Output bayesianische CFA (N28, Kreuzvalidierung, FW)	197
A.57.Output bayesianische CFA (N13, Kreuzvalidierung, FW)	198
A.58.Output bayesianische CFA (N15, Kreuzvalidierung, FW)	199

A. Tabellen

Tabelle A.1.: Itemübersicht und Modellzuordnung finales Testheft

Nummer	Bezeichnung	Seite	FDW-Subfacette bzw. fachliches Niveau	Itemkodierung
FDPHm_03_k	Flugbahn	8	Schk	0, 1, 2
FDPHm_05_i	Geschwindigkeit	9	E & R	0, 1
FDPHe_02a_k	Parallelsch. a	10	Schk	0, 1
FDPHe_02b_k	Parallelsch. b	10	E & R	0, 1
FDPHe_05_i	Turnhalle	11	E & R	0, 1, 2
FDPHe_12_me	Steckbrett	12	M & E	0, ..., 3
FDPHe_14_me	Zitr.batterie	13	M & E	0, 1, 2
FDPHe_01_i	Stromverbrauch	14	E & R	0, ..., 4
FDPHe_11_i	Spann.begriff	15	E & R	0, ..., 5
FDPHo_02_me	Spiegelbild	16	M & E	0, 1, 2
FDPHo_01_k	Sehvorgang	17	Schk	0, ..., 3
FDPHw_03_me	Temperatur	18	M & E	0, 1, 2
FDPHw_05a_k	Wärmeausbr.	19	Schk	0, 1
FDPHw_01_k	Wärmeleitung	19	Schk	0, ..., 3
FWPHm_11	E_{kin} Fall	21	Schulwissen	0, 1
FWPHm_12	Geschw. Fall	21	Schulwissen	0, 1
FWPHm_05	Kugel	21	Schulwissen	0, 1
FWPHm_02	senkr. Wurf	22	vert. Wissen	0, 1
FWPHe_21	NTC	22	vert. Wissen	0, ..., 2
FWPHe_04	Reihensch.	23	Schulwissen	0, 1
FWPHe_14	Trafo	23	Schulwissen	0, ..., 2
FWPHe_16a	Diode	23	vert. Wissen	0, 0,5, 1
FWPHe_20b	Induktion	24	Schulwissen	0, ..., 4
FWPHe_13b1	Wagenschein	25	vert. Wissen	0, 1
FWPHe_12	Dotierung	25	Schulwissen	0, ..., 2
FWPHe_06	Kochplatte	26	vert. Wissen	0, 1, 2
FWPHo_07	Spiegel Blume	26	Schulwissen	0, 1
FWPHo_06	Spiegel	26	Schulwissen	0, 1
FWPHo_03	Abbildung	27	vert. Wissen	0, 1, 2

Fortsetzung auf der nächsten Seite

Tabelle A.1.: (Fortsetzung von S. 111)

FWPHw_09	Wärmeübertr.	27	Schulwissen	0, ..., 3
FWPHw_07	Kochtopf	28	Schulwissen	0, ..., 2
FWPHw_05	Druck	28	vert. Wissen	0, 1

Tabelle A.2.

Fachwissen	Mögliches Scoring	Prozent Score				M	SD	Trennschärfe	Augenschein- validität	Interrater- reliabilität
		0	0,5	1	1,5					
Elkin Fall	0, 1	24,6		75,4		0,75	0,43	0,36	3,57	—
Geschw. Fall	0, 1	14,2		85,8		0,86	0,35	0,16	3,32	—
Kugel	0, 1	47		53		0,53	0,5	0,37	3,57	1
Senkrechter Wurf	0, 1	51,9		48,1		0,48	0,5	0,44	3,78	1
NTC	0, ..., 2	31,7	2,7	7,1	11,5	47		1,2	0,57	0,94
Reihenschaltung	0, 1	14,2		85,8		0,89	0,35	0,34	3,8	—
Trafo	0, ..., 2	44,8	2,7	7,7	22,4	22,4		0,48	3,5	0,86
Diode	0, 0,5, 1	36,6	44,3	19,1		0,41	0,36	0,34	3,5	0,97
Induktion	0, ..., 4	17,5		8,7	14,2	36,6	23	2,39	0,55	0,99
Wagenschein	0, 1	66,7		33,3		0,33	0,47	0,42	3	0,97
Dotierung	0, ..., 2	51,4	1,6	30,6	0	16,4		0,57	2,86	0,94
Kochplatte	0, 1, 2	69,9		15,8	14,2		0,44	0,56	3,3	0,96
Spiegel Blume	0, 1	39,9		60,1		0,6	0,49	0,16	3,52	—
Spiegel	0, 1	71		29		0,29	0,45	0,46	3,57	—
Abbildung	0, 1, 2	83,6		13,7	2,7		0,19	0,38	3,23	0,86
Wärmeübertragung	0, ..., 3	59,6		19,1	10,4	10,9		0,52	3,89	0,91
Kochtopf	0, ..., 2	49,2	16,4	17,5	9,8	7,1		0,49	3,74	0,85

Druck	0, 1	78,7	21,3	0,21	0,41	0,32	3,37	—
Fachdidaktisches Wissen	Subskala	Skala FDW						
		Subskala						
Geschwindigkeit	0, 1	69,4	30,6	0,31	0,46	0,19	3,39	0,92
Parallelschaltung b	0, 1	48,6	51,4	0,51	0,5	0,29	3,67	0,93
Turnhalle	0, 1, 2	36,1	35,5	28,4	0,92	0,8	3,64	0,87
Stromverbrauch	0, ..., 4	36,6	43,2	16,9	2,7	0,5	3,83	0,87
Spannungsbegriff	0, ..., 4	24,6	37,7	29	8,7	0	3,9	0,91
Flugbahn	0, 1, 2	30,1	44,3	25,7	0,96	0,75	3,46	0,83
Parallelschaltung a	0, 1	21,9	78,1	0,78	0,41	0,2	3,67	0,83
Sehvorgang	0, ..., 3	36,6	52,5	8,7	2,2	0,35	3,62	0,96
Wärmeausbreitung	0, 1	48,6	51,4	0,51	0,5	0,27	3,17	0,83
Wärmeleitung	0, ..., 3	54,6	36,1	8,2	1,1	0,27	3,44	0,95
Steckbrett	0, ..., 3	10,9	38,8	40,4	9,8	0,36	3,91	0,9
Zitronenbatterie	0, 1, 2	84,7	15,3	0	0,15	0,36	3,09	0,87
Spiegelbild	0, 1, 2	68,3	5,5	26,2	0,58	0,88	3,74	0,91
Temperatur	0, 1, 2	30,6	18,6	50,8	1,2	0,88	3,67	0,77

Tabelle A.3.: Prozentuale Verteilung der Antwortscores, arithmetische Mittelwerte (M), Standardabweichungen (SD) nach beruflichem Status getrennt sowie Gruppenunterschiede auf Itemebene. Die Effektstärke d berechnet sich aus der Mittelwertdifferenz beider Gruppen dividiert durch die gepoolte Standardabweichung. Nach Cohen (1992) entspricht $d = 0,20$ einem kleinen, $d = 0,50$ einem mittleren und $d = 0,80$ einem großen Effekt. Signifikanter Unterschied: * $p \leq 0,05$; ** $p \leq 0,01$.

Fachwissen	Mögliches Scoring	Lehrer: Prozent Score ($N = 65$)				Studierende: Prozent Score ($N = 118$)				Lehrer	Stud.	Gruppendiff.					
		0	1	2	3	4	0	1	2				3	4			
E _{kin} Fall	0, 1	15,4	84,6			29,7	70,3			0,85	0,36	0,7	0,46	0,34	0,02		
Geschw. Fall	0, 1	12,3	87,7			15,3	84,7			0,88	0,33	0,85	0,36	0,09	0,59		
Kugel	0, 1	20	80			61,9	38,1			0,8	0,4	0,38	0,49	0,91	<0,01		
Senkrechter Wurf	0, 1	40	60			58,5	41,5			0,6	0,49	0,42	0,49	0,37	0,02		
NTC	0, ..., 2	9,2	6,2	3,1	13,8	67,7	44,1	0,8	9,3	10,2	35,6	1,62	0,66	0,96	0,91	0,8	<0,01
Reihenschaltung	0, 1	10,8	89,2			16,1	83,9			0,89	0,31	0,84	0,37	0,15	0,3		
Trafo	0, ..., 2	21,5	4,6	30,8	61,5	57,6	1,7	9,3	17,8	13,6	1,3	0,78	0,64	0,8	0,83	<0,01	
Diode	0, 0,5, 1	26,2	43,1	30,8		42,2	44,9	12,7		0,52	0,38	0,35	0,34	0,48	<0,01		
Induktion	0, ..., 4	9,2	3,1	7,7	38,5	41,5	22	11,9	17,8	35,6	12,7	3	1,21	2,05	1,37	0,72	<0,01
Wagenschein	0, 1	53,8	46,2			73,7	26,3			0,46	0,5	0,26	0,44	0,43	<0,01		
Dotierung	0, ..., 2	33,8	1,5	35,4	0	29,2	61	1,7	28	0	9,3	0,95	0,8	0,48	0,66	0,66	<0,01
Kochplatte	0, 1, 2	47,7	23,1	29,2		82,2	11,9	5,9		0,82	0,86	0,24	0,55	0,85	<0,01		
Spiegel Blume	0, 1	41,5	58,5			39	61			0,59	0,5	0,61	0,49	-0,05	0,74		
Spiegel	0, 1	66,2	33,8			73,7	26,3			0,34	0,48	0,26	0,44	0,17	0,29		
Abbildung	0, 1, 2	75,4	18,5	6,2		88,1	11	0,8		0,31	0,58	0,13	0,36	0,4	0,03		
Wärmeübertr.	0, ..., 3	29,2	24,6	21,5		76,3	16,1	2,5	5,1	1,39	1,13	0,36	0,77	1,12	<0,01		
Kochtopf	0, ..., 2	36,9	4,6	26,2	15,4	16,9	55,9	22,9	12,7	6,8	1,7	0,85	0,76	0,38	0,51	0,78	<0,01
Druck	0, 1	61,5	38,5			88,1	11,9			0,39	0,49	0,12	0,32	0,75	<0,01		

Fachdidaktisches Subskala
Wissen

Geschwindigkeit	0, 1	64,6	35,4	72	28	0,35	0,48	0,28	0,45	0,16	0,3						
Parallelschalt. b	0, 1	41,5	58,5	52,5	47,5	0,59	0,5	0,48	0,5	0,22	0,16						
Turnhalle	0, 1, 2	29,2	32,3	38,5	37,3	1,09	0,82	0,83	0,78	0,33	0,03						
Stromverbrauch	0, ..., 4	21,5	50,8	23,1	3,1	1,5	44,9	39	13,6	2,5	0	1,12	0,84	0,74	0,79	0,48	<0,01
Spannungsbegriff	0, ..., 4	12,3	32,3	41,5	13,8	0	31,4	40,7	22	5,9	0	1,57	0,88	1,03	0,88	0,62	<0,01
Flugbahn	0, 1, 2	26,2	46,2	27,7	32,2	43,2	24,6	1,02	0,74	0,92	0,75	0,12	0,43				
Parallelsch. a	0, 1	26,2	73,8	19,5	80,5	0,74	0,44	0,81	0,4	-0,16	0,32						
Selbvorgang	0, ..., 3	27,7	60	10,8	1,5	41,5	48,3	7,6	2,5	0,86	0,67	0,71	0,72	0,22	0,17		
Wärmeausbr.	0, 1	43,1	56,9	51,7	48,3	0,57	0,5	0,48	0,5	0,17	0,27						
Wärmeleitung	0, ..., 3	60	27,7	10,8	1,5	51,7	40,7	6,8	0,8	0,54	0,75	0,57	0,66	-0,04	0,79		
Steckbrett	0, ..., 3	4,6	35,4	41,5	18,5	14,4	40,7	39,8	5,1	1,74	0,82	1,36	0,79	0,48	<0,01		
Zitronenbatterie	0, 1, 2	69,2	30,8	0	93,2	6,8	0	0,31	0,47	0,07	0,25	0,7	<0,01				
Spiegelbild	0, 1, 2	44,6	12,3	43,1	81,4	1,7	16,9	0,99	0,94	0,36	0,76	0,76	<0,01				
Temperatur	0, 1, 2	26,2	23,1	50,8	33,1	16,1	50,8	1,25	0,85	1,18	0,9	0,08	0,62				

Tabelle A.4.: Mittelwerte der Augenscheinvaliditäten der Experten und erfahrenen Lehrkräfte aus der Prä-Pilotstudie 1

Deskriptive Statistik							
	<i>N</i>	range	Min	Max	MW	St.abw.	Varianz
FDPHe_01_i_ein	11	3	1	4	2,909	0,9439	0,891
FDPHe_01_i_unt	11	1	3	4	3,818	0,4045	0,164
FDPHe_01_i_lr	11	1	3	4	3,545	0,5222	0,273
FDPHe_02a_k_ein	11	2	2	4	3,364	0,809	0,655
FDPHe_02a_k_unt	11	1	3	4	3,909	0,3015	0,091
FDPHe_02a_k_lr	11	1	3	4	3,727	0,4671	0,218
FDPHe_02b_i_ein	11	1	3	4	3,909	0,3015	0,091
FDPHe_02b_i_unt	11	1	3	4	3,909	0,3015	0,091
FDPHe_02b_i_lr	11	2	2	4	3,727	0,6467	0,418
FDPHe_03_k_ein	1	0	3	3	3		
FDPHe_03_k_unt	1	0	4	4	4		
FDPHe_03_k_lr	1	0	4	4	4		
FDPHe_05_i_ein	1	0	4	4	4		
FDPHe_05_i_unt	1	0	3	3	3		
FDPHe_05_i_lr	1	0	1	1	1		
FDPHe_06_me_ein	1	0	3	3	3		
FDPHe_06_me_unt	1	0	4	4	4		
FDPHe_06_me_lr	1	0	3	3	3		
FDPHe_07_i_ein	10	1	3	4	3,7	0,483	0,233
FDPHe_07_i_unt	10	1	3	4	3,6	0,5164	0,267
FDPHe_07_i_lr	10	2	2	4	3,4	0,6992	0,489
FDPHe_08_k_ein	1	0	4	4	4		
FDPHe_08_k_unt	1	0	4	4	4		
FDPHe_08_k_lr	1	0	4	4	4		
FDPHe_09_i_ein	10	2	2	4	3,65	0,6687	0,447
FDPHe_09_i_unt	10	2	2	4	3,45	0,8317	0,692
FDPHe_09_i_lr	10	2	2	4	3,6	0,6992	0,489
FDPHe_10_me_ein	1	0	3	3	3		
FDPHe_10_me_unt	1	0	4	4	4		
FDPHe_10_me_lr	1	0	2	2	2		
FDPHe_11_i_ein	9	1	3	4	3,667	0,5	0,25
FDPHe_11_i_unt	9	2	2	4	3,778	0,6667	0,444
FDPHe_11_i_lr	9	2	2	4	3,667	0,7071	0,5
FDPHe_12_me_ein	1	0	4	4	4		
FDPHe_12_me_unt	1	0	4	4	4		
FDPHe_12_me_lr	1	0	4	4	4		
FDPHe_13_k_ein	1	0	4	4	4		

Fortsetzung auf der nächsten Seite

Tabelle A.4.: (Fortsetzung)

FDPHe_13_k_unt	1	0	4	4	4		
FDPHe_13_k_lr	1	0	3	3	3		
FDPHe_14_me_ein	1	0	4	4	4		
FDPHe_14_me_unt	1	0	3	3	3		
FDPHe_14_me_lr	1	0	1	1	1		
FWPHe_02_ein	2	3	1	4	2,5	2,1213	4,5
FWPHe_02_unt	2	0	4	4	4	0	0
FWPHe_02_lr	2	0	4	4	4	0	0
FWPHe_03_ein	8	2	2	4	3,625	0,744	0,554
FWPHe_03_unt	8	3	1	4	3,125	1,3562	1,839
FWPHe_03_lr	8	3	1	4	3,125	1,3562	1,839
FWPHe_04_ein	9	2	2	4	3,667	0,7071	0,5
FWPHe_04_unt	9	0	4	4	4	0	0
FWPHe_04_lr	9	1	3	4	3,889	0,3333	0,111
FWPHe_05_ein	9	2	2	4	3,389	0,7817	0,611
FWPHe_05_unt	9	1	3	4	3,556	0,527	0,278
FWPHe_05_lr	9	3	1	4	3,167	1	1
FWPHe_06_ein	1	0	3	3	3		
FWPHe_06_unt	1	0	3	3	3		
FWPHe_06_lr	1	0	3	3	3		
FWPHe_07_ein	9	1	3	4	3,778	0,441	0,194
FWPHe_07_unt	9	3	1	4	2,889	1,1667	1,361
FWPHe_07_lr	9	3	1	4	2,444	1,236	1,528
FWPHe_09b_ein	7	2	2	4	3,571	0,7868	0,619
FWPHe_09b_unt	7	1	3	4	3,571	0,5345	0,286
FWPHe_09b_lr	7	1	3	4	3,286	0,488	0,238
FWPHe_10_ein	0						
FWPHe_10_unt	1	0	4	4	4		
FWPHe_10_lr	1	0	4	4	4		
FWPHe_11_ein	8	1	3	4	3,75	0,4629	0,214
FWPHe_11_unt	9	2	2	4	3,667	0,7071	0,5
FWPHe_11_lr	9	1	3	4	3,778	0,441	0,194
FWPHe_12_ein	9	3	1	4	3,556	1,0138	1,028
FWPHe_12_unt	9	3	1	4	3,444	1,0138	1,028
FWPHe_12_lr	9	3	1	4	3,333	1	1
FWPHe_13a_ein	9	3	1	4	3,111	1,1667	1,361
FWPHe_13a_unt	9	3	1	4	2,667	1,118	1,25
FWPHe_13a_lr	9	3	1	4	2,278	0,8333	0,694
FWPHe_13b_ein	8	3	1	4	3,625	1,0607	1,125
FWPHe_13b_unt	8	3	1	4	3	1,1952	1,429
FWPHe_13b_lr	8	2	2	4	3,125	0,991	0,982

Fortsetzung auf der nächsten Seite

Tabelle A.4.: (Fortsetzung)

FWPHe_14_ein	9	1	3	4	3,889	0,3333	0,111
FWPHe_14_unt	9	1	3	4	3,778	0,441	0,194
FWPHe_14_lr	9	1	3	4	3,778	0,441	0,194
FWPHe_16a_ein	8	2	2	4	3,375	0,9161	0,839
FWPHe_16a_unt	8	2	2	4	3,5	0,9258	0,857
FWPHe_16a_lr	8	2	2	4	3,375	0,744	0,554
FWPHe_17_ein	8	3	1	4	3,375	1,0607	1,125
FWPHe_17_unt	8	3	1	4	2,75	1,165	1,357
FWPHe_17_lr	8	3	1	4	2,75	1,0351	1,071
FWPHe_18_ein	1	0	4	4	4		
FWPHe_18_unt	1	0	4	4	4		
FWPHe_18_lr	1	0	3	3	3		
FWPHe_19_ein	9	2	2	4	3,111	0,928	0,861
FWPHe_19_unt	9	2	2	4	3,444	0,7265	0,528
FWPHe_19_lr	9	3	1	4	3,333	1	1
FWPHe_20b_ein	9	3	1	4	3,333	1	1
FWPHe_20b_unt	9	1	3	4	3,778	0,441	0,194
FWPHe_20b_lr	9	1	3	4	3,556	0,527	0,278
FWPHe_21_ein	1	0	1	1	1		
FWPHe_21_unt	1	0	4	4	4		
FWPHe_21_lr	1	0	4	4	4		
Gültige Werte (Listenweise)	13						

Tabelle A.5.: Mittelwerte der Augenscheinvaliditäten der Experten und erfahrenen Lehrkräfte aus der Prä-Pilotstudie 2

Deskriptive Statistik							
	<i>N</i>	range	Min	Max	MW	St.abw.	Varianz
FDPHm_02_k_ein	13	3	1	4	3,462	0,9674	0,936
FDPHm_02_k_ber	11	3	1	4	3,182	1,0787	1,164
FDPHm_02_k_lr	12	2	2	4	3,417	0,6686	0,447
FDPHm_03_k_ein	13	3	1	4	3,538	0,8771	0,769
FDPHm_03_k_ber	11	2	2	4	3,455	0,8202	0,673
FDPHm_03_k_lr	12	1	3	4	3,583	0,5149	0,265
FDPHm_05_i_ein	13	2	2	4	3,692	0,6304	0,397
FDPHm_05_i_ber	11	3	1	4	3,455	0,9342	0,873
FDPHm_05_i_lr	12	1	3	4	3,833	0,3892	0,152
FDPHm_06_i_ein	13	1,5	2,5	4	3,5	0,5774	0,333
FDPHm_06_i_ber	11	3	1	4	3,545	0,9342	0,873
FDPHm_06_i_lr	12	2	2	4	3,583	0,6686	0,447
FDPHo_01_k_ein	11	2	2	4	3,409	0,7355	0,541
FDPHo_01_k_ber	10	1	3	4	3,4	0,5164	0,267
FDPHo_01_k_lr	11	2	2	4	3,364	0,809	0,655
FDPHo_02_me_ein	13	2	2	4	3,654	0,6253	0,391
FDPHo_02_me_ber	11	1	3	4	3,727	0,4671	0,218
FDPHo_02_me_lr	12	2	2	4	3,667	0,6513	0,424
FDPHo_03_me_ein	11	1	3	4	3,591	0,4908	0,241
FDPHo_03_me_ber	9	1	3	4	3,722	0,441	0,194
FDPHo_03_me_lr	10	2	2	4	3,55	0,6852	0,469
FDPHo_05_i_ein	13	3	1	4	2,654	1,2142	1,474
FDPHo_05_i_ber	10	3	1	4	2,95	1,0124	1,025
FDPHo_05_i_lr	11	3	1	4	3	1,0247	1,05
FDPHw_01_k_ein	13	2	2	4	3,385	0,7679	0,59
FDPHw_01_k_ber	11	2	2	4	3,545	0,6876	0,473
FDPHw_01_k_lr	12	2	2	4	3,458	0,7217	0,521
FDPHw_02_k_ein	13	3	1	4	3,269	1,0919	1,192
FDPHw_02_k_ber	11	3	1	4	3,318	1,0068	1,014
FDPHw_02_k_lr	12	3	1	4	3,25	0,9886	0,977
FDPHw_03_me_ein	13	2	2	4	3,692	0,6304	0,397
FDPHw_03_me_ber	11	3	1	4	3,5	0,922	0,85
FDPHw_03_me_lr	11	3	1	4	3,5	0,922	0,85
FDPHw_05_k_ein	13	1,5	2,5	4	3,423	0,5718	0,327
FDPHw_05_k_ber	11	3	1	4	3,273	1,009	1,018
FDPHw_05_k_lr	12	3	1	4	3,25	0,9653	0,932
FDPHw_06_i_ein	10	3	1	4	3,25	1,0341	1,069

Fortsetzung auf der nächsten Seite

Tabelle A.5.: (Fortsetzung)

FDPHw_06_i_ber	9	1	3	4	3,722	0,441	0,194
FDPHw_06_i_lr	9	1	3	4	3,722	0,441	0,194
FWPHm_01_ein	12	3	1	4	2,917	1,1645	1,356
FWPHm_01_ber	10	1	3	4	3,35	0,4743	0,225
FWPHm_01_lr	11	1	3	4	3,5	0,5	0,25
FWPHm_02_ein	13	2	2	4	3,538	0,6602	0,436
FWPHm_02_ber	11	1	3	4	3,909	0,3015	0,091
FWPHm_02_lr	12	1	3	4	3,917	0,2887	0,083
FWPHm_04_ein	13	2	2	4	3,654	0,6253	0,391
FWPHm_04_ber	10	3	1	4	3,55	0,956	0,914
FWPHm_04_lr	11	1	3	4	3,773	0,4101	0,168
FWPHm_05_ein	13	1	3	4	3,808	0,384	0,147
FWPHm_05_ber	11	3	1	4	3,5	0,922	0,85
FWPHm_05_lr	12	3	1	4	3,542	0,8908	0,794
FWPHw_08_ein	13	2	2	4	3,423	0,8623	0,744
FWPHw_08_ber	11	2	2	4	3,682	0,6431	0,414
FWPHw_08_lr	12	3	1	4	3,625	0,8823	0,778
FWPHm_09_ein	13	2	2	4	3,692	0,6304	0,397
FWPHm_09_ber	11	1	3	4	3,727	0,4671	0,218
FWPHm_09_lr	12	2	2	4	3,667	0,6513	0,424
FWPHm_11_ein	13	1,5	2,5	4	3,577	0,5718	0,327
FWPHm_11_ber	11	0,5	3,5	4	3,955	0,1508	0,023
FWPHm_11_lr	12	2	2	4	3,792	0,5823	0,339
FWPHm_12_ein	13	2	2	4	3,769	0,5991	0,359
FWPHm_12_ber	11	2	2	4	3,273	0,9045	0,818
FWPHm_12_lr	12	2	2	4	3,5	0,9045	0,818
FWPHo_02_ein	13	1	3	4	3,769	0,4385	0,192
FWPHo_02_ber	11	1	3	4	3,909	0,3015	0,091
FWPHo_02_lr	12	1	3	4	3,917	0,2887	0,083
FWPHo_03_ein	13	2	2	4	3,577	0,7596	0,577
FWPHo_03_ber	11	2	2	4	3,409	0,8006	0,641
FWPHo_03_lr	12	2	2	4	3,458	0,7821	0,612
FWPHo_04_ein	13	1	3	4	3,846	0,3755	0,141
FWPHo_04_ber	11	2	2	4	3,636	0,809	0,655
FWPHo_04_lr	12	2	2	4	3,583	0,793	0,629
FWPHo_05_ein	12	3	1	4	3,5	0,9045	0,818
FWPHo_05_ber	10	2	2	4	3,4	0,8433	0,711
FWPHo_05_lr	11	3	1	4	3,273	1,2721	1,618
FWPHo_06_ein	13	3	1	4	3,538	0,8771	0,769
FWPHo_06_ber	10	2	2	4	3,6	0,6992	0,489
FWPHo_06_lr	11	2	2	4	3,364	0,9244	0,855

Fortsetzung auf der nächsten Seite

Tabelle A.5.: (Fortsetzung)

FWPHo_07_ein	13	2	2	4	3,731	0,5991	0,359
FWPHo_07_ber	11	1	3	4	3,682	0,4622	0,214
FWPHo_07_lr	12	2	2	4	3,375	0,8823	0,778
FWPHw_01_ein	13	2	2	4	3,654	0,6253	0,391
FWPHw_01_ber	11	1	3	4	3,864	0,3233	0,105
FWPHw_01_lr	12	2	2	4	3,708	0,6201	0,384
FWPHw_04_ein	13	1,5	2,5	4	3,731	0,525	0,276
FWPHw_04_ber	11	3	1	4	3,409	0,97	0,941
FWPHw_04_lr	12	1,5	2,5	4	3,625	0,5691	0,324
FWPHw_05_ein	13	1,5	2,5	4	3,885	0,416	0,173
FWPHw_05_ber	11	2	2	4	3,227	0,8174	0,668
FWPHw_05_lr	12	2	2	4	3,375	0,8292	0,688
FWPHw_07_ein	13	1	3	4	3,769	0,4385	0,192
FWPHw_07_ber	11	1	3	4	3,636	0,5045	0,255
FWPHw_07_lr	12	2	2	4	3,667	0,6513	0,424
FWPHw_09_ein	12	2	2	4	3,75	0,6216	0,386
FWPHw_09_ber	10	0	4	4	4	0	0
FWPHw_09_lr	11	0	4	4	4	0	0
Gültige Werte (Listenweise)	13						

Tabelle A.6.: Schwierigkeitsindices der Items aus der Pilotstudie 1; Hinweis: Insgesamt lagen 56 Datensätze vor. Allerdings konnten aufgrund einiger früherer Abbrüche nur 44 vollständige Datensätze ausgewertet werden!

Item	FDPHe_01_i	FDPHe_02a_k	FDPHe_02b_i	FDPHe_03_k	FDPHe_05_i
m_item	1	0,6	0,4	2	1
SD_item	1,05	0,5	0,49	0,89	0,89
Max_item	4	4	1	3	2
Index [%]	25,57	14,2	38,64	68,18	52,27
Item	FDPHe_06_me	FDPHe_07_i	FDPHe_08_k	FDPHe_09_i	FDPHe_10_me
m_item	1,3	0,3	1,2	0,5	0,3
SD_item	0,74	0,47	0,81	0,44	0,48
Max_item	2	1	2	3	1
Index [%]	63,64	31,82	59,09	18,18	34,09
Item	FDPHe_11_i	FDPHe_12_me	FDPHe_13_k	FDPHe_14_me	FWPHe_02
m_item	1	1,3	0,7	1,6	0,7
SD_item	0,86	1,03	0,79	0,95	0,88
Max_item	5	3	4	4	2
Index [%]	20	44,7	17,61	40,34	35,23
Item	FWPHe_03	FWPHe_04	FWPHe_05	FWPHe_06	FWPHe_07
m_item	0,3	0,6	2,1	0,4	0,5
SD_item	0,47	0,5	1,37	0,72	0,5
Max_item	1	1	4	2	1
Index [%]	31,82	56,82	52,84	19,32	54,55
Item	FWPHe_09b	FWPHe_10	FWPHe_11	FWPHe_12	FWPHe_13a
m_item	0,7	0,7	0,5	0,5	0,2
SD_item	0,88	0,94	0,53	0,78	0,58
Max_item	3	2	1	2	2
Index [%]	22,73	34,66	52,27	26,14	9,09
Item	FWPHe_13b	FWPHe_14	FWPHe_16a	FWPHe_17	FWPHe_18
m_item	0,7	1,2	0,6	0,2	1,9
SD_item	0,8	0,84	0,49	0,51	0,3
Max_item	2	2	1	2	2
Index [%]	35,8	57,95	57,95	10,23	94,32
Item	FWPHe_19	FWPHe_20b	FWPHe_21		
m_item	0,4	2	1,5		
SD_item	0,47	1,54	0,71		
Max_item	1	4	2		
Index [%]	36,36	50	74,43		

Tabelle A.7.: Schwierigkeitsindices der Items aus der Pilotstudie 2; Hinweis: Insgesamt lagen 66 Datensätze vor. Allerdings konnten aufgrund einiger früherer Abbrüche nur 35 (FDW) bzw. 46 (FW) vollständige Datensätze ausgewertet werden!

Item	FDPHm_02_k	FDPHm_03_k	FDPHm_05_i	FDPHm_06_i	FDPHo_01_k
m_item	0,23	1,34	0,34	0,2	1,03
SD_item	0,43	0,76	0,48	0,41	0,89
Max_item	1	2	1	2	3
Index [%]	22,86	67,14	34,29	10	34,29
Item	FDPHo_02_me	FDPHo_03_me	FDPHo_05_i	FDPHw_01_k	FDPHw_02_k
m_item	1,06	1,37	0,34	0,89	0,94
SD_item	0,91	0,84	0,54	0,87	0,73
Max_item	2	2	2	3	2
Index [%]	52,86	68,57	17,14	29,52	47,14
Item	FDPHw_03_me	FDPHw_05_k	FDPHw_06_i	FWPHm_01	FWPHm_02
m_item	1,17	0,49	0,31	0,3	0,33
SD_item	0,92	0,51	0,58	0,47	0,47
Max_item	2	1	2	1	1
Index [%]	58,57	48,57	15,71	30,43	32,61
Item	FWPHm_04	FWPHm_05	FWPHm_08	FWPHm_09	FWPHm_11
m_item	1,39	0,52	0,8	0,76	0,78
SD_item	0,68	0,51	0,45	0,43	0,42
Max_item	2	1	1	1	1
Index [%]	69,57	52,17	80,43	76,09	78,26
Item	FWPHm_12	FWPHo_02	FWPHo_03	FWPHo_04	FWPHo_05
m_item	0,67	0,65	0,48	0,76	0,63
SD_item	0,47	0,48	0,72	0,43	0,49
Max_item	1	1	2	1	1
Index [%]	67,39	65,22	23,91	76,09	63,04
Item	FWPHo_06	FWPHo_07	FWPHw_01	FWPHw_04	FWPHw_05
m_item	0,46	0,5	1,24	0,33	0,41
SD_item	0,5	0,51	0,57	0,47	0,5
Max_item	1	1	2	1	1
Index [%]	45,65	50	61,96	32,61	41,3
Item	FWPHw_07	FWPHw_09			
m_item	0,96	1,35			
SD_item	0,87	1,18			
Max_item	2	3			
Index [%]	47,83	44,93			

Tabelle A.8.: Trennschärpen auf Einzelitemebene in Bezug auf das FDW (es wurden 36 vollständige Datensätze für die Berechnungen herangezogen)

Item-Skala-Statistiken				
	Skalenmw., ohne Item	Skalenvar., ohne Item	Korr. Item- Skala-Korr.	Cronbachs α , ohne Item
FDPHm_03_k_score	20,319	73,959	0,357	0,844
FDPHm_05_i_score	21,347	77,026	0,237	0,847
FDPHm_02_k_score	21,431	78,131	0,117	0,849
FDPHm_06_i_score	21,431	77,074	0,267	0,847
FDPHe_02a_k_score	21,097	74,769	0,468	0,842
FDPHe_02b_i_score	21,153	73,04	0,673	0,837
FDPHe_05_i_score	20,431	69,474	0,625	0,834
FDPHe_12_me_score	20,431	69,702	0,517	0,838
FDPHe_14_me_score	21,347	76,426	0,271	0,846
FDPHe_01_i_score	20,569	67,674	0,538	0,838
FDPHe_11_i_score	20,847	71,497	0,493	0,839
FDPHe_03_k_score	19,736	72,564	0,365	0,844
FDPHe_06_me_score	20,333	70,271	0,7	0,833
FDPHe_07_i_score	21,236	73,607	0,621	0,839
FDPHe_08_k_score	20,347	76,455	0,131	0,853
FDPHe_09_i_score	21,125	73,663	0,616	0,839
FDPHe_10_me_score	21,375	77,405	0,197	0,848
FDPHe_13_k_score	20,819	71,231	0,419	0,843
FDPHo_02_me_score	20,569	72,645	0,357	0,845
FDPHo_01_k_score	20,625	73,92	0,317	0,846
FDPHo_03_me_score	20,347	73,283	0,335	0,845
FDPHo_05_i_score	21,236	75,007	0,401	0,843
FDPHw_03_me_score	20,347	73,026	0,37	0,844
FDPHw_05a_k_score	21,125	75,591	0,371	0,844
FDPHw_01_k_score	20,792	74,191	0,289	0,847
FDPHw_02_k_score	20,597	75,526	0,238	0,848
FDPHw_06_i_score	21,236	75,321	0,331	0,845

Tabelle A.9.: Trennschärpen auf Einzelitemebene in Bezug auf das FW (es wurden 36 vollständige Datensätze für die Berechnungen herangezogen)

Item-Skala-Statistiken				
	Skalenmw., ohne Item	Skalenvar., ohne Item	Korr. Item- Skala-Korr.	Cronbachs α , ohne Item
FWPHm_11_score	28,167	150,757	0,543	0,907
FWPHm_12_score	28,333	151,271	0,397	0,908
FWPHm_05_score	28,472	146,728	0,756	0,904
FWPHm_02_score	28,694	149,804	0,563	0,907
FWPHm_01_score	28,639	152,652	0,287	0,909
FWPHm_04_score	27,306	156,061	-0,006	0,912
FWPHm_08_score	28,111	154,116	0,233	0,91
FWPHm_09_score	28,111	154,802	0,154	0,91
FWPHe_21_score	27,653	142,926	0,646	0,904
FWPHe_04_score	28,306	148,233	0,671	0,906
FWPHe_14_score	27,708	142,92	0,684	0,904
FWPHe_16a_score	28,528	150,042	0,632	0,906
FWPHe_20b_score	26,806	129,861	0,695	0,905
FWPHe_13b1_score	28,806	151,261	0,523	0,907
FWPHe_06_score	28,306	141,947	0,635	0,904
FWPHe_02_score	28,111	142,344	0,611	0,905
FWPHe_03_score	28,639	149,423	0,566	0,907
FWPHe_05_score	26,708	146,963	0,201	0,918
FWPHe_07_score	28,389	149,916	0,499	0,907
FWPHe_09b_score	28,167	141,814	0,559	0,906
FWPHe_10_score	28,264	139,05	0,721	0,903
FWPHe_11_score	28,417	145,421	0,824	0,904
FWPHe_13a_score	28,694	147,09	0,511	0,906
FWPHe_17_score	28,694	146,575	0,58	0,906
FWPHe_18_score	27,167	151,957	0,332	0,909
FWPHe_19_score	28,569	150,816	0,433	0,908
FWPHo_07_score	28,417	151,521	0,362	0,908
FWPHo_06_score	28,417	152,036	0,32	0,909
FWPHo_03_score	28,722	152,849	0,205	0,91
FWPHo_02_score	28,222	150,563	0,512	0,907
FWPHo_04_score	28,194	152,518	0,343	0,909
FWPHo_05_score	28,472	151,985	0,322	0,909
FWPHw_09_score	27,722	134,849	0,698	0,903
FWPHw_07_score	28,153	143,012	0,741	0,903
FWPHw_05_score	28,5	151,071	0,397	0,908
FWPHw_01_score	27,75	156,893	-0,066	0,913
FWPHw_04_score	28,667	155,386	0,057	0,911

Tabelle A.10.: Mittelwerte der Augenscheinvaliditäten aus der Pilotstudie 1

Deskriptive Statistik							
	<i>N</i>	range	Min	Max	MW	St.abw.	Varianz
FDPHe_01_i_ein	58	3	1	4	3,466	0,7063	0,499
FDPHe_01_i_unt	58	1	3	4	3,759	0,4317	0,186
FDPHe_01_i_lr	58	2	2	4	3,724	0,4882	0,238
FDPHe_02a_k_ein	60	3	1	4	3,45	0,8115	0,658
FDPHe_02a_k_unt	60	3	1	4	3,617	0,6662	0,444
FDPHe_02a_k_lr	60	2	2	4	3,65	0,5771	0,333
FDPHe_02b_i_ein	58	3	1	4	3,448	0,8413	0,708
FDPHe_02b_i_unt	58	3	1	4	3,569	0,7282	0,53
FDPHe_02b_i_lr	58	3	1	4	3,621	0,6709	0,45
FDPHe_03_k_ein	49	2	2	4	3,694	0,5479	0,3
FDPHe_03_k_unt	49	3	1	4	3,714	0,6124	0,375
FDPHe_03_k_lr	49	2	2	4	3,673	0,5548	0,308
FDPHe_05_i_ein	44	3	1	4	3,5	0,7924	0,628
FDPHe_05_i_unt	45	3	1	4	3,533	0,6606	0,436
FDPHe_05_i_lr	45	3	1	4	3,556	0,659	0,434
FDPHe_06_me_ein	45	1	3	4	3,711	0,4584	0,21
FDPHe_06_me_unt	45	3	1	4	3,4	0,802	0,643
FDPHe_06_me_lr	45	3	1	4	3,422	0,783	0,613
FDPHe_07_i_ein	55	1	3	4	3,618	0,4903	0,24
FDPHe_07_i_unt	56	2	2	4	3,554	0,6006	0,361
FDPHe_07_i_lr	56	2	2	4	3,482	0,6603	0,436
FDPHe_08_k_ein	46	3	1	4	3,5	0,7528	0,567
FDPHe_08_k_unt	47	3	1	4	3,404	0,7419	0,55
FDPHe_08_k_lr	47	3	1	4	3,426	0,773	0,598
FDPHe_09_i_ein	53	2	2	4	3,594	0,6283	0,395
FDPHe_09_i_unt	53	3	1	4	3,057	0,8915	0,795
FDPHe_09_i_lr	53	3	1	4	3,189	0,8783	0,771
FDPHe_10_me_ein	44	2	2	4	3,636	0,5743	0,33
FDPHe_10_me_unt	44	3	1	4	3,227	0,803	0,645
FDPHe_10_me_lr	44	3	1	4	3,273	0,7884	0,622
FDPHe_11_i_ein	48	1	3	4	3,708	0,4593	0,211
FDPHe_11_i_unt	48	2	2	4	3,792	0,5035	0,254
FDPHe_11_i_lr	48	2	2	4	3,813	0,4451	0,198
FDPHe_12_me_ein	41	2	2	4	3,756	0,4889	0,239
FDPHe_12_me_unt	41	1	3	4	3,756	0,4348	0,189
FDPHe_12_me_lr	41	1,5	2,5	4	3,695	0,4858	0,236
FDPHe_13_k_ein	26	2	2	4	3,654	0,6288	0,395
FDPHe_13_k_unt	26	2	2	4	3,077	0,7961	0,634

Fortsetzung auf der nächsten Seite

Tabelle A.10.: (Fortsetzung)

FDPHe_13_k_lr	26	2	2	4	3,231	0,7646	0,585
FDPHe_14_me_ein	30	2	2	4	3,7	0,535	0,286
FDPHe_14_me_unt	30	2	2	4	3,15	0,8003	0,641
FDPHe_14_me_lr	30	3	1	4	3,2	0,8469	0,717
FWPHe_02_ein	45	3	1	4	3,733	0,58	0,336
FWPHe_02_unt	45	3	1	4	3,044	0,9034	0,816
FWPHe_02_lr	45	3	1	4	3,244	0,8021	0,643
FWPHe_03_ein	51	2	2	4	3,627	0,6312	0,398
FWPHe_03_unt	51	3	1	4	3,235	0,9075	0,824
FWPHe_03_lr	51	3	1	4	3,314	0,8122	0,66
FWPHe_04_ein	52	2	2	4	3,769	0,5093	0,259
FWPHe_04_unt	52	3	1	4	3,673	0,6174	0,381
FWPHe_04_lr	52	2	2	4	3,615	0,5655	0,32
FWPHe_05_ein	53	2	2	4	3,632	0,6366	0,405
FWPHe_05_unt	53	3	1	4	3,443	0,7316	0,535
FWPHe_05_lr	53	3	1	4	3,377	0,808	0,653
FWPHe_06_ein	42	3	1	4	3,548	0,7055	0,498
FWPHe_06_unt	42	3	1	4	3,155	0,769	0,591
FWPHe_06_lr	42	2	2	4	3,262	0,6648	0,442
FWPHe_07_ein	53	2	2	4	3,755	0,4766	0,227
FWPHe_07_unt	53	3	1	4	3,151	0,9281	0,861
FWPHe_07_lr	53	3	1	4	3,057	0,9887	0,978
FWPHe_09b_ein	48	3	1	4	3,5	0,8251	0,681
FWPHe_09b_unt	48	3	1	4	3,021	0,9998	1
FWPHe_09b_lr	49	3	1	4	3	0,8898	0,792
FWPHe_10_ein	35	2	2	4	3,657	0,5392	0,291
FWPHe_10_unt	36	3	1	4	2,861	1,0185	1,037
FWPHe_10_lr	36	3	1	4	3,083	0,7973	0,636
FWPHe_11_ein	49	3	1	4	3,776	0,5502	0,303
FWPHe_11_unt	51	3	1	4	3,059	1,0278	1,056
FWPHe_11_lr	51	3	1	4	3,294	0,9229	0,852
FWPHe_12_ein	44	3	1	4	3,727	0,6599	0,436
FWPHe_12_unt	45	3	1	4	2,789	1,1604	1,346
FWPHe_12_lr	45	3	1	4	2,967	1,1201	1,255
FWPHe_13a_ein	47	3	1	4	3,489	0,8565	0,734
FWPHe_13a_unt	47	3	1	4	3	0,9089	0,826
FWPHe_13a_lr	47	3	1	4	2,989	0,9118	0,831
FWPHe_13b_ein	48	3	1	4	3,5	0,8251	0,681
FWPHe_13b_unt	48	3	1	4	3,062	0,9319	0,868
FWPHe_13b_lr	48	3	1	4	3,208	0,8495	0,722
FWPHe_14_ein	46	3	1	4	3,826	0,5293	0,28

Fortsetzung auf der nächsten Seite

Tabelle A.10.: (Fortsetzung)

FWPHe_14_unt	47	3	1	4	3,543	0,6662	0,444
FWPHe_14_lr	47	3	1	4	3,617	0,6098	0,372
FWPHe_16a_ein	48	2	2	4	3,792	0,5035	0,254
FWPHe_16a_unt	48	2	2	4	3,458	0,7426	0,551
FWPHe_16a_lr	48	3	1	4	3,417	0,7672	0,589
FWPHe_17_ein	43	3	1	4	3,605	0,7283	0,53
FWPHe_17_unt	43	3	1	4	2,581	1,0055	1,011
FWPHe_17_lr	43	3	1	4	2,744	1,0022	1,004
FWPHe_18_ein	40	1	3	4	3,95	0,2207	0,049
FWPHe_18_unt	40	2	2	4	3,85	0,4267	0,182
FWPHe_18_lr	40	2	2	4	3,8	0,5164	0,267
FWPHe_19_ein	46	2	2	4	3,609	0,6824	0,466
FWPHe_19_unt	47	2	2	4	3,468	0,7178	0,515
FWPHe_19_lr	47	3	1	4	3,596	0,712	0,507
FWPHe_20b_ein	48	3	1	4	3,521	0,7143	0,51
FWPHe_20b_unt	48	2	2	4	3,427	0,7508	0,564
FWPHe_20b_lr	48	2	2	4	3,583	0,6131	0,376
FWPHe_21_ein	37	3	1	4	3,757	0,6414	0,411
FWPHe_21_unt	38	3	1	4	3,132	0,9056	0,82
FWPHe_21_lr	38	3	1	4	3,342	0,8785	0,772
Gültige Werte (Listenweise)	13						

Tabelle A.11.: Mittelwerte der Augenscheinvaliditäten aus der Pilotstudie 2

Deskriptive Statistik							
	<i>N</i>	range	Min	Max	MW	St.abw.	Varianz
FDPHm_02_k_ein	40	3	1	4	3,5	0,8165	0,667
FDPHm_02_k_ber	38	3	1	4	3,263	0,76	0,578
FDPHm_02_k_lr	39	2	2	4	3,436	0,718	0,516
FDPHm_03_k_ein	40	3	1	4	3,725	0,6789	0,461
FDPHm_03_k_ber	38	2	2	4	3,421	0,7929	0,629
FDPHm_03_k_lr	39	3	1	4	3,487	0,7564	0,572
FDPHm_05_i_ein	37	2	2	4	3,649	0,5877	0,345
FDPHm_05_i_ber	35	3	1	4	3,371	0,8432	0,711
FDPHm_05_i_lr	36	3	1	4	3,583	0,6918	0,479
FDPHm_06_i_ein	34	3	1	4	3,574	0,7191	0,517
FDPHm_06_i_ber	33	3	1	4	3,455	0,9045	0,818
FDPHm_06_i_lr	34	3	1	4	3,471	0,7876	0,62
FDPHo_01_k_ein	37	2	2	4	3,635	0,6524	0,426
FDPHo_01_k_ber	36	2	2	4	3,667	0,5345	0,286
FDPHo_01_k_lr	36	2	2	4	3,694	0,5767	0,333
FDPHo_02_me_ein	40	2	2	4	3,887	0,3838	0,147
FDPHo_02_me_ber	38	2	2	4	3,658	0,5825	0,339
FDPHo_02_me_lr	39	2	2	4	3,692	0,5691	0,324
FDPHo_03_me_ein	38	3	1	4	3,75	0,5899	0,348
FDPHo_03_me_ber	36	1	3	4	3,903	0,2883	0,083
FDPHo_03_me_lr	37	2	2	4	3,824	0,4444	0,197
FDPHo_05_i_ein	38	3	1	4	3,145	1,0195	1,039
FDPHo_05_i_ber	36	3	1	4	3,222	0,8738	0,763
FDPHo_05_i_lr	37	3	1	4	3,216	0,8623	0,744
FDPHw_01_k_ein	39	3	1	4	3,436	0,7879	0,621
FDPHw_01_k_ber	37	3	1	4	3,459	0,7301	0,533
FDPHw_01_k_lr	38	3	1	4	3,382	0,73	0,533
FDPHw_02_k_ein	40	3	1	4	3,162	1,1457	1,313
FDPHw_02_k_ber	38	3	1	4	3,329	0,9025	0,815
FDPHw_02_k_lr	39	3	1	4	3,385	0,8845	0,782
FDPHw_03_me_ein	39	2	2	4	3,821	0,4514	0,204
FDPHw_03_me_ber	37	3	1	4	3,689	0,6163	0,38
FDPHw_03_me_lr	37	3	1	4	3,689	0,6163	0,38
FDPHw_05_k_ein	39	3	1	4	3,5	0,6589	0,434
FDPHw_05_k_ber	37	3	1	4	3,27	0,8383	0,703
FDPHw_05_k_lr	38	3	1	4	3,237	0,8833	0,78
FDPHw_06_i_ein	36	3	1	4	3,125	0,9811	0,962
FDPHw_06_i_ber	34	3	1	4	3,309	0,7589	0,576

Fortsetzung auf der nächsten Seite

Tabelle A.11.: (Fortsetzung)

FDPHw_06_i_lr	34	3	1	4	3,338	0,7659	0,587
FWPHm_01_ein	46	3	1	4	3,348	0,9479	0,899
FWPHm_01_ber	41	2	2	4	3,524	0,5472	0,299
FWPHm_01_lr	42	2	2	4	3,583	0,5835	0,34
FWPHm_02_ein	46	3	1	4	3,587	0,7173	0,514
FWPHm_02_ber	41	2	2	4	3,634	0,6984	0,488
FWPHm_02_lr	42	2	2	4	3,714	0,5537	0,307
FWPHm_04_ein	48	3	1	4	3,635	0,6664	0,444
FWPHm_04_ber	43	3	1	4	3,267	0,902	0,814
FWPHm_04_lr	44	2	2	4	3,375	0,7163	0,513
FWPHm_05_ein	49	3	1	4	3,745	0,5602	0,314
FWPHm_05_ber	44	3	1	4	3,443	0,8971	0,805
FWPHm_05_lr	45	3	1	4	3,5	0,8118	0,659
FWPHw_08_ein	48	3	1	4	3,74	0,6357	0,404
FWPHw_08_ber	43	3	1	4	3,5	0,8238	0,679
FWPHw_08_lr	44	3	1	4	3,602	0,72	0,518
FWPHm_09_ein	49	3	1	4	3,571	0,7638	0,583
FWPHm_09_ber	44	3	1	4	3,341	0,8611	0,742
FWPHm_09_lr	45	3	1	4	3,489	0,7268	0,528
FWPHm_11_ein	49	2	2	4	3,745	0,5412	0,293
FWPHm_11_ber	44	2	2	4	3,557	0,6578	0,433
FWPHm_11_lr	45	2	2	4	3,611	0,573	0,328
FWPHm_12_ein	48	2	2	4	3,854	0,412	0,17
FWPHm_12_ber	43	3	1	4	3,372	0,8735	0,763
FWPHm_12_lr	44	2	2	4	3,545	0,6973	0,486
FWPHo_02_ein	47	3	1	4	3,66	0,6351	0,403
FWPHo_02_ber	42	2	2	4	3,595	0,627	0,393
FWPHo_02_lr	43	2	2	4	3,698	0,5578	0,311
FWPHo_03_ein	46	3	1	4	3,511	0,7187	0,517
FWPHo_03_ber	43	3	1	4	3,244	0,8119	0,659
FWPHo_03_lr	44	3	1	4	3,307	0,8228	0,677
FWPHo_04_ein	49	3	1	4	3,786	0,559	0,312
FWPHo_04_ber	44	3	1	4	3,568	0,7894	0,623
FWPHo_04_lr	45	3	1	4	3,6	0,7508	0,564
FWPHo_05_ein	47	3	1	4	3,681	0,7831	0,613
FWPHo_05_ber	42	2	2	4	3,524	0,6339	0,402
FWPHo_05_lr	43	3	1	4	3,465	0,8266	0,683
FWPHo_06_ein	48	3	1	4	3,583	0,679	0,461
FWPHo_06_ber	42	2	2	4	3,571	0,5903	0,348
FWPHo_06_lr	43	2	2	4	3,558	0,6288	0,395
FWPHo_07_ein	48	3	1	4	3,677	0,7182	0,516

Fortsetzung auf der nächsten Seite

Tabelle A.11.: (Fortsetzung)

FWPHo_07_ber	44	2	2	4	3,477	0,6901	0,476
FWPHo_07_lr	45	2	2	4	3,489	0,6866	0,471
FWPHw_01_ein	49	3	1	4	3,745	0,6624	0,439
FWPHw_01_ber	44	3	1	4	3,625	0,7163	0,513
FWPHw_01_lr	45	3	1	4	3,678	0,6321	0,399
FWPHw_04_ein	48	3	1	4	3,76	0,6101	0,372
FWPHw_04_ber	43	3	1	4	3,337	0,9304	0,866
FWPHw_04_lr	44	3	1	4	3,466	0,8026	0,644
FWPHw_05_ein	46	2	2	4	3,815	0,5095	0,26
FWPHw_05_ber	41	3	1	4	3,305	0,8282	0,686
FWPHw_05_lr	42	3	1	4	3,464	0,8144	0,663
FWPHw_07_ein	46	1	3	4	3,87	0,3405	0,116
FWPHw_07_ber	41	2	2	4	3,683	0,5674	0,322
FWPHw_07_lr	42	2	2	4	3,643	0,5768	0,333
FWPHw_09_ein	46	2	2	4	3,826	0,4855	0,236
FWPHw_09_ber	41	1	3	4	3,78	0,4191	0,176
FWPHw_09_lr	42	1	3	4	3,81	0,3974	0,158
Gültige Werte (Listenweise)	17						

Tabelle A.12.: Schwierigkeitsindizes der Items aus der Hauptstudie; Hinweis: Insgesamt lagen 183 Datensätze (Hauptstichprobe) vor.

Item	FDPHm_03_k	FDPHm_05_i	FDPHe_02a_k	FDPHe_02b_i	FDPHe_05_i
m_item	1	0,3	0,8	0,5	0,9
SD_item	0,75	0,46	0,41	0,5	0,8
Max_item	2	1	1	1	2
Index [%]	47,81	30,6	78,14	51,37	46,17
Item	FDPHe_12_me	FDPHe_14_me	FDPHe_01_i	FDPHe_11_i	FDPHo_02_me
m_item	1,5	0,2	0,9	1,2	0,6
SD_item	0,82	0,36	0,83	0,92	0,88
Max_item	3	2	4	5	2
Index [%]	49,73	7,65	21,86	24,37	28,96
Item	FDPHo_01_k	FDPHw_03_me	FDPHw_05a_k	FDPHw_01_k	FWPHe_11
m_item	0,8	1,2	0,5	0,6	0,8
SD_item	0,7	0,88	0,5	0,69	0,43
Max_item	3	2	1	3	1
Index [%]	25,5	60,11	51,37	18,58	75,41
Item	FWPHe_12	FWPHe_05	FWPHe_02	FWPHe_21	FWPHe_04
m_item	0,9	0,5	0,5	1,2	0,9
SD_item	0,35	0,5	0,5	0,89	0,35
Max_item	1	1	1	2	1
Index [%]	85,79	53,01	48,09	59,84	85,79
Item	FWPHe_14	FWPHe_16a	FWPHe_20b	FWPHe_13b1	FWPHe_12
m_item	0,9	0,4	2,4	0,3	0,6
SD_item	0,85	0,36	1,39	0,47	0,75
Max_item	2	1	4	1	2
Index [%]	43,72	41,26	59,7	33,33	32,1
Item	FWPHe_06	FWPHe_07	FWPHe_06	FWPHe_03	FWPHe_09
m_item	0,4	0,6	0,3	0,2	0,7
SD_item	0,73	0,49	0,45	0,46	1,03
Max_item	2	1	1	2	3
Index [%]	22,13	60,11	28,96	9,56	24,23
Item	FWPHe_07	FWPHe_05			
m_item	0,5	0,2			
SD_item	0,65	0,41			
Max_item	2	1			
Index [%]	27,32	21,31			

Tabelle A.13.: Geschlechtsverteilung, Altersstruktur sowie Anzahl der Berufsjahre der Experten und erfahrenen Lehrkräfte aus der Prä-Pilotstudie 1

Statistiken						
		sex	Alter2014	Berj		
N	Gültig	11	11	10		
	Fehlend	0	0	1		
Mittelwert		1,182	50,2727	22,000		
Standardabweichung		,4045	12,17449	11,3725		
Varianz		,164	148,218	129,333		
Spannweite		1,0	42,00	31,0		
Minimum		1,0	32,00	6,0		
Maximum		2,0	74,00	37,0		
sex						
		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente	
Gültig	1,0	9	81,8	81,8	81,8	
	2,0	2	18,2	18,2	100,0	
Gesamt		11	100,0	100,0		

Tabelle A.14.: Geschlechtsverteilung, Altersstruktur sowie Anzahl der Berufsjahre der Experten und erfahrenen Lehrkräfte aus der Prä-Pilotstudie 2

Statistiken						
		sex	Alter2014	Berj		
N	Gültig	13	12	11		
	Fehlend	0	1	2		
Mittelwert		1,154	45,250	214,455		
Standardabweichung		,3755	12,5127	13,3144		
Varianz		,141	156,568	177,273		
Spannweite		1,0	36,0	37,0		
Minimum		1,0	29,0	,0		
Maximum		2,0	65,0	37,0		
sex						
		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente	
Gültig	1,0	11	84,6	84,6	84,6	
	2,0	2	15,4	15,4	100,0	
Gesamt		13	100,0	100,0		

Tabelle A.15.: Stichprobenzusammensetzung der Pilot-Studie 1

Art		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Pro- zente
Gültig	HS, MS	33	44,0	44,0	44,0
	RS	30	40,0	40,0	84,0
	Gy	4	5,3	5,3	89,3
	Uni	7	9,3	9,3	98,7
	Sonst	1	1,3	1,3	100,0
	Gesamt	75	100,0	100,0	
Status		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Pro- zente
Gültig	Student	14	18,7	18,7	18,7
	Leh. U-Fach	24	32,0	32,0	50,7
	Leh. D-Fach	1	1,3	1,3	52,0
	Leh. n. stud.	28	37,7	37,7	89,3
	Uni	7	9,3	9,3	98,7
	Sonst	1	1,3	1,3	100,0
	Gesamt	75	100,0	100,0	
Statistiken		sex	Alter2014	BerJ	
N	Gültig	75	75	69	
	Fehlend	0	0	6	
Mittelwert		1,360	37,8000	10,942	
Standardabweichung		,4832	11,33519	10,1230	
Varianz		,234	128,486	102,475	
Spannweite		1,0	52,00	37,0	
Minimum		1,0	22,0	,0	
Maximum		2,0	74,00	37,0	

Tabelle A.16.: Stichprobenzusammensetzung der Pilot-Studie 2

Art		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Pro- zente
Gültig	HS, MS	19	34,5	34,5	34,5
	RS	12	21,8	21,8	56,4
	Gy	7	12,7	12,7	69,1
	Uni	8	14,5	14,5	83,6
	Uni fachfr.	9	16,4	16,4	100,0
	Gesamt	55	100,0	100,0	
Status		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Pro- zente
Gültig	Student	6	10,9	10,9	10,9
	Leh. U-Fach	14	25,5	25,5	36,4
	Leh. D-Fach	2	3,6	3,6	40,0
	Leh. n. stud.	16	29,1	29,1	69,1
	Uni	8	14,5	14,5	83,6
	Uni fachfr.	9	16,4	16,4	100,0
	Gesamt	55	100,0	100,0	
Statistiken		sex	Alter2014	BerJ	
N	Gültig	55	55	54	
	Fehlend	0	0	1	
Mittelwert		1,418	40,455	11,972	
Standardabweichung		,4978	12,4273	11,6000	
Varianz		,248	154,438	134,561	
Spannweite		1,0	50,0	40,0	
Minimum		1,0	25,0	,0,	
Maximum		2,0	75,00	40,0	

Tabelle A.17.: Interraterreliabilität (Spearman's ρ) auf Einzelitemebene, Pilot 1 + 2

	Pilot 1, N = 28	Pilot 2, N = 35		
FDW:	FDPHe_01_i_score	0,894	FDPHm_02_k_score	1,000
	FDPHe_02a_k_score	0,645	FDPHm_03_k_score	0,894
	FDPHe_02b_i_score	0,893	FDPHm_05_i_score	0,894
	FDPHe_03_k_score	0,822	FDPHm_06_i_score	0,920
	FDPHe_05_i_score	0,807	FDPHo_01_k_score	0,901
	FDPHe_06_me_score	0,909	FDPHo_02_me_score	0,928
	FDPHe_07_i_score	0,718	FDPHo_03_me_score	0,797
	FDPHe_08_k_score	0,864	FDPHo_05_i_score	0,797
	FDPHe_01_i_score	0,894	FDPHm_02_k_score	1,000
	FDPHe_09_i_score	0,806	FDPHw_01_k_score	0,929
	FDPHe_10_me_score	0,828	FDPHw_02_k_score	0,900
	FDPHe_11_i_score	0,760	FDPHw_03_me_score	0,960
	FDPHe_13_k_score	0,718	FDPHw_05_k_score	0,891
	FDPHe_14_me_score	0,900	FDPHw_06_i_score	0,860
FW:	FWPHe_05_score	0,957	FWPHm_02_score	0,854
	FWPHe_16a_score	0,930	FWPHm_05_score	0,957
	FWPHe_02_score	0,850	FWPHo_03_score	0,839
	FWPHe_06_score	0,734	FWPHw_01_score	0,887
	FWPHe_09b_score	0,225	FWPHw_07_score	0,764
	FWPHe_18_score	0,895	FWPHw_09_score	0,903
	FWPHe_20b_score	0,989		
	FWPHe_10_score	1,000		
	FWPHe_17_score	0,488		
	FWPHe_13b_score	0,676		
	FWPHe_11_score	0,768		
	FWPHe_12_score	1,000		
	FWPHe_21_score	0,791		
	FWPHe_19_score	0,426		
	FWPHe_14_score	0,762		

Tabelle A.18.: Stichprobenzusammensetzung der Hauptstudie, aufgeteilt nach „beruflicher Stellung“, „Schulart“, „Status des getesteten Fachs“, „Geschlecht“, „Alter“, „Berufsjahre“ und „Semesteranzahl“

Berufliche Stellung		Häufigkeit	Prozent	Gült. Proz.	Kum. Proz.	
Gültig	Studierende	118	43,9	43,9	43,9	
	Referendar/in	7	2,6	2,6	46,5	
	Lehrkraft	83	30,9	30,9	77,3	
	Leh. fachfr.	32	11,9	11,9	89,2	
	Fachwiss.	29	10,8	10,8	100,0	
	Gesamt	269	100,0	100,0		
Schulart		Häufigkeit	Prozent	Gült. Proz.	Kum. Proz.	
Gültig	Mittelschule	33	12,3	12,3	12,3	
	Realschule	81	30,1	30,1	42,4	
	Leh. D-Fach	2	3,6	3,6	40,0	
	Gymnasium	126	46,8	46,8	89,2	
	Sonstiges	29	10,8	10,8	100,0	
	Gesamt	269	100,0	100,0		
Status des getesteten Fachs		Häufigkeit	Prozent	Gült. Proz.	Kum. Proz.	
Gültig	nicht stud.	54	20,1	20,1	20,1	
	Didaktikfach	4	1,5	1,5	21,6	
	Unt.fach	179	66,5	66,5	88,1	
	D-F. NWT	3	1,1	1,1	89,2	
	Fachphysik	29	10,8	10,8	100,0	
	Gesamt	269	100,0	100,0		
Geschlechtsangabe		Häufigkeit	Prozent	Gült. Proz.	Kum. Proz.	
Gültig	weiblich	90	33,5	33,5	33,5	
	männlich	179	66,5	66,5	100,0	
	Gesamt	269	100,0	100,0		
Deskriptive Statistik		N	Minimum	Maximum	Mittelwert	St.abw.
Alter_2016	267	20,00	74,00	32,5318	10,94217	

Fortsetzung auf der nächsten Seite

Tabelle A.18.: (Fortsetzung von S. 139)

Ber.j. (Leh.)	147	1	40	11,22	8,500
Sem. (Stud.)	118	2	12	6,35	2,827

Tabelle A.19.: t-Test zur Begründung der Zusammenlegung der RS- und MS-Lehrkräfte; MS-Lehrkräfte mit Unterrichts- und Didaktikfach Physik wurden in einer Stichprobe zusammengefasst

Gruppenstatistiken						
	Schulart	N	Mittelwert	Standardabweichung	Std.feh. d. MW	
FDWFWScores	Mittelschule	6	18,8333	12,019 43	4,906 91	
	Realschule	24	29,6458	8,087 13	1,650 78	
FDWScores	Mittelschule	6	9,1667	4,915 96	2,006 93	
	Realschule	24	12,8750	4,396 76	0,897 49	
FWScores	Mittelschule	6	9,6667	7,600 44	3,102 87	
	Realschule	24	16,7708	5,012 98	1,023 27	
FDW_ges_ohneME	Mittelschule	6	5,8333	3,371 45	1,376 93	
	Realschule	24	8,4583	3,361 92	0,686 25	

Test bei unabhängigen Stichproben

		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit						
	F	Signifi- kanz	T	df	Sig. (2-sei.)	Mittl. Diff.	St.f. d. Diff.	95 % Konfidenzintervall der Differenz		
								untere	obere	
PCKCK_Score_ges	Var. =	1,082	0,307	-2,656	28	0,013	-10,81250	4,07023	-19,14999	-2,47501
PCKCK_Score_ges	Var. ≠			-2,089	6,179	0,080	-10,81250	5,17715	-23,39228	1,76728
PCK_Score_ges	Var. =	0,256	0,617	-1,808	28	0,081	-3,70833	2,05117	-7,90996	0,49329
PCK_Score_ges	Var. ≠			-1,687	7,138	0,135	-3,70833	2,19847	-8,88662	1,46995
CK_Score_ges	Var. =	0,541	0,468	-2,797	28	0,009	-7,10417	2,53960	-12,30630	-1,90203
CK_Score_ges	Var. ≠			-2,174	6,131	0,072	-7,10417	3,26724	-15,05761	0,84928
PCKCK_ges_ohneME	Var. =	0,010	0,921	-1,710	28	0,098	-2,62500	1,53528	-5,76988	0,51988
PCKCK_ges_ohneME	Var. ≠			-1,707	7,692	0,128	-2,62500	1,53798	-6,19650	0,94650

Tabelle A.20.: t-Test zur Begründung der Zusammenlegung der RS- und MS-Lehrkräfte; MS-Lehrkräfte mit Unterrichtsfach wurden zu einer Gesamtgruppe zusammengefasst

Gruppenstatistiken						
	Schulart	N	Mittelwert	Standardabweichung	Std.feh. d. MW	
FDWFWScores	Mittelschule	2	31,7500	11,667 26	8,250 00	
	Realschule	24	29,6458	8,087 13	1,650 78	
FDWScores	Mittelschule	2	14,0000	2,828 43	2,000 00	
	Realschule	24	12,8750	4,396 76	0,897 49	
FWScores	Mittelschule	2	17,7500	8,838 83	6,250 00	
	Realschule	24	16,7708	5,012 98	1,023 27	
FDW_ges_ohneME	Mittelschule	2	9,0000	2,828 43	2,000 00	
	Realschule	24	8,4583	3,361 92	0,686 25	

Test bei unabhängigen Stichproben

		Levene-Test der Varianzgleichheit						T-Test für die Mittelwertgleichheit			
	F	Signifi- kanz	T	df	Sig. (2-sei.)	Mittl. Diff.	St.f. d. Diff.	95 % Konfidenzintervall der Differenz			
								untere	obere		
PCKCK_Score_ges	Var. = 0,253	0,620	0,264	24	0,732	2,10417	6,08458	-10,45379	14,66212		
PCKCK_Score_ges	Var. ≠		0,350	1,082	0,842	2,10417	8,41354	-87,50329	91,71163		
PCK_Score_ges	Var. = 0,955	0,338	0,352	24	0,728	1,12500	3,19617	-5,47157	7,72157		
PCK_Score_ges	Var. ≠		0,513	1,441	0,675	1,12500	2,19214	-12,85146	15,10146		
CK_Score_ges	Var. = 1,407	0,247	0,254	24	0,801	0,97917	3,84813	-6,96299	8,92133		
CK_Score_ges	Var. ≠		0,155	1,054	0,901	0,97917	6,33321	-70,32995	72,28828		
PCKCK_ges_ohneME	Var. = 0,159	0,694	0,220	24	0,828	0,54167	2,45920	-4,53388	5,61721		
PCKCK_ges_ohneME	Var. ≠		0,256	1,249	0,834	0,54167	2,11446	-16,46750	17,55083		

Tabelle A.21.: t-Test zur Begründung der Zusammenlegung der RS- und MS-Lehrkräfte; MS-Lehrkräfte mit Didaktikfach wurden zu einer Gesamtgruppe zusammengefasst

Gruppenstatistiken						
	Schulart	N	Mittelwert	Standardabweichung	Std.feh. d. MW	
FDWFWScores	Mittelschule	4	12,3750	5,344 39	2,672 19	
	Realschule	24	29,6458	8,087 13	1,650 78	
FDWScores	Mittelschule	4	6,7500	3,774 92	1,887 46	
	Realschule	24	12,8750	4,396 76	0,897 49	
FWScores	Mittelschule	4	5,6250	2,212 65	1,106 33	
	Realschule	24	16,7708	5,012 98	1,023 27	
FDW_ges_ohneME	Mittelschule	4	4,2500	2,500 00	1,2500	
	Realschule	24	8,4583	3,361 92	0,686 25	

Test bei unabhängigen Stichproben

		T-Test für die Mittelwertgleichheit								
		Levene-Test der Varianzgleichheit			95 % Konfidenzintervall der Differenz					
		F	Signifi- kanz	T	df	Sig. (2-sei.)	Mittl. Diff.	St.f. d. Diff.	untere	obere
PCKCK_Score_ges	Var. =	1,580	0,220	-4,089	26	0,000	-17,27083	4,22324	-25,95182	-8,58985
PCKCK_Score_ges	Var. ≠			-5,499	5,620	0,002	-17,27083	3,14097	-25,08428	-9,45739
PCK_Score_ges	Var. =	0,518	0,478	-2,619	26	0,015	-6,12500	2,33824	-10,93131	-1,31869
PCK_Score_ges	Var. ≠			-2,931	4,480	0,037	-6,12500	2,08997	-11,69050	-0,55950
CK_Score_ges	Var. =	5,828	0,023	-4,323	26	0,000	-11,14583	2,57849	-16,44600	-5,84567
CK_Score_ges	Var. ≠			-7,396	9,428	0,000	-11,14583	1,50700	-14,53143	-7,76024
PCKCK_ges_ohneME	Var. =	0,604	0,444	-2,380	26	0,025	-4,20833	1,76820	-7,84292	-0,57374
PCKCK_ges_ohneME	Var. ≠			-2,951	5,021	0,032	-4,20833	1,42599	-7,86925	-0,54742

Tabelle A.22.: Reliabilitäten (Cronbachs α) der gebildeten Skalen zur Überprüfung der Repräsentativitätshypothese

$N = 183$	Mechanik	E-Lehre	Optik	Wärmelehre
FDW + FW(32 Items)	0,59 (6 Items)	0,86 (15 Items)	0,51 (5 Items)	0,56 (5 Items)
FDW(14 Items)	0,17 (2 Items)	0,72 (7 Items)	0,36 (2 Items)	0,42 (3 Items)
FW(18 Items)	0,55 (4 Items)	0,80 (8 Items)	0,39 (3 Items)	0,56 (3 Items)

Tabelle A.23.: Mittelwerte der Augenscheinvaliditäten der Studienteilnehmer aus der Hauptstudie

	Deskriptive Statistik				
	<i>N</i>	Minimum	Maximum	Mittelwert	Standard- abweichung
Augenscheinvalidität PCK Aufgabe 1 Eindeutigkeit	25	2	4	3,76	0,5228
Augenscheinvalidität PCK Aufgabe 1 Ausbildungsrelevanz	24	1	4	3,5	0,7802
Augenscheinvalidität PCK Aufgabe 2 Eindeutigkeit	24	2	4	3,708	0,55
Augenscheinvalidität PCK Aufgabe 2 Ausbildungsrelevanz	23	3	4	3,696	0,4705
Augenscheinvalidität PCK Aufgabe 3a Eindeutigkeit	12	2	4	3,583	0,6686
Augenscheinvalidität PCK Aufgabe 3a Ausbildungsrelevanz	12	2	4	3,667	0,6513
Augenscheinvalidität PCK Aufgabe 3b Eindeutigkeit	12	2	4	3,5	0,6742
Augenscheinvalidität PCK Aufgabe 3b Ausbildungsrelevanz	12	2	4	3,667	0,6513
Augenscheinvalidität PCK Aufgabe 4 Eindeutigkeit	11	2	4	3,545	0,6876
Augenscheinvalidität PCK Aufgabe 4 Ausbildungsrelevanz	11	1	4	3,364	0,9244
Augenscheinvalidität PCK Aufgabe 5 Eindeutigkeit	11	2	4	3,727	0,6467
Augenscheinvalidität PCK Aufgabe 5 Ausbildungsrelevanz	11	3	4	3,909	0,3015
Augenscheinvalidität PCK Aufgabe 6 Eindeutigkeit	11	3	4	3,727	0,4671
Augenscheinvalidität PCK Aufgabe 6 Ausbildungsrelevanz	11	1	4	3	1
Augenscheinvalidität PCK Aufgabe 7 Eindeutigkeit	12	3	4	3,5	0,5222
Augenscheinvalidität PCK Aufgabe 7 Ausbildungsrelevanz	12	3	4	3,75	0,4523
Augenscheinvalidität PCK Aufgabe 8 Eindeutigkeit	11	3	4	3,636	0,5045
Augenscheinvalidität PCK Aufgabe 8 Ausbildungsrelevanz	11	4	4	4	0
Augenscheinvalidität PCK Aufgabe 9 Eindeutigkeit	25	2	4	3,86	0,4453
Augenscheinvalidität PCK Aufgabe 9 Ausbildungsrelevanz	24	2	4	3,708	0,55
Augenscheinvalidität PCK Aufgabe 10 Eindeutigkeit	22	2	4	3,659	0,6794

Fortsetzung auf der nächsten Seite

Tabelle A.23.: (Fortsetzung)

Augenscheinvalidität PCK Aufgabe 10 Ausbildungsrelevanz	22	2	4	3,727	0,5505
Augenscheinvalidität PCK Aufgabe 11 Eindeutigkeit	25	2	4	3,8	0,5
Augenscheinvalidität PCK Aufgabe 11 Ausbildungsrelevanz	23	1	4	3,674	0,7008
Augenscheinvalidität PCK Aufgabe 12 Eindeutigkeit	25	1	4	3,46	0,7348
Augenscheinvalidität PCK Aufgabe 12 Ausbildungsrelevanz	24	1	4	3,125	0,947
Augenscheinvalidität PCK Aufgabe 13 Eindeutigkeit	25	1	4	3,44	0,8699
Augenscheinvalidität PCK Aufgabe 13 Ausbildungsrelevanz	24	1	4	3,396	0,7937
Augenscheinvalidität CK Aufgabe 1 Eindeutigkeit	34	2	4	3,838	0,4723
Augenscheinvalidität CK Aufgabe 1 Ausbildungsrelevanz	30	2	4	3,65	0,5438
Augenscheinvalidität CK Aufgabe 2 Eindeutigkeit	33	2	4	3,848	0,4417
Augenscheinvalidität CK Aufgabe 2 Ausbildungsrelevanz	29	2	4	3,517	0,7378
Augenscheinvalidität CK Aufgabe 3 Eindeutigkeit	34	1	4	3,691	0,6278
Augenscheinvalidität CK Aufgabe 3 Ausbildungsrelevanz	30	1	4	3,583	0,8103
Augenscheinvalidität CK Aufgabe 4 Eindeutigkeit	33	1	4	3,576	0,7084
Augenscheinvalidität CK Aufgabe 4 Ausbildungsrelevanz	29	3	4	3,862	0,3509
Augenscheinvalidität CK Aufgabe 5 Eindeutigkeit	10	1	4	3,6	0,9661
Augenscheinvalidität CK Aufgabe 5 Ausbildungsrelevanz	10	2	4	3,3	0,6749
Augenscheinvalidität CK Aufgabe 6 Eindeutigkeit	10	2	4	3,8	0,6325
Augenscheinvalidität CK Aufgabe 6 Ausbildungsrelevanz	10	3	4	3,7	0,483
Augenscheinvalidität CK Aufgabe 7 Eindeutigkeit	10	1	4	3,6	0,9661
Augenscheinvalidität CK Aufgabe 7 Ausbildungsrelevanz	10	3	4	3,6	0,5164
Augenscheinvalidität CK Aufgabe 8 Eindeutigkeit	10	3	4	3,8	0,4216
Augenscheinvalidität CK Aufgabe 8 Ausbildungsrelevanz	10	1	4	3,1	1,1005

Fortsetzung auf der nächsten Seite

Tabelle A.23.: (Fortsetzung)

Augenscheinvalidität CK Aufgabe 9 Eindeutigkeit	11	2	4	3,273	0,6467
Augenscheinvalidität CK Aufgabe 9 Ausbildungsrelevanz	11	2	4	3,364	0,6742
Augenscheinvalidität CK Aufgabe 10 Eindeutigkeit	11	1	4	3,364	1,0269
Augenscheinvalidität CK Aufgabe 10 Ausbildungsrelevanz	11	2	4	2,818	0,7508
Augenscheinvalidität CK Aufgabe 11 Eindeutigkeit	11	1	4	3,364	0,9244
Augenscheinvalidität CK Aufgabe 11 Ausbildungsrelevanz	11	1	4	3,045	0,9606
Augenscheinvalidität CK Aufgabe 12 Eindeutigkeit	10	3	4	3,7	0,483
Augenscheinvalidität CK Aufgabe 12 Ausbildungsrelevanz	10	3	4	3,4	0,5164
Augenscheinvalidität CK Aufgabe 13 Eindeutigkeit	34	1	4	3,574	0,8177
Augenscheinvalidität CK Aufgabe 13 Ausbildungsrelevanz	30	2	4	3,433	0,716
Augenscheinvalidität CK Aufgabe 14 Eindeutigkeit	33	2	4	3,576	0,6139
Augenscheinvalidität CK Aufgabe 14 Ausbildungsrelevanz	29	2	4	3,552	0,6317
Augenscheinvalidität CK Aufgabe 15 Eindeutigkeit	32	1	4	3,484	0,7565
Augenscheinvalidität CK Aufgabe 15 Ausbildungsrelevanz	30	1	4	3,283	0,9067
Augenscheinvalidität CK Aufgabe 16 Eindeutigkeit	31	2	4	3,774	0,5603
Augenscheinvalidität CK Aufgabe 16 Ausbildungsrelevanz	27	3	4	3,889	0,3203
Augenscheinvalidität CK Aufgabe 17 Eindeutigkeit	32	3	4	3,875	0,336
Augenscheinvalidität CK Aufgabe 17 Ausbildungsrelevanz	28	2	4	3,714	0,5345
Augenscheinvalidität CK Aufgabe 18 Eindeutigkeit	31	2	4	3,758	0,5898
Augenscheinvalidität CK Aufgabe 18 Ausbildungsrelevanz	27	2	4	3,574	0,6608
Gültige Werte (Listenweise)	0				

Tabelle A.24.: Manifeste Korrelationen (nach Pearson) der vier separaten Skalen nach Inhalt (FDW)

	Mechanik (2 Items)	E-Lehre (7 Items)	Optik (2 Items)	Wärmelehre (3 Items)
Mechanik (2 Items)	1,00			
E-Lehre (7 Items)	0,30	1,00		
Optik (2 Items)	0,26	0,39	1,00	
Wärmelehre (3 Items)	0,23	0,55	0,34	1,00

Tabelle A.25.: Manifeste Korrelationen (nach Pearson) der vier separaten Skalen nach Inhalt (FW)

	Mechanik (4 Items)	E-Lehre (8 Items)	Optik (3 Items)	Wärmelehre (3 Items)
Mechanik (4 Items)	1,00			
E-Lehre (8 Items)	0,50	1,00		
Optik (3 Items)	0,29	0,44	1,00	
Wärmelehre (3 Items)	0,35	0,55	0,34	1,00

Tabelle A.26.: Stichprobenzusammensetzung der Teilstichprobe aus gymnasialen und nicht-gymnasialen Physiklehrkräften ($N = 65$)

Berufliche Stellung						
		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente	
Gültig	Ref./-in	4	6,2	6,2	6,2	
	Lehrkraft	61	93,8	93,8	100,0	
	Gesamt	65	100,0	100,0		
Schulart						
		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente	
Gültig	Mittelschule	6	9,2	9,2	9,2	
	Realschule	24	36,9	36,9	46,2	
	Gymnasium	35	53,8	53,8	100,0	
	Gesamt	65	100,0	100,0		
Geschlechtsangabe						
		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente	
Gültig	weiblich	21	32,3	32,3	32,3	
	männlich	44	67,7	67,7	100,0	
	Gesamt	65	100,0	100,0		
Deskriptive Statistik						
		N	Minimum	Maximum	Mittelwert	Standardab- weichung
Berufsjahre	65	1	40	13,54	9,582	
Alter_2016	65	26	70	42,58	10,648	

Tabelle A.27.: Stichprobenzusammensetzung der Teilstichprobe aus gymnasialen Physiklehrkräften ($N = 35$)

Berufliche Stellung						
		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente	
Lehrkraft		35	100,0	100,0	100,0	
Schulart						
		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente	
Gymnasium		35	100,0	100,0	100,0	
Geschlechtsangabe						
		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente	
Gültig	weiblich	7	20,0	20,0	20,0	
	männlich	28	80,0	80,0	100,0	
	Gesamt	35	100,0	100,0		
Deskriptive Statistik						
		N	Minimum	Maximum	Mittelwert	Standardab- weichung
Berufsjahre	35	5	40	15,94	7,806	
Alter_2016	35	33	70	46,11	8,397	

Tabelle A.28.: Stichprobenzusammensetzung der Teilstichprobe aus nicht-gymnasialen Physiklehrkräften ($N = 30$)

Berufliche Stellung					
		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	Ref./-in	4	13,3	13,3	13,3
	Lehrkraft	26	86,7	86,7	100,0
	Gesamt	30	100,0	100,0	
Schulart					
		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
	Mittelschule	6	20,0	20,0	20,0
	Realschule	24	80,0	80,0	100,0
	Gesamt	30	100,0	100,0	
Geschlechtsangabe					
		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	weiblich	14	46,7	46,7	46,7
	männlich	16	53,3	53,3	100,0
	Gesamt	30	100,0	100,0	
Deskriptive Statistik					
	N	Minimum	Maximum	Mittelwert	Standardab- weichung
Berufsahre	30	1	40	10,73	10,773
Alter_2016	30	26	64	38,47	11,617

Tabelle A.29.: Stichprobenzusammensetzung der Teilstichprobe aus fachfremden Mittelschullehrkräften ($N = 25$)

Berufliche Stellung					
		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	Ref./-in	3	12,0	12,0	12,0
	Lehrkraft	22	88,0	88,0	100,0
	Gesamt	25	100,0	100,0	
Schulart					
		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
	Mittelschule	25	100,0	100,0	100,0
Geschlechtsangabe					
		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	weiblich	10	40,0	40,0	40,0
	männlich	15	60,0	60,0	100,0
	Gesamt	25	100,0	100,0	
Deskriptive Statistik					
N		Minimum	Maximum	Mittelwert	Standardab- weichung
Berufsahre	25	1	35	11,68	8,750
Alter_2016	25	26	62	40,92	9,768

Tabelle A.30.: Stichprobenzusammensetzung der Teilstichprobe aus fachfremden Gymnasial- und Realschullehrkräften ($N = 32$)

Berufliche Stellung						
		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente	
	Leh. fachfr.	32	100,0	100,0	100,0	
Schulart						
		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente	
	Realschule	27	84,4	84,4	84,4	
	Gymnasium	5	15,6	15,6	100,0	
	Gesamt	32	100,0	100,0		
Geschlechtsangabe						
		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente	
Gültig	weiblich	21	65,6	65,6	65,6	
	männlich	11	34,4	34,4	100,0	
	Gesamt	32	100,0	100,0		
Deskriptive Statistik						
		N	Minimum	Maximum	Mittelwert	Standardab- weichung
Berufsjahre		32	4	17	7,88	2,756
Alter_2016		32	29	44	35,22	4,338

Tabelle A.31.: Stichprobenzusammensetzung der Teilstichprobe aus Fachphysikern ($N = 29$)

Berufliche Stellung						
		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente	
	Fachw.schaft	29	100,0	100,0	100,0	
Schulart						
		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente	
	Sonstiges	29	100,0	100,0	100,0	
Geschlechtsangabe						
		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente	
Gültig	weiblich	5	17,2	17,2	17,2	
	männlich	24	82,8	82,8	100,0	
	Gesamt	29	100,0	100,0		
Deskriptive Statistik						
		N	Minimum	Maximum	Mittelwert	Standardab- weichung
Berufsjahre	25	1	35	9,04	8,580	
Alter_2016	29	22	74	32,48	12,549	

Tabelle A.32.: Stichprobenzusammensetzung der Teilstichprobe für die Kreuzvalidierung, sowohl der gesamten Stichprobe ($n = 28$) als auch für die Sub-Gruppe der Gy- ($n = 13$) und nGy-Studierenden ($n = 15$)

Deskriptive Statistik - Gy + nGy ($n_{\text{ges}} = 28$)					
	N	Minimum	Maximum	Mittelwert	St.abw.
Schulart	28	2	4	2,50	0,577
Geschlecht	28	0	1	0,64	0,488
Alter1516FAL	28	20	28	23,46	2,134
SemFALKreuz	28	1	11	6,57	2,937
Alter 2016	28	21	29	24,14	2,384
Sem. Kreuz	28	2	13	8,32	3,163
Schulart					
		Häufigkeit	Prozent	Gült. Proz.	Kum. Proz.
Gültig	Realschule	15	53,6	53,6	53,6
	Gymnasium	12	42,9	42,9	96,4
	Sonstiges	1	3,6	3,6	100,0
	Gesamt	28	100,0	100,0	
Geschlechtsangabe					
		Häufigkeit	Prozent	Gült. Proz.	Kum. Proz.
Gültig	weiblich	10	35,7	35,7	35,7
	männlich	18	64,3	64,3	100,0
	Gesamt	28	100,0	100,0	
Deskriptive Statistik - Gy ($n = 13$)					
	N	Minimum	Maximum	Mittelwert	St.abw.
Geschlecht	13	0	1	0,69	0,480
Alter1516FAL	13	20	28	24,00	2,582
SemFALKreuz	13	6	11	8,62	1,758
Alter 2016	13	21	29	24,85	2,577
Sem. Kreuz	13	9	13	10,46	1,613
Schulart					
		Häufigkeit	Prozent	Gült. Proz.	Kum. Proz.
Gültig	Gymnasium	12	92,3	92,3	92,3
	Sonstiges	1	7,7	7,7	100,0
	Gesamt	13	100,0	100,0	
Geschlechtsangabe					

Fortsetzung auf der nächsten Seite

Tabelle A.32.: (Fortsetzung von S. 158)

		Häufigkeit	Prozent	Gült. Proz.	Kum. Proz.
Gültig	weiblich	4	30,8	30,8	30,8
	männlich	9	69,2	69,2	100,0
	Gesamt	13	100,0	100,0	
Deskriptive Statistik - nGy ($n = 15$)					
	N	Minimum	Maximum	Mittelwert	St.abw.
Geschlecht	15	0	1	0,60	0,507
Alter1516FAL	15	21	27	23,00	1,604
SemFALKreuz	15	1	11	4,80	2,597
Alter 2016	15	21	29	23,53	2,100
Sem. Kreuz	15	2	13	6,47	3,021
	Sonstiges	1	7,7	7,7	100,0
	Gesamt	13	100,0	100,0	
Geschlechtsangabe					
		Häufigkeit	Prozent	Gült. Proz.	Kum. Proz.
Gültig	weiblich	6	40,0	40,0	40,0
	männlich	9	60,0	60,0	100,0
	Gesamt	15	100,0	100,0	

Tabelle A.33.: Manifeste Korrelationen (nach Pearson) der sechs Skalen der sechs Professionswissensskalen (FDW/FW, FDW, FW), Gy-Stichprobe ($n = 13$)

	FDW/- FW_{FA} (28 Items)	FDW/- FW_{Pro} (40 Items)	FDW_{FA} (10 Items)	FDW_{Pro} (20 Items)	FW_{FA} (18 Items)	FW_{Pro} (20 Items)
FDW/FW _{FA}	1,00	0,82				
FDW/FW _{Pro}	0,82	1,00				
FDW _{FA}			1,00	0,76		
FDW _{Pro}			0,76	1,00		
FW _{FA}					1,00	0,76
FW _{Pro}					0,76	1,00

Bem.: _{FA}: FALKO, _{Pro}: ProfiLe-PTabelle A.34.: Manifeste Korrelationen (nach Pearson) der sechs Skalen der sechs Professionswissensskalen (FDW/FW, FDW, FW), nGy-Stichprobe ($n = 15$)

	FDW/- FW_{FA} (28 Items)	FDW/- FW_{Pro} (40 Items)	FDW_{FA} (10 Items)	FDW_{Pro} (20 Items)	FW_{FA} (18 Items)	FW_{Pro} (20 Items)
FDW/FW _{FA}	1,00	0,60				
FDW/FW _{Pro}	0,60	1,00				
FDW _{FA}			1,00	0,59		
FDW _{Pro}			0,59	1,00		
FW _{FA}					1,00	0,46
FW _{Pro}					0,46	1,00

Bem.: _{FA}: FALKO, _{Pro}: ProfiLe-P

Tabelle A.35.: Manifeste Korrelationen (nach Pearson) der FDW Subskalen sowie der FW-EM-Subskala, gesamte Stichprobe ($n_{ges} = 28$)

	FA_{E&R} (5 Items)	Pro_{ibid} (6 Items)	FA_{Schk} (5 Items)	Pro_s (9 Items)	FA_{M&E} (4 Items)	Pro_{ex} (5 Items)	Pro_{EM} (5 Items)
FA _{E&R}	1,00	0,59					
Pro _{ibid}	0,59	1,00					
FA _{Schk}			1,00	0,45			
Pro _s			0,45	1,00			
FA _{M&E}					1,00	0,47	0,30
Pro _{ex}					0,47	1,00	1,00
Pro _{EM}					0,30	0,10	1,00

Bem.: FA: FALKO, Pro: ProfiLe-P, _{E&R}: „Erklären & Repräsentieren“, _{ibid}: „Instruktionsstrategien“, _{Schk}: „Schülerkognition“, _s: „Schülervorstellungen“, _{M&E}: „Messen & Experimentieren“, _{ex}: „Experimente“, _{EM}: „Experiment-Raum“.

Tabelle A.36.: Manifeste Korrelationen (nach Pearson) der sechs Professionswissensskalen (FDW/FW, FDW, FW), Gy-Stichprobe ($n = 13$)

	FA_{E&R} (5 Items)	Pro_{ibid} (6 Items)	FA_{Schk} (5 Items)	Pro_s (9 Items)	FA_{M&E} (4 Items)	Pro_{ex} (5 Items)	Pro_{EM} (5 Items)
FA _{E&R}	1,00	0,66					
Pro _{ibid}	0,66	1,00					
FA _{Schk}			1,00	0,57			
Pro _s			0,57	1,00			
FA _{M&E}					1,00	0,55	0,02
Pro _{ex}					0,55	1,00	-0,09
Pro _{EM}					0,02	-0,09	1,00

Bem.: FA: FALKO, Pro: ProfiLe-P, _{E&R}: „Erklären & Repräsentieren“, _{ibid}: „Instruktionsstrategien“, _{Schk}: „Schülerkognition“, _s: „Schülervorstellungen“, _{M&E}: „Messen & Experimentieren“, _{ex}: „Experimente“, _{EM}: „Experiment-Raum“.

Tabelle A.37.: Manifeste Korrelationen (nach Pearson) der sechs Professionswissensskalen (FDW/FW, FDW, FW), nGy-Stichprobe ($n = 15$)

	FA_{E&R} (5 Items)	Pro_{ibid} (6 Items)	FA_{Schk} (5 Items)	Pro_s (9 Items)	FA_{M&E} (4 Items)	Pro_{ex} (5 Items)	Pro_{EM} (5 Items)
FA _{E&R}	1,00	0,60					
Pro _{ibid}	0,60	1,00					
FA _{Schk}			1,00	0,36			
Pro _s			0,36	1,00			
FA _{M&E}					<i>1,00</i>	<i>0,48</i>	<i>0,53</i>
Pro _{ex}					<i>0,48</i>	<i>1,00</i>	<i>0,28</i>
Pro _{EM}					<i>0,53</i>	<i>0,28</i>	<i>1,00</i>

Bem.: FA: FALKO, Pro: ProfiLe-P, _{E&R}: „Erklären & Repräsentieren“, _{ibid}: „Instruktionsstrategien“, _{Schk}: „Schülerkognition“, _s: „Schülervorstellungen“, _{M&E}: „Messen & Experimentieren“, _{ex}: „Experimente“, _{EM}: „Experiment-Raum“.

Tabelle A.38.: Gruppenunterschiede Kreuzvalidierung; Betrachtung Gesamt- ($n_{ges} = 28$) und Teilstichproben (Gy: $n = 13$, RS: $n = 15$)

Unterschied zwischen den Gy- und nGy-Studierenden ($n_{ges} = 28$) im FALKO- und ProfiLeP-Test, Betrachtung Professionswissen gesamt (FDW- + FW-Skala), t-test

Gruppenstatistiken

	Kr.val.teiln.	N	Mittelwert	St.abw.	St.f. d. MW
Scorefdwfw_ fapro_proz.	TN FALKO GynGy	28	33,285	15,3628	2,9033
	TN ProfiLe-P GynGy	28	39,777	12,2382	2,3128

Unterschied zwischen den Gy- und nGy-Studierenden ($n_{ges} = 28$) im FALKO- und ProfiLeP-Test, Betrachtung Professionswissen gesamt (FDW- + FW-Skala) unter Kontrolle des Fachsemesters, ANCOVA

Kreuzvalteilnahme – geschätzte Gruppenmittelwerte

Abhängige Variable: Score_fdw_fw_falko_profilep_prozent

			95 %-K.int.	
Kr.val.teil.	Mittelwert	Standardfehler	Untergrenze	Obergrenze
TN FALKO GynGy	35,485 ^a	2,255	30,963	40,007
TN ProfiLeP GynGy	37,576 ^a	2,255	33,054	42,099

a. Die Kovariaten im Modell werden anhand der folgenden Werte berechnet:
Semester = 7,45.

Unterschied zwischen den Gy-Studierenden ($n_{ges} = 13$) im FALKO- und ProfiLeP-Test, Betrachtung Professionswissen gesamt (FDW- + FW-Skala), t-test

Gruppenstatistiken

	Kr.val.teiln.	N	Mittelwert	St.abw.	St.f. d. MW
Scorefdwfw_ fapro_proz.	TN FALKO Gy	13	40,227	13,8517	3,8418
	TN ProfiLe-P Gy	13	44,519	11,1633	3,0961

Unterschied zwischen den Gy-Studierenden ($n_{ges} = 13$) im FALKO- und ProfiLeP-Test, Betrachtung Professionswissen gesamt (FDW- + FW-Skala) unter Kontrolle des Fachsemesters, ANCOVA

Kreuzvalteilnahme – geschätzte Gruppenmittelwerte

Abhängige Variable: Score_fdw_fw_falko_profilep_prozent

			95 %-K.int.	
Kr.val.teil.	Mittelwert	Standardfehler	Untergrenze	Obergrenze

Fortsetzung auf der nächsten Seite

Tabelle A.38.: (Fortsetzung von S. 163)

TN Gy	FALKO	44,721 ^a	2,910	38,702	50,741
TN Gy	ProfiLeP	40,025 ^a	2,910	34,005	46,044

a. Die Kovariaten im Modell werden anhand der folgenden Werte berechnet:
Semester = 9,54.

Unterschied zwischen den nGy-Studierenden ($n_{\text{ges}} = 15$) im FALKO- und ProfiLeP-Test, Betrachtung Professionswissen gesamt (FDW- + FW-Skala), t-test

Gruppenstatistiken

	Kr.val.teilh.	N	Mittelwert	St.abw.	St.f. d. MW
Scorefdwfw_ fapro_proz.	TN FALKO nGy	15	27,268	14,3934	3,7164
	TN ProfiLe-P nGy	15	35,667	11,9604	3,0882

Unterschied zwischen den nGy-Studierenden ($n_{\text{ges}} = 15$) im FALKO- und ProfiLeP-Test, Betrachtung Professionswissen gesamt (FDW- + FW-Skala) unter Kontrolle des Fachsemesters, ANCOVA

Kreuzvalteilnahme – geschätzte Gruppenmittelwerte

Abhängige Variable: Score_fdw_fw_falko_profilep_prozent

			95 %-K.int.		
Kr.val.teilh.	Mittelwert	Standardfehler	Untergrenze	Obergrenze	
TN FALKO nGy	28,549 ^a	3,364	21,647	35,452	
TN ProfiLeP Gy	34,385 ^a	3,364	27,483	41,288	

a. Die Kovariaten im Modell werden anhand der folgenden Werte berechnet:
Semester = 5,63.

Unterschied zwischen den Gy- und nGy-Studierenden ($n_{\text{ges}} = 28$) im FALKO- und ProfiLeP-Test, Betrachtung FDW-Skala (FALKO: Ohne Subfacette „M & E“), t-test

Gruppenstatistiken

	Kr.val.teilh.	N	Mittelwert	St.abw.	St.f. d. MW
Scorefdw_fa pro_proz.	TN FALKO GynGy	28	27,0186	12,37453	2,33857
	TN ProfiLe-P GynGy	28	50,3348	17,31256	3,27177

Unterschied zwischen den Gy und nGy-Studierenden ($n_{\text{ges}} = 28$) im FALKO- und ProfiLeP-Test, Betrachtung FDW-Skala

Fortsetzung auf der nächsten Seite

Tabelle A.38.: (Fortsetzung von S. 163)

**(FALKO: Ohne Subfacette „M & E“) unter Kontrolle des Fachsemesters, ANCOVA
Kreuzvalteilnahme – geschätzte Gruppenmittelwerte**

Abhängige Variable: Score_fdw_falko_profilep_prozent

Kr.val.teil.		Mittelwert	Standardfehler	95 %-K.int.	
				Untergrenze	Obergrenze
TN FALKO	GynGy	29,120 ^a	2,560	23,985	34,254
TN ProfiLeP	GynGy	48,234 ^a	2,560	43,099	53,368

a. Die Kovariaten im Modell werden anhand der folgenden Werte berechnet:
Semester = 7,45.

**Unterschied zwischen den Gy-Studierenden ($n_{ges} = 13$) im
FALKO- und ProfiLeP-Test, Betrachtung FDW-Skala
(FALKO: Ohne Subfacette „M & E“), t-test**

Gruppenstatistiken

	Kr.val.teiln.	N	Mittelwert	St.abw.	St.f. d. MW
Scorefdw_fa pro_proz.	TN FALKO Gy	13	31,7726	9,94386	2,75793
	TN ProfiLe-P Gy	13	55,5288	16,83608	4,66949

**Unterschied zwischen den Gy-Studierenden ($n_{ges} = 13$) im
FALKO- und ProfiLeP-Test, Betrachtung FDW-Skala
(FALKO: Ohne Subfacette „M & E“) unter Kontrolle des Fachsemesters, ANCOVA**

Kreuzvalteilnahme – geschätzte Gruppenmittelwerte

Abhängige Variable: Score_fdw_falko_profilep_prozent

Kr.val.teil.		Mittelwert	Standardfehler	95 %-K.int.	
				Untergrenze	Obergrenze
TN FALKO	Gy	35,496 ^a	3,676	27,892	43,100
TN ProfiLeP	Gy	51,805 ^a	3,676	44,201	59,409

a. Die Kovariaten im Modell werden anhand der folgenden Werte berechnet:
Semester = 9,54.

**Unterschied zwischen den nGy-Studierenden ($n_{ges} = 15$) im
FALKO- und ProfiLeP-Test, Betrachtung FDW-Skala
(FALKO: Ohne Subfacette „M & E“), t-test**

Gruppenstatistiken

	Kr.val.teiln.	N	Mittelwert	St.abw.	St.f. d. MW
--	---------------	---	------------	---------	-------------

Fortsetzung auf der nächsten Seite

Tabelle A.38.: (Fortsetzung von S. 163)

Scorefdw_fa pro_proz.	TN nGy	FALKO 15	22,8986	13,09170	3,38026
	TN nGy	ProfiLe-P 15	45,8333	16,97994	4,38420

Unterschied zwischen den nGy-Studierenden ($n_{ges} = 15$) im FALKO- und ProfiLeP-Test, Betrachtung FDW-Skala (FALKO: Ohne Subfacette „M & E“) unter Kontrolle des Fachsemesters, ANCOVA

Kreuzvalteilnahme – geschätzte Gruppenmittelwerte

Abhängige Variable: Score_fdw_falko_profilep_prozent

			95 %-K.int.		
Kr.val.teil.	Mittelwert	Standardfehler	Untergrenze	Obergrenze	
TN FALKO nGy	24,499 ^a	3,810	16,681	32,317	
TN ProfiLeP nGy	44,233 ^a	3,810	36,415	52,051	

a. Die Kovariaten im Modell werden anhand der folgenden Werte berechnet:
Semester = 5,63.

Unterschied zwischen den Gy- und nGy-Studierenden ($n_{ges} = 28$) im FALKO- und ProfiLeP-Test, Betrachtung FDW-Subskala „E & R“, (bzw. „ib/id“), t-test

Gruppenstatistiken

	Kr.val.teiln.	N	Mittelwert	St.abw.	St.f. d. MW
ScofdwER_fa- pro_proz.	TN FALKO GynGy	28	23,9011	13,13116	2,48156
	TN ProfiLe-P GynGy	28	35,3896	17,90005	3,38279

Unterschied zwischen den Gy und nGy-Studierenden ($n_{ges} = 28$) im FALKO- und ProfiLeP-Test, Betrachtung FDW-Subskala „E & R“ (bzw. „ib/id“) unter Kontrolle des Fachsemesters, ANCOVA

Kreuzvalteilnahme – geschätzte Gruppenmittelwerte

Abhängige Variable: Score_fdw_ER_falko_profilep_prozent

			95 %-K.int.		
Kr.val.teil.	Mittelwert	Standardfehler	Untergrenze	Obergrenze	
TN FALKO GynGy	25,790 ^a	2,775	20,224	31,357	
TN ProfiLeP GynGy	33,501 ^a	2,775	27,934	39,067	

a. Die Kovariaten im Modell werden anhand der folgenden Werte berechnet:

Fortsetzung auf der nächsten Seite

Tabelle A.38.: (Fortsetzung von S. 163)

Semester = 7,45.

Unterschied zwischen den Gy-Studierenden ($n_{ges} = 13$) im FALKO- und ProfiLeP-Test, Betrachtung FDW-Subskala „E & R“, (bzw. „ib/id“), t-test

Gruppenstatistiken

	Kr.val.teiln.	N	Mittelwert	St.abw.	St.f. d. MW
ScofdwER_fa- pro_proz.	TN FALKO Gy	13	30,1775	9,13494	2,53358
	TN ProfiLe-P Gy	13	37,7622	19,58458	5,43179

Unterschied zwischen den Gy-Studierenden ($n_{ges} = 13$) im FALKO- und ProfiLeP-Test, Betrachtung FDW-Subskala „E & R“, (bzw. „ib/id“) unter Kontrolle des Fachsemesters, ANCOVA

Kreuzvalteilnahme – geschätzte Gruppenmittelwerte

Abhängige Variable: Score_fdw_ER_falko_profilep_prozent

			95 %-K.int.		
Kr.val.teil.		Mittelwert	Standardfehler	Untergrenze	Obergrenze
TN Gy	FALKO	33,247 ^a	4,341	24,266	42,227
TN Gy	ProfiLeP	34,693 ^a	4,341	25,712	43,674

a. Die Kovariaten im Modell werden anhand der folgenden Werte berechnet:
Semester = 9,54.

Unterschied zwischen den nGy-Studierenden ($n_{ges} = 15$) im FALKO- und ProfiLeP-Test, Betrachtung FDW-Subskala „E & R“, (bzw. „ib/id“), t-test

Gruppenstatistiken

	Kr.val.teiln.	N	Mittelwert	St.abw.	St.f. d. MW
ScofdwER_fa- pro_proz.	TN FALKO nGy	15	18,4615	13,88274	3,58451
	TN ProfiLe-P nGy	15	33,3333	16,71579	4,31600

Unterschied zwischen den nGy-Studierenden ($n_{ges} = 15$) im FALKO- und ProfiLeP-Test, Betrachtung FDW-Subskala „E & R“, (bzw. „ib/id“) unter Kontrolle des Fachsemesters, ANCOVA

Kreuzvalteilnahme – geschätzte Gruppenmittelwerte

Abhängige Variable: Score_fdw_ER_falko_profilep_prozent

			95 %-K.int.		
--	--	--	-------------	--	--

Fortsetzung auf der nächsten Seite

Tabelle A.38.: (Fortsetzung von S. 163)

Kr.val.teil.	Mittelwert	Standardfehler	Untergrenze	Obergrenze
TN FALKO nGy	20,053 ^a	3,872	12,108	27,998
TN ProfiLeP nGy	31,742 ^a	3,872	23,797	39,687

a. Die Kovariaten im Modell werden anhand der folgenden Werte berechnet:
Semester = 5,63.

Unterschied zwischen den Gy- und nGy-Studierenden ($n_{\text{ges}} = 28$) im FALKO- und ProfiLeP-Test, Betrachtung FDW-Subskala „Schk“, (bzw. „s“), t-test

Gruppenstatistiken

	Kr.val.teiln.	N	Mittelwert	St.abw.	St.f. d. MW
ScofdwSchk- _fapro_proz.	TN FALKO GynGy	28	31,0714	17,91736	3,38606
	TN ProfiLe-P GynGy	28	57,6190	21,63782	4,08916

Unterschied zwischen den Gy und nGy-Studierenden ($n_{\text{ges}} = 28$) im FALKO- und ProfiLeP-Test, Betrachtung FDW-Subskala „Schk“, (bzw. „s“) unter Kontrolle des Fachsemesters, ANCOVA

Kreuzvalteilnahme – geschätzte Gruppenmittelwerte

Abhängige Variable: Score_fdw_Schk_falko_profilep_prozent

Kr.val.teil.	Mittelwert	Standardfehler	Untergrenze	Obergrenze
TN FALKO GynGy	33,323 ^a	3,554	26,195	40,452
TN ProfiLeP GynGy	55,367 ^a	3,554	48,239	62,496

a. Die Kovariaten im Modell werden anhand der folgenden Werte berechnet:
Semester = 7,45.

Unterschied zwischen den Gy-Studierenden ($n_{\text{ges}} = 13$) im FALKO- und ProfiLeP-Test, Betrachtung FDW-Subskala „Schk“, (bzw. „s“), t-test

Gruppenstatistiken

	Kr.val.teiln.	N	Mittelwert	St.abw.	St.f. d. MW
ScofdwSchk- _fapro_proz.	TN FALKO Gy	13	33,8462	13,86750	3,84615
	TN ProfiLe-P Gy	13	65,6410	20,15609	5,59029

Unterschied zwischen den Gy-Studierenden ($n_{\text{ges}} = 13$) im

Fortsetzung auf der nächsten Seite

Tabelle A.38.: (Fortsetzung von S. 163)

FALKO- und ProfiLeP-Test, Betrachtung FDW-Subskala „Schk“, (bzw. „s“) unter Kontrolle des Fachsemesters, ANCOVA

Kreuzvalteilnahme – geschätzte Gruppenmittelwerte

Abhängige Variable: Score_fdw_Schk_falko_profilep_prozent

			95 %-K.int.	
Kr.val.teil.	Mittelwert	Standardfehler	Untergrenze	Obergrenze
TN FALKO Gy	37,703 ^a	4,825	27,721	47,685
TN ProfiLeP Gy	61,784 ^a	4,825	51,802	71,767

a. Die Kovariaten im Modell werden anhand der folgenden Werte berechnet:
Semester = 9,54.

Unterschied zwischen den nGy-Studierenden ($n_{ges} = 15$) im FALKO- und ProfiLeP-Test, Betrachtung FDW-Subskala „Schk“, (bzw. „s“), t-test

Gruppenstatistiken

	Kr.val.teiln.	N	Mittelwert	St.abw.	St.f. d. MW
ScofdwSchk- _fapro_proz.	TN FALKO nGy	15	28,6667	20,99887	5,42188
	TN ProfiLe-P nGy	15	50,6667	21,05171	5,43553

Unterschied zwischen den nGy-Studierenden ($n_{ges} = 15$) im FALKO- und ProfiLeP-Test, Betrachtung FDW-Subskala „Schk“, (bzw. „s“) unter Kontrolle des Fachsemesters, ANCOVA

Kreuzvalteilnahme – geschätzte Gruppenmittelwerte

Abhängige Variable: Score_fdw_Schk_falko_profilep_prozent

			95 %-K.int.	
Kr.val.teil.	Mittelwert	Standardfehler	Untergrenze	Obergrenze
TN FALKO nGy	30,372 ^a	5,440	19,211	41,534
TN ProfiLeP nGy	48,961 ^a	5,440	37,800	60,122

a. Die Kovariaten im Modell werden anhand der folgenden Werte berechnet:
Semester = 5,63.

Unterschied zwischen den Gy- und nGy-Studierenden ($n_{ges} = 28$) im FALKO- und ProfiLeP-Test, Betrachtung FDW-Subskala „M & E“, (bzw. „ex“), t-test

Gruppenstatistiken

Fortsetzung auf der nächsten Seite

Tabelle A.38.: (Fortsetzung von S. 163)

	Kr.val.teiln.	N	Mittelwert	St.abw.	St.f. d. MW
Scorefdw_ME_fapro_proz.	TN FALKO GynGy	28	35,7143	19,21104	3,63054
	TN ProfiLe-P GynGy	28	22,3214	8,74480	1,65261

Unterschied zwischen den Gy und nGy-Studierenden ($n_{\text{ges}} = 28$) im FALKO- und ProfiLeP-Test, Betrachtung FDW-Subskala „M & E“, (bzw. „ex“) unter Kontrolle des Fachsemesters, ANCOVA

Kreuzvalteilnahme – geschätzte Gruppenmittelwerte

Abhängige Variable: Score_fdw_ME_falko_profilep_prozent

			95 %-K.int.		
Kr.val.teil.	Mittelwert	Standardfehler	Untergrenze	Obergrenze	
TN FALKO GynGy	36,740 ^a	2,822	31,079	42,401	
TN ProfiLeP GynGy	21,296 ^a	2,822	15,635	26,957	

a. Die Kovariaten im Modell werden anhand der folgenden Werte berechnet:
Semester = 7,45.

Unterschied zwischen den Gy-Studierenden ($n_{\text{ges}} = 13$) im FALKO- und ProfiLeP-Test, Betrachtung FDW-Subskala „M & E“, (bzw. „ex“), t-test

Gruppenstatistiken

	Kr.val.teiln.	N	Mittelwert	St.abw.	St.f. d. MW
ScofdwME_fa-pro_proz.	TN FALKO Gy	13	42,7350	14,93894	4,14332
	TN ProfiLe-P Gy	13	23,5577	8,13055	2,25501

Unterschied zwischen den Gy-Studierenden ($n_{\text{ges}} = 13$) im FALKO- und ProfiLeP-Test, Betrachtung FDW-Subskala „M & E“, (bzw. „ex“) unter Kontrolle des Fachsemesters, ANCOVA

Kreuzvalteilnahme – geschätzte Gruppenmittelwerte

Abhängige Variable: Score_fdw_ME_falko_profilep_prozent

			95 %-K.int.		
Kr.val.teil.	Mittelwert	Standardfehler	Untergrenze	Obergrenze	
TN FALKO Gy	46,659 ^a	2,949	40,559	52,759	
TN ProfiLeP Gy	19,634 ^a	2,949	13,534	25,733	

Fortsetzung auf der nächsten Seite

Tabelle A.38.: (Fortsetzung von S. 163)

a. Die Kovariaten im Modell werden anhand der folgenden Werte berechnet:
Semester = 9,54.

Unterschied zwischen den nGy-Studierenden ($n_{\text{ges}} = 15$) im FALKO- und ProfiLeP-Test, Betrachtung FDW-Subskala „M & E“, (bzw. „ex“), t-test

Gruppenstatistiken

	Kr.val.teiln.	N	Mittelwert	St.abw.	St.f. d. MW
ScofdwME_fa- TN FALKO	15	29,6269	20,85757	5,38540	
pro_proz. nGy					
TN ProfiLe-P	15	21,2500	9,38987	2,42445	
nGy					

Unterschied zwischen den nGy-Studierenden ($n_{\text{ges}} = 15$) im FALKO- und ProfiLeP-Test, Betrachtung FDW-Subskala „M & E“, (bzw. „ex“) unter Kontrolle des Fachsemesters, ANCOVA

Kreuzvalteilnahme – geschätzte Gruppenmittelwerte

Abhängige Variable: Score_fdw_ME_falko_profilep_prozent

			95 %-K.int.		
Kr.val.teil.	Mittelwert	Standardfehler	Untergrenze	Obergrenze	
TN FALKO	29,204 ^a	4,334	20,312	38,097	
nGy					
TN ProfiLeP	21,675 ^a	4,334	12,782	30,568	
nGy					

a. Die Kovariaten im Modell werden anhand der folgenden Werte berechnet:
Semester = 5,63.

Unterschied zwischen den Gy- und nGy-Studierenden ($n_{\text{ges}} = 28$) im FALKO- und ProfiLeP-Test, Betrachtung FW-Skala

Gruppenstatistiken

	Kr.val.teiln.	N	Mittelwert	St.abw.	St.f. d. MW
Scorefw_fapro TN FALKO	28	37,5000	19,82342	3,74627	
_proz. GynGy					
TN ProfiLe-P	28	58,2011	19,64624	3,71279	
GynGy					

Unterschied zwischen den Gy und nGy-Studierenden ($n_{\text{ges}} = 28$) im FALKO- und ProfiLeP-Test, Betrachtung FW-Skala unter Kontrolle des Fachsemesters, ANCOVA

Kreuzvalteilnahme – geschätzte Gruppenmittelwerte

Abhängige Variable: Score_fw_falko_profilep_prozent

			95 %-K.int.		
--	--	--	-------------	--	--

Fortsetzung auf der nächsten Seite

Tabelle A.38.: (Fortsetzung von S. 163)

Kr.val.teil.	Mittelwert	Standardfehler	Untergrenze	Obergrenze
TN FALKO GynGy	40,561 ^a	3,233	34,077	47,045
TN ProfiLeP GynGy	55,140 ^a	3,233	48,656	61,624

a. Die Kovariaten im Modell werden anhand der folgenden Werte berechnet:
Semester = 7,45.

Unterschied zwischen den Gy-Studierenden ($n_{\text{ges}} = 13$) im FALKO- und ProfiLeP-Test, Betrachtung FW-Skala

Gruppenstatistiken

	Kr.val.teiln.	N	Mittelwert	St.abw.	St.f. d. MW
Scorefw_fapro _proz.	TN FALKO Gy	13	46,1538	20,41428	5,66190
	TN ProfiLe-P Gy	13	66,0969	17,81671	4,94147

Unterschied zwischen den Gy-Studierenden ($n_{\text{ges}} = 13$) im FALKO- und ProfiLeP-Test, Betrachtung FW-Skala unter Kontrolle des Fachsemesters, ANCOVA

Kreuzvalteilnahme – geschätzte Gruppenmittelwerte

Abhängige Variable: Score_fw_falko_profilep_prozent

				95 %-K.int.
Kr.val.teil.	Mittelwert	Standardfehler	Untergrenze	Obergrenze
TN FALKO Gy	52,600 ^a	4,615	43,054	62,146
TN ProfiLeP Gy	59,650 ^a	4,615	50,104	69,197

a. Die Kovariaten im Modell werden anhand der folgenden Werte berechnet:
Semester = 9,54.

Unterschied zwischen den nGy-Studierenden ($n_{\text{ges}} = 15$) im FALKO- und ProfiLeP-Test, Betrachtung FW-Skala

Gruppenstatistiken

	Kr.val.teiln.	N	Mittelwert	St.abw.	St.f. d. MW
Scorefw_fapro _proz.	TN FALKO nGy	15	29,6269	20,85757	5,38540
	TN ProfiLe-P nGy	15	30,0000	16,45755	4,24932

Unterschied zwischen den nGy-Studierenden ($n_{\text{ges}} = 15$) im FALKO- und ProfiLeP-Test, Betrachtung FW-Skala

Fortsetzung auf der nächsten Seite

Tabelle A.38.: (Fortsetzung von S. 163)

unter Kontrolle des Fachsemesters, ANCOVA

Kreuzvalteilnahme – geschätzte Gruppenmittelwerte

Abhängige Variable: Score_fw_falko_profilep_prozent

Kr.val.teil.		Mittelwert	Standardfehler	95 %-K.int.	
				Untergrenze	Obergrenze
TN FALKO nGy		31,664 ^a	4,549	22,330	40,998
TN ProfiLeP nGy		49,694 ^a	4,549	40,360	59,028

a. Die Kovariaten im Modell werden anhand der folgenden Werte berechnet:
Semester = 5,63.

Unterschied zwischen den Gy- und nGy-Studierenden ($n_{ges} = 28$) im FALKO- und ProfiLeP-Test, Betrachtung FDW-Subskala „M & E“, (bzw. „EM“), t-test

Gruppenstatistiken

	Kr.val.teiln.	N	Mittelwert	St.abw.	St.f. d. MW
ScofdwME- EM_fapro_- proz.	TN FALKO GynGy	28	35,7143	19,21104	3,63054
	TN ProfiLe-P GynGy	28	37,1429	28,13422	5,31687

Unterschied zwischen den Gy und nGy-Studierenden ($n_{ges} = 28$) im FALKO- und ProfiLeP-Test, Betrachtung FDW-Subskala „M & E“, (bzw. „EM“) unter Kontrolle des Fachsemesters, ANCOVA

Kreuzvalteilnahme – geschätzte Gruppenmittelwerte

Abhängige Variable: Score_fdw_ME_EM_falko_profilep_prozent

Kr.val.teil.		Mittelwert	Standardfehler	95 %-K.int.	
				Untergrenze	Obergrenze
TN FALKO GynGy		38,202 ^a	4,377	29,422	46,981
TN ProfiLeP GynGy		34,655 ^a	4,377	25,876	43,435

a. Die Kovariaten im Modell werden anhand der folgenden Werte berechnet:
Semester = 7,45.

Unterschied zwischen den Gy-Studierenden ($n_{ges} = 13$) im FALKO- und ProfiLeP-Test, Betrachtung FDW-Subskala „M & E“, (bzw. „EM“), t-test

Gruppenstatistiken

Fortsetzung auf der nächsten Seite

Tabelle A.38.: (Fortsetzung von S. 163)

	Kr.val.teiln.	N	Mittelwert	St.abw.	St.f. d. MW
ScofdwME- EM_fapro_- proz.	TN FALKO Gy	13	42,7350	14,93894	4,14332
	TN ProfiLe-P Gy	13	44,6154	31,78453	8,81544

Unterschied zwischen den Gy-Studierenden ($n_{\text{ges}} = 13$) im FALKO- und ProfiLeP-Test, Betrachtung FDW-Subskala „M & E“, (bzw. „EM“) unter Kontrolle des Fachsemesters, ANCOVA

Kreuzvalteilnahme – geschätzte Gruppenmittelwerte

Abhängige Variable: Score_fdw_ME_EM_falko_profilep_prozent

			95 %-K.int.	
Kr.val.teil.	Mittelwert	Standardfehler	Untergrenze	Obergrenze
TN FALKO Gy	47,621 ^a	7,078	32,980	62,262
TN ProfiLeP Gy	39,730 ^a	7,078	25,089	54,371

a. Die Kovariaten im Modell werden anhand der folgenden Werte berechnet:
Semester = 9,54.

Unterschied zwischen den nGy-Studierenden ($n_{\text{ges}} = 15$) im FALKO- und ProfiLeP-Test, Betrachtung FDW-Subskala „M & E“, (bzw. „EM“), t-test

Gruppenstatistiken

	Kr.val.teiln.	N	Mittelwert	St.abw.	St.f. d. MW
ScofdwME- EM_fapro_- proz.	TN FALKO nGy	15	29,6269	20,85757	5,38540
	TN ProfiLe-P nGy	15	30,6667	23,74467	6,13085

Unterschied zwischen den nGy-Studierenden ($n_{\text{ges}} = 15$) im FALKO- und ProfiLeP-Test, Betrachtung FDW-Subskala „M & E“, (bzw. „EM“) unter Kontrolle des Fachsemesters, ANCOVA

Kreuzvalteilnahme – geschätzte Gruppenmittelwerte

Abhängige Variable: Score_fdw_ME_EM_falko_profilep_prozent

			95 %-K.int.	
Kr.val.teil.	Mittelwert	Standardfehler	Untergrenze	Obergrenze
TN FALKO nGy	30,875 ^a	5,905	18,760	42,990

Fortsetzung auf der nächsten Seite

Tabelle A.38.: (Fortsetzung von S. 163)

TN	ProfiLeP	29,421 ^a	5,905	17,306	41,536
----	----------	---------------------	-------	--------	--------

nGy

a. Die Kovariaten im Modell werden anhand der folgenden Werte berechnet:
Semester = 5,63.

Tabelle A.39.: Output der konfirmatorischen Faktorenanalyse des Fach- und fachdidaktischen Wissens (mit „M & E“; zweifaktorielle Modellierung, Hauptstichprobe N = 183)

<i>Notes for Model (Default model), Computation of degrees of freedom (Def. m.)</i>							
Number of distinct sample moments:						21	
Number of distinct parameters to be estimated:						13	
Degrees of freedom (21 - 13):						8	
Min. was achieved, Chi-square = 11,695, Degr. of free. = 8, Prob. level = 0,165							
<i>Estimates (Gr. n. 1 - Def. m.), Scalar Estim. (Gr. n. 1 - Def. m.), Max. Like. Estim.</i>							
<i>Regression Weights: (Group number 1 - Default model)</i>							
			Estimate	S.E.	C.R.	P	Label
fdw m e	<—	FDW	0,313	0,035	8,994	***	
fdw schk	<—	FDW	0,189	0,027	7,069	***	
fdw e r	<—	FDW	0,256	0,031	8,220	***	
fw par1	<—	FW	0,324	0,025	13,066	***	
fw par2	<—	FW	0,285	0,027	10,658	***	
fw par3	<—	FW	0,298	0,024	12,505	***	
<i>Standardized Regression Weights: (Group number 1 - Default model)</i>							
			Estimate				
fdw m e	<—	FDW	0,660				
fdw schk	<—	FDW	0,532				
fdw e r	<—	FDW	0,609				
fw par1	<—	FW	0,836				
fw par2	<—	FW	0,722				
fw par3	<—	FW	0,810				
<i>Covariances: (Group number 1 - Default model)</i>							
			Estimate	S.E.	C.R.	P	Label
FDW	<->	FDW	0,974	0,048	20,363	***	
<i>Model Fit Summary, Baseline Comparisons</i>							
Model			NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Default Model			0,972	0,947	0,991	0,983	0,991
Saturated Model			1,000		1,000		1,000
Independence Model			0,000	0,000	0,000	0,000	0,00
<i>RMSEA</i>							

Fortsetzung auf der nächsten Seite

Tabelle A.39.: (Fortsetzung)

Model	RMSEA	LO 90	HI 90	PCLOSE
Default Model	0,050	0,000	0,108	0,436
Independence Model	0,384	0,352	0,416	0,000

Tabelle A.40.: Output der konfirmatorischen Faktorenanalyse des Fach- und fachdidaktischen Wissens (ohne „M & E“) (zweifaktorielle Modellierung auf Basis der Hauptstichprobe N = 183)

lavaan (0.5-23.1097) converged normally after 17 iterations				
Number of observations				183
Estimator				<i>ML</i>
Minimum Function Test Statistic				4,977
Degrees of freedom				4
P-value (Chi-square)				0,290
Model test baseline model:				
Minimum Function Test Statistic				333,072
Degrees of freedom				10
P-value				0,000
User model versus baseline model:				
Comparative Fit Index (CFI)				0,997
Tucker-Lewis Index (TLI)				0,992
Loglikelihood and Information Criteria:				
Loglikelihood user model (H0)				-1131,774
Loglikelihood unrestricted model (H1)				-1129,286
Number of free parameters				11
Akaike (AIC)				2285,549
Bayesian (BIC)				2320,853
Sample-size adjusted Bayesian (BIC)				2286,14
Root Mean Square Error of Approximation:				
RMSEA				0,037
90 Percent Confidence Interval				0,000
				0,112
P-value RMSEA $\leq 0,05$				0,503
Standardized Root Mean Square Residual:				
SRMR				0,02
Parameter Estimates:				
Information				<i>Expected</i>
Standard Errors				<i>Standard</i>
Latent Variables:				
	Estimate	Std.Err	z-value	P(> z)
fdw =~				
N_fd_e_ges	0,69	0,082	8,433	0,000
N_fd_s_ges	0,586	0,08	7,364	0,000
fw =~				
N_fw_p1	0,793	0,066	12,01	0,000
N_fw_p2	0,728	0,068	10,731	0,000
N_fw_p3	0,841	0,065	13,015	0,000
Covariances:				

	Estimate	Std.Err	z-value	P(> z)
fdw ~~				
fw	0,855	0,069	12,459	0,000
Variances:				
	Estimate	Std.Err	z-value	P(> z)
.N_fd_e_ges	0,519	0,088	5,878	0,000
.N_fd_s_ges	0,651	0,085	7,688	0,000
.N_fw_p1	0,366	0,055	6,606	0,000
.N_fw_p2	0,464	0,061	7,662	0,000
.N_fw_p3	0,287	0,053	5,412	0,000
fdw	1,000			
fw	1,000			

Tabelle A.41.: Output der CFA (Bayes-Modellierung) ohne „M & E“ (zweifaktorielle Modellierung auf Basis der Hauptstichprobe, $N = 183$)

two.f.cfa.b <- bcfa(model2, data, std.ov=T, std.lv = T)

Number of observations						183	
Number of missing patterns						1	
Statistic					MargLogLik		PPP
Value					-222,546		0,464
Latent Variables:							
	Estimate	Post.SD	HPD.025	HPD.975	PSRF		Prior
fdw= \sim							
Score_FDW_ER	0,703	0,233	0,251	1,166	1,002		dnorm(0,1e-2)T(0,)
Score_FDW_Schk	0,711	0,231	0,272	1,184	1,001		dno(0,1e-2)
fw= \sim							
Score_FW_Prcl1	0,952	0,183	0,606	1,322	1,002		dnorm(0,1e-2)T(0,)
Score_FW_Prcl2	0,883	0,194	0,527	1,281	1,001		dno(0,1e-2)
Score_FW_Prcl3	0,920	0,188	0,569	1,298	1,002		dno(0,1e-2)
Covariances:							
	Estimate	Post.SD	HPD.025	HPD.975	PSRF		Prior
fdw $\sim\sim$ fw	0,858	0,119	0,619	1	1,002		dbeta(1,1)
Intercepts:							
	Estimate	Post.SD	HPD.025	HPD.975	PSRF		Prior
Score_FDW_ER	-0,008	0,206	-0,42	0,379	1,001		dno(0,1e-3)
Score_FDW_Schk	-0,009	0,204	-0,429	0,378	1,000		dno(0,1e-3)
Score_FW_Prcl1	-0,011	0,207	-0,43	0,376	1,000		dno(0,1e-3)
Score_FW_Prcl2	-0,011	0,206	-0,419	0,399	1,000		dno(0,1e-3)
Score_FW_Prcl3	-0,010	0,204	-0,426	0,372	1,000		dno(0,1e-3)
fdw	0,000	fw	0,000				
Variances:							
	Estimate	Post.SD	HPD.025	HPD.975	PSRF		Prior
Score_FDW_ER	0,628	0,241	0,189	1,13	1,011		dgamma(1,5)
Score_FDW_Schk	0,615	0,235	0,193	1,1	1,001		dgamma(1,5)
Score_FW_Prcl1	0,280	0,114	0,089	0,499	1,001		dgamma(1,5)
Score_FW_Prcl2	0,390	0,144	0,15	0,674	1,022		dgamma(1,5)
Score_FW_Prcl3	0,328	0,129	0,119	0,586	1,001		dgamma(1,5)
fdw	1,000	fw	1,000				

Tabelle A.42.: Output der konfirmatorischen Faktorenanalyse des Fach- und fachdidaktischen Wissens (ohne „M & E“) (einfaktorielle Modellierung auf Basis der Studierendenstichprobe der Hauptstudie, $N = 118$)

lavaan (0.5-23.1097), cfa(modell1, data, std.ov=T, std.lv=t)				
Number of observations				118
Estimator				ML
Minimum Function Test Statistic				12,015
Degrees of freedom				5
P-value (Chi-square)				0,035
Model test baseline model:				
Minimum Function Test Statistic				173,268
Degrees of freedom				10
P-value				0,000
User model versus baseline model:				
Comparative Fit Index (CFI)				0,975
Tucker-Lewis Index (TLI)				0,914
Loglikelihood and Information Criteria:				
Loglikelihood user model (H0)				-754,037
Loglikelihood unrestricted model (H1)				-748,029
Number of free parameters				10
Akaike (AIC)				1528,073
Bayesian (BIC)				1555,780
Sample-size adjusted Bayesian (BIC)				1524,168
Root Mean Square Error of Approximation:				
RMSEA				0,109
90 Percent Confidence Interval				0,027
				0,190
P-value RMSEA $\leq 0,05$				0,095
Standardized Root Mean Square Residual:				
SRMR				0,046
Parameter Estimates:				
Information				Expected
Standard Errors				Standard
Latent Variables:				
	Estimate	Std.Err	z-value	P(> z)
fdw =~				
N_fd_e_ges	0,598	0,092	6,529	0,000
N_fd_s_ges	0,612	0,091	6,718	0,000
fw =~				
N_fw_p1	0,751	0,087	8,648	0,000
N_fw_p2	0,544	0,093	5,835	0,000

Fortsetzung auf der nächsten Seite

Tabelle A.42.: (Fortsetzung von S. 181)

	Estimate	Std.Err	z-value	P(> z)
N_fw_p3	0,781	0,086	9,089	0,000
Covariances:				
fdw ~ ~				
fw	1,000			
Variances:				
.N_fd_e_ges	0,634	0,095	6,694	0,000
.N_fd_s_ges	0,617	0,093	6,615	0,000
.N_fw_p1	0,428	0,081	5,286	0,000
.N_fw_p2	0,695	0,100	6,942	0,000
.N_fw_p3	0,381	0,079	4,802	0,000
fdw	1,000			
fw	1,000			

Tabelle A.43.: Output der konfirmatorischen Faktorenanalyse des Fach- und fachdidaktischen Wissens (ohne „M & E“) (zweifaktorielle Modellierung auf Basis der Studierendenstichprobe der Hauptstudie, $N = 118$)

lavaan (0.5-23.1097), cfa(model2, data, std.ov=T, std.lv=t)				
Number of observations				118
Estimator				ML
Minimum Function Test Statistic				11,620
Degrees of freedom				4
P-value (Chi-square)				0,020
Model test baseline model:				
Minimum Function Test Statistic				173,268
Degrees of freedom				10
P-value				0,000
User model versus baseline model:				
Comparative Fit Index (CFI)				0,953
Tucker-Lewis Index (TLI)				0,883
Loglikelihood and Information Criteria:				
Loglikelihood user model (H0)				-753,839
Loglikelihood unrestricted model (H1)				-748,029
Number of free parameters				11
Akaike (AIC)				1529,678
Bayesian (BIC)				1560,156
Sample-size adjusted Bayesian (BIC)				1525,382
Root Mean Square Error of Approximation:				
RMSEA				0,127
90 Percent Confidence Interval				0,045
				0,215
P-value RMSEA $\leq 0,05$				0,058
Standardized Root Mean Square Residual:				
SRMR				0,045
Parameter Estimates:				
Information				Expected
Standard Errors				Standard
Latent Variables:				
	Estimate	Std.Err	z-value	P(> z)
fdw =~				
N_fd_e_ges	0,622	0,099	6,305	0,000
N_fd_s_ges	0,638	0,099	6,445	0,000
fw =~				
N_fw_p1	0,754	0,087	8,641	0,000
N_fw_p2	0,550	0,093	5,892	0,000

Fortsetzung auf der nächsten Seite

Tabelle A.43.: (Fortsetzung von S. 183)

N_fw_p3	0,783	0,087	9,038	0,000
Covariances:				
	Estimate	Std.Err	z-value	P(> z)
fdw ~ ~				
fw	0,943	0,088	10,708	0,000
Variances:				
	Estimate	Std.Err	z-value	P(> z)
.N_fd_e_ges	0,605	0,104	5,827	0,000
.N_fd_s_ges	0,585	0,104	5,609	0,000
.N_fw_p1	0,423	0,082	5,158	0,000
.N_fw_p2	0,689	0,100	6,903	0,000
.N_fw_p3	0,379	0,081	4,669	0,000
fdw	1,000			
fw	1,000			
ANOVA (Chi-Quadr.-Diff.test) ein- und zweifakt. Modell, $N = 118$, ohne „M & E“				
Chi Square Difference Test		Chisq diff	DF diff	Pr(>-Chisq)
		0,39531	1	0,5295

Tabelle A.44.: Output der CFA (Bayes-Modellierung) ohne „M & E“ (zweifaktorielle Modellierung auf Basis der Gy+nGy-Studierendenstichprobe, $N = 118$)

two.f.cfa.b <- bcfa(model2, data, std.ov=T, std.lv = T)

Number of observations	118	Number of missing patterns	1
Statistic		MargLogLik	PPP
Value		-823,314	0,192

Latent Variables:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
Score_FDW_ER	0,642	0,104	0,442	0,851	1,002	dnorm(0,1e-2)T(0,)
Score_FDW_Schk	0,667	0,105	0,462	0,874	1,002	dno(0,1e-2)
Score_FW_Prcl1	0,777	0,093	0,596	0,963	1,003	dnorm(0,1e-2)T(0,)
Score_FW_Prcl2	0,567	0,098	0,373	0,755	1,000	dno(0,1e-2)
Score_FW_Prcl3	0,792	0,093	0,612	0,977	1,003	dno(0,1e-2)

Covariances:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
fdw~~fw	0,894	0,119	0,074	0,758	1,004	dbeta(1,1)

Intercepts:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
Score_FDW_ER	-0,000	0,094	-0,185	0,182	1,001	dno(0,1e-3)
Score_FDW_Schk	0,000	0,094	-0,183	0,184	1,002	dno(0,1e-3)
Score_FW_Prcl1	0,000	0,093	-0,181	0,186	1,003	dno(0,1e-3)
Score_FW_Prcl2	-0,001	0,093	-0,178	0,184	1,001	dno(0,1e-3)
Score_FW_Prcl3	-0,000	0,094	-0,181	0,186	1,003	dno(0,1e-3)
fdw	0,000	fw	0,000			

Variances:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
Score_FDW_ER	0,614	0,111	0,398	0,834	1,002	dgamma(1,5)
Score_FDW_Schk	0,584	0,110	0,377	0,808	1,003	dgamma(1,5)
Score_FW_Prcl1	0,431	0,093	0,256	0,619	1,005	dgamma(1,5)
Score_FW_Prcl2	0,705	0,105	0,504	0,912	1,001	dgamma(1,5)
Score_FW_Prcl3	0,402	0,092	0,219	0,58	1,007	dgamma(1,5)
fdw	1,000	fw	1,000			

Tabelle A.45.: Output der CFA (Bayes-Modellierung) ohne „M & E“ (zweifaktorielle Modellierung auf Basis der Gy-Studierendenstichprobe, $N = 86$)

```
two.f.cfa.b <- bcfa(model2, data, std.ov=T, std.lv = T)
```

Number of observations	86	Number of missing patterns	1
Statistic		MargLogLik	PPP
Value		-621,822	0,156

Latent Variables:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
Score_FDW_ER	0,633	0,140	0,367	0,916	1,002	dnorm(0,1e-2)T(0,)
Score_FDW_Schk	0,693	0,142	0,42	0,975	1,001	dno(0,1e-2)
Score_FW_Prcl1	0,784	0,120	0,542	1,014	1,003	dnorm(0,1e-2)T(0,)
Score_FW_Prcl2	0,572	0,120	0,342	0,811	1,000	dno(0,1e-2)
Score_FW_Prcl3	0,771	0,121	0,539	1,011	1,002	dno(0,1e-2)

Covariances:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
fdw~~fw	0,767	0,114	0,538	0,964	1,034	dbeta(1,1)

Intercepts:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
Score_FDW_ER	0,001	0,109	-0,211	0,221	1,000	dno(0,1e-3)
Score_FDW_Schk	0,002	0,110	-0,217	0,215	1,000	dno(0,1e-3)
Score_FW_Prcl1	0,003	0,111	-0,222	0,212	1,000	dno(0,1e-3)
Score_FW_Prcl2	0,002	0,110	-0,212	0,223	1,000	dno(0,1e-3)
Score_FW_Prcl3	0,002	0,111	-0,216	0,218	1,000	dno(0,1e-3)
fdw	0,000	fw	0,000			

Variances:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
Score_FDW_ER	0,632	0,156	0,318	0,949	1,002	dgamma(1,5)
Score_FDW_Schk	0,557	0,161	0,243	0,884	1,001	dgamma(1,5)
Score_FW_Prcl1	0,436	0,127	0,191	0,689	1,006	dgamma(1,5)
Score_FW_Prcl2	0,711	0,127	0,48	0,973	1,000	dgamma(1,5)
Score_FW_Prcl3	0,454	0,127	0,21	0,711	1,004	dgamma(1,5)
fdw	1,000	fw	1,000			

Tabelle A.46.: Output der CFA (Bayes-Modellierung) ohne „M & E“ (zweifaktorielle Modellierung auf Basis der nGy-Studierendenstichprobe, $N = 32$)

two.f.cfa.b <- bcfa(model2, data, std.ov=T, std.lv = T)

Number of observations	32	Number of missing patterns	1
Statistic		MargLogLik	PPP
Value		-256,688	0,398

Latent Variables:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
Score_FDW_ER	0,756	0,204	0,358	1,17	1,000	dnorm(0,1e-2)T(0,)
Score_FDW_Schk	0,710	0,204	0,317	1,124	1,002	dno(0,1e-2)
Score_FW_Prcl1	0,836	0,187	0,489	1,224	1,001	dnorm(0,1e-2)T(0,)
Score_FW_Prcl2	0,559	0,206	0,162	0,972	1,001	dno(0,1e-2)
Score_FW_Prcl3	0,888	0,183	0,529	1,252	1,002	dno(0,1e-2)

Covariances:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
fdw~~fw	0,931	0,071	0,785	1	1,016	dbeta(1,1)

Intercepts:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
Score_FDW_ER	0,006	0,191	-0,376	0,378	1,002	dno(0,1e-3)
Score_FDW_Schk	0,003	0,192	-0,375	0,38	1,002	dno(0,1e-3)
Score_FW_Prcl1	0,007	0,195	-0,37	0,389	1,004	dno(0,1e-3)
Score_FW_Prcl2	0,005	0,190	-0,375	0,374	1,002	dno(0,1e-3)
Score_FW_Prcl3	0,007	0,195	-0,389	0,379	1,004	dno(0,1e-3)

fdw	0,000	fw	0,000
-----	-------	----	-------

Variances:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
Score_FDW_ER	0,546	0,198	0,215	0,967	1,000	dgamma(1,5)
Score_FDW_Schk	0,608	0,210	0,247	1,035	1,002	dgamma(1,5)
Score_FW_Prcl1	0,440	0,160	0,168	0,765	1,007	dgamma(1,5)
Score_FW_Prcl2	0,782	0,224	0,404	1,226	1,001	dgamma(1,5)
Score_FW_Prcl3	0,359	0,151	0,108	0,656	1,005	dgamma(1,5)

fdw	1,000	fw	1,000
-----	-------	----	-------

Tabelle A.47.: Output der CFA (Bayes-Modellierung) ohne „M & E“ (zweifaktorielle Modellierung auf Basis der Gy-nGy-Lehrkräftestichprobe, $N = 65$)

```
two.f.cfa.b <- bcfa(model2, data, std.ov=T, std.lv = T)
```

Number of observations	65	Number of missing patterns	1
Statistic		MargLogLik	PPP
Value		-474,599	0,488

Latent Variables:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
Score_FDW_ER	0,611	0,157	0,315	0,933	1,000	dnorm(0,1e-2)T(0,)
Score_FDW_Schk	0,747	0,154	0,444	1,046	1,000	dno(0,1e-2)
Score_FW_Prcl1	0,720	0,129	0,474	0,979	1,000	dnorm(0,1e-2)T(0,)
Score_FW_Prcl2	0,809	0,126	0,565	1,056	1,003	dno(0,1e-2)
Score_FW_Prcl3	0,815	0,125	0,567	1,06	1,000	dno(0,1e-2)

Covariances:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
fdw~fw	0,723	0,131	0,48	0,972	1,001	dbeta(1,1)

Intercepts:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
Score_FDW_ER	0,004	0,128	-0,249	0,254	1,001	dno(0,1e-3)
Score_FDW_Schk	0,005	0,128	-0,245	0,259	1,001	dno(0,1e-3)
Score_FW_Prcl1	0,005	0,129	-0,254	0,251	1,000	dno(0,1e-3)
Score_FW_Prcl2	0,005	0,128	-0,255	0,251	1,001	dno(0,1e-3)
Score_FW_Prcl3	0,006	0,128	-0,244	0,254	1,001	dno(0,1e-3)
fdw	0,000	fw	0,000			

Variances:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
Score_FDW_ER	0,670	0,172	0,331	1,018	1,001	dgamma(1,5)
Score_FDW_Schk	0,493	0,173	0,158	0,826	1,001	dgamma(1,5)
Score_FW_Prcl1	0,542	0,126	0,315	0,808	1,001	dgamma(1,5)
Score_FW_Prcl2	0,414	0,119	0,196	0,651	1,003	dgamma(1,5)
Score_FW_Prcl3	0,404	0,118	0,195	0,648	1,001	dgamma(1,5)
fdw	1,000	fw	1,000			

Tabelle A.48.: Output der CFA (Bayes-Modellierung) ohne „M & E“ (zweifaktorielle Modellierung auf Basis der Gy-Lehrkräftestichprobe, $N = 35$)

two.f.cfa.b <- bcfa(model2, data, std.ov=T, std.lv = T)

Number of observations	35	Number of missing patterns	1
Statistic		MargLogLik	PPP
Value		-279,078	0,546

Latent Variables:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
Score_FDW_ER	0,492	0,224	0,051	0,918	1,001	dnorm(0,1e-2)T(0,)
Score_FDW_Schk	0,803	0,219	0,363	1,229	1,001	dno(0,1e-2)
Score_FW_Prcl1	0,778	0,188	0,431	1,166	1,000	dnorm(0,1e-2)T(0,)
Score_FW_Prcl2	0,894	0,180	0,554	1,263	1,001	dno(0,1e-2)
Score_FW_Prcl3	0,806	0,186	0,444	1,178	1,001	dno(0,1e-2)

Covariances:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
fdw~~fw	0,676	0,187	0,327	0,997	1,005	dbeta(1,1)

Intercepts:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
Score_FDW_ER	0,004	0,180	-0,355	0,357	1,002	dno(0,1e-3)
Score_FDW_Schk	0,003	0,182	-0,365	0,352	1,005	dno(0,1e-3)
Score_FW_Prcl1	0,007	0,182	-0,342	0,374	1,003	dno(0,1e-3)
Score_FW_Prcl2	0,006	0,184	-0,346	0,376	1,005	dno(0,1e-3)
Score_FW_Prcl3	0,006	0,184	-0,361	0,359	1,003	dno(0,1e-3)

fdw

0,000 fw 0,000

Variances:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
Score_FDW_ER	0,831	0,253	0,377	1,394	1,010	dgamma(1,5)
Score_FDW_Schk	0,472	0,231	0,103	0,917	1,001	dgamma(1,5)
Score_FW_Prcl1	0,525	0,167	0,232	0,866	1,002	dgamma(1,5)
Score_FW_Prcl2	0,355	0,146	0,107	0,643	1,001	dgamma(1,5)
Score_FW_Prcl3	0,490	0,167	0,192	0,828	1,001	dgamma(1,5)

fdw

1,000 fw 1,000

Tabelle A.49.: Output der CFA (Bayes-Modellierung) ohne „M & E“ (zweifaktorielle Modellierung auf Basis der nGy-Lehrkräftestichprobe, $N = 30$)

two.f.cfa.b <- bcfa(model2, data, std.ov=T, std.lv = T)

Number of observations	30	Number of missing patterns	1
Statistic		MargLogLik	PPP
Value		-241,583	0,484

Latent Variables:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
Score_FDW_ER	0,778	0,228	0,321	1,224	1,001	dnorm(0,1e-2)T(0,)
Score_FDW_Schk	0,713	0,232	0,248	1,165	1,001	dno(0,1e-2)
Score_FW_Prcl1	0,763	0,205	0,377	1,178	1,000	dnorm(0,1e-2)T(0,)
Score_FW_Prcl2	0,970	0,183	0,629	1,343	1,002	dno(0,1e-2)
Score_FW_Prcl3	0,840	0,199	0,461	1,233	1,001	dno(0,1e-2)

Covariances:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
fdw~fw	0,721	0,168	0,407	0,992	1,002	dbeta(1,1)

Intercepts:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
Score_FDW_ER	-0,014	0,200	-0,411	0,358	1,002	dno(0,1e-3)
Score_FDW_Schk	-0,012	0,199	-0,401	0,381	1,001	dno(0,1e-3)
Score_FW_Prcl1	-0,013	0,196	-0,404	0,372	1,004	dno(0,1e-3)
Score_FW_Prcl2	-0,019	0,203	-0,421	0,378	1,006	dno(0,1e-3)
Score_FW_Prcl3	-0,016	0,199	-0,419	0,369	1,005	dno(0,1e-3)
fdw	0,000	fw	0,000			

Variances:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
Score_FDW_ER	0,527	0,230	0,12	0,965	1,000	dgamma(1,5)
Score_FDW_Schk	0,611	0,239	0,161	1,073	1,005	dgamma(1,5)
Score_FW_Prcl1	0,564	0,181	0,253	0,92	1,001	dgamma(1,5)
Score_FW_Prcl2	0,254	0,113	0,073	0,475	1,002	dgamma(1,5)
Score_FW_Prcl3	0,457	0,161	0,178	0,779	1,001	dgamma(1,5)
fdw	1,000	fw	1,000			

Tabelle A.50.: Output der CFA (Bayes-Modellierung) ohne „M & E“ (zweifaktorielle Modellierung auf Basis der ff. RS-Gy-Lehrkräftestichprobe, $N = 32$)

two.f.cfa.b <- bcfa(model2, data, std.ov=T, std.lv = T)

Number of observations	32	Number of missing patterns	1
Statistic		MargLogLik	PPP
Value		-271,102	0,514

Latent Variables:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
Score_FDW_ER	0,540	0,311	0	1,058	1,003	dnorm(0,1e-2)T(0,)
Score_FDW_Schk	0,375	0,438	-0,54	1,242	1,002	dno(0,1e-2)
Score_FW_Prcl1	0,467	0,316	0	1,035	1,002	dnorm(0,1e-2)T(0,)
Score_FW_Prcl2	0,201	0,428	-0,676	1,084	1,014	dno(0,1e-2)
Score_FW_Prcl3	0,256	0,437	-0,664	1,119	1,007	dno(0,1e-2)

Covariances:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
fdw~~fw	0,307	0,453	-0,58	0,999	1,007	dbeta(1,1)

Intercepts:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
Score_FDW_ER	-0,002	0,190	-0,372	0,377	1,000	dno(0,1e-3)
Score_FDW_Schk	-0,002	0,188	-0,372	0,369	1,000	dno(0,1e-3)
Score_FW_Prcl1	0,002	0,188	-0,355	0,383	1,000	dno(0,1e-3)
Score_FW_Prcl2	0,001	0,187	-0,359	0,374	1,000	dno(0,1e-3)
Score_FW_Prcl3	0,002	0,189	-0,367	0,373	1,000	dno(0,1e-3)

fdw

0,000 fw 0,000

Variances:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
Score_FDW_ER	0,754	0,343	0,134	1,367	1,001	dgamma(1,5)
Score_FDW_Schk	0,801	0,343	0,103	1,398	1,003	dgamma(1,5)
Score_FW_Prcl1	0,827	0,348	0,125	1,444	1,002	dgamma(1,5)
Score_FW_Prcl2	0,907	0,327	0,157	1,479	1,002	dgamma(1,5)
Score_FW_Prcl3	0,881	0,326	0,211	1,506	1,003	dgamma(1,5)

fdw

1,000 fw 1,000

Tabelle A.51.: Output der CFA (Bayes-Modellierung) ohne „M & E“ (zweifaktorielle Modellierung auf Basis der Fachphysiker, $N = 29$)

two.f.cfa.b <- bcfa(model2, data, std.ov=T, std.lv = T)

Number of observations	29	Number of missing patterns	1
Statistic		MargLogLik	PPP
Value		-243,483	0,257

Latent Variables:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
Score_FDW_ER	0,761	0,267	0,197	1,252	1,005	dnorm(0,1e-2)T(0,)
Score_FDW_Schk	0,552	0,295	-0,005	1,143	1,003	dno(0,1e-2)
Score_FW_Prcl1	0,938	0,206	0,546	1,353	1,000	dnorm(0,1e-2)T(0,)
Score_FW_Prcl2	0,891	0,212	0,506	1,336	1,000	dno(0,1e-2)
Score_FW_Prcl3	0,417	0,236	-0,047	0,888	1,000	dno(0,1e-2)

Covariances:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
fdw~~fw	0,417	0,236	-0,047	0,888	1,000	dbeta(1,1)

Intercepts:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
Score_FDW_ER	-0,000	0,202	-0,4	0,395	1,000	dno(0,1e-3)
Score_FDW_Schk	-0,002	0,202	-0,404	0,39	1,000	dno(0,1e-3)
Score_FW_Prcl1	0,000	0,210	-0,415	0,41	1,001	dno(0,1e-3)
Score_FW_Prcl2	-0,000	0,208	-0,416	0,406	1,001	dno(0,1e-3)
Score_FW_Prcl3	-0,001	0,200	-0,397	0,389	1,000	dno(0,1e-3)
fdw	0,000	fw	0,000			

Variances:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
Score_FDW_ER	0,540	0,288	0,101	1,091	1,010	dgamma(1,5)
Score_FDW_Schk	0,768	0,305	0,145	1,34	1,002	dgamma(1,5)
Score_FW_Prcl1	0,334	0,160	0,082	0,651	1,001	dgamma(1,5)
Score_FW_Prcl2	0,403	0,174	0,095	0,738	1,001	dgamma(1,5)
Score_FW_Prcl3	0,923	0,270	0,465	1,451	1,000	dgamma(1,5)
fdw	1,000	fw	1,000			

Tabelle A.52.: Output der CFA (Bayes-Modellierung) ohne „M & E“ (zweifaktorielle Modellierung auf Basis der ff. MS-Lehrkräftestichprobe, $N = 25$)

two.f.cfa.b <- bcfa(model2, data, std.ov=T, std.lv = T)

Number of observations	25	Number of missing patterns	1
Statistic		MargLogLik	PPP
Value		-204,647	0,248

Latent Variables:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
Score_FDW_ER	0,899	0,227	0,474	1,367	1,002	dnorm(0,1e-2)T(0,)
Score_FDW_Schk	0,692	0,243	0,242	1,199	1,001	dno(0,1e-2)
Score_FW_Prcl1	1,026	0,199	0,65	1,418	1,004	dnorm(0,1e-2)T(0,)
Score_FW_Prcl2	0,896	0,216	0,494	1,333	1,003	dno(0,1e-2)
Score_FW_Prcl3	0,809	0,226	0,385	0,271	1,002	dno(0,1e-2)

Covariances:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
fdw~~fw	0,816	0,140	0,543	1	1,001	dbeta(1,1)

Intercepts:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
Score_FDW_ER	0,001	0,229	-0,462	0,447	1,003	dno(0,1e-3)
Score_FDW_Schk	-0,000	0,222	-0,441	0,43	1,001	dno(0,1e-3)
Score_FW_Prcl1	0,001	0,234	-0,475	0,45	1,003	dno(0,1e-3)
Score_FW_Prcl2	0,001	0,230	-0,459	0,448	1,002	dno(0,1e-3)
Score_FW_Prcl3	0,001	0,227	-0,456	0,438	1,003	dno(0,1e-3)

fdw

0,000 fw 0,000

Variances:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
Score_FDW_ER	0,395	0,190	0,094	0,759	1,001	dgamma(1,5)
Score_FDW_Schk	0,669	0,249	0,223	1,172	1,001	dgamma(1,5)
Score_FW_Prcl1	0,203	0,095	0,063	0,392	1,002	dgamma(1,5)
Score_FW_Prcl2	0,411	0,156	0,154	0,721	1,000	dgamma(1,5)
Score_FW_Prcl3	0,534	0,193	0,219	0,911	1,001	dgamma(1,5)

fdw

1,000 fw 1,000

Tabelle A.53.: Output der CFA (Bayes-Modellierung) für das FDW beider Projekte (FALKO, ProfiLe-P) für die gesamte Teilstichprobe der Kreuzvalidierung, $N = 28$

```
two.f.cfa.b <- bcfa(model2, data, std.ov=T, std.lv = T)
```

Number of observations	28	Number of missing patterns	1
Statistic		MargLogLik	PPP
Value		-231,395	0,386

Latent Variables:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
Score_FDW_ER	0,840	0,228	0,386	1,28	1,002	dnorm(0,1e-2)T(0,)
Score_FDW_Schk	0,514	0,251	0,031	1,016	1,001	dno(0,1e-2)
Score_FW_Prcl1	0,897	0,214	0,486	1,332	1,003	dnorm(0,1e-2)T(0,)
Score_FW_Prcl2	0,729	0,227	0,309	1,203	1,001	dno(0,1e-2)
Score_FW_Prcl3	0,842	0,216	0,425	0,281	1,003	dno(0,1e-2)

Covariances:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
fdw~~fw	0,805	0,157	0,493	0,999	1,008	dbeta(1,1)

Intercepts:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
Score_FDW_ER	0,001	0,208	-0,393	0,426	1,004	dno(0,1e-3)
Score_FDW_Schk	0,001	0,206	-0,413	0,395	1,002	dno(0,1e-3)
Score_FW_Prcl1	-0,001	0,208	-0,403	0,423	1,004	dno(0,1e-3)
Score_FW_Prcl2	-0,000	0,205	-0,4	0,412	1,002	dno(0,1e-3)
Score_FW_Prcl3	0,001	0,207	-0,414	0,41	1,004	dno(0,1e-3)
fdw	0,000	fw	0,000			

Variances:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
Score_FDW_ER	0,457	0,220	0,092	0,88	1,003	dgamma(1,5)
Score_FDW_Schk	0,843	0,279	0,365	1,413	1,000	dgamma(1,5)
Score_FW_Prcl1	0,393	0,175	0,1	0,731	1,000	dgamma(1,5)
Score_FW_Prcl2	0,616	0,224	0,205	1,043	1,000	dgamma(1,5)
Score_FW_Prcl3	0,470	0,187	0,145	0,835	1,000	dgamma(1,5)
fdw	1,000	fw	1,000			

Tabelle A.54.: Output der CFA (Bayes-Modellierung) für das FDW beider Projekte (FALKO, ProfiLe-P) für die Gy-Teilstichprobe der Kreuzvalidierung, $N = 13$)

two.f.cfa.b <- bcfa(model2, data, std.ov=T, std.lv = T)

Number of observations	13	Number of missing patterns	1
Statistic		MargLogLik	PPP
Value		-127,730	0,532

Latent Variables:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
Score_FDW_ER	0,893	0,414	0,115	1,695	1,007	dnorm(0,1e-2)T(0,)
Score_FDW_Schk	1,071	0,404	0,337	1,906	1,015	dno(0,1e-2)
Score_FW_Prcl1	1,105	0,395	0,416	1,919	1,011	dnorm(0,1e-2)T(0,)
Score_FW_Prcl2	0,854	0,418	0,077	1,711	1,007	dno(0,1e-2)
Score_FW_Prcl3	0,774	0,427	-0,004	1,676	1,003	dno(0,1e-2)

Covariances:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
fdw~fw	0,843	0,176	0,47	1	1,005	dbeta(1,1)

Intercepts:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
Score_FDW_ER	-0,013	0,345	-0,711	0,671	1,004	dno(0,1e-3)
Score_FDW_Schk	-0,015	0,355	-0,73	0,693	1,007	dno(0,1e-3)
Score_FW_Prcl1	-0,014	0,351	-0,736	0,655	1,007	dno(0,1e-3)
Score_FW_Prcl2	-0,014	0,338	-0,696	0,649	1,007	dno(0,1e-3)
Score_FW_Prcl3	-0,007	0,341	-0,72	0,639	1,003	dno(0,1e-3)

fdw	0,000	fw	0,000
-----	-------	----	-------

Variances:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
Score_FDW_ER	0,639	0,353	0,127	1,313	1,001	dgamma(1,5)
Score_FDW_Schk	0,401	0,260	0,068	0,893	1,001	dgamma(1,5)
Score_FW_Prcl1	0,368	0,230	0,08	0,812	1,001	dgamma(1,5)
Score_FW_Prcl2	0,684	0,366	0,179	1,401	1,002	dgamma(1,5)
Score_FW_Prcl3	0,771	0,405	0,201	1,527	1,000	dgamma(1,5)

fdw	1,000	fw	1,000
-----	-------	----	-------

Tabelle A.55.: Output der CFA (Bayes-Modellierung) für das FDW beider Projekte (FALKO, ProfiLe-P) für die nGy-Teilstichprobe der Kreuzvalidierung, $N = 15$

```
two.f.cfa.b <- bcfa(model2, data, std.ov=T, std.lv = T)
```

Number of observations	15	Number of missing patterns	1
Statistic		MargLogLik	PPP
Value		-142,316	0,268

Latent Variables:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
Score_FDW_ER	0,866	0,390	0,002	1,516	1,007	dnorm(0,1e-2)T(0,)
Score_FDW_Schk	0,291	0,417	-0,647	1,264	1,004	dno(0,1e-2)
Score_FW_Prcl1	1,043	0,359	0,421	1,805	1,002	dnorm(0,1e-2)T(0,)
Score_FW_Prcl2	0,716	0,388	-0,041	1,461	1,001	dno(0,1e-2)
Score_FW_Prcl3	1,092	0,349	-0,468	1,779	1,001	dno(0,1e-2)

Covariances:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
fdw~~fw	0,638	0,327	-0,049	1	1,058	dbeta(1,1)

Intercepts:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
Score_FDW_ER	0,011	0,317	-0,602	0,654	1,004	dno(0,1e-3)
Score_FDW_Schk	0,004	0,296	-0,587	0,589	1,001	dno(0,1e-3)
Score_FW_Prcl1	0,011	0,325	-0,629	0,661	1,005	dno(0,1e-3)
Score_FW_Prcl2	0,007	0,313	-0,625	0,621	1,003	dno(0,1e-3)
Score_FW_Prcl3	0,010	0,328	-0,649	0,658	1,006	dno(0,1e-3)
fdw	0,000	fw	0,000			

Variances:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
Score_FDW_ER	0,569	0,375	0,069	1,274	1,006	dgamma(1,5)
Score_FDW_Schk	1,016	0,475	0,22	1,953	1,016	dgamma(1,5)
Score_FW_Prcl1	0,393	0,236	0,084	0,854	1,000	dgamma(1,5)
Score_FW_Prcl2	0,784	0,366	0,233	1,514	1,000	dgamma(1,5)
Score_FW_Prcl3	0,319	0,191	0,074	0,688	1,000	dgamma(1,5)
fdw	1,000	fw	1,000			

Tabelle A.56.: Output der CFA (Bayes-Modellierung) für das FW beider Projekte (FALKO, ProfiLe-P) für die gesamte Teilstichprobe der Kreuzvalidierung, $N = 28$

two.f.cfa.b <- bcfa(model2, data, std.ov=T, std.lv = T)

Number of observations					28	
Number of missing patterns					1	
Statistic					MargLogLik	PPP
Value					-264,549	0,711
Latent Variables:						
	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
fwFALKO =~						
S_FW_FALKO_P1	1,009	0,198	0,654	1,417	1,001	dnorm(0,1e-2)T(0,)
S_FW_FALKO_P2	0,920	0,207	0,531	1,34	1,001	dno(0,1e-2)
S_FW_FALKO_P3	0,915	0,207	0,541	1,34	1,001	dnorm(0,1e-2)T(0,)
fwProfiLeP =~						
Sc_FW_PLP_P1	0,849	0,219	0,426	1,29	1,001	dno(0,1e-2)
Sc_FW_PLP_P2	0,720	0,224	0,301	1,189	1,001	dno(0,1e-2)
Sc_FW_PLP_P3	0,942	0,208	0,546	1,355	1,001	dno(0,1e-2)
Covariances:						
	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
fwFALKO~~fwProf	0,768	0,125	0,511	0,966	1,006	dbeta(1,1)
Intercepts:						
	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
.S_FW_FALKO_P1	-0,009	0,226	-0,449	0,448	1,002	dno(0,1e-3)
.S_FW_FALKO_P2	-0,010	0,223	-0,45	0,428	1,002	dno(0,1e-3)
.S_FW_FALKO_P3	-0,010	0,222	-0,434	0,449	1,002	dno(0,1e-3)
.Sc_FW_PLP_P1	-0,007	0,218	-0,457	0,415	1,002	dno(0,1e-3)
.Sc_FW_PLP_P2	-0,007	0,213	-0,424	0,415	1,002	dno(0,1e-3)
.Sc_FW_PLP_P3	-0,009	0,222	-0,457	0,43	1,003	dno(0,1e-3)
fwFALKO	0,000	fwProf	0,000			
Variances:						
	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
.S_FW_FALKO_P1	0,236	0,099	0,081	0,43	1,001	dgamma(1,5)
.S_FW_FALKO_P2	0,381	0,136	0,155	1,649	1,000	dgamma(1,5)
.S_FW_FALKO_P3	0,386	0,138	0,153	1,657	1,000	dgamma(1,5)
.Sc_FW_PLP_P1	0,479	0,178	0,182	0,84	1,000	dgamma(1,5)
.Sc_FW_PLP_P2	0,644	0,211	0,288	1,061	1,000	dgamma(1,5)
.Sc_FW_PLP_P3	0,343	0,149	0,105	0,641	1,001	dgamma(1,5)
fwFALKO	1,000	fwProf	1,000			

Tabelle A.57.: Output der CFA (Bayes-Modellierung) für das FW beider Projekte (FALKO, ProfiLe-P) für die Gy-Teilstichprobe der Kreuzvalidierung, $N = 13$)

two.f.cfa.b <- bcfa(model2, data, std.ov=T, std.lv = T)

Number of observations		13	
Number of missing patterns		1	
Statistic		MargLogLik	PPP
Value		-150,708	0,672

Latent Variables:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
fwFALKO =~						
S_FW_FALKO_P1	1,201	0,429	0,488	2,075	1,004	dnorm(0,1e-2)T(0,)
S_FW_FALKO_P2	1,151	0,427	0,424	2,007	1,007	dno(0,1e-2)
S_FW_FALKO_P3	1,067	0,433	0,299	1,942	1,005	dnorm(0,1e-2)T(0,)
fwProfiLeP =~						
Sc_FW_PLP_P1	0,947	0,442	0,111	1,799	1,004	dno(0,1e-2)
Sc_FW_PLP_P2	0,737	0,469	-0,147	1,696	1,001	dno(0,1e-2)
Sc_FW_PLP_P3	1,070	0,444	0,245	1,97	1,002	dno(0,1e-2)

Covariances:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
fwFALKO~~fwProf	0,816	0,181	0,442	1	1,004	dbeta(1,1)

Intercepts:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
.S_FW_FALKO_P1	-0,014	0,386	-0,779	0,764	1,007	dno(0,1e-3)
.S_FW_FALKO_P2	-0,011	0,380	-0,771	0,748	1,007	dno(0,1e-3)
.S_FW_FALKO_P3	-0,015	0,373	-0,753	0,742	1,006	dno(0,1e-3)
.Sc_FW_PLP_P1	-0,011	0,364	-0,735	0,708	1,006	dno(0,1e-3)
.Sc_FW_PLP_P2	-0,013	0,349	-0,72	0,679	1,003	dno(0,1e-3)
.Sc_FW_PLP_P3	-0,012	0,372	-0,788	0,718	1,006	dno(0,1e-3)

fwFALKO 0,000 fwProf 0,000

Variances:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
.S_FW_FALKO_P1	0,327	0,195	0,074	0,697	1,001	dgamma(1,5)
.S_FW_FALKO_P2	0,395	0,227	0,081	0,82	1,000	dgamma(1,5)
.S_FW_FALKO_P3	0,500	0,271	0,124	1,023	1,000	dgamma(1,5)
.Sc_FW_PLP_P1	0,629	0,347	0,119	1,29	1,000	dgamma(1,5)
.Sc_FW_PLP_P2	0,835	0,437	0,216	1,676	1,000	dgamma(1,5)
.Sc_FW_PLP_P3	0,494	0,308	0,086	1,09	1,002	dgamma(1,5)

fwFALKO 1,000 fwProf 1,000

Tabelle A.58.: Output der CFA (Bayes-Modellierung) für das FW beider Projekte (FALKO, ProfiLe-P) für die nGy-Teilstichprobe der Kreuzvalidierung, $N = 15$)

two.f.cfa.b <- bcfa(model2, data, std.ov=T, std.lv = T)

Number of observations	15		
Number of missing patterns	1		
Statistic	MargLogLik	PPP	
Value	-169,111	0,547	

Latent Variables:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
fwFALKO =~						
S_FW_FALKO_P1	1,061	0,348	0,439	1,788	1,001	dnorm(0,1e-2)T(0,)
S_FW_FALKO_P2	0,863	0,371	0,18	1,641	1,001	dno(0,1e-2)
S_FW_FALKO_P3	1,079	0,347	0,465	1,806	1,002	dnorm(0,1e-2)T(0,)
fwProfiLeP =~						
Sc_FW_PLP_P1	1,053	0,356	0,372	1,761	1,000	dno(0,1e-2)
Sc_FW_PLP_P2	0,830	0,378	0,106	1,608	1,000	dno(0,1e-2)
Sc_FW_PLP_P3	1,039	0,365	0,379	1,786	1,000	dno(0,1e-2)

Covariances:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
fwFALKO~~fwProf	0,495	0,280	-0,055	0,945	1,006	dbeta(1,1)

Intercepts:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
.S_FW_FALKO_P1	-0,003	0,337	-0,675	0,693	1,002	dno(0,1e-3)
.S_FW_FALKO_P2	-0,003	0,323	-0,659	0,631	1,001	dno(0,1e-3)
.S_FW_FALKO_P3	-0,006	0,336	-0,69	0,648	1,002	dno(0,1e-3)
.Sc_FW_PLP_P1	0,007	0,331	-0,638	0,675	1,002	dno(0,1e-3)
.Sc_FW_PLP_P2	0,007	0,317	-0,629	0,627	1,002	dno(0,1e-3)
.Sc_FW_PLP_P3	0,007	0,335	-0,658	0,674	1,002	dno(0,1e-3)
fwFALKO	0,000	fwProf	0,000			

Variances:

	Estimate	Post.SD	HPD.025	HPD.975	PSRF	Prior
.S_FW_FALKO_P1	0,394	0,223	0,09	0,816	1,000	dgamma(1,5)
.S_FW_FALKO_P2	0,642	0,309	0,188	1,26	1,000	dgamma(1,5)
.S_FW_FALKO_P3	0,363	0,205	0,079	0,755	1,001	dgamma(1,5)
.Sc_FW_PLP_P1	0,408	0,234	0,079	0,859	1,001	dgamma(1,5)
.Sc_FW_PLP_P2	0,682	0,333	0,183	1,332	1,001	dgamma(1,5)
.Sc_FW_PLP_P3	0,428	0,250	0,089	0,902	1,001	dgamma(1,5)
fwFALKO	1,000	fwProf	1,000			

B. Testheft



Universität Regensburg



FALKO
Fachspezifische Lehrerkompetenzen

FALKO

Fachspezifische Lehrerkompetenzen

Fragebogen zur Erhebung fachspezifischer Lehrerkompetenzen

Testheft

Physik

Fachdidaktisches Wissen
Fachwissen



FALKO
Fachspezifische Lehrerkompetenzen

Forschungsgruppe der Universitäten Regensburg und Augsburg:

Allgemeine Pädagogik:	Prof. Dr. Regina Mulder Susanne Sauer Franziska Kempka
Deutsch-Didaktik:	Prof. Dr. Anita Schilcher Dr. Markus Pissarek
Englisch-Didaktik:	Prof. Dr. Petra Kirchhoff
Evangelische Theologie:	Prof. Dr. Michael Fricke
Latein-Didaktik:	Alfred Lindl Harald Kloiber
Mathematik-Didaktik:	Prof. Dr. Stefan Krauss
Musikpädagogik:	Prof. Dr. Bernhard Hofmann Dr. Gabriele Puffer
Nat.wiss. und Technik, Physik:	Anja Schödl Dr. Anja Göhring M. A.



FALKO
Fachspezifische Lehrerkompetenzen

Sehr geehrte Lehrkraft,

Sie haben sich dazu bereit erklärt, bei der Entwicklung der Untersuchungsinstrumente zur Erfassung von Professionswissen von Lehrkräften mitzuwirken. Das FALKO-Team der Universität Regensburg bedankt sich ganz herzlich für Ihre Hilfe!

Dieses Testheft gliedert sich in drei Abschnitte:

Im 1. Teil stellen wir Ihnen einige **Fragen zu Ihrer Person**. Diese Befragung verläuft anonymisiert, indem Sie anhand bestimmter Kriterien ein Kennwort festlegen. Dieses kann nur von Ihnen nachvollzogen werden, erlaubt uns aber eine strukturierte Auswertung Ihrer Antworten. Es werden keine Rückschlüsse auf Personen oder Schulen gezogen. Die erhaltenen Daten werden vertraulich behandelt und ausschließlich im Rahmen unserer Forschungsgruppe zur Optimierung der Testinstrumente verwendet. Bitte antworten Sie daher ehrlich und vollständig.

Es folgt Teil 2 zur Erhebung des **fachdidaktischen Wissens (ca. 60 Minuten Bearbeitungszeit)**.

Gegenstand des 3. Teils ist das **fachwissenschaftliche Wissen (ca. 30 Minuten Bearbeitungszeit)**.

Bitte „beißen“ Sie sich bei der Bearbeitung nicht an einzelnen Aufgaben fest. Versuchen Sie, möglichst alle Fragen zu beantworten, auch wenn Sie sich bei einer Antwort evtl. nicht ganz sicher sind. Fehlende Antworten stellen ein Auswertungsproblem dar und erschweren die Weiterentwicklung unserer Testinstrumente. Daher bitten wir Sie, sich in Zweifelsfällen für die Ihnen am plausibelsten erscheinende Antwort zu entscheiden. Formulieren Sie diese bei offenen Antworten möglichst in ganzen Sätzen und verwenden Sie keine Formelsammlungen, Lexika oder Ähnliches.

Sollte Ihnen der vorgesehene Platz zur Beantwortung nicht ausreichen, können Sie Ihre Antworten auf dem zusätzlichen Leerblatt am Ende des Testhefts fortführen.

Möchten Sie eine Korrektur Ihrer Antwort vornehmen, streichen Sie diese durch und führen Sie sie, falls nötig, ebenfalls auf dem zusätzlichen Leerblatt, fort. Zur Korrektur von Multiple-Choice-Aufgaben malen Sie bitte das falsch angekreuzte Kästchen komplett aus und kreuzen Sie das richtige an. Falls Sie sich erneut für das entwertete Kästchen entscheiden, kreuzen Sie einfach links neben dem entwerteten Kästchen an.

Zum Schluss ein Hinweis: Aus Gründen der Lesbarkeit wird in den Aufgabenstellungen, Personen betreffend, immer nur ein Geschlecht genannt, z. B. Schüler oder Schülerin.

Herzlichen Dank für Ihre Mithilfe!



FALKO

Fachspezifische Lehrerkompetenzen

Bitte tragen Sie hier ihr persönliches Kennwort ein:

(Umlaute bitte als „AE“ für „Ä“, usw. eintragen dabei bitte das „E“ nicht berücksichtigen → siehe Beispiel!)

Tragen Sie hier bitte die ersten beiden Buchstaben des **Vornamens Ihrer Mutter** ein, z. B. Sofie = .

Tragen Sie hier bitte die ersten beiden Buchstaben des aktuellen **Nachnamens Ihrer Mutter** ein, z. B. Önal = .

Tragen Sie hier bitte die letzten beiden Stellen **Ihres eigenen Geburtsjahres** ein, z. B. ; wenn Sie am 24.01.1968 geboren sind.

Teil I: Personenbezogene Daten

1. Sie sind? weiblich männlich
 2. In welchem Jahr wurden Sie geboren? 19_____
 3. Wie viele Jahre sind Sie bereits im Beruf (inkl. evtl. Vorbereitungsdienst, ohne evtl. Elternzeit)?:

- Frage 3. trifft nicht zu.*



FALKO

Fachspezifische Lehrerkompetenzen

4. Sie sind zum jetzigen Zeitpunkt... → bitte a), b), c), d) oder e) wählen!

- a) Lehramtsstudent/in Bundesland: _____ Fachsemester: _____
Regelstudienzeit (in Semestern): _____
 Gymnasium Fach 1 _____ Fach 2 _____ Fach 3 _____
 Realschule Fach 1 _____ Fach 2 _____ Fach 3 _____
 HS/MS: Studierte Unterrichts-/Hauptfach/-fächer: _____
Studierte Didaktik-/Nebenfächer: _____
- b) Student/in Bundesland: _____ Fachsemester: _____
Regelstudienzeit (in Semestern): _____
Welchen Abschluss streben Sie an?
Diplom im Fach: _____
Bachelor im Fach: _____
Master im Fach: _____
Sonstiges: _____
- c) Referendar/in, LAA Bundesland: _____
im 1. Halbjahr 2. HJ 3. HJ 4. HJ Sonst.: _____
 Gymnasium Fach 1 _____ Fach 2 _____ Fach 3 _____
 Realschule Fach 1 _____ Fach 2 _____ Fach 3 _____
 HS/MS: Studierte Unterrichts-/Hauptfach/-fächer: _____
Studierte Didaktik-/Nebenfächer: _____
- d) Lehrer/in Bundesland: _____
 Gymnasium Fach 1 _____ Fach 2 _____ Fach 3 _____
 Realschule Fach 1 _____ Fach 2 _____ Fach 3 _____
 HS/MS: Studierte Unterrichts-/Hauptfach/-fächer: _____
Studierte Didaktik-/Nebenfächer: _____
Wie wurden Sie im Vorbereitungsdienst ausgebildet?
 Lehrerbildungsanstalt, z. B. München-Pasing Referendariat/Vorbereitungsdienst
Sonstiges: _____
- e) Sonstiges: Bundesland: _____
Tätigkeit bitte kurz beschreiben:

5. Welches ist der höchste Schul-/Studien- und/oder Berufsabschluss Ihrer Eltern?

Vater: _____ Mutter: _____

6. An welchem Hochschultyp studieren/studierten Sie für Ihren jetzigen Beruf?

- Universität Technische Hochschule Institut für Lehrerbildung
 Pädagogische Hochschule Anderer Hochschultyp: _____



FALKO

Fachspezifische Lehrerkompetenzen

7. Welche Ausbildung(en) absolvierten Sie, bevor Sie Ihr Lehramtsstudium aufnahmen und wo?

Abitur Bundesland: _____

Fachabitur Bundesland: _____

Sonstiges (z. B. Ausbildungsberuf, anderes (Fach-)Hochschulstudium, ...).

Bitte nennen: _____

8. Bitte nennen Sie Ihre Noten(durchschnitte), falls zutreffend:

Abitur/Fachabitur: _____

Erstes und zweites Staatsexamen: _____

Letzte (Schul)Note im Fach Physik: _____

9. *Frage 9. trifft nicht zu.*

a) Wie viele Fortbildungen (Ph, PCB, ...) besuchten Sie im letzten und laufenden (Schul)Jahr? .

b) Zählen Sie bitte kurz die Themengebiete auf, die in diesen Fortbildungen behandelt wurden.

10. Bitte denken Sie ab dem heutigen Datum **ein Kalenderjahr** zurück:

Frage 10. trifft nicht zu.

a) In welchen Klassenstufen unterrichteten Sie im Fach Physik/PCB/MINT usw. innerhalb dieses Jahres?

5 6 7

8 9 10

b) Wie viele Stunden pro Woche (Wochenstunden) unterrichteten Sie im Fach Physik/PCB/MINT usw. in den einzelnen Jahrgangsstufen innerhalb dieses Jahres?

5. _____ Stunden 6. _____ Stunden 7. _____ Stunden

8. _____ Stunden 9. _____ Stunden 10. _____ Stunden

c) Wie viele Stunden pro Woche (Wochenstunden) unterrichteten Sie im Fach Physik/PCB/MINT usw. in den einzelnen Jahrgangsstufen innerhalb dieses Jahres?

7. _____ Stunden 8. _____ Stunden 9. _____ Stunden

10. _____ Stunden 11. _____ Stunden 12. _____ Stunden

d) Welche Themengebiete im Fach Physik/PCB/MINT usw. unterrichteten Sie in welcher Jahrgangsstufe innerhalb dieses Jahres?

Mechanik 5 6 7 8 9 10

Elektrizitätslehre 5 6 7 8 9 10

Wärmelehre 5 6 7 8 9 10

Optik 5 6 7 8 9 10

Akustik 5 6 7 8 9 10

Atom-/Kernphysik 5 6 7 8 9 10

11. Notieren Sie bitte (pädagogische, didaktische, fachliche) „Besonderheiten“ Ihrer Person, z. B. Seminarleiter/-Lehrer(in), Fachleiter(in), Betreuungslehrkraft für Referendare/Studierende, Schulbuchautor(in), regelmäßiges Erteilen von Nachhilfe, Tutor, etc.

Frage 11. trifft nicht zu.



Teil II: Fachdidaktik Physik

Bitte tragen Sie rechts die Zeit ein, die Sie für die fachdidaktischen Aufgaben insgesamt benötigen.

oder

Bitte tragen Sie rechts die Uhrzeit ein, zu der Sie mit den fachdidaktischen Aufgaben beginnen und enden.

Beginn:

Ende:



Teil III: Fachwissen Physik

Bitte tragen Sie rechts die Zeit ein, die Sie für die fachwissenschaftlichen Aufgaben insgesamt benötigen.

oder

Bitte tragen Sie rechts die Uhrzeit ein, zu der Sie mit den fachwissenschaftlichen Aufgaben beginnen und enden.

Beginn:

Ende:

Bisher erschienene Bände der Reihe „*Studien zum Physik- und Chemielernen*“

ISSN 1614-8967 (vormals *Studien zum Physiklernen* ISSN 1435-5280)

- 1 Helmut Fischler, Jochen Peuckert (Hrsg.): Concept Mapping in fachdidaktischen Forschungsprojekten der Physik und Chemie
ISBN 978-3-89722-256-4 40.50 EUR
- 2 Anja Schoster: Bedeutungsentwicklungsprozesse beim Lösen algorithmischer Physikaufgaben. *Eine Fallstudie zu Lernprozessen von Schülern im Physiknachhilfeunterricht während der Bearbeitung algorithmischer Physikaufgaben*
ISBN 978-3-89722-045-4 40.50 EUR
- 3 Claudia von Aufschnaiter: Bedeutungsentwicklungen, Interaktionen und situatives Erleben beim Bearbeiten physikalischer Aufgaben
ISBN 978-3-89722-143-7 40.50 EUR
- 4 Susanne Haeberlen: Lernprozesse im Unterricht mit Wasserstromkreisen. *Eine Fallstudie in der Sekundarstufe I*
ISBN 978-3-89722-172-7 40.50 EUR
- 5 Kerstin Haller: Über den Zusammenhang von Handlungen und Zielen. *Eine empirische Untersuchung zu Lernprozessen im physikalischen Praktikum*
ISBN 978-3-89722-242-7 40.50 EUR
- 6 Michaela Horstendahl: Motivationale Orientierungen im Physikunterricht
ISBN 978-3-89722-227-4 50.00 EUR
- 7 Stefan Deylitz: Lernergebnisse in der Quanten-Atomphysik. *Evaluation des Bremer Unterrichtskonzepts*
ISBN 978-3-89722-291-5 40.50 EUR
- 8 Lorenz Hucke: Handlungsregulation und Wissenserwerb in traditionellen und computergestützten Experimenten des physikalischen Praktikums
ISBN 978-3-89722-316-5 50.00 EUR
- 9 Heike Theyßen: Ein Physikpraktikum für Studierende der Medizin. *Darstellung der Entwicklung und Evaluation eines adressatenspezifischen Praktikums nach dem Modell der Didaktischen Rekonstruktion*
ISBN 978-3-89722-334-9 40.50 EUR
- 10 Annette Schick: Der Einfluß von Interesse und anderen selbstbezogenen Kognitionen auf Handlungen im Physikunterricht. *Fallstudien zu Interessenhandlungen im Physikunterricht*
ISBN 978-3-89722-380-6 40.50 EUR
- 11 Roland Berger: Moderne bildgebende Verfahren der medizinischen Diagnostik. *Ein Weg zu interessanterem Physikunterricht*
ISBN 978-3-89722-445-2 40.50 EUR

- 12 Johannes Werner: Vom Licht zum Atom. *Ein Unterrichtskonzept zur Quantenphysik unter Nutzung des Zeigermodells*
ISBN 978-3-89722-471-1 40.50 EUR
- 13 Florian Sander: Verbindung von Theorie und Experiment im physikalischen Praktikum. *Eine empirische Untersuchung zum handlungsbezogenen Vorverständnis und dem Einsatz grafikorientierter Modellbildung im Praktikum*
ISBN 978-3-89722-482-7 40.50 EUR
- 14 Jörn Gerdes: Der Begriff der physikalischen Kompetenz. *Zur Validierung eines Konstruktes*
ISBN 978-3-89722-510-7 40.50 EUR
- 15 Malte Meyer-Arndt: Interaktionen im Physikpraktikum zwischen Studierenden und Betreuern. *Feldstudie zu Bedeutungsentwicklungsprozessen im physikalischen Praktikum*
ISBN 978-3-89722-541-1 40.50 EUR
- 16 Dietmar Höttecke: Die Natur der Naturwissenschaften historisch verstehen. *Fachdidaktische und wissenschaftshistorische Untersuchungen*
ISBN 978-3-89722-607-4 40.50 EUR
- 17 Gil Gabriel Mavanga: Entwicklung und Evaluation eines experimentell- und phänomenorientierten Optikcurriculums. *Untersuchung zu Schülervorstellungen in der Sekundarstufe I in Mosambik und Deutschland*
ISBN 978-3-89722-721-7 40.50 EUR
- 18 Meike Ute Zastrow: Interaktive Experimentieranleitungen. *Entwicklung und Evaluation eines Konzeptes zur Vorbereitung auf das Experimentieren mit Messgeräten im Physikalischen Praktikum*
ISBN 978-3-89722-802-3 40.50 EUR
- 19 Gunnar Friege: Wissen und Problemlösen. *Eine empirische Untersuchung des wissenszentrierten Problemlösens im Gebiet der Elektrizitätslehre auf der Grundlage des Experten-Novizen-Vergleichs*
ISBN 978-3-89722-809-2 40.50 EUR
- 20 Erich Starauschek: Physikunterricht nach dem Karlsruher Physikkurs. *Ergebnisse einer Evaluationsstudie*
ISBN 978-3-89722-823-8 40.50 EUR
- 21 Roland Paatz: Charakteristika analogiebasierten Denkens. *Vergleich von Lernprozessen in Basis- und Zielbereich*
ISBN 978-3-89722-944-0 40.50 EUR
- 22 Silke Mikelskis-Seifert: Die Entwicklung von Metakzepten zur Teilchenvorstellung bei Schülern. *Untersuchung eines Unterrichts über Modelle mithilfe eines Systems multipler Repräsentationsebenen*
ISBN 978-3-8325-0013-9 40.50 EUR
- 23 Brunhild Landwehr: Distanzen von Lehrkräften und Studierenden des Sachunterrichts zur Physik. *Eine qualitativ-empirische Studie zu den Ursachen*
ISBN 978-3-8325-0044-3 40.50 EUR

- 24 Lydia Murmann: Physiklernen zu Licht, Schatten und Sehen. *Eine phänomenografische Untersuchung in der Primarstufe*
ISBN 978-3-8325-0060-3 40.50 EUR
- 25 Thorsten Bell: Strukturprinzipien der Selbstregulation. *Komplexe Systeme, Elementarisierungen und Lernprozessstudien für den Unterricht der Sekundarstufe II*
ISBN 978-3-8325-0134-1 40.50 EUR
- 26 Rainer Müller: Quantenphysik in der Schule
ISBN 978-3-8325-0186-0 40.50 EUR
- 27 Jutta Roth: Bedeutungsentwicklungsprozesse von Physikerinnen und Physikern in den Dimensionen Komplexität, Zeit und Inhalt
ISBN 978-3-8325-0183-9 40.50 EUR
- 28 Andreas Saniter: Spezifika der Verhaltensmuster fortgeschrittener Studierender der Physik
ISBN 978-3-8325-0292-8 40.50 EUR
- 29 Thomas Weber: Kumulatives Lernen im Physikunterricht. *Eine vergleichende Untersuchung in Unterrichtsgängen zur geometrischen Optik*
ISBN 978-3-8325-0316-1 40.50 EUR
- 30 Markus Rehm: Über die Chancen und Grenzen moralischer Erziehung im naturwissenschaftlichen Unterricht
ISBN 978-3-8325-0368-0 40.50 EUR
- 31 Marion Budde: Lernwirkungen in der Quanten-Atom-Physik. *Fallstudien über Resonanzen zwischen Lernangeboten und SchülerInnen-Vorstellungen*
ISBN 978-3-8325-0483-0 40.50 EUR
- 32 Thomas Reyer: Oberflächenmerkmale und Tiefenstrukturen im Unterricht. *Exemplarische Analysen im Physikunterricht der gymnasialen Sekundarstufe*
ISBN 978-3-8325-0488-5 40.50 EUR
- 33 Christoph Thomas Müller: Subjektive Theorien und handlungsleitende Kognitionen von Lehrern als Determinanten schulischer Lehr-Lern-Prozesse im Physikunterricht
ISBN 978-3-8325-0543-1 40.50 EUR
- 34 Gabriela Jonas-Ahrend: Physiklehrvorstellungen zum Experiment im Physikunterricht
ISBN 978-3-8325-0576-9 40.50 EUR
- 35 Dimitrios Stavrou: Das Zusammenspiel von Zufall und Gesetzmäßigkeiten in der nicht-linearen Dynamik. *Didaktische Analyse und Lernprozesse*
ISBN 978-3-8325-0609-4 40.50 EUR
- 36 Katrin Engeln: Schülerlabors: authentische, aktivierende Lernumgebungen als Möglichkeit, Interesse an Naturwissenschaften und Technik zu wecken
ISBN 978-3-8325-0689-6 40.50 EUR
- 37 Susann Hartmann: Erklärungsvielfalt
ISBN 978-3-8325-0730-5 40.50 EUR

- 38 Knut Neumann: Didaktische Rekonstruktion eines physikalischen Praktikums für Physiker
ISBN 978-3-8325-0762-6 40.50 EUR
- 39 Michael Späth: Kontextbedingungen für Physikunterricht an der Hauptschule. *Möglichkeiten und Ansatzpunkte für einen fachübergreifenden, handlungsorientierten und berufsorientierten Unterricht*
ISBN 978-3-8325-0827-2 40.50 EUR
- 40 Jörg Hirsch: Interesse, Handlungen und situatives Erleben von Schülerinnen und Schülern beim Bearbeiten physikalischer Aufgaben
ISBN 978-3-8325-0875-3 40.50 EUR
- 41 Monika Hüther: Evaluation einer hypermedialen Lernumgebung zum Thema Gasgesetze. *Eine Studie im Rahmen des Physikpraktikums für Studierende der Medizin*
ISBN 978-3-8325-0911-8 40.50 EUR
- 42 Maike Tesch: Das Experiment im Physikunterricht. *Didaktische Konzepte und Ergebnisse einer Videostudie*
ISBN 978-3-8325-0975-0 40.50 EUR
- 43 Nina Nicolai: Skriptgeleitete Eltern-Kind-Interaktion bei Chemiehausaufgaben. *Eine Evaluationsstudie im Themenbereich Säure-Base*
ISBN 978-3-8325-1013-8 40.50 EUR
- 44 Antje Leisner: Entwicklung von Modellkompetenz im Physikunterricht
ISBN 978-3-8325-1020-6 40.50 EUR
- 45 Stefan Rumann: Evaluation einer Interventionsstudie zur Säure-Base-Thematik
ISBN 978-3-8325-1027-5 40.50 EUR
- 46 Thomas Wilhelm: Konzeption und Evaluation eines Kinematik/Dynamik-Lehrgangs zur Veränderung von Schülervorstellungen mit Hilfe dynamisch ikonischer Repräsentationen und graphischer Modellbildung – mit CD-ROM
ISBN 978-3-8325-1046-6 45.50 EUR
- 47 Andrea Maier-Richter: Computerunterstütztes Lernen mit Lösungsbeispielen in der Chemie. *Eine Evaluationsstudie im Themenbereich Löslichkeit*
ISBN 978-3-8325-1046-6 40.50 EUR
- 48 Jochen Peuckert: Stabilität und Ausprägung kognitiver Strukturen zum Atombegriff
ISBN 978-3-8325-1104-3 40.50 EUR
- 49 Maik Walpuski: Optimierung von experimenteller Kleingruppenarbeit durch Strukturierungshilfen und Feedback
ISBN 978-3-8325-1184-5 40.50 EUR
- 50 Helmut Fischler, Christiane S. Reiners (Hrsg.): Die Teilchenstruktur der Materie im Physik- und Chemieunterricht
ISBN 978-3-8325-1225-5 34.90 EUR
- 51 Claudia Eysel: Interdisziplinäres Lehren und Lernen in der Lehrerbildung. *Eine empirische Studie zum Kompetenzerwerb in einer komplexen Lernumgebung*
ISBN 978-3-8325-1238-5 40.50 EUR

- 52 Johannes Günther: Lehrerfortbildung über die Natur der Naturwissenschaften. *Studien über das Wissenschaftsverständnis von Grundschullehrkräften*
ISBN 978-3-8325-1287-3 40.50 EUR
- 53 Christoph Neugebauer: Lernen mit Simulationen und der Einfluss auf das Problemlösen in der Physik
ISBN 978-3-8325-1300-9 40.50 EUR
- 54 Andreas Schnirch: Gendergerechte Interessen- und Motivationsförderung im Kontext naturwissenschaftlicher Grundbildung. *Konzeption, Entwicklung und Evaluation einer multimedial unterstützten Lernumgebung*
ISBN 978-3-8325-1334-4 40.50 EUR
- 55 Hilde Köster: Freies Explorieren und Experimentieren. *Eine Untersuchung zur selbstbestimmten Gewinnung von Erfahrungen mit physikalischen Phänomenen im Sachunterricht*
ISBN 978-3-8325-1348-1 40.50 EUR
- 56 Eva Heran-Dörr: Entwicklung und Evaluation einer Lehrerfortbildung zur Förderung der physikdidaktischen Kompetenz von Sachunterrichtslehrkräften
ISBN 978-3-8325-1377-1 40.50 EUR
- 57 Agnes Szabone Varnai: Unterstützung des Problemlösens in Physik durch den Einsatz von Simulationen und die Vorgabe eines strukturierten Kooperationsformats
ISBN 978-3-8325-1403-7 40.50 EUR
- 58 Johannes Rethfeld: Aufgabenbasierte Lernprozesse in selbstorganisationsoffenem Unterricht der Sekundarstufe I zum Themengebiet ELEKTROSTATIK. *Eine Feldstudie in vier 10. Klassen zu einer kartenbasierten Lernumgebung mit Aufgaben aus der Elektrostatik*
ISBN 978-3-8325-1416-7 40.50 EUR
- 59 Christian Henke: Experimentell-naturwissenschaftliche Arbeitsweisen in der Oberstufe. *Untersuchung am Beispiel des HIGHSEA-Projekts in Bremerhaven*
ISBN 978-3-8325-1515-7 40.50 EUR
- 60 Lutz Kasper: Diskursiv-narrative Elemente für den Physikunterricht. *Entwicklung und Evaluation einer multimedialen Lernumgebung zum Erdmagnetismus*
ISBN 978-3-8325-1537-9 40.50 EUR
- 61 Thorid Rabe: Textgestaltung und Aufforderung zu Selbsterklärungen beim Physiklernen mit Multimedia
ISBN 978-3-8325-1539-3 40.50 EUR
- 62 Ina Glemnitz: Vertikale Vernetzung im Chemieunterricht. *Ein Vergleich von traditionellem Unterricht mit Unterricht nach Chemie im Kontext*
ISBN 978-3-8325-1628-4 40.50 EUR
- 63 Erik Einhaus: Schülerkompetenzen im Bereich Wärmelehre. *Entwicklung eines Testinstruments zur Überprüfung und Weiterentwicklung eines normativen Modells fachbezogener Kompetenzen*
ISBN 978-3-8325-1630-7 40.50 EUR

- 64 Jasmin Neuroth: Concept Mapping als Lernstrategie. *Eine Interventionsstudie zum Chemielernen aus Texten*
ISBN 978-3-8325-1659-8 40.50 EUR
- 65 Hans Gerd Hegeler-Burkhardt: Zur Kommunikation von Hauptschülerinnen und Hauptschülern in einem handlungsorientierten und fächerübergreifenden Unterricht mit physikalischen und technischen Inhalten
ISBN 978-3-8325-1667-3 40.50 EUR
- 66 Karsten Rincke: Sprachentwicklung und Fachlernen im Mechanikunterricht. *Sprache und Kommunikation bei der Einführung in den Kraftbegriff*
ISBN 978-3-8325-1699-4 40.50 EUR
- 67 Nina Strehle: Das Ion im Chemieunterricht. *Alternative Schülervorstellungen und curriculare Konsequenzen*
ISBN 978-3-8325-1710-6 40.50 EUR
- 68 Martin Hopf: Problemorientierte Schülerexperimente
ISBN 978-3-8325-1711-3 40.50 EUR
- 69 Anne Beerenwinkel: Fostering conceptual change in chemistry classes using expository texts
ISBN 978-3-8325-1721-2 40.50 EUR
- 70 Roland Berger: Das Gruppenpuzzle im Physikunterricht der Sekundarstufe II. *Eine empirische Untersuchung auf der Grundlage der Selbstbestimmungstheorie der Motivation*
ISBN 978-3-8325-1732-8 40.50 EUR
- 71 Giuseppe Colicchia: Physikunterricht im Kontext von Medizin und Biologie. *Entwicklung und Erprobung von Unterrichtseinheiten*
ISBN 978-3-8325-1746-5 40.50 EUR
- 72 Sandra Winheller: Geschlechtsspezifische Auswirkungen der Lehrer-Schüler-Interaktion im Chemieanfangsunterricht
ISBN 978-3-8325-1757-1 40.50 EUR
- 73 Isabel Wahser: Training von naturwissenschaftlichen Arbeitsweisen zur Unterstützung experimenteller Kleingruppenarbeit im Fach Chemie
ISBN 978-3-8325-1815-8 40.50 EUR
- 74 Claus Brell: Lernmedien und Lernerfolg - reale und virtuelle Materialien im Physikunterricht. *Empirische Untersuchungen in achten Klassen an Gymnasien (Laborstudie) zum Computereinsatz mit Simulation und IBE*
ISBN 978-3-8325-1829-5 40.50 EUR
- 75 Rainer Wackermann: Überprüfung der Wirksamkeit eines Basismodell-Trainings für Physiklehrer
ISBN 978-3-8325-1882-0 40.50 EUR
- 76 Oliver Tepner: Effektivität von Aufgaben im Chemieunterricht der Sekundarstufe I
ISBN 978-3-8325-1919-3 40.50 EUR

- 77 Claudia Geyer: Museums- und Science-Center-Besuche im naturwissenschaftlichen Unterricht aus einer motivationalen Perspektive. *Die Sicht von Lehrkräften und Schülerinnen und Schülern*
ISBN 978-3-8325-1922-3 40.50 EUR
- 78 Tobias Leonhard: Professionalisierung in der Lehrerbildung. *Eine explorative Studie zur Entwicklung professioneller Kompetenzen in der Lehrererstausbildung*
ISBN 978-3-8325-1924-7 40.50 EUR
- 79 Alexander Kauertz: Schwierigkeitserzeugende Merkmale physikalischer Leistungstestaufgaben
ISBN 978-3-8325-1925-4 40.50 EUR
- 80 Regina Hübinger: Schüler auf Weltreise. *Entwicklung und Evaluation von Lehr-/Lernmaterialien zur Förderung experimentell-naturwissenschaftlicher Kompetenzen für die Jahrgangsstufen 5 und 6*
ISBN 978-3-8325-1932-2 40.50 EUR
- 81 Christine Waltner: Physik lernen im Deutschen Museum
ISBN 978-3-8325-1933-9 40.50 EUR
- 82 Torsten Fischer: Handlungsmuster von Physiklehrkräften beim Einsatz neuer Medien. *Fallstudien zur Unterrichtspraxis*
ISBN 978-3-8325-1948-3 42.00 EUR
- 83 Corinna Kieren: Chemiehausaufgaben in der Sekundarstufe I des Gymnasiums. *Fragebogenerhebung zur gegenwärtigen Praxis und Entwicklung eines optimierten Hausaufgabendesigns im Themenbereich Säure-Base*
978-3-8325-1975-9 37.00 EUR
- 84 Marco Thiele: Modelle der Thermohalinen Zirkulation im Unterricht. *Eine empirische Studie zur Förderung des Modellverständnisses*
ISBN 978-3-8325-1982-7 40.50 EUR
- 85 Bernd Zinn: Physik lernen, um Physik zu lehren. *Eine Möglichkeit für interessanteren Physikunterricht*
ISBN 978-3-8325-1995-7 39.50 EUR
- 86 Esther Klaes: Außerschulische Lernorte im naturwissenschaftlichen Unterricht. *Die Perspektive der Lehrkraft*
ISBN 978-3-8325-2006-9 43.00 EUR
- 87 Marita Schmidt: Kompetenzmodellierung und -diagnostik im Themengebiet Energie der Sekundarstufe I. *Entwicklung und Erprobung eines Testinventars*
ISBN 978-3-8325-2024-3 37.00 EUR
- 88 Gudrun Franke-Braun: Aufgaben mit gestuften Lernhilfen. *Ein Aufgabenformat zur Förderung der sachbezogenen Kommunikation und Lernleistung für den naturwissenschaftlichen Unterricht*
ISBN 978-3-8325-2026-7 38.00 EUR
- 89 Silke Klos: Kompetenzförderung im naturwissenschaftlichen Anfangsunterricht. *Der Einfluss eines integrierten Unterrichtskonzepts*
ISBN 978-3-8325-2133-2 37.00 EUR

- 90 Ulrike Elisabeth Burkard: Quantenphysik in der Schule. *Bestandsaufnahme, Perspektiven und Weiterentwicklungsmöglichkeiten durch die Implementation eines Medienservers*
ISBN 978-3-8325-2215-5 43.00 EUR
- 91 Ulrike Gromadecki: Argumente in physikalischen Kontexten. *Welche Geltungsgründe halten Physikanfänger für überzeugend?*
ISBN 978-3-8325-2250-6 41.50 EUR
- 92 Jürgen Bruns: Auf dem Weg zur Förderung naturwissenschaftsspezifischer Vorstellungen von zukünftigen Chemie-Lehrenden
ISBN 978-3-8325-2257-5 43.50 EUR
- 93 Cornelius Marsch: Räumliche Atomvorstellung. *Entwicklung und Erprobung eines Unterrichtskonzeptes mit Hilfe des Computers*
ISBN 978-3-8325-2293-3 82.50 EUR
- 94 Maja Brückmann: Sachstrukturen im Physikunterricht. *Ergebnisse einer Videostudie*
ISBN 978-3-8325-2272-8 39.50 EUR
- 95 Sabine Fechner: Effects of Context-oriented Learning on Student Interest and Achievement in Chemistry Education
ISBN 978-3-8325-2343-5 36.50 EUR
- 96 Clemens Nagel: eLearning im Physikalischen Anfängerpraktikum
ISBN 978-3-8325-2355-8 39.50 EUR
- 97 Josef Riese: Professionelles Wissen und professionelle Handlungskompetenz von (angehenden) Physiklehrkräften
ISBN 978-3-8325-2376-3 39.00 EUR
- 98 Sascha Bernholt: Kompetenzmodellierung in der Chemie. *Theoretische und empirische Reflexion am Beispiel des Modells hierarchischer Komplexität*
ISBN 978-3-8325-2447-0 40.00 EUR
- 99 Holger Christoph Stawitz: Auswirkung unterschiedlicher Aufgabenprofile auf die Schülerleistung. *Vergleich von Naturwissenschafts- und Problemlöseaufgaben der PISA 2003-Studie*
ISBN 978-3-8325-2451-7 37.50 EUR
- 100 Hans Ernst Fischer, Elke Sumfleth (Hrsg.): nwu-essen – 10 Jahre Essener Forschung zum naturwissenschaftlichen Unterricht
ISBN 978-3-8325-3331-1 40.00 EUR
- 101 Hendrik Härtig: Sachstrukturen von Physikschulbüchern als Grundlage zur Bestimmung der Inhaltsvalidität eines Tests
ISBN 978-3-8325-2512-5 34.00 EUR
- 102 Thomas Grüß-Niehaus: Zum Verständnis des Löslichkeitskonzeptes im Chemieunterricht. *Der Effekt von Methoden progressiver und kollaborativer Reflexion*
ISBN 978-3-8325-2537-8 40.50 EUR
- 103 Patrick Bronner: Quantenoptische Experimente als Grundlage eines Curriculums zur Quantenphysik des Photons

- ISBN 978-3-8325-2540-8 36.00 EUR
- 104 Adrian Voßkühler: Blickbewegungsmessung an Versuchsaufbauten. *Studien zur Wahrnehmung, Verarbeitung und Usability von physikbezogenen Experimenten am Bildschirm und in der Realität*
ISBN 978-3-8325-2548-4 47.50 EUR
- 105 Verena Tobias: Newton'sche Mechanik im Anfangsunterricht. *Die Wirksamkeit einer Einführung über die zweidimensionale Dynamik auf das Lehren und Lernen*
ISBN 978-3-8325-2558-3 54.00 EUR
- 106 Christian Rogge: Entwicklung physikalischer Konzepte in aufgabenbasierten Lernumgebungen
ISBN 978-3-8325-2574-3 45.00 EUR
- 107 Mathias Ropohl: Modellierung von Schülerkompetenzen im Basiskonzept Chemische Reaktion. *Entwicklung und Analyse von Testaufgaben*
ISBN 978-3-8325-2609-2 36.50 EUR
- 108 Christoph Kulgemeyer: Physikalische Kommunikationskompetenz. *Modellierung und Diagnostik*
ISBN 978-3-8325-2674-0 44.50 EUR
- 109 Jennifer Olszewski: The Impact of Physics Teachers' Pedagogical Content Knowledge on Teacher Actions and Student Outcomes
ISBN 978-3-8325-2680-1 33.50 EUR
- 110 Annika Ohle: Primary School Teachers' Content Knowledge in Physics and its Impact on Teaching and Students' Achievement
ISBN 978-3-8325-2684-9 36.50 EUR
- 111 Susanne Mannel: Assessing scientific inquiry. *Development and evaluation of a test for the low-performing stage*
ISBN 978-3-8325-2761-7 40.00 EUR
- 112 Michael Plomer: Physik physiologisch passend praktiziert. *Eine Studie zur Lernwirksamkeit von traditionellen und adressatenspezifischen Physikpraktika für die Physiologie*
ISBN 978-3-8325-2804-1 34.50 EUR
- 113 Alexandra Schulz: Experimentierspezifische Qualitätsmerkmale im Chemieunterricht. *Eine Videostudie*
ISBN 978-3-8325-2817-1 40.00 EUR
- 114 Franz Boczianowski: Eine empirische Untersuchung zu Vektoren im Physikunterricht der Mittelstufe
ISBN 978-3-8325-2843-0 39.50 EUR
- 115 Maria Ploog: Internetbasiertes Lernen durch Textproduktion im Fach Physik
ISBN 978-3-8325-2853-9 39.50 EUR

- 116 Anja Dhein: Lernen in Explorier- und Experimentiersituationen. *Eine explorative Studie zu Bedeutungsentwicklungsprozessen bei Kindern im Alter zwischen 4 und 6 Jahren*
ISBN 978-3-8325-2859-1 45.50 EUR
- 117 Irene Neumann: Beyond Physics Content Knowledge. *Modeling Competence Regarding Nature of Scientific Inquiry and Nature of Scientific Knowledge*
ISBN 978-3-8325-2880-5 37.00 EUR
- 118 Markus Emden: Prozessorientierte Leistungsmessung des naturwissenschaftlich-experimentellen Arbeitens. *Eine vergleichende Studie zu Diagnoseinstrumenten zu Beginn der Sekundarstufe I*
ISBN 978-3-8325-2867-6 38.00 EUR
- 119 Birgit Hofmann: Analyse von Blickbewegungen von Schülern beim Lesen von physikbezogenen Texten mit Bildern. *Eye Tracking als Methodenwerkzeug in der physikdidaktischen Forschung*
ISBN 978-3-8325-2925-3 59.00 EUR
- 120 Rebecca Knobloch: Analyse der fachinhaltlichen Qualität von Schüleräußerungen und deren Einfluss auf den Lernerfolg. *Eine Videostudie zu kooperativer Kleingruppenarbeit*
ISBN 978-3-8325-3006-8 36.50 EUR
- 121 Julia Hostenbach: Entwicklung und Prüfung eines Modells zur Beschreibung der Bewertungskompetenz im Chemieunterricht
ISBN 978-3-8325-3013-6 38.00 EUR
- 122 Anna Windt: Naturwissenschaftliches Experimentieren im Elementarbereich. *Evaluation verschiedener Lernsituationen*
ISBN 978-3-8325-3020-4 43.50 EUR
- 123 Eva Kölbach: Kontexteinflüsse beim Lernen mit Lösungsbeispielen
ISBN 978-3-8325-3025-9 38.50 EUR
- 124 Anna Lau: Passung und vertikale Vernetzung im Chemie- und Physikunterricht
ISBN 978-3-8325-3021-1 36.00 EUR
- 125 Jan Lamprecht: Ausbildungswege und Komponenten professioneller Handlungskompetenz. *Vergleich von Quereinsteigern mit Lehramtsabsolventen für Gymnasien im Fach Physik*
ISBN 978-3-8325-3035-8 38.50 EUR
- 126 Ulrike Böhm: Förderung von Verstehensprozessen unter Einsatz von Modellen
ISBN 978-3-8325-3042-6 41.00 EUR
- 127 Sabrina Dollny: Entwicklung und Evaluation eines Testinstruments zur Erfassung des fachspezifischen Professionswissens von Chemielehrkräften
ISBN 978-3-8325-3046-4 37.00 EUR

- 128 Monika Zimmermann: Naturwissenschaftliche Bildung im Kindergarten. *Eine integrative Längsschnittstudie zur Kompetenzentwicklung von Erzieherinnen*
ISBN 978-3-8325-3053-2 54.00 EUR
- 129 Ulf Saballus: Über das Schlussfolgern von Schülerinnen und Schülern zu öffentlichen Kontroversen mit naturwissenschaftlichem Hintergrund. *Eine Fallstudie*
ISBN 978-3-8325-3086-0 39.50 EUR
- 130 Olaf Krey: Zur Rolle der Mathematik in der Physik. *Wissenschaftstheoretische Aspekte und Vorstellungen Physiklernender*
ISBN 978-3-8325-3101-0 46.00 EUR
- 131 Angelika Wolf: Zusammenhänge zwischen der Eigenständigkeit im Physikunterricht, der Motivation, den Grundbedürfnissen und dem Lernerfolg von Schülern
ISBN 978-3-8325-3161-4 45.00 EUR
- 132 Johannes Börlin: Das Experiment als Lerngelegenheit. *Vom interkulturellen Vergleich des Physikunterrichts zu Merkmalen seiner Qualität*
ISBN 978-3-8325-3170-6 45.00 EUR
- 133 Olaf Uhden: Mathematisches Denken im Physikunterricht. *Theorieentwicklung und Problemanalyse*
ISBN 978-3-8325-3170-6 45.00 EUR
- 134 Christoph Gut: Modellierung und Messung experimenteller Kompetenz. *Analyse eines large-scale Experimentiertests*
ISBN 978-3-8325-3213-0 40.00 EUR
- 135 Antonio Rueda: Lernen mit ExploMultimedial in kolumbianischen Schulen. *Analyse von kurzzeitigen Lernprozessen und der Motivation beim länderübergreifenden Einsatz einer deutschen computergestützten multimedialen Lernumgebung für den naturwissenschaftlichen Unterricht*
ISBN 978-3-8325-3218-5 45.50 EUR
- 136 Krisztina Berger: Bilder, Animationen und Notizen. *Empirische Untersuchung zur Wirkung einfacher visueller Repräsentationen und Notizen auf den Wissenserwerb in der Optik*
ISBN 978-3-8325-3238-3 41.50 EUR
- 137 Antony Crossley: Untersuchung des Einflusses unterschiedlicher physikalischer Konzepte auf den Wissenserwerb in der Thermodynamik der Sekundarstufe I
ISBN 978-3-8325-3275-8 40.00 EUR
- 138 Tobias Viering: Entwicklung physikalischer Kompetenz in der Sekundarstufe I. *Validierung eines Kompetenzentwicklungsmodells für das Energiekonzept im Bereich Fachwissen*
ISBN 978-3-8325-3277-2 37.00 EUR
- 139 Nico Schreiber: Diagnostik experimenteller Kompetenz. *Validierung technologiegestützter Testverfahren im Rahmen eines Kompetenzstrukturmodells*
ISBN 978-3-8325-3284-0 39.00 EUR

- 140 Sarah Hundertmark: Einblicke in kollaborative Lernprozesse. *Eine Fallstudie zur reflektierenden Zusammenarbeit unterstützt durch die Methoden Concept Mapping und Lernbegleitbogen*
ISBN 978-3-8325-3251-2 43.00 EUR
- 141 Ronny Scherer: Analyse der Struktur, Messinvarianz und Ausprägung komplexer Problemlösekompetenz im Fach Chemie. *Eine Querschnittstudie in der Sekundarstufe I und am Übergang zur Sekundarstufe II*
ISBN 978-3-8325-3312-0 43.00 EUR
- 142 Patricia Heitmann: Bewertungskompetenz im Rahmen naturwissenschaftlicher Problemlöseprozesse. *Modellierung und Diagnose der Kompetenzen Bewertung und analytisches Problemlösen für das Fach Chemie*
ISBN 978-3-8325-3314-4 37.00 EUR
- 143 Jan Fleischhauer: Wissenschaftliches Argumentieren und Entwicklung von Konzepten beim Lernen von Physik
ISBN 978-3-8325-3325-0 35.00 EUR
- 144 Nermin Özcan: Zum Einfluss der Fachsprache auf die Leistung im Fach Chemie. *Eine Förderstudie zur Fachsprache im Chemieunterricht*
ISBN 978-3-8325-3328-1 36.50 EUR
- 145 Helena van Vorst: Kontextmerkmale und ihr Einfluss auf das Schülerinteresse im Fach Chemie
ISBN 978-3-8325-3321-2 38.50 EUR
- 146 Janine Cappell: Fachspezifische Diagnosekompetenz angehender Physiklehrkräfte in der ersten Ausbildungsphase
ISBN 978-3-8325-3356-4 38.50 EUR
- 147 Susanne Bley: Förderung von Transferprozessen im Chemieunterricht
ISBN 978-3-8325-3407-3 40.50 EUR
- 148 Cathrin Blaes: Die übungsgestützte Lehrerpräsentation im Chemieunterricht der Sekundarstufe I. *Evaluation der Effektivität*
ISBN 978-3-8325-3409-7 43.50 EUR
- 149 Julia Suckut: Die Wirksamkeit von piko-OWL als Lehrerfortbildung. Eine Evaluation zum Projekt *Physik im Kontext* in Fallstudien
ISBN 978-3-8325-3440-0 45.00 EUR
- 150 Alexandra Dorschu: Die Wirkung von Kontexten in Physikkompetenztestaufgaben
ISBN 978-3-8325-3446-2 37.00 EUR
- 151 Jochen Scheid: Multiple Repräsentationen, Verständnis physikalischer Experimente und kognitive Aktivierung: *Ein Beitrag zur Entwicklung der Aufgabenkultur*
ISBN 978-3-8325-3449-3 49.00 EUR
- 152 Tim Plasa: Die Wahrnehmung von Schülerlaboren und Schülerforschungszentren
ISBN 978-3-8325-3483-7 35.50 EUR

- 153 Felix Schoppmeier: Physikkompetenz in der gymnasialen Oberstufe. *Entwicklung und Validierung eines Kompetenzstrukturmodells für den Kompetenzbereich Umgang mit Fachwissen*
ISBN 978-3-8325-3502-5 36.00 EUR
- 154 Katharina Groß: Experimente alternativ dokumentieren. *Eine qualitative Studie zur Förderung der Diagnose- und Differenzierungskompetenz in der Chemielehrerbildung*
ISBN 978-3-8325-3508-7 43.50 EUR
- 155 Barbara Hank: Konzeptwandelprozesse im Anfangsunterricht Chemie. *Eine quasixperimentelle Längsschnittstudie*
ISBN 978-3-8325-3519-3 38.50 EUR
- 156 Katja Freyer: Zum Einfluss von Studieneingangsvoraussetzungen auf den Studienerfolg Erstsemesterstudierender im Fach Chemie
ISBN 978-3-8325-3544-5 38.00 EUR
- 157 Alexander Rachel: Auswirkungen instruktionaler Hilfen bei der Einführung des (Ferro-)Magnetismus. *Eine Vergleichsstudie in der Primar- und Sekundarstufe*
ISBN 978-3-8325-3548-3 43.50 EUR
- 158 Sebastian Ritter: Einfluss des Lerninhalts Nanogrößeneffekte auf Teilchen- und Teilchenmodellvorstellungen von Schülerinnen und Schülern
ISBN 978-3-8325-3558-2 36.00 EUR
- 159 Andrea Harbach: Problemorientierung und Vernetzung in kontextbasierten Lernaufgaben
ISBN 978-3-8325-3564-3 39.00 EUR
- 160 David Obst: Interaktive Tafeln im Physikunterricht. *Entwicklung und Evaluation einer Lehrerfortbildung*
ISBN 978-3-8325-3582-7 40.50 EUR
- 161 Sophie Kirschner: Modellierung und Analyse des Professionswissens von Physiklehrkräften
ISBN 978-3-8325-3601-5 35.00 EUR
- 162 Katja Stief: Selbstregulationsprozesse und Hausaufgabenmotivation im Chemieunterricht
ISBN 978-3-8325-3631-2 34.00 EUR
- 163 Nicola Meschede: Professionelle Wahrnehmung der inhaltlichen Strukturierung im naturwissenschaftlichen Grundschulunterricht. *Theoretische Beschreibung und empirische Erfassung*
ISBN 978-3-8325-3668-8 37.00 EUR
- 164 Johannes Maximilian Barth: Experimentieren im Physikunterricht der gymnasialen Oberstufe. *Eine Rekonstruktion übergeordneter Einbettungsstrategien*
ISBN 978-3-8325-3681-7 39.00 EUR
- 165 Sandra Lein: Das Betriebspraktikum in der Lehrerbildung. *Eine Untersuchung zur Förderung der Wissenschafts- und Technikbildung im allgemeinbildenden Unterricht*
ISBN 978-3-8325-3698-5 40.00 EUR

- 166 Veranika Maiseyenko: Modellbasiertes Experimentieren im Unterricht. *Praxistauglichkeit und Lernwirkungen*
ISBN 978-3-8325-3708-1 38.00 EUR
- 167 Christoph Stolzenberger: Der Einfluss der didaktischen Lernumgebung auf das Erreichen geforderter Bildungsziele am Beispiel der W- und P-Seminare im Fach Physik
ISBN 978-3-8325-3708-1 38.00 EUR
- 168 Pia Altenburger: Mehrebenenregressionsanalysen zum Physiklernen im Sachunterricht der Primarstufe. *Ergebnisse einer Evaluationsstudie.*
ISBN 978-3-8325-3717-3 37.50 EUR
- 169 Nora Ferber: Entwicklung und Validierung eines Testinstruments zur Erfassung von Kompetenzentwicklung im Fach Chemie in der Sekundarstufe I
ISBN 978-3-8325-3727-2 39.50 EUR
- 170 Anita Stender: Unterrichtsplanung: Vom Wissen zum Handeln. Theoretische Entwicklung und empirische Überprüfung des Transformationsmodells der Unterrichtsplanung
ISBN 978-3-8325-3750-0 41.50 EUR
- 171 Jenna Koenen: Entwicklung und Evaluation von experimentunterstützten Lösungsbeispielen zur Förderung naturwissenschaftlich-experimenteller Arbeitsweisen
ISBN 978-3-8325-3785-2 43.00 EUR
- 172 Teresa Henning: Empirische Untersuchung kontextorientierter Lernumgebungen in der Hochschuldidaktik. *Entwicklung und Evaluation kontextorientierter Aufgaben in der Studieneingangsphase für Fach- und Nebenfachstudierende der Physik*
ISBN 978-3-8325-3801-9 43.00 EUR
- 173 Alexander Pusch: Fachspezifische Instrumente zur Diagnose und individuellen Förderung von Lehramtsstudierenden der Physik
ISBN 978-3-8325-3829-3 38.00 EUR
- 174 Christoph Vogelsang: Validierung eines Instruments zur Erfassung der professionellen Handlungskompetenz von (angehenden) Physiklehrkräften. *Zusammenhangsanalysen zwischen Lehrerkompetenz und Lehrerperformanz*
ISBN 978-3-8325-3846-0 50.50 EUR
- 175 Ingo Brebeck: Selbstreguliertes Lernen in der Studieneingangsphase im Fach Chemie
ISBN 978-3-8325-3859-0 37.00 EUR
- 176 Axel Eghtessad: Merkmale und Strukturen von Professionalisierungsprozessen in der ersten und zweiten Phase der Chemielehrerbildung. *Eine empirisch-qualitative Studie mit niedersächsischen Fachleiter_innen der Sekundarstufenlehrämter*
ISBN 978-3-8325-3861-3 45.00 EUR
- 177 Andreas Nehring: Wissenschaftliche Denk- und Arbeitsweisen im Fach Chemie. Eine kompetenzorientierte Modell- und Testentwicklung für den Bereich der Erkenntnisgewinnung
ISBN 978-3-8325-3872-9 39.50 EUR

- 178 Maike Schmidt: Professionswissen von Sachunterrichtslehrkräften. Zusammenhangsanalyse zur Wirkung von Ausbildungshintergrund und Unterrichtserfahrung auf das fachspezifische Professionswissen im Unterrichtsinhalt „Verbrennung“
ISBN 978-3-8325-3907-8 38.50 EUR
- 179 Jan Winkelmann: Auswirkungen auf den Fachwissenszuwachs und auf affektive Schülermerkmale durch Schüler- und Demonstrationsexperimente im Physikunterricht
ISBN 978-3-8325-3915-3 41.00 EUR
- 180 Iwen Kobow: Entwicklung und Validierung eines Testinstrumentes zur Erfassung der Kommunikationskompetenz im Fach Chemie
ISBN 978-3-8325-3927-6 34.50 EUR
- 181 Yvonne Gramzow: Fachdidaktisches Wissen von Lehramtsstudierenden im Fach Physik. Modellierung und Testkonstruktion
ISBN 978-3-8325-3931-3 42.50 EUR
- 182 Evelin Schröter: Entwicklung der Kompetenzerwartung durch Lösen physikalischer Aufgaben einer multimedialen Lernumgebung
ISBN 978-3-8325-3975-7 54.50 EUR
- 183 Inga Kallweit: Effektivität des Einsatzes von Selbsteinschätzungsbögen im Chemieunterricht der Sekundarstufe I. *Individuelle Förderung durch selbstreguliertes Lernen*
ISBN 978-3-8325-3965-8 44.00 EUR
- 184 Andrea Schumacher: Paving the way towards authentic chemistry teaching. *A contribution to teachers' professional development*
ISBN 978-3-8325-3976-4 48.50 EUR
- 185 David Woitkowski: Fachliches Wissen Physik in der Hochschulausbildung. *Konzeptualisierung, Messung, Niveaubildung*
ISBN 978-3-8325-3988-7 53.00 EUR
- 186 Marianne Korner: Cross-Age Peer Tutoring in Physik. *Evaluation einer Unterrichtsmethode*
ISBN 978-3-8325-3979-5 38.50 EUR
- 187 Simone Nakoinz: Untersuchung zur Verknüpfung submikroskopischer und makroskopischer Konzepte im Fach Chemie
ISBN 978-3-8325-4057-9 38.50 EUR
- 188 Sandra Anus: Evaluation individueller Förderung im Chemieunterricht. *Adaptivität von Lerninhalten an das Vorwissen von Lernenden am Beispiel des Basiskonzeptes Chemische Reaktion*
ISBN 978-3-8325-4059-3 43.50 EUR
- 189 Thomas Roßbegalle: Fachdidaktische Entwicklungsforschung zum besseren Verständnis atmosphärischer Phänomene. *Treibhauseffekt, saurer Regen und stratosphärischer Ozonabbau als Kontexte zur Vermittlung von Basiskonzepten der Chemie*
ISBN 978-3-8325-4059-3 45.50 EUR
- 190 Kathrin Steckenmesser-Sander: Gemeinsamkeiten und Unterschiede physikbezogener Handlungs-, Denk- und Lernprozesse von Mädchen und Jungen
ISBN 978-3-8325-4066-1 38.50 EUR

- 191 Cornelia Geller: Lernprozessorientierte Sequenzierung des Physikunterrichts im Zusammenhang mit Fachwissenserwerb. *Eine Videostudie in Finnland, Deutschland und der Schweiz*
ISBN 978-3-8325-4082-1 35.50 EUR
- 192 Jan Hofmann: Untersuchung des Kompetenzaufbaus von Physiklehrkräften während einer Fortbildungsmaßnahme
ISBN 978-3-8325-4104-0 38.50 EUR
- 193 Andreas Dickhäuser: Chemiespezifischer Humor. *Theoriebildung, Materialentwicklung, Evaluation*
ISBN 978-3-8325-4108-8 37.00 EUR
- 194 Stefan Korte: Die Grenzen der Naturwissenschaft als Thema des Physikunterrichts
ISBN 978-3-8325-4112-5 57.50 EUR
- 195 Carolin Hülsmann: Kurswahlmotive im Fach Chemie. Eine Studie zum Wahlverhalten und Erfolg von Schülerinnen und Schülern in der gymnasialen Oberstufe
ISBN 978-3-8325-4144-6 49.00 EUR
- 196 Caroline Körbs: Mindeststandards im Fach Chemie am Ende der Pflichtschulzeit
ISBN 978-3-8325-4148-4 34.00 EUR
- 197 Andreas Vorholzer: Wie lassen sich Kompetenzen des experimentellen Denkens und Arbeitens fördern? *Eine empirische Untersuchung der Wirkung eines expliziten und eines impliziten Instruktionsansatzes*
ISBN 978-3-8325-4194-1 37.50 EUR
- 198 Anna Katharina Schmitt: Entwicklung und Evaluation einer Chemielehrerfortbildung zum Kompetenzbereich Erkenntnisgewinnung
ISBN 978-3-8325-4228-3 39.50 EUR
- 199 Christian Maurer: Strukturierung von Lehr-Lern-Sequenzen
ISBN 978-3-8325-4247-4 36.50 EUR
- 200 Helmut Fischler, Elke Sumfleth (Hrsg.): Professionelle Kompetenz von Lehrkräften der Chemie und Physik
ISBN 978-3-8325-4523-9 34.00 EUR
- 201 Simon Zander: Lehrerfortbildung zu Basismodellen und Zusammenhänge zum Fachwissen
ISBN 978-3-8325-4248-1 35.00 EUR
- 202 Kerstin Arndt: Experimentierkompetenz erfassen.
Analyse von Prozessen und Mustern am Beispiel von Lehramtsstudierenden der Chemie
ISBN 978-3-8325-4266-5 45.00 EUR
- 203 Christian Lang: Kompetenzorientierung im Rahmen experimentalchemischer Praktika
ISBN 978-3-8325-4268-9 42.50 EUR
- 204 Eva Cauet: Testen wir relevantes Wissen? *Zusammenhang zwischen dem Professionswissen von Physiklehrkräften und gutem und erfolgreichem Unterrichten*
ISBN 978-3-8325-4276-4 39.50 EUR

- 205 Patrick Löffler: Modellanwendung in Problemlöseaufgaben. *Wie wirkt Kontext?*
ISBN 978-3-8325-4303-7 35.00 EUR
- 206 Carina Gehlen: Kompetenzstruktur naturwissenschaftlicher Erkenntnisgewinnung im Fach Chemie
ISBN 978-3-8325-4318-1 43.00 EUR
- 207 Lars Oettinghaus: Lehrerüberzeugungen und physikbezogenes Professionswissen. *Vergleich von Absolventinnen und Absolventen verschiedener Ausbildungswege im Physikreferendariat*
ISBN 978-3-8325-4319-8 38.50 EUR
- 208 Jennifer Petersen: Zum Einfluss des Merkmals Humor auf die Gesundheitsförderung im Chemieunterricht der Sekundarstufe I. *Eine Interventionsstudie zum Thema Sonnenschutz*
ISBN 978-3-8325-4348-8 40.00 EUR
- 209 Philipp Straube: Modellierung und Erfassung von Kompetenzen naturwissenschaftlicher Erkenntnisgewinnung bei (Lehramts-) Studierenden im Fach Physik
ISBN 978-3-8325-4351-8 35.50 EUR
- 210 Martin Dickmann: Messung von Experimentierfähigkeiten. *Validierungsstudien zur Qualität eines computerbasierten Testverfahrens*
ISBN 978-3-8325-4356-3 41.00 EUR
- 211 Markus Bohlmann: Science Education. Empirie, Kulturen und Mechanismen der Didaktik der Naturwissenschaften
ISBN 978-3-8325-4377-8 44.00 EUR
- 212 Martin Draude: Die Kompetenz von Physiklehrkräften, Schwierigkeiten von Schülerinnen und Schülern beim eigenständigen Experimentieren zu diagnostizieren
ISBN 978-3-8325-4382-2 37.50 EUR
- 213 Henning Rode: Prototypen evidenzbasierten Physikunterrichts. *Zwei empirische Studien zum Einsatz von Feedback und Blackboxes in der Sekundarstufe*
ISBN 978-3-8325-4389-1 42.00 EUR
- 214 Jan-Henrik Kechel: Schülerschwierigkeiten beim eigenständigen Experimentieren. *Eine qualitative Studie am Beispiel einer Experimentieraufgabe zum Hooke'schen Gesetz*
ISBN 978-3-8325-4392-1 55.00 EUR
- 215 Katharina Fricke: Classroom Management and its Impact on Lesson Outcomes in Physics. *A multi-perspective comparison of teaching practices in primary and secondary schools*
ISBN 978-3-8325-4394-5 40.00 EUR
- 216 Hannes Sander: Orientierungen von Jugendlichen beim Urteilen und Entscheiden in Kontexten nachhaltiger Entwicklung. *Eine rekonstruktive Perspektive auf Bewertungskompetenz in der Didaktik der Naturwissenschaft*
ISBN 978-3-8325-4434-8 46.00 EUR

- 217 Inka Haak: Maßnahmen zur Unterstützung kognitiver und metakognitiver Prozesse in der Studieneingangsphase. *Eine Design-Based-Research-Studie zum universitären Lernzentrum Physiktreff*
ISBN 978-3-8325-4437-9 46.50 EUR
- 218 Martina Brandenburger: Was beeinflusst den Erfolg beim Problemlösen in der Physik? *Eine Untersuchung mit Studierenden*
ISBN 978-3-8325-4409-6 42.50 EUR
- 219 Corinna Helms: Entwicklung und Evaluation eines Trainings zur Verbesserung der Erklärqualität von Schülerinnen und Schülern im Gruppenpuzzle
ISBN 978-3-8325-4454-6 42.50 EUR
- 220 Viktoria Rath: Diagnostische Kompetenz von angehenden Physiklehrkräften. *Modellierung, Testinstrumentenentwicklung und Erhebung der Performanz bei der Diagnose von Schülervorstellungen in der Mechanik*
ISBN 978-3-8325-4456-0 42.50 EUR
- 221 Janne Krüger: Schülerperspektiven auf die zeitliche Entwicklung der Naturwissenschaften
ISBN 978-3-8325-4457-7 45.50 EUR
- 222 Stefan Mutke: Das Professionswissen von Chemiereferendarinnen und -referendaren in Nordrhein-Westfalen. *Eine Längsschnittstudie*
ISBN 978-3-8325-4458-4 37.50 EUR
- 223 Sebastian Habig: Systematisch variierte Kontextaufgaben und ihr Einfluss auf kognitive und affektive Schülerfaktoren
ISBN 978-3-8325-4467-6 40.50 EUR
- 224 Sven Liepertz: Zusammenhang zwischen dem Professionswissen von Physiklehrkräften, dem sachstrukturellen Angebot des Unterrichts und der Schülerleistung
ISBN 978-3-8325-4480-5 34.00 EUR
- 225 Elina Platova: Optimierung eines Laborpraktikums durch kognitive Aktivierung
ISBN 978-3-8325-4481-2 39.00 EUR
- 226 Tim Reschke: Lesegeschichten im Chemieunterricht der Sekundarstufe I zur Unterstützung von situationalem Interesse und Lernerfolg
ISBN 978-3-8325-4487-4 41.00 EUR
- 227 Lena Mareike Walper: Entwicklung der physikbezogenen Interessen und selbstbezogenen Kognitionen von Schülerinnen und Schülern in der Übergangsphase von der Primar- in die Sekundarstufe. *Eine Längsschnittanalyse vom vierten bis zum siebten Schuljahr*
ISBN 978-3-8325-4495-9 43.00 EUR
- 228 Stefan Anthofer: Förderung des fachspezifischen Professionswissens von Chemielehramtsstudierenden
ISBN 978-3-8325-4498-0 39.50 EUR
- 229 Marcel Bullinger: Handlungsorientiertes Physiklernen mit instruierten Selbsterklärungen in der Primarstufe. *Eine experimentelle Laborstudie*
ISBN 978-3-8325-4504-8 44.00 EUR

- 230 Thomas Amenda: Bedeutung fachlicher Elementarisierungen für das Verständnis der Kinematik
ISBN 978-3-8325-4531-4 43.50 EUR
- 231 Sabrina Milke: Beeinflusst *Priming* das Physiklernen?
Eine empirische Studie zum Dritten Newtonschen Axiom
ISBN 978-3-8325-4549-4 42.00 EUR
- 232 Corinna Erfmann: Ein anschaulicher Weg zum Verständnis der elektromagnetischen Induktion. *Evaluation eines Unterrichtsvorschlags und Validierung eines Leistungsdiagnoseinstruments*
ISBN 978-3-8325-4550-5 49.50 EUR
- 233 Hanne Rautenstrauch: Erhebung des (Fach-)Sprachstandes bei Lehramtsstudierenden im Kontext des Faches Chemie
ISBN 978-3-8325-4556-7 40.50 EUR
- 234 Tobias Klug: Wirkung kontextorientierter physikalischer Praktikumsversuche auf Lernprozesse von Studierenden der Medizin
ISBN 978-3-8325-4558-1 37.00 EUR
- 235 Mareike Bohrmann: Zur Förderung des Verständnisses der Variablenkontrolle im naturwissenschaftlichen Sachunterricht
ISBN 978-3-8325-4559-8 52.00 EUR
- 236 Anja Schödl: FALKO-Physik – Fachspezifische Lehrerkompetenzen im Fach Physik. *Entwicklung und Validierung eines Testinstruments zur Erfassung des fachspezifischen Professionswissens von Physiklehrkräften*
ISBN 978-3-8325-4553-6 40.50 EUR
- 237 Hilda Scheuermann: Entwicklung und Evaluation von Unterstützungsmaßnahmen zur Förderung der Variablenkontrollstrategie beim Planen von Experimenten
ISBN 978-3-8325-4568-0 39.00 EUR
- 238 Christian G. Strippel: Naturwissenschaftliche Erkenntnisgewinnung an chemischen Inhalten vermitteln. *Konzeption und empirische Untersuchung einer Ausstellung mit Experimentierstation*
ISBN 978-3-8325-4577-2 41.50 EUR
- 239 Sarah Rau: Durchführung von Sachunterricht im Vorbereitungsdienst. *Eine längsschnittliche, videobasierte Unterrichtsanalyse*
ISBN 978-3-8325-4579-6 46.00 EUR
- 240 Thomas Plotz: Lernprozesse zu nicht-sichtbarer Strahlung. *Empirische Untersuchungen in der Sekundarstufe 2*
ISBN 978-3-8325-4624-3 39.50 EUR

- 241 Wolfgang Aschauer: Elektrische und magnetische Felder. *Eine empirische Studie zu Lernprozessen in der Sekundarstufe II*
ISBN 978-3-8325-4625-0 50.00 EUR
- 242 Anna Donhauser: Didaktisch rekonstruierte Materialwissenschaft. *Aufbau und Konzeption eines Schülerlabors für den Exzellenzcluster Engineering of Advanced Materials*
ISBN 978-3-8325-4636-6 39.00 EUR
- 243 Katrin Schübler: Lernen mit Lösungsbeispielen im Chemieunterricht. *Einflüsse auf Lernerfolg, kognitive Belastung und Motivation*
ISBN 978-3-8325-4640-3 42.50 EUR

Alle erschienenen Bücher können unter der angegebenen ISBN direkt online (<http://www.logos-verlag.de>) oder per Fax (030 - 42 85 10 92) beim Logos Verlag Berlin bestellt werden.

Studien zum Physik- und Chemielernen

Herausgegeben von Hans Niedderer, Helmut Fischler und Elke Sumfleth

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 in Deutschland.

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:

Prof. Dr. Hans Niedderer
Institut für Didaktik der Naturwissenschaften,
Abt. Physikdidaktik, FB Physik/Elektrotechnik,
Universität Bremen,
Postfach 33 04 40, 28334 Bremen
Tel. 0421-218 2484/4695, e-mail:
niedderer@physik.uni-bremen.de

Prof. Dr. Helmut Fischler
Didaktik der Physik, FB Physik, Freie Universität Berlin,
Arnimallee 14, 14195 Berlin
Tel. 030-838 56712/55966, e-mail:
fischler@physik.fu-berlin.de

Prof. Dr. Elke Sumfleth
Didaktik der Chemie,
Fachbereich Chemie,
Universität Duisburg-Essen,
Schützenbahn 70, 45127 Essen
Tel. 0201-183 3757/3761, e-mail:
elke.sumfleth@uni-essen.de

Ziel des hier vorgestellten Projekts war es, ein Testinstrument zur objektiven, reliablen und validen Messung des fachdidaktischen Wissens (FDW) und des Fachwissens (FW) von Physiklehrkräften der Sekundarstufe zu entwickeln. FALKO-Physik folgte bei der Testkonstruktion einem projektübergreifenden Rahmenmodell, das sich an die COACTIV-Studie anlehnt sowie die Wissenstaxonomie Shulmans berücksichtigt. Dabei wurde im Unterschied zu anderen Testkonstruktionen im Fach Physik besonders darauf Wert gelegt, mehrere zentrale physikalische Themengebiete sowohl im FW-Testteil als auch im FDW-Testteil zu implementieren.

Während der Validierungsphase erfolgte eine Untersuchung von Physiklehrkräften sowie angehenden Physiklehrkräften (Studierenden) der Sekundarstufe I. Die Reliabilität des FW-Testteils war zufriedenstellend ($,82$), das FDW konnte dagegen nur eingeschränkt reliabel gemessen werden ($,65$). Weiterhin wurden Konstruktvalidierungsstudien mit Konstrastgruppen durchgeführt. Während beispielsweise Gymnasiallehrkräfte und Fachphysiker im FW aufgrund der ähnlichen fachlichen universitären Ausbildung erwartungsgemäß vergleichbar abschnitten ($d = 0,10$; $p = ,69$), war der Vorteil der Lehrkräfte im FDW dagegen erheblich größer ($d = 0,72$; $p < ,01$).

Logos Verlag Berlin

ISBN 978-3-8325-4553-6