

Struktur und Entwicklung des Professions- wissens angehender Physiklehrkräfte

Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Christian-Albrechts-Universität zu Kiel

vorgelegt von

Jochen Kröger

Kiel, 2019

Erster Gutachter:
Prof. Dr. Knut Neumann

Zweiter Gutachter:
Prof. Dr. Dietmar Block

Tag der mündlichen Prüfung:
05.02.2019

Teile der vorliegenden Arbeit sind in den folgenden Veröffentlichungen enthalten:

- Kröger, J.**, Euler, M., Neumann, K., Härtig, H., und Petersen, S. (2012). Messung Professioneller Kompetenz im Fach Physik. in S. Bernholt (Hrsg.): *Konzepte fachdidaktischer Strukturierung für den Unterricht: GDGP-Jahrestagung Oldenburg 2011*. (Band 32, S. 616-618). Berlin: Lit.
- Kröger, J.**, Neumann, K., und Petersen, S. (2013). Messung professioneller Kompetenz im Fach Physik. in S. Bernholt (Hrsg.): *Inquiry-based Learning - Forschendes Lernen: Gesellschaft für Didaktik der Chemie und Physik, Jahrestagung in Hannover 2012* (S. 533-535).
- Kröger, J.**, Neumann, K., & Petersen, S. (2014). Erfassung des Professionswissens angehender Physiklehrkräfte im Rahmen des Projekts KiL. in S. Bernholt (Hrsg.): *Naturwissenschaftliche Bildung zwischen Science- und Fachunterricht: Gesellschaft für Didaktik der Chemie und Physik, Jahrestagung in München 2013* (S. 117-119).
- Kleickmann, T., Großschedl, J., Harms, Z., Heinze, A., Herzog, S., Hohenstein, F., Köller, O., **Kröger, J.**, Lindmeier, A., Loch, C., Mahler, D., Möller, J., Neumann, K., Parchmann, I., Steffensky, M., Taskin, V. und Zimmermann, F. (2014). Professionswissen von Lehramtsstudierenden der mathematisch-naturwissenschaftlichen Fächer-Testentwicklung im Rahmen des Projekts KiL. *Unterrichtswissenschaft*, 42(3), 280-288.
- Sorge, S., **Kröger, J.**, Petersen, S. & Neumann, K. (2017). Structure and development of pre-service physics teachers' professional knowledge. *International Journal of Science Education*, 1-28.

Zusammenfassung

Die *professionelle Kompetenz* von Lehrkräften spielt für die Qualität von Unterricht und damit für den Lernerfolg von Schülerinnen und Schülern eine wichtige Rolle. Das professionelle Wissen von Lehrkräften, das sogenannte *Professionswissen*, stellt dabei einen Kernbereich der professionellen Kompetenz dar. Dementsprechend besteht ein zentrales Ziel der Ausbildung angehender Lehrerinnen und Lehrer darin, ein umfangreiches und praxisrelevantes Professionswissen zu vermitteln. Eine systematische Evaluation des Professionswissens angehender Lehrkräfte findet jedoch bisher nicht statt. Ferner ist für das Fach Physik unklar, wie Professionswissen strukturiert ist, wie es sich entwickelt und welche Faktoren eine positive Entwicklung begünstigen können. Über die naturwissenschaftlichen Fächer hinweg fehlt es zudem an etablierten Instrumenten, mit denen die Struktur und Ausprägung des Professionswissens in den einzelnen Fächern untersucht werden könnte. Im Rahmen der vorliegenden Arbeit wurde ein solches Messinstrument für das Professionswissen angehender Physiklehrkräfte entwickelt und erprobt. Die damit verbundenen Forschungsziele sind, jeweils für angehende Physiklehrkräfte, die Untersuchung der Struktur des Professionswissens, die Messung der Ausprägung des Professionswissens in verschiedenen Studienjahren sowie die Untersuchung des Zusammenhangs zwischen Personenmerkmalen und der Ausprägung des Professionswissens.

Das im Rahmen der vorliegenden Arbeit schrittweise entwickelte und überprüfte Instrument zur Messung des Professionswissens angehender Physiklehrkräfte wurde in einer bundesweiten Studie an zwölf Universitäten eingesetzt. Die Daten dieser Studie wurden sowohl klassisch als auch mit Hilfe einer Raschanalyse ausgewertet. Die Auswertung legt nahe, dass das Professionswissen angehender Physiklehrkräfte mithilfe des neuentwickelten Instruments mit guter Reliabilität und Validität erfasst wird.

Die Ergebnisse der vorliegenden Studie zeigen, dass beim Professionswissen im Fach Physik die Bereiche *Fachwissen*, *fachdidaktisches Wissen* und *pädagogisches Wissen* voneinander abgrenzbar sind und dass ihre Ausprägungen unabhängig voneinander gemessen werden können. Dabei zeigen sich starke Zusammenhänge sowohl zwischen Fachwissen und fachdidaktischem Wissen als auch zwischen fachdidaktischem Wissen und pädagogischem Wissen. Ein schwächerer Zusammenhang besteht zwischen dem Fachwissen und dem pädagogischen Wissen. In den Bereichen Fachwissen und fachdidaktisches Wissen konnte für die ersten vier Studienjahre eine positive Entwicklung entsprechend der Anordnung der Lerngelegenheiten im Studium nachgewiesen werden, was für die generelle Wirksamkeit des Lehramtsstudiums spricht. Personenmerkmale, die diese positive Entwicklung begünstigen können sind zum einen schulpraktische Erfahrungen, wie Unterrichtsplanungen und Hospitationen, und zum anderen eine positive Einstellung und Motivation der angehenden Lehrkräfte im Hinblick auf den Studienerfolg und das Unterrichten.

Abstract

Teachers' *professional competence* plays an important role for the quality of teaching and therefore for learning success in school. On this occasion, teachers' *professional knowledge* is a core area of professional competence. Accordingly, one central goal of teachers' education is to impart comprehensive and practice-oriented professional knowledge. However, up to now a systematic evaluation of pre-service teachers' professional knowledge is not carried out. Furthermore, for the school subject physics, it is not clear how professional knowledge is structured, how it evolves and which factors are favouring its evolvement. Across natural sciences, there is also a lack of established instruments for an investigation of the structure and evolvement of professional knowledge in individual subjects. Within the present thesis project, such a measuring instrument for the professional knowledge of pre-service physics teachers was developed and tested. The related research objectives are to investigate the structure of professional knowledge, to measure the degree of professional knowledge in different academic years and to investigate relations between individual personal characteristics and the degree of professional knowledge.

The measuring instrument for pre-service teachers' professional knowledge that was developed within this thesis project was applied at twelve German universities. The collected data was evaluated using classical analysis and Rasch modelling. Results of the evaluation suggest good reliability and validity of the new developed instrument.

The results of the present thesis imply for the subject physics that the areas of *content knowledge*, *pedagogical content knowledge* and *pedagogical knowledge* are distinct areas and their degree can be measured independently. High correlations between content knowledge and pedagogical content knowledge as well as between pedagogical knowledge and pedagogical content knowledge have been found. In contrast, the correlation between content knowledge and pedagogical knowledge is less high. For the areas of content knowledge and pedagogical content knowledge a positive development during the first four academic years, in accordance with learning opportunities, could be observed, which suggests a general effectiveness of pre-service physics teachers' university education. Individual personal characteristics that favour the evolvement of pre-service physics teachers' professional knowledge are on the one hand school experiences like lessons planning or school attendance and on the other hand a positive attitude towards academic success and teaching physics.

Inhaltsverzeichnis

1. Einleitung	1
2. Theoretischer Hintergrund	4
2.1 Professionswissen in Mathematik und den Naturwissenschaften	8
2.1.1 Konzeptualisierungen des Professionswissens	8
2.1.2 Befunde zum Professionswissen	12
2.2 Facetten des Professionswissens im Fach Physik	18
2.2.1 Fachwissen	18
2.2.2 Fachdidaktisches Wissen	28
2.2.3 Pädagogisches Wissen	33
2.3 Wissensarten	33
2.4 Korrelate des Professionswissens	34
3. Forschungsfragen	37
3.1 Erste Forschungsfrage	37
3.2 Zweite Forschungsfrage	37
3.3 Dritte Forschungsfrage	38
4. Methode	39
4.1 Das KiL-Projekt	39
4.2 Design	40
4.3 Definition und Modellierung des Professionswissens	43
4.3.1 Inhaltsbereiche des Fachwissens	44
4.3.2 Inhaltsbereiche des fachdidaktischen Wissens	46
4.3.3 Inhaltsbereiche des pädagogischen Wissens	47
4.3.4 Wissensarten	48
4.3.5 Modellierung des Fachwissens und fachdidaktischen Wissens	49
4.4 Konstruktion und Validierung der Messinstrumente	51
4.4.1 Auswahl, Adaption, Entwicklung und Präpilotierung von Aufgaben	51
4.4.2 Pilotierungsstudie	58
4.4.3 Expertenratings	64
4.4.4 Fragebogenstudie	69
4.4.5 Überarbeitung und Neukonstruktion von Aufgaben	74
4.4.6 Lautes-Denken-Studie	75
4.4.7 Hauptstudie	80
5. Ergebnisse	87
5.1 Kennwerte des Messinstruments	87
5.1.1 Reliabilität	87
5.1.2 Validität	90
5.2 Struktur des Professionswissens	93
5.2.1 Modellprüfung	94
5.2.2 Binnenstruktur des Professionswissens	95
5.3 Entwicklung des Fachwissens und fachdidaktischen Wissens	97
5.4 Korrelate und Determinanten des Professionswissens	104

Diskussion	111
5.5 Güte der gewonnenen Daten	111
5.6 Forschungsfragen	113
6. Zusammenfassung und Ausblick	116
7. Literaturverzeichnis	118

1. Einleitung

Die professionelle Kompetenz von Lehrkräften spielt eine zentrale Rolle für die Qualität von Unterricht und damit für den Lernerfolg von Schülerinnen und Schülern. Dieser positive Einfluss auf die Unterrichtsqualität und auf den Lernerfolg ist gut belegt (Ball et al., 2005; Hill et al., 2005; Neubrand, 2006; Magnusson et al., 1999; Hattie, 2003; Abell, 2007; Baumert und Kunter, 2011b; Ohle et al., 2011; Sadler et al., 2013; Ergönenc et al., 2014). Das professionelle Wissen von Lehrkräften, das sogenannte Professionswissen, stellt dabei neben Überzeugungen und Werthaltungen, motivationalen Orientierungen und selbstregulativen Fähigkeiten einen Kernbereich der professionellen Kompetenz dar (Baumert und Kunter, 2006). Dementsprechend besteht ein zentrales Ziel der Lehrerinnen- und Lehrerausbildung darin, ein möglichst hohes, praxisrelevantes Professionswissen zu vermitteln.

Besonders seit der Durchführung der ersten Studie zum „Programme for International Student Assessment“ (PISA) im Jahr 2000 (Baumert et al., 2001) und dem folgenden „Pisa-Schock“, der 2001 mit der Veröffentlichung der Studienergebnisse durch die Presse einherging, ist die Qualität der Schulbildung sowohl in Deutschland als auch international zunehmend in den Fokus bildungspolitischer Diskussionen gerückt (vgl. Terhart, 2000; Oser, 2001; 2002; Fahrholz, 2002; Oelkers, 2003). Seitdem werden in verschiedenen Ländern unterschiedliche Maßnahmen, wie zum Beispiel die Durchsetzung von Bildungsstandards (Schlesinger, 2013), Vergleiche zwischen Bundesländern (Kerstan, 2011; Pant et al., 2013) oder die Förderung von Kindern aus sozial schwachen Verhältnissen und mit Migrationshintergrund diskutiert (Bruckner und Günther, 2013). Das Ziel dabei liegt stets in der Verbesserung der Unterrichtsqualität, insbesondere im Fach Mathematik und in den Naturwissenschaften (vgl. auch Baumert et al., 2004; Kunter et al., 2005; Helmke, 2006; Klieme et al., 2010). Zentrale Stellgröße für die Unterrichtsqualität ist der Lehrer (Kennedy et al., 2008; Hattie, 2003; 2008; Hascher, 2011). Als naheliegende Maßnahme zur Verbesserung der Unterrichtsqualität rückt damit die Verbesserung der Ausbildung angehender Lehrkräfte in den Fokus. Eine systematische, bundesländerübergreifende Evaluation der universitären Lehrerinnen- und Lehrerausbildung liegt jedoch bisher nicht vor, obwohl Unterschiede im fachbezogenen Professionswissen ihre Ursprünge im Studium haben (Brunner et al., 2006). Über die naturwissenschaftlichen Fächer hinweg fehlt es zudem an etablierten Instrumenten, mit denen die Entwicklung und Ausprägung des Professionswissens in den einzelnen Fächern gemessen werden könnte. Auch international wurde die Notwendigkeit solcher Messinstrumente formuliert (Abell, 2007). Die vorliegende Arbeit soll deshalb dazu beitragen, entsprechende Messinstrumente für das Professionswissen angehender Physiklehrkräfte zu entwickeln und zu erproben.

Darüber, in welchen Bereichen Lehrkräfte Wissen erwerben sollten, besteht zunehmender Konsens. National wie international hat sich eine Einteilung des professionellen Wissens von Lehrkräften in die Bereiche Fachwissen, fachdidaktisches Wissen und pädagogisches Wissen, aufbauend auf den Arbeiten von Shulman (1986, 1987), etabliert (vgl. Baumert und Kunter, 2006; Abell, 2007; Borowski et al., 2010). International werden für die drei angesprochenen Bereiche des Professionswissens meist die Termini CK für „Content Knowledge“, manchmal auch SMK für „Subject Matter Knowledge“ für das Fachwissen, PCK für „Pedagogical Content Knowledge“ für das fachdidaktische Wissen und PK für „Pedagogical Knowledge“ für das pädagogische Wissen einer Lehrkraft verwendet. Die englischen und deutschen Bezeichnungen werden zumeist synonym verwendet, wenngleich Gramzow (2014) anmerkt, dass sich die Begriffe „PCK“ und „fachdidaktisches Wissen“ aus unterschiedlichen theoretischen Ansätzen ableiten. Die Grundideen von fachdidaktischem Wissen und PCK seien jedoch erstaunlich ähnlich (Gramzow, 2014, S.65). Lee Shulman (1986, 1987), der den Begriff des PCK geprägt hat, verwendet ihn als „Amalgam aus Inhalt und Pädagogik“, also dasjenige Wissen, das eine Verbin-

dung von Wissen über das Fach und Wissen über Pädagogik darstellt. Der im deutschsprachigen Raum übliche Begriff des fachdidaktischen Wissens ist weiter gefasst und beinhaltet neben PCK auch Wissen über Lernende, Lerninhalte und Lernziele. Daraus resultiert, dass das fachdidaktische Wissen eher einen eigenständigen Bereich neben dem Fachwissen und dem pädagogischen Wissen bilde. Daher wird für die vorgestellten Studien im Rahmen der vorliegenden Arbeit der umfassendere, deutschsprachige Begriff des fachdidaktischen Wissens verwendet. Soweit auf Arbeiten zurückgegriffen wird, die die englischsprachigen Begriffe CK, PCK und PK verwenden, werden diese synonym zu Fachwissen, fachdidaktischem Wissen und pädagogischem Wissen.

Aufbauend auf dieser (potentiellen) Struktur des Professionswissens wurde verschiedentlich damit begonnen, Messverfahren und Instrumente zu entwickeln und zu erproben. Im deutschsprachigen Raum wurde zunächst im Projekt "Professionswissen von Lehrkräften, kognitiv aktivierender Mathematikunterricht und die Entwicklung mathematischer Kompetenz" (COACTIV) die professionelle Kompetenz und insbesondere das Professionswissen von Mathematiklehrkräften untersucht und mit Unterrichtsmerkmalen in Beziehung gesetzt. Dabei wurde ein hohes Fachwissen als notwendige, aber nicht hinreichende Bedingung für die Entwicklung hohen fachdidaktischen Wissens identifiziert. Ein hohes fachdidaktisches Wissen wiederum führte laut den Ergebnissen der Studie zu einer höheren kognitiven Aktivierung der Schülerinnen und Schüler im Unterricht (Kunter et al., 2011a). Vergleichbare Untersuchungen für die Naturwissenschaften Biologie, Chemie und Physik wurden mit einem eigens entwickelten Testinstrument im Projekt "Professionswissen in den Naturwissenschaften" (ProwiN) durchgeführt. Für das Fach Physik konnte dabei die empirische Trennbarkeit von CK, PCK und PK bestätigt und die Ausprägung dieser Bereiche für Physiklehrer gemessen werden (Kirschner, 2013).

Ob sich die hierbei gefundenen Strukturen auch in der Ausbildungsphase wiederfinden, ist jedoch unklar, da man davon ausgehen muss, dass in der Ausbildungsphase ein Großteil des später angewendeten Wissens überhaupt erst erworben wird (vgl. Shulman, 1986; Brunner, 2006). Eine Untersuchung über den Aufbau des Professionswissens angehender Lehrkräfte sollte daher am Beginn der Ausbildungsphase ansetzen und hinsichtlich der Entwicklung verschiedene Phasen der universitären Ausbildung berücksichtigen. Für angehende Physiklehrkräfte wurde mit den Arbeiten von Riese (2009) sowie Riese und Reinhold (2010) ein wichtiger Schritt in diese Richtung getan. Riese (2009) bestätigte dabei einen Teil der in den oben genannten Studien erzielten Befunde. So konnten zum Beispiel die empirische Trennbarkeit wie auch vergleichbare substantielle Zusammenhänge zwischen den drei Dimensionen professionellen Wissens, dem Fachwissen, dem fachdidaktischen Wissen und dem pädagogischen Wissen, bestätigt werden. Die Generalisierbarkeit der Ergebnisse von Riese ist allerdings dadurch eingeschränkt, dass das verwendete Messinstrument auf einzelne inhaltliche Aspekte professionellen Wissens fokussierte, wie zum Beispiel das Fachwissen im Bereich Mechanik (vgl. auch Woitkowski et al., 2011; Woitkowski, 2015), und das fachdidaktische Wissen zum Experimentieren im Schulunterricht (Riese, 2009). Riese (2009) bemerkt dazu, dass eine Untersuchung über die von ihm gewählten inhaltlichen Schwerpunkte hinaus einen Beitrag zur systematischen Erforschung des Professionswissens liefern würde. Das Ziel der vorliegenden Dissertation besteht daher unter anderem darin, zu prüfen, ob sich die gemachten Befunde zur Entwicklung und zur Struktur des Professionswissens angehender Physiklehrkräfte auf der Basis einer inhaltlich umfassend angelegten Operationalisierung bestätigen lassen (siehe Kapitel 3: *Forschungsfragen*).

Die Entwicklung von Messinstrumente zur Erfassung des Professionswissens angehender Lehrkräfte in den Naturwissenschaften und Mathematik auf Basis einer umfassend angelegten Operationalisie-

rung, war Gegenstand des im Jahr 2011 am Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik (IPN) begonnenen Projekts "Messung professioneller Kompetenzen in mathematischen und naturwissenschaftlichen Lehramtsstudiengängen" (KiL) (Kleickmann et al., 2013). Im Rahmen dieses Projekts ist auch die vorliegende Dissertation entstanden. Um ein entsprechendes Testinstrument für das Professionswissen angehender Physiklehrkräfte zu entwickeln, wurde zunächst der Forschungsstand zum Professionswissen von Lehrkräften und angehenden Lehrkräften detailliert aufgearbeitet (siehe Kapitel 2: *Theoretischer Hintergrund*). Dabei wurden Konzeptionen des Professionswissens aus vorhergehenden Forschungsprojekten analysiert und adaptiert. Zusätzlich wurden Lehrpläne und Lehrbücher analysiert, um unterrichtsrelevante Inhaltsbereiche der Physik zu identifizieren und damit dem Anspruch nach einer inhaltlich breiten Modellierung des zu erfassenden Konstrukts gerecht zu werden. Von den Befunden zum Professionswissen ausgehend werden Forschungsfragen und zugehörige Hypothesen abgeleitet (siehe Kapitel 3: *Forschungsfragen*). Auf den Konzeptionen des Professionswissens aufbauend wurde ein eigenes Modell des Professionswissens angehender Physiklehrkräfte, speziell des Fachwissens und des fachdidaktischen Wissens begründet. Anschließend wurden bestehende Testinstrumente gesichtet, um potentiell geeignete Aufgaben zur Abdeckung des entwickelten Modells zu identifizieren. Lücken bei der Abdeckung des Modells wurden durch gezielte Aufgabenentwicklungen durch Experten geschlossen. Für die Aufgabenentwicklung wurde ein detaillierter projektinterner Leitfaden erarbeitet, der eine hohe Qualität neuentwickelter Testaufgaben gewährleisten soll. Der so entwickelte Aufgabenpool wurde in einem schrittweisen Verfahren geprüft und bewertet, ehe für die bundesweit durchgeführte Haupterhebung dieser Arbeit ein finales Messinstrument aus den geeignetsten Aufgaben zusammengestellt wurde (siehe Kapitel 4: *Methode*).

Die Daten der Haupterhebung wurden sowohl klassisch als auch mit Hilfe einer Raschanalyse ausgewertet. Die Auswertung legt nahe, dass das Professionswissen angehender Physiklehrkräfte mit dem eingesetzten Instrument mit guter Reliabilität und Validität erfasst wird, so dass man davon ausgehen kann, dass die Schlussfolgerungen der vorliegenden Arbeit das Professionswissen angehender Physiklehrkräfte in verschiedenen Phasen des Studiums auf verlässlichen Ergebnissen beruhen (siehe Kapitel 5: *Ergebnisse*). Bezogen auf die Wirksamkeit des Lehramtsstudiums weisen die Daten in den ersten vier Studienjahren auf einen kontinuierlichen Anstieg des Professionswissens hin. Erst im fünften Studienjahr ergibt sich eine Stagnation bei der Wissensentwicklung, für die in der vorliegenden Arbeit verschiedene Deutungsmöglichkeiten diskutiert werden. Bezogen auf die Struktur des Professionswissens zeigt sich eine qualitative Bestätigung der Ergebnisse von Riese (2009) und Kirschner (2013) insoweit, als dass durch Faktorenanalysen eine empirische Trennbarkeit von Fachwissen, fachdidaktischem Wissen und pädagogischem Wissen bestätigt werden konnte. Dabei zeigen sich starke Zusammenhänge zwischen Fachwissen und fachdidaktischem Wissen und zwischen pädagogischem Wissen und fachdidaktischem Wissen. Zwischen Fachwissen und pädagogischem Wissen wurde dagegen nur ein mittlerer Zusammenhang gemessen. Die Ergebnisse der Untersuchung der vorliegenden Arbeit werden ausführlich diskutiert und bewertet (siehe Kapitel 6: *Diskussion*). Abschließend folgt eine Zusammenfassung der wichtigsten Erkenntnisse, sowie eine Diskussion der Einschränkungen zur vorliegenden Arbeit (Kapitel 7: *Zusammenfassung und Ausblick*).

2. Theoretischer Hintergrund

Die Lehrkraft nimmt eine zentrale Rolle für die Leistungs- und Motivationsentwicklung ihrer Schülerinnen und Schüler ein (Hashweh, 1987; Hattie, 2003; Hill et al., 2005; Baumert und Kunter, 2011b; Sadler et al., 2013). Was dabei einen „guten“ Lehrer ausmacht ist allerdings bisher nicht abschließend geklärt. Die Metastudien von Hattie (2003; 2005; 2009) legen jedoch nahe, dass die Kompetenzen von Lehrkräften, darunter auch das Professionswissen, eine wichtige Rolle spielen. Professionswissen stellt definitionsgemäß das Wissen dar, das Menschen brauchen, um ihren Beruf ausüben zu können (Kirschner, 2013). Bezogen auf Lehrkräfte ergeben sich hierbei zwei Probleme: Zum einen stellt der Lehrerberuf, zum Beispiel in Abhängigkeit von Schulform, Unterrichtsfach und individuellen Eigenarten der unterrichteten Schülerinnen und Schüler, gänzlich unterschiedliche Anforderungen. Zum anderen ist nicht klar, über welche Kompetenzen Lehrkräfte verfügen müssen, um ihre jeweiligen beruflichen Tätigkeiten optimal ausüben zu können. Was dies für das Professionswissen für Lehrkräfte bedeutet, ist noch nicht abschließend geklärt, wenngleich auf diesem Forschungsgebiet beginnend mit den Arbeiten von Shulman (1986; 1987) beachtliche Fortschritte erzielt wurden.

Während die Bildungsforschung in den 60er und 70er Jahren noch hauptsächlich auf einfache Persönlichkeitsmerkmale fokussierte, die das erzieherische Wirken von Lehrkräften beeinflussen (Getzels und Jackson, 1963; vgl. auch Bromme, 1997; Fischer und Borowski, 2009), trat in den 80er und 90er Jahren zunehmend eine differenziertere Betrachtung der professionellen Kompetenz und Expertise von Lehrkräften in den Fokus bildungswissenschaftlicher und fachdidaktischer Forschung (Elbaz, 1983; Shulman, 1986; 1987; Hashweh, 1987; Bromme, 1992; 1994; Cochran, DeRuiter und King, 1993; Adams und Krockover, 1997; Douvdevany, Dreyfus und Jungwirth, 1997; Weinert, 1996; 2001; Berliner, 2001). In den letzten Jahren wurden Zusammenhänge zwischen Lehrermerkmalen und Prozessen im Schulunterricht modelliert, die der Komplexität der Tätigkeiten von Lehrkräften zunehmend gerecht werden sollen (Doll und Prenzel, 2001; vgl. auch Fischer et al., 2010). Ein theoretisches Rahmenmodell zur Wirkung von Lehrermerkmalen im Unterricht aus der COACTIV-Studie (Krauss et al., 2004) basierend auf dem Mehrebenenmodell von Doll und Prenzel (2001) ist in Abbildung 2.1 dargestellt. Das Modell beschreibt auf vier Ebenen den Zusammenhang verschiedener Faktoren (Ebenen II-IV), welche jeweils Einfluss auf die Zielkriterien des Unterrichts (Ebene I) nehmen. Dabei stehen die Merkmale der Lehrkraft auf Stufe III und wirken im Klassenkontext direkt auf die Unterrichtsprozesse, die unter Einfluss der individuellen Lernvoraussetzungen der Schülerinnen und Schüler die kognitiven Verarbeitungsprozesse bestimmen. Die Lehrermerkmale enthalten dabei neben epistemologischen Überzeugungen und allgemeinen Merkmalen das für die vorliegende Arbeit bedeutsame theoretische Konstrukt der Expertise. Dieses Konstrukt geht laut Krauss et al. (2004) auf die etablierten Konzepte von Shulman (1986, 1987), Schön (1983, 1987) und Bromme (1992, 1997) zurück.

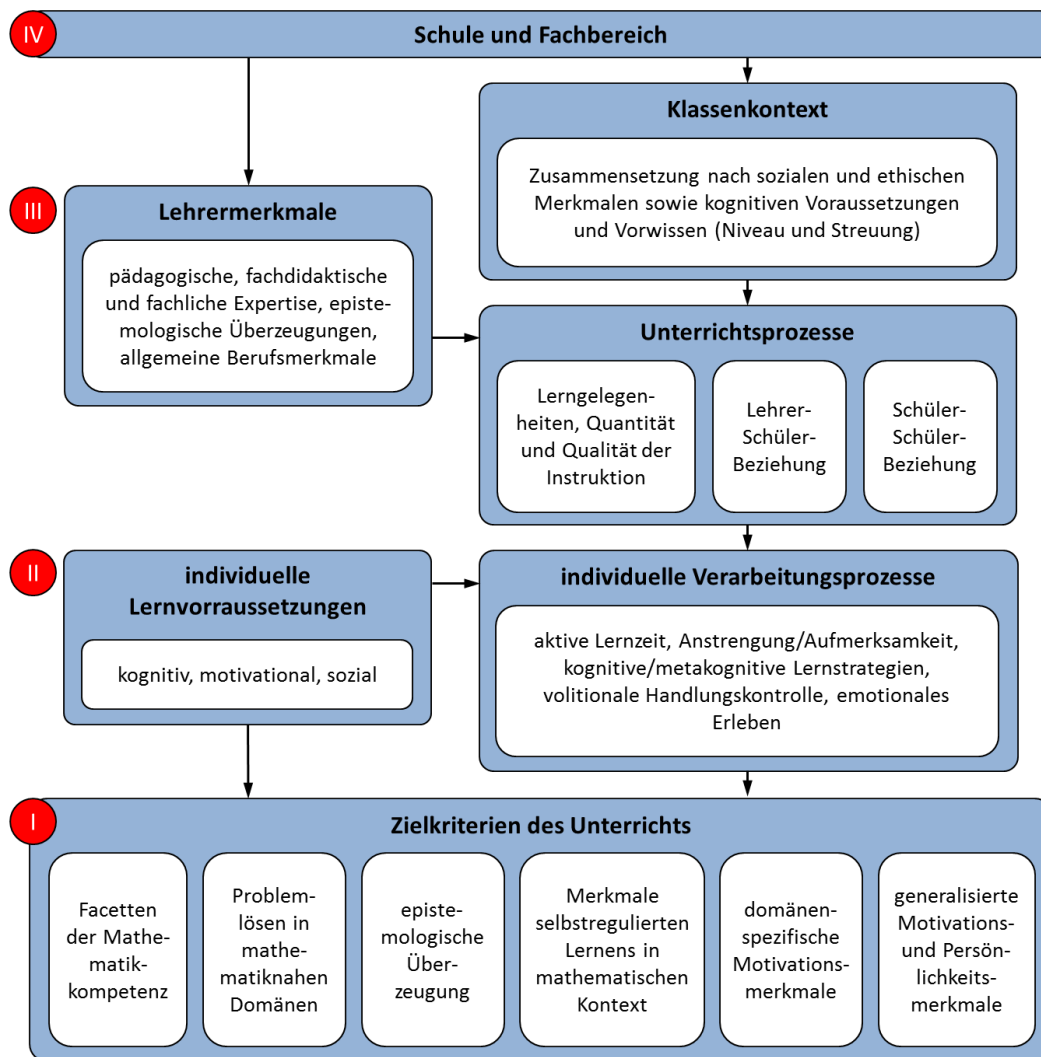


Abbildung 2.1: Mehrebenenanalytisches Rahmenmodell der COACTIV-Studie (Krauss et al., 2004).

Die Expertise bzw. das Professionswissen der Lehrkraft wird dabei im obigen Modell als pädagogische, fachdidaktische und fachliche Expertise bezeichnet. Um im Rahmen dieser Arbeit zu einer anschlussfähigen Konzeptualisierung des Professionswissens zu kommen, soll im folgenden Abschnitt geklärt werden, welche Konzeptualisierungen des Professionswissens sich im Forschungsfeld etabliert haben. Die in den 60er und 70er Jahren untersuchten Persönlichkeitsmerkmale von Lehrkräften werden heute nur noch als ein Teil professioneller Kompetenz betrachtet (Baumert und Kunter, 2006). In dem vielzitierten Überblicksartikel von Baumert und Kunter (2006) wird daher ein neues Modell für die professionelle Kompetenz vorgestellt (vgl. auch Baumert und Kunter, 2011a), welches sämtliche Aspekte berücksichtigt. Hiernach umfasst die professionelle Kompetenz einer Lehrkraft ihre Überzeugungen und Werthaltungen, ihre motivationalen Orientierungen, selbstregulative Fähigkeiten sowie ihr Professionswissen (siehe Abbildung 2.2).

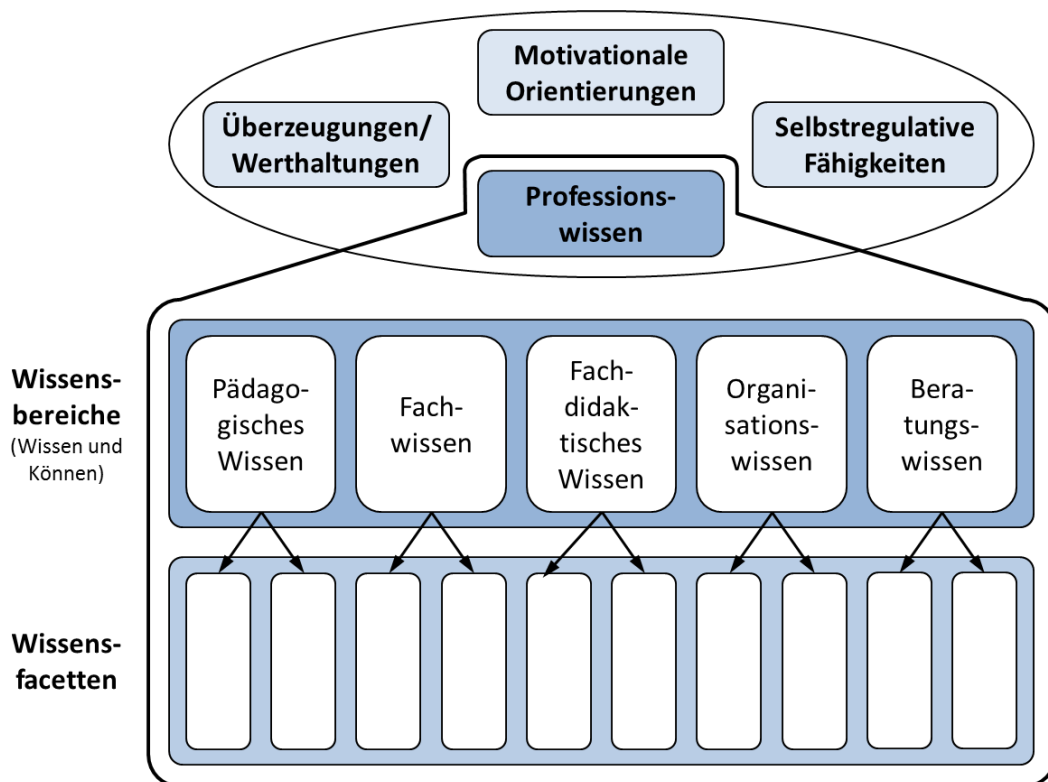


Abbildung 2.2: Modell professioneller Handlungskompetenz und des Professionswissens (Baumert und Kunter, 2006; 2011a).

Innerhalb des Professionswissens unterscheiden Baumert und Kunter (2006) verschiedene Wissensbereiche: *Pädagogisches Wissen*, *Fachwissen* und *fachdidaktisches Wissen*. Systematisch werden sie, aufbauend auf Shulman (1987) um die Bereiche *Organisationswissen* und *Beratungswissen* ergänzt. Die beiden zuletzt genannten Bereiche seien jedoch nur sinnvoll im Rahmen institutionalisierter und sozial verteilter Wissensbestände behandelbar und werden daher nicht weiter diskutiert (Baumert und Kunter, 2006, S. 482). Das Professionswissen stellt nach Baumert und Kunter (2006) einen Kernbereich der professionellen Handlungskompetenz dar, wobei wiederum die Wissensbereiche pädagogisches Wissen, Fachwissen und fachdidaktisches Wissen Kernbereiche des Professionswissens darstellen. Innerhalb der Wissensbereiche werden zusätzlich Wissensfacetten angedeutet (siehe Abbildung 2.2). Diese stellen feinere Kategorisierungen innerhalb des pädagogischen, des Fach- und des fachdidaktischen Wissens dar. Kunter et al. (2011a) führten diese Facetten für das Fach Mathematik konkret aus: Beim fachdidaktischen Wissen werden beispielsweise Erklärungswissen, Wissen über das Denken von Schülern und Wissen über Aufgaben unterschieden (Baumert und Kunter, 2011a).

Man geht heute davon aus, dass ein ausgeprägtes Professionswissen, wie auch andere Persönlichkeitsmerkmale, für Lehrkräfte ein notwendiges aber nicht hinreichendes Kriterium für erfolgreiches Unterrichten darstellt (Baumert und Kunter, 2006; Fischer et al., 2012; vgl. auch Kirschner, 2013). Wenngleich noch keine differenzierte und allgemein anerkannte Konzeptualisierung dieses Professionswissens von Lehrkräften existiert (Abell, 2007; Tepner et al., 2012), so gibt es doch gewisse Gemeinsamkeiten hinsichtlich üblicherweise verwendeter Konzepte (Park und Oliver, 2008; Carlson, J., Stokes, L., Helms, J., Gess-Newsome, J., und Gardner, A.; 2015). Mit dem Ziel einer gemeinsamen Modellierung fachdidaktischen Wissens fand 2012 ein internationaler Gipfel statt, zu dem Professionswissensforscher mit besonderer Expertise im Bereich PCK eingeladen waren (Carlson et al., 2015).

In Vorbereitung dazu wurden die Forschungsgruppen aufgefordert, standardisierte Zusammenfassungen zu Konzeption und Befundlage von PCK vorzubereiten, die dann im Rahmen des Gipfels vorgestellt und diskutiert wurden. Auf dieser Grundlage wurde das in Abbildung 2.3 dargestellte Konsens-Modell für das Professionswissen von Lehrkräften entwickelt.

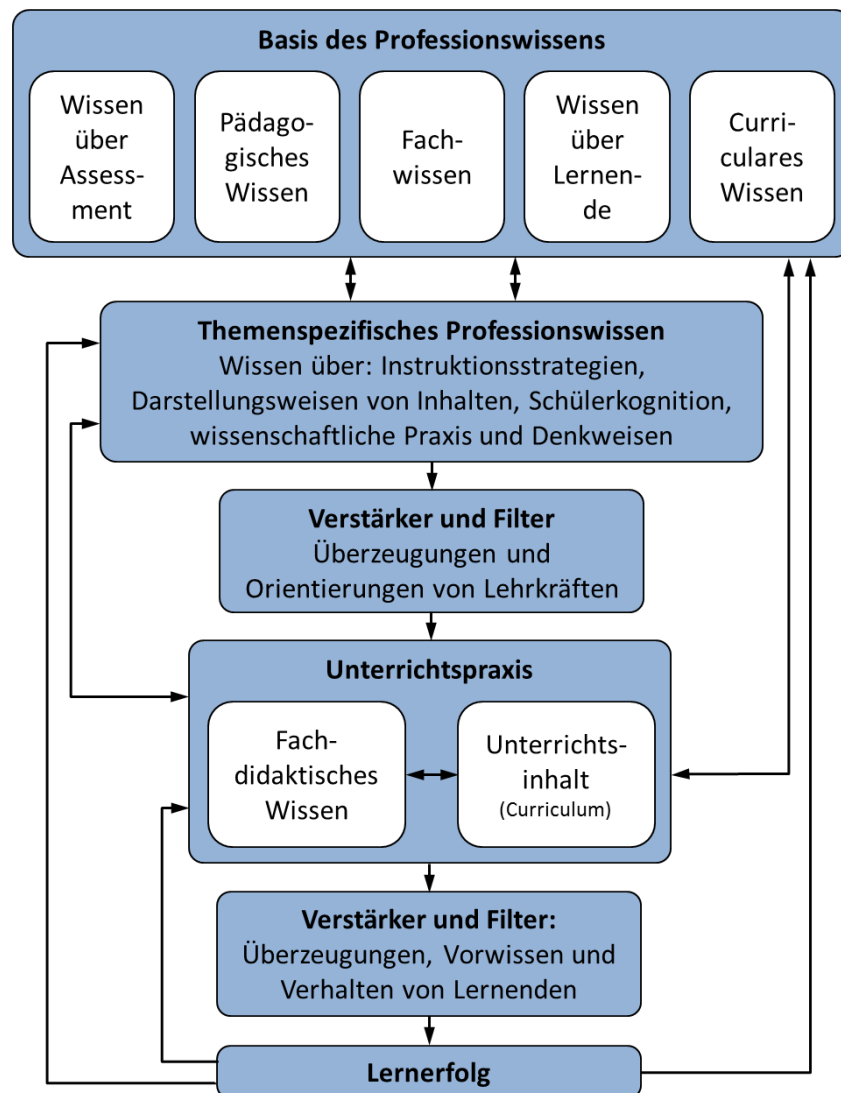


Abbildung 2.3: Konsens-Modell des Professionswissens (Gess-Newsome und Carlson, 2013; Gess-Newsome 2015).

Dieses unterscheidet sich insoweit von dem Modell von Magnusson et al. (1999), als dass es sich auf das gesamte professionelle Wissen im Unterrichtskontext bezieht und das fachdidaktische Wissen als zentrale Facette darin verortet. Dabei wird in der obersten Ebene des Konsens-Modells (siehe Abbildung 2.3) von einer allgemeinen Basis des Professionswissens ausgegangen, bestehend aus *Fachwissen*, *pädagogischem Wissen* sowie den fachdidaktischen Facetten *Wissen über Assessment*, *Wissen über Lernende* und *Wissen über Curricula*. Die Basis des Professionswissens bezieht sich dabei auf allgemeines, also nicht themenspezifisches, professionelles Wissen von Lehrkräften. Ein Beispiel dafür ist die allgemeine Kenntnis über das Format des formativen Assessments als Instrument zur Informationsgewinnung mit dem Ziel als Lehrkraft lehr- und lernverbessernde Rückmeldungen an die Lernenden geben zu können (Gess-Newsome, 2015; Black und William, 2006).

Im Folgenden wird sich die vorliegende Arbeit an die in Abbildung 2.2 dargestellte und im deutschsprachigen Raum etablierte Benennung von Baumert und Kunter (2006) anschließen und deren Nomenklatur verwenden. Das bedeutet, dass Fachwissen, fachdidaktisches Wissen und pädagogisches Wissen als drei *Wissensbereiche* des Professionswissens bezeichnet werden. Ferner werden Inhaltsbereiche des Fachwissens und fachdidaktische Inhaltsbereiche als *Wissensfacetten* bezeichnet.

Um Anschlussfähigkeit an das Feld der Professionswissensforschung herzustellen werden im folgenden Kapitel zunächst verschiedene Konzeptualisierungen des Professionswissens von Lehrkräften vorgestellt. Dabei wird ergründet, welche Wissensbereiche und Facetten das Professionswissen in der Betrachtung der fachdidaktischen Forschung typischerweise beinhaltet und es werden verschiedene fachdidaktische Forschungsprojekte zum Professionswissen berichtet (siehe Abschnitt 2.1). Dabei werden zunächst allgemeine Konzeptualisierungen und Befunde zum Professionswissen, also solche von Lehrkräften verschiedener Fachrichtungen, aufgegriffen. Anschließend wird das spezielle Professionswissen von Lehrkräften im Fach Physik untersucht (siehe Abschnitt 2.2). Hier wird insbesondere auf die Facetten des Professionswissens näher eingegangen. Danach wird gezeigt, welche verschiedenen Arten von Professionswissen innerhalb der fachspezifischen Wissensbereiche des Professionswissens unterschieden werden können (siehe Abschnitt 2.3). Zuletzt werden Korrelate des Professionswissen mit anderen Konstrukten beschrieben (siehe Abschnitt 2.4).

2.1 Professionswissen in Mathematik und den Naturwissenschaften

Im Folgenden werden Konzeptualisierungen des Professionswissens werden im Abschnitt 2.1.1 dargestellt, während im Abschnitt 2.1.2 Befunde zum Professionswissen berichtet werden.

2.1.1 Konzeptualisierungen des Professionswissens

Erste konkrete Angaben darüber, was das Professionswissen einer Lehrkraft umfassen sollte, gibt Shulman (1986). Er schlägt die getrennte Betrachtung dreier Kategorien des Professionswissens vor, die er jedoch alle als Teile des Professionswissens einer Lehrkraft versteht: (a) *fachinhaltsbezogenes Fachwissen*, (b) *fachdidaktisches Wissen* und (c) *Wissen über das Curriculum*. Unter fachinhaltsbezogenem Fachwissen einer Lehrkraft versteht Shulman nicht nur die Kenntnis von fachlichen Fakten und Konzepten, sondern ein darüber hinaus gehendes Verständnis darüber, wie dieses fachbezogene Wissen strukturiert ist (Shulman, 1986, S.9). Dieses Verständnis sollte dabei bereichsbezogen betrachtet werden, da es je nach Inhaltsbereich innerhalb eines Faches unterschiedlich stark ausgeprägt sein könne. Als Besonderheit des Fachwissens einer Lehrkraft hebt Shulman (1986) hervor, dass das Fachwissen von Lehrern nicht nur dem aktuellen Stand des Wissens entsprechen sollte, sondern auch Begründungswissen enthalten müsse. Das Begründungswissen sei nötig, um Schülerinnen und Schülern vermitteln zu können, welche Werte und Anwendungsbereiche bestimmte Wissensdomänen haben und welche Zusammenhänge es zu anderen – auch fachfremden – Domänen gibt. Zudem sollten Lehrkräfte verschiedene Arten der Strukturierung ihres Fachwissens kennen, was ihnen eine adressatengerechte Auswahl der Fachstruktur, in Abhängigkeit von der jeweiligen Zielgruppe, erlaubt. Unter fachdidaktischem Wissen versteht Shulman (1986) die Art von Fachwissen, die eine Lehrkraft zum Unterrichten benötigt. Es beinhaltet Kenntnisse über unterschiedliche Repräsentationsformen, Kenntnisse über typische Lernschwierigkeiten und Vorstellungen bzw. Fehlvorstellungen von Schülerinnen und Schülern und Kenntnisse über Strategien, wie man solchen Lernschwierigkeiten und Fehlvorstellungen begegnen kann. Das fachdidaktische Wissen ist damit für Shulman (1986)

eng mit dem Fachwissen verbunden. Wissen über das Curriculum beinhaltet nach Shulman (1986) Kenntnisse typischer Lerneinheiten zu spezifischen Fachinhaltsbereichen und Wissen über geeignetes dazugehöriges Lernmaterial. Innerhalb dieser Bereiche des Professionswissens betont Shulman (1986) die herausragende Bedeutung des Fachwissens, wozu er auch das fachdidaktische Wissen zählt, da dieses bisher sowohl in der Bildungs- und fachdidaktischen Forschung als auch als Ausbildungs- und Prüfungskriterium für angehende Lehrkräfte häufig vernachlässigt worden wäre.

Neben den obigen drei *Wissensbereichen* unterscheidet Shulman (1986) auch drei unterschiedliche *Wissensarten*: *Propositionales Wissen* stellt Faktenwissen dar, das im Fach Physik zum Beispiel die Kenntnis konkreter Gesetze oder Formeln beinhaltet. *Prozedurales Fallwissen* meint Wissen über typische Handlungsabläufe und Prozesse. Dazu gehörten beispielsweise im Unterricht ablaufende Handlungsskripte (vgl. Stender, 2014), also zum Beispiel Wissen darüber, wie man in einer bestimmten Lerneinheit einer spezifischen Schülerfehlvorstellung durch Auslösung eines passenden kognitiven Konflikts begegnen kann. *Strategisches Wissen* stellt die dritte Wissensart dar, die laut Shulman (1986) stärker vernetztes Wissen darstellt. Ein Beispiel wäre Wissen darüber, unter welchen Voraussetzungen bzw. warum ein bestimmtes Gesetz Anwendung finden kann. Eine solche Unterscheidung von Wissensarten wurde zum Beispiel auch von Paris et al. (1983) und von Shavelson et al. (2005) vorgenommen, die ebenfalls zwischen deklarativem Faktenwissen, prozeduralem Wissen über typische Handlungsabläufe und übergeordnetem, vernetztem konditionalen (Paris et al., 1983) bzw. schematischem und strategischem Wissen (Shavelson et al., 2005) unterscheiden. Die Ansicht, dass es sich hierbei um zentrale Komponenten professioneller Handlungskompetenz handelt, hat sich in der Professionswissensforschung weitestgehend durchgesetzt (Baumert und Kunter, 2006, S. 481; Shavelson et al., 2005; Tepner et al., 2012).

Später charakterisiert Shulman (1987) professionelles Wissen, indem er konkrete inhaltliche Kategorien beschreibt, die das gesamte Professionswissen einer Lehrkraft notwendigerweise umfassen sollten. Er benennt dabei: (a) das Fachwissen (CK) einer Lehrkraft, (b) allgemeines pädagogisches Wissen (PK), bezogen auf allgemeine Prinzipien und Strategien, also zum Beispiel Organisationswissen oder Wissen über das Anleiten von Schülergruppen, (c) curriculares Wissen über Anordnung von Lehrinhalten, Lehrpläne und Unterrichtsmaterialien, (d) das fachdidaktische Wissen (PCK) als schulspezifisches Amalgam aus Wissen über das Fach und Wissen über Pädagogik, (e) Wissen über die Kognition von Schülerinnen und Schülern, also zum Beispiel Wissen über typische Lernschwierigkeiten und Fehlvorstellungen (f) spezielles Organisationswissen zu Bildungskontexten, Sozialformen im Unterricht und der Schulverwaltung und (g) Wissen über Lernziele, also Wissen, warum bestimmte Lerninhalte besonders wichtig sind sowie erziehungsphilosophisches, bildungstheoretisches und bildungshistorisches Wissen. Bezogen auf diese sieben Kategorien des Professionswissens unterstreicht Shulman (1987) die zentrale Rolle der Lehrkraft für den schulischen Lernerfolg, da die Ausprägung des Professionswissens in den Kategorien eine Eigenschaft der Lehrkraft sei und den Lernerfolg der Schülerinnen und Schüler maßgeblich beeinflusst. Seine Charakterisierung des Punktes (d), des PCK als Amalgam aus Wissen über das Fach und Wissen über Pädagogik, in dem Fachwissen und pädagogisches Wissen zusammentreffen, wurde von Fachdidaktikern seitdem immer wieder aufgegriffen (Grossmann, 1990; Bromme, 2004). Laut Shulman (1987) handele es sich beim PCK um streng lehrkraftspezifisches Professionswissen. Ein Beispiel hierfür wäre Wissen über typische Schülerfehlvorstellungen in spezifischen Fachinhaltsbereichen. Solches Wissen erfordert einerseits Wissen zu typischen Schülervorstellungen und andererseits Kenntnisse zur Bewertung der fachlichen Angemessenheit solcher Schülervorstellungen. Über diese Kombination von Fachwissen und pädagogischem Wis-

sen verfügen üblicherweise weder Fachphysiker, da Ihnen das pädagogische Wissen fehlt, noch Allgemeinpädagogen, da Sie nicht über das nötige Fachwissen verfügen.

Die von Shulman (1986) eingeführte Unterscheidung zwischen Fachwissen (CK) und fachdidaktischem Wissen (PCK), sowie die Charakterisierung von PCK wird von Bromme (1992; 1994; 1997) aufgegriffen, der ebenfalls die hervorgehobene zentrale Rolle der Lehrkraft für die Leistungsentwicklung von Schülerinnen und Schülern betont. Er leitet theoretisch her, was eine gute Lehrkraft ausmacht und sieht das Professionswissen von Lehrkräften als wichtige und nötige Voraussetzung für erfolgreiches Unterrichten (Bromme, 1994). Gleichzeitig weist er auf die mangelnde aufgeklärte Varianz über die Erklärung von Leistungen der Schülerinnen und Schülern durch das Handeln der Lehrkraft im Unterricht hin. Bisherige Studien zeigten zum Beispiel, dass ein hohes CK die Qualität von Erläuterungen im Unterricht positiv beeinflusst (Bromme, 1997). Zudem sind Lehrkräfte mit hohem CK eher in der Lage, nicht intendierte bzw. unvorhergesehene Beiträge von Schülerinnen und Schülern in den Unterricht einzubinden (Hashweh, 1987; Hattie, 2003). Es lässt sich feststellen, dass Lehrkräfte mit niedrigem CK eher zu einem eng angeleiteten Unterricht neigen (Hashweh, 1987, S.115) und weniger Verbindungen zwischen verschiedenen Abschnitten des unterrichteten Stoffes herstellen (Bromme, 1997), wohingegen ein hohes CK mit einem kognitiv stärker herausfordernden Unterricht einhergeht (Baumert et al., 2004; Baumert & Kunter, 2006, S.496). Als nötigen Schritt einer weiteren Untersuchung des Professionswissens von Lehrkräften sieht Bromme (1997) die begriffliche Aufgliederung in einzelne Bereiche, die nicht nur analytisch begründet, sondern möglichst empirisch gerechtfertigt sind. Aus den von Shulman (1986) angegebenen Anforderungen an das Professionswissen von Lehrkräften, dem fachunabhängigen *pädagogischen Wissen (PK)* und den drei fachabhängigen Domänen *Fachwissen (CK)*, *fachdidaktisches Wissen (PCK)* und *curriculares Wissen*, leitet Bromme eine eigene Kategorisierung ab, die er Topologie des professionellen Wissens nennt. Dabei fügt er Shulmans Kategorien die der *Philosophie des Schulfaches* hinzu. Darunter versteht Bromme (1997) die Auffassungen einer Lehrkraft darüber, wofür ein Fachinhalt nützlich ist und wie er mit anderen Bereichen in Beziehung steht. Seine Topologie des professionellen Wissens enthält damit die folgenden Kategorien: (a) Das Fachwissen (CK), (b) curriculares Wissen, (c) die Philosophie des Schulfaches, (d) das pädagogische Wissen (PK) und (e) fachdidaktisches Wissen (PCK). Damit ergibt sich gegenüber der von Shulman (1986) eingeführten Kategorisierung eine Erhöhung auf fünf Dimensionen des Professionswissens, wobei alle bis auf das pädagogische Wissen als fachspezifisch gelten können.

Während Bromme (1994) auf dem ersten Ansatz von Shulman (1986) aufbaut, beziehen sich Krauss et al. (2004) sowie Baumert und Kunter (2006) bei der Konzeptualisierung der Lehrermerkmale im Rahmen des Forschungsprogramms COACTIV auf die erweiterte Topologie von Shulman (1987). Krauss et al. (2004) konzeptualisieren das Professionswissen einer Lehrkraft in ihrem Vierebenenmodell (siehe Abbildung 2.1) daher entsprechend als pädagogische, fachdidaktische und fachliche Expertise. Baumert und Kunter (2006) stellen ferner fest, dass sich in praktisch allen Konzeptualisierungen des Professionswissens (vgl. Grossmann, 1990; Bromme, 1994; Borko und Putnam, 1996; Cochran, 1997) die von Shulman (1986; 1987) vorgeschlagene getrennte Betrachtung von pädagogischem Wissen, fachdidaktischem Wissen und Fachwissen durchgesetzt habe (vgl. auch Terhart, 2002; Terhart, 2005, S.276). In der COACTIV-Studie wird daher diese Einteilung für die Untersuchung der professionellen Kompetenz von Mathematiklehrkräften verwendet (siehe Abbildung 2.2). Dieselbe Kategorisierung bezüglich der Inhaltsbereiche des Professionswissens wird auch im Projekt *Mathematics Teaching in the 21st Century (MT21)* (Blömeke et al., 2008a; 2008b) vorgenommen, das die Wirksamkeit der Lehrerbildung untersucht. Bezüglich der Repräsentation verschiedener Wissenstypen stellen

Baumert und Kunter (2006) zusammenfassend fest: (1) Professionelles Wissen ist domänenspezifisch und ausbildungs- bzw. trainingsabhängig. (2) Expertenwissen ist sehr gut vernetzt und hierarchisch organisiert. (3) In professionellen Domänen ist Expertenwissen um Schlüsselkonzepte und eine begrenzte Zahl von Ereignisschemata arrangiert, an die Einzelfälle, episodische Einheiten oder Sequenzen von Episoden (Skripts) angedockt sind. (4) Professionelles Expertenwissen integriert Kontexte und erlaubt variantenreicheres „opportunistisches Verhalten“. (5) Basisprozeduren sind automatisiert, aber gleichwohl flexibel an die spezifischen Bedingungen des Einzelfalles und des Kontextes adaptierbar“ (Baumert und Kunter, 2006, S.483).

Die Abhängigkeit des Professionswissens vom jeweiligen Inhaltsbereich und Kontext (Park und Chen, 2012) macht eine differenzierte Betrachtung verschiedener Inhalte des Fachwissens und des fachdidaktischen Wissens nötig. Die dabei als relevant identifizierten Domänen müssen bei der Konstruktion eines Messinstruments entsprechend berücksichtigt werden durch die zitierte Identifizierung von stark vernetztem Expertenwissen und die Automatisierung von Basisprozeduren ergibt sich zudem die Notwendigkeit, über das domänenspezifische hinaus auch unterschiedliche Wissensarten bzw. mentale Repräsentationen des Wissens zu berücksichtigen. Als unterschiedliche Wissensarten werden bei Baumert und Kunter (2006) weiterhin formales Faktenwissen, welches zumeist dem Fachwissen zuzuordnen ist und praktisches, anwendungsbezogenes Wissen diskutiert, welches erfahrungsbasiert in spezifische Kontexte und Problemstellungen eingebettet ist.

Ein explizites Modell für das Professionswissen von Lehrkräften in den Naturwissenschaften wurde im Projekt ProwiN von Tepner (2012) entwickelt (siehe Abbildung 2.4). Dieses umfasst je nach Wissensbereich zwei oder drei Dimensionen. Für das Fachwissen ergibt sich ein zweidimensionales Modell, welches fachliche *Inhaltsbereiche* und *Wissensarten* umfasst. Für das fachdidaktische Wissen umfasst das Modell drei Dimensionen: fachliche *Inhaltsbereiche*, *PCK-Facetten* und *Wissensarten*. Für das pädagogische Wissen ergeben sich die zwei Dimensionen *PK-Facetten* und *Wissensarten*.

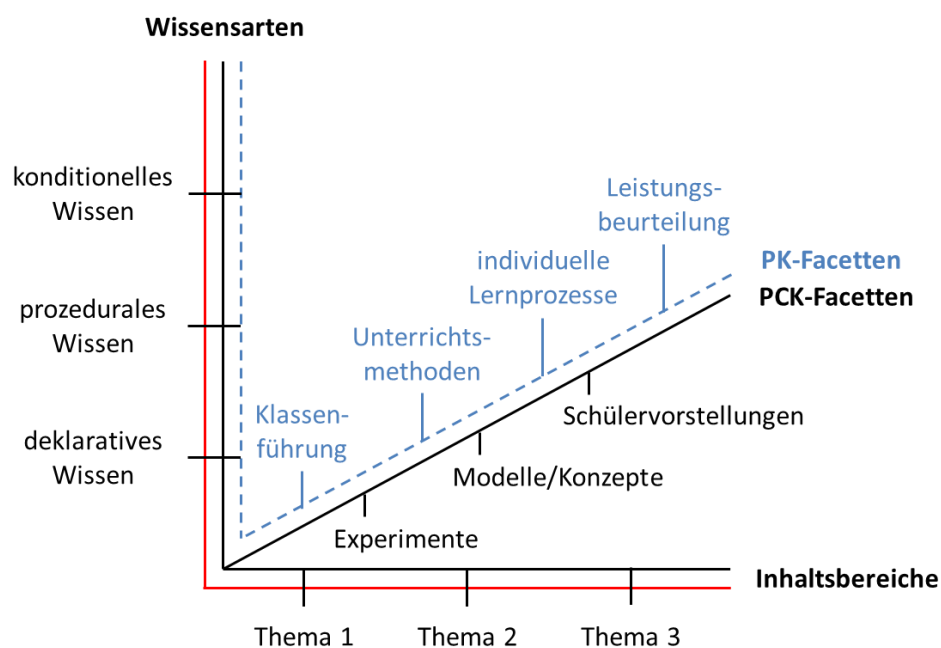


Abbildung 2.4: Modell des Professionswissens für Lehrkräfte aus dem ProwiN-Projekt (Tepner et al., 2012).

Die Unterscheidung der Wissensbereiche CK, PCK und PK liegt auch der Modellierung des Professionswissens von Physik Lehrkräften nach Kirschner (2013) zugrunde. Hier werden, wie bei Tepner (2012), in Anlehnung an Paris et al. (1983) die Wissensarten *deklaratives, prozedurales* und *konditionales Wissen* unterschieden und auf die drei Wissensarten bezogen. Ferner unterscheidet Kirschner (2013, S.27) in Anlehnung an Riese (2009) und Baumert et. al (2010) noch Fachwissen auf verschiedenen Stufen: *Schulwissen, vertieftes Schulwissen, universitäres Wissen* und *fachliches Alltagswissen*.

Zusammenfassend lässt sich anhand der zuvor vorgestellten Studien der letzten Jahre feststellen, dass eine Unterteilung der separaten Bereiche Fachwissen, fachdidaktisches Wissen und pädagogisches Wissen für Lehrkräfte und Referendare in Mathematik und den Naturwissenschaften als theoretisch und empirisch fundiert gilt (vgl. Bromme, 1992; 1994; Baumert und Kunter, 2006; Riese, 2009; Kirschner, 2013; Gramzow, 2014). Der Schwerpunkt der im Forschungsfeld betrachteten Projekte liegt damit unzweifelhaft auf den Wissensbereichen Fachwissen, fachdidaktisches Wissen und pädagogisches Wissen, welche ursprünglich auf die Konzeption von Shulman (1986; 1987) zurückgehen. Diese drei Bereiche sind besonders geeignet, da sie auch die drei Kernbereiche der Lehrerinnen- und Lehrerausbildung in Deutschland repräsentieren: fachlich, fachdidaktisch und pädagogisch (vgl. Tepner et al., 2012). Aus der Abhängigkeit des Wissens von Domäne, Kontext und Vernetztheit ergibt sich jedoch zusätzlich die Notwendigkeit eine differenzierte Betrachtung innerhalb der drei Bereiche vorzunehmen. Dies sollte sinnvollerweise durch die Differenzierung sowohl inhaltlicher Facetten als auch Wissensarten bzw. mentaler Repräsentationen geschehen. Die systematische inhaltliche Differenzierung der Wissensarten des Professionswissens ist deshalb ein zentrales Ziel der vorliegenden Arbeit.

2.1.2 Befunde zum Professionswissen

Im vorigen Abschnitt wurde dargestellt, dass sich die Gliederung des Professionswissens in die Wissensbereiche Fachwissen, fachdidaktisches Wissen und pädagogisches Wissen etabliert hat. Wie die Ausprägung dieser Merkmale jeweils den Unterrichtserfolg einer Lehrkraft beeinflussen, inwieweit sie im mit anderen Unterrichtsvariablen wechselwirken und ob sie überhaupt zuverlässig gemessen werden können, wird durch die vorherigen, theoretischen Betrachtungen noch nicht deutlich. Ferner ist nicht klar, welche Facetten der einzelnen Wissensbereiche für das Unterrichten relevant sind. Im Folgenden werden daher empirische Befunde zum Professionswissen, insbesondere zu seiner Struktur, Wirksamkeit und Entwicklung, näher betrachtet.

(a) Struktur

Ob sich Fachwissen, fachdidaktisches Wissen und pädagogisches Wissen voneinander getrennt erfassen lassen ist eine empirische Frage. Erste Befunde konnten für das Fach Mathematik in der COACTIV-Studie generiert werden. Zwischen pädagogischem Wissen und Fachwissen wurde eine latente Korrelation von $.24$ ($p < .05$) und zwischen pädagogischem Wissen und fachdidaktischem Wissen von $.42$ ($p < .05$) gemessen (Voss und Kunter, 2011). Krauss et al. (2008; 2011) berichten Ergebnisse der Messung des Fachwissens und des fachdidaktischen Wissens von $n = 198$ Lehrkräften, die im Rahmen des Projekts erhoben wurden. Zentral war dabei die Fragestellung, ob Fachwissen und fachdidaktisches Wissen zwei verschiedene Wissensbereiche oder einen gemeinsamen Wissenskomplex darstellen (Krauss et al., 2011, S.147). Auf Basis eines konfirmatorischen Faktorenmodells wurde dabei zwischen dem Fachwissen und dem fachdidaktischen Wissen eine hohe Korrelation von $r = .79$ ($p < .05$),

die sich jedoch signifikant von einer Korrelation von 1 unterscheidet, was für eine empirische Trennbarkeit von Fachwissen und fachdidaktischem Wissen spricht. In einer internationalen Vergleichsstudie im Rahmen des MT21-Projekts wurden entsprechende Zusammenhänge zwischen Fachwissen und fachdidaktischem Wissen bestimmt. Dabei ergaben sich für die an der Studie beteiligten Länder teilweise recht deutliche Unterschiede in der Korrelation zwischen Fachwissen und fachdidaktischem Wissen (Blömeke et al., 2010, S.44), wobei die Werte für deutsche Mathematiklehrkräfte international im Mittelfeld liegen. Die Ergebnisse sind in Tabelle 2.1 dargestellt.

Tabelle 2.1: Korrelationen zwischen CK und PCK im internationalen Vergleich (Blömeke et al., 2010, S.44).

Land	Südkorea	Taiwan	Bulgarien	Deutschland	USA
Korrelation r zwischen CK und PCK	.56 ($p < .001$)	.63 ($p < .001$)	.80 ($p < .001$)	.74 ($p < .001$)	.74 ($p < .001$)

Im internationalen Vergleich fällt der Zusammenhang zwischen CK und PCK in Südkorea und Taiwan also besonders niedrig aus. Die Autoren verweisen darauf, dass der Anteil fachbezogener Lerngelegenheiten in der Lehrerinnen- und Lehrerausbildung in diesen Ländern besonders hoch sei. Für Deutschland kann man davon ausgehen, dass der Anteil fachbezogener Lerngelegenheiten im gymnasialen Lehramtsstudium ebenfalls höher ist als im nichtgymnasialen Lehramtsstudium. Hier ergibt sich jedoch ein umgekehrtes Bild. Für Nichtgymnasiallehrkräfte ergibt sich hier ein korrelativer Zusammenhang von $r = .61$ ($p < .05$), für Gymnasiallehrkräfte dagegen von $r = .96$ ($p < .05$). Die Art und der spezielle Aufbau des von den Lehrkräften durchlaufenen Ausbildungsprogramms scheint also einen komplexeren Einfluss auf den Zusammenhang zwischen Fachwissen und fachdidaktischem Wissen auszuüben. Man kann jedoch anhand der Befunde von MT21 resümieren, dass auch in Ausbildungsprogrammen anderer Länder ähnliche strukturelle Zusammenhänge zwischen Fachwissen und fachdidaktischem Wissen für das Fach Mathematik beobachtet werden können. Die Größenordnung der korrelativen Zusammenhänge aus der MT21-Studie decken sich in mit denen in der COACTIV-Studie ermittelten Korrelationsspanne für Gymnasial- bzw. Nichtgymnasiallehrkräfte (COACTIV: .61 bis .96 vs. MT21: .56 bis .80, p jeweils $< .05$).

Eine Untersuchung der Struktur des Professionswissens für Lehrkräfte der Naturwissenschaften wurde im Rahmen des Projekts ProWiN durchgeführt. Für Biologie konnten Jüttner et al. (2013) mithilfe eines neuentwickelten Testinstruments zur Messung von PCK (vgl. Schmelzing et al., 2010; 2013; Jüttner und Neuhaus, 2013) bei einer Stichprobengröße von $n = 158$ Lehrkräften eine Korrelation von $r = .22$ ($p < .01$) zwischen CK und PCK messen. Auch für das Fach Chemie liegen Befunde zum Zusammenhang von Fachwissen und fachdidaktischem Wissen vor. Hier konnten Tepner und Dollny (2011) eine mittlere Korrelation von $r = .36$ ($p < .001$) zwischen CK und PCK messen. Auch hier ergaben sich ein, wenn auch zur MT21-Studie für angehende Mathematiklehrkräfte umgekehrter, Unterschied zwischen angehenden Gymnasial- und Nichtgymnasiallehrkräften. Für angehende Gymnasiallehrkräfte ergibt sich eine Korrelation von $r = .17$ ($p < .05$) (MT21: $r = .96$, $p < .05$) und für angehende Nichtgymnasiallehrkräfte von $r = .44$ ($p < .05$) (MT21: $r = .61$, $p < .05$). Das Professionswissen von Lehrkräften wurde im ProWiN-Projekt auch für das Fach Physik untersucht. Auf Basis eines entsprechenden Modells (Tepner et al., 2012) wurden dabei, auf der Basis des Modells professioneller Handlungskompetenz (Baumert und Kunter, 2006) ebenfalls die drei Wissensbereiche Fachwissen, fachdidaktisches Wissen und pädagogisches Wissen unterschieden. Bei einer Strukturanalyse konnte eine drei-

dimensionale Struktur des Professionswissens für Physiklehrkräfte bestätigt werden (Kirschner 2013). Dabei wurden die folgenden korrelativen Zusammenhänge zwischen PK, CK und PCK gefunden: CK-PCK: $r = .45, p < .001$; PCK-PK: $r = .27, p < .001$; CK-PK: $r = .17, p < .01$ (Kirschner, 2013).

Ein Modell des Professionswissens von Studierenden des Lehramts Physik wurde von Riese (2009) vorgeschlagen und geprüft. Bezüglich der Strukturanalyse ergab sich auch hier der stärkste Zusammenhang zwischen den Bereichen CK und PCK, wohingegen der Zusammenhang zwischen CK und PK wiederum am geringsten ausfiel: CK-PCK: $r = .49, p < .001$; PCK-PK: $r = .45, p < .001$; CK-PK: $r = .29, p < .001$ (Riese, 2009).

Die Korrelationen geben Aufschluss über die qualitativ unterschiedlichen Zusammenhänge zwischen den Wissensbereichen. Dabei ist festzustellen, dass sowohl Kirschner (2013) für Physiklehrkräfte als auch Riese (2009) für angehende Physiklehrkräfte die jeweils stärksten Zusammenhänge zwischen CK und PCK beobachteten, die schwächsten Zusammenhang dagegen zwischen CK und PK. Dies erscheint plausibel, denn es entspricht der Annahme, dass das fachdidaktische Wissen im Sinne der ursprünglichen Charakterisierung von PCK durch Shulman (1986) eine spezielle Art von CK darstellt, in dem Inhalte des CK mit solchen des PCK verschmelzen.

Eingeschränkt werden die Ergebnisse der zitierten Studien dadurch, dass bei den Untersuchungen von Kirschner (2013) und Riese (2009) beim Fachwissen eine Fokussierung auf den Inhaltsbereich Mechanik vorgenommen wurde, obwohl die Ausprägung der fachabhängigen Wissensbereiche als themen- und kontextspezifisch angenommen wird (Park und Chen, 2014). Die Validierung dieser Ergebnisse anhand einer umfassenden Konzeptualisierungen des Fachwissens und des fachdidaktischen Wissens stellen somit bislang eine Forschungslücke dar, die durch die Beantwortung der Forschungsfrage der vorliegenden Arbeit geschlossen werden soll.

Zusammenfassend kann die Gliederung des Professionswissens in die Wissensbereiche in Fachwissen, fachdidaktisches Wissen und pädagogisches Wissen also durch die berichteten Studien empirisch belegt werden. Beim Vergleich von Korrelationen zwischen den Wissensbereichen muss jedoch berücksichtigt werden, dass diese offenbar stark von der Operationalisierung im Rahmen der jeweiligen Testkonstruktion abhängen und gegebenenfalls von Curriculum und Organisation des Lehramtsstudiums im jeweiligen Land abhängen. Damit lassen sich die Ergebnisse zwischen den verschiedenen Studien nicht ohne weiteres übertragen.

(b) Wirksamkeit

Bezogen auf die Wirksamkeit des Professionswissens besteht in der fachdidaktischen Forschung Konsens darüber, dass ein ausgeprägtes Professionswissen, vor allem in den fachbezogenen Wissensbereichen, eine notwendige Voraussetzung für erfolgreiches Unterrichten darstellt (Hattie, 2003; Ball et al., 2005; Hill et al., 2005; Abell, 2007; Baumert und Kunter, 2011b). In welcher Weise sich ein ausgeprägtes Professionswissen entsprechend des Mehrebenenmodells in Abbildung 2.1 im Klassenkontext Unterrichtsprozesse auf kognitive Verarbeitungsprozesse der Schülerinnen und Schüler und damit letztlich auf die Zielkriterien des Unterrichts auswirkt, ist noch weitestgehend ungeklärt (Abell, 2007, S.1134). Im Folgenden werden daher Studien betrachtet, die Zusammenhänge zwischen Merkmalen der Lehrkraft und Merkmalen der Unterrichtsqualität identifizieren konnten. Weinert und Helmke (1996) berichten dazu die Ergebnisse einer Studie, in der die Leistungsentwicklung von Schülerinnen und Schülern in 47 Unterrichtsklassen im Fach Mathematik untersucht wurde. Anhand

dieser Leistungsentwicklung wurden die beteiligten Lehrkräfte in drei Gruppen, eine Negativgruppe, eine Mittelgruppe und eine Positivgruppe, aufgeteilt. Anschließend wurde die Abhängigkeit verschiedener Unterrichtsmerkmale von der Gruppenzugehörigkeit untersucht, um Erkenntnisse über die Wirksamkeit dieser Unterrichtsmerkmale zu generieren. Die Unterrichtsmerkmale sind in dem Mehrebenenmodell in Abbildung 2.1 der Stufe III zuzuordnen und gehören somit zu den Lehrermerkmalen oder resultieren direkt aus diesen. Die betrachteten Merkmale waren die Klarheit und Strukturiertheit des Unterrichts, die Effektivität der Klassenführung durch die Lehrkraft, das Aktivitätsniveau der Klasse, die individuelle Unterstützung der Schülerinnen und Schüler durch die Lehrkraft, die Adaptivität der Lehrkraft, das soziale Klassenklima und die Variabilität der Unterrichtsformen. Dabei zeigte sich, dass der Unterricht einer erfolgreichen Lehrkraft sich in besonderem Maße durch die Klarheit und Strukturiertheit des Unterrichts, durch die Effektivität der Klassenführung, durch das Aktivitätsniveau der Klasse und durch die Variabilität der Unterrichtsform auszeichnet (Weinert und Helmke, 1996, S.230). Damit ist offensichtlich, dass Variablen der Lehrkraft direkt oder indirekt auf den Unterrichtserfolg einwirken. Eine entsprechende Analyse breitgestreuter potentieller Einflussfaktoren veröffentlichte Hattie (2003). In einem ersten Schritt wurden dabei die Faktoren ‚Schüler‘, ‚Elternhaus‘, ‚Schule‘, ‚Schulleitung‘, ‚Freundeskreis‘ und ‚Lehrer‘ untersucht. Dabei konnte 30% der Varianz der Schülerleistungen durch Eigenschaften der Lehrkraft erklärt werden. Nach der Fähigkeit der Schülerinnen und Schüler selbst steht die Lehrkraft damit an zweiter Stelle der Faktoren für Schülerleistung. In einem zweiten Schritt wurden Eigenschaften analysiert, die einen „guten Lehrer“ ausmachen. Die folgenden fünf Eigenschaften konnten dabei als die bedeutsamsten identifiziert werden:

(1) Gute Lehrer verfügen über ein hohes Fachwissen und sind in der Lage, repräsentative Fachinhalte für ihre Schülerinnen und Schüler individuell angemessen zu organisieren und mit vorhandenem Wissen zu verknüpfen. (2) Sie schaffen ein gutes Lernklima in ihrer Klasse, das den Schülerinnen und Schülern ein effektives Lernen erlaubt. (3) Sie können Wissen und Fähigkeiten bewerten, wissen wie man Lernschwierigkeiten, wie z.B. Fehlvorstellungen, begegnet und sind in der Lage, ihren Schülerinnen und Schüler fachlich und pädagogisch angemessene Rückmeldungen zu geben. (4) Sie behandeln Schülerinnen und Schüler respektvoll, bringen selbst Begeisterung für Ihren Unterricht auf und vermitteln so positive Emotionen in Verbindung mit dem Lernprozess. (5) Sie ermuntern ihre Schülerinnen und Schüler durch angemessen schwierige Lernziele, die ermutigend und dadurch motivierend wirken.

Erste Ergebnisse einer Untersuchung des Zusammenhangs zwischen dem Professionswissen einer Lehrkraft und der Schülerleistung präsentieren Hill et al. (2005). Dabei zeigte sich, dass schulrelevantes Mathematikwissen den Lernfortschritt von Schülerinnen und Schülern in der ersten und dritten Klassenstufe positiv beeinflusst. Dies überrascht insofern, weil die Anforderungen hinsichtlich des fachlichen Niveaus in der Grundschule als eher gering angenommen werden dürfen. Der Befund bestätigt dennoch die Annahme, dass eine – wie auch immer geartete – Wirksamkeit des Fachwissens angenommen werden kann. Eine systematische Konzeptualisierung des Fachwissens und fachdidaktischen Wissens, wurden für Mathematiklehrkräfte in der COACTIV-Studie vorgenommen. Die Studienergebnisse liefern empirische Befunde zum Zusammenhang zwischen Fach- und fachdidaktischem Wissen und der Unterrichtsqualität, wie auch dem Lernfortschritt von Schülerinnen und Schülern (Kunter et al., 2011a). Dabei wurde festgestellt, dass für das kognitive Aktivierungspotential des Unterrichts vor allem das fachdidaktische Wissen und nicht das Fachwissen entscheidend ist. Das Fachwissen ist aber natürlich insofern nicht ohne Bedeutung, als dass es als Voraussetzung für die

Entwicklung fachdidaktischen Wissens gilt. Zudem konnte belegt werden, dass das curriculare Niveau von Lernaufgaben die Leistungsentwicklung der Schülerinnen und Schüler positiv beeinflusst. Ein höheres curriculares Niveau geht wiederum mit höherem Fachwissen einher (Baumert und Kunter, 2011b, S.183).

Diese Befunde verdeutlichen, dass Lehrkräfte insgesamt über eine breitgestreute Palette von notwendigen Eigenschaften verfügen müssen, um erfolgreich unterrichten zu können. Während einige dieser Eigenschaften, z.B. das respektvolle Verhalten gegenüber Schülerinnen und Schülern, womöglich stark mit der Persönlichkeit einer Lehrkraft verknüpft sind, lassen sich andere Eigenschaften, wie z.B. das Fachwissen und die Fähigkeit dieses adressatengerecht aufzubereiten, als erwerbbar Komponenten professionellen Wissens im Rahmen der Lehrerausbildung vermitteln. Die Befundlage erlaubt damit vorerst keine klare Zuordnung der Wirkung von Merkmalen der Lehrkraft zu Merkmalen des Unterrichts, zumal hierzu nur wenige quantitative Studien existieren, wie auch von Baumert und Kunter (2011b) angemerkt wurde. Zudem beziehen sich die Studien fast ausnahmslos auf das Fach Mathematik. Ein Problem im Forschungsfeld scheint in der Problematik der Erfassung handlungsrelevanten Wissens zu bestehen.

Für das Fach Physik konnte in einer Studie zur Wirkung physikalischen Fachwissens auf die Unterrichtsqualität gezeigt werden, dass das Fachwissen von Grundschullehrkräften die Unterrichtsqualität positiv beeinflusst (Ohle et al., 2011). Dabei wurde der Physikunterricht von $n = 58$ Primarstufenlehrkräften an der Grundschule videographiert. Die Unterrichtsqualität wurde anhand der Sequenzierung von Lernprozessen und der inhaltlichen Struktur des Unterrichts operationalisiert. Anhand dieses Ergebnisses kann angenommen werden, dass das Fachwissen auch den Lernerfolg der Schülerinnen und Schüler, moderiert durch die Unterrichtsqualität, positiv beeinflusst (Ohle et al., 2011). In einer Studie von $n = 20$ Physiklehrkräften an Gymnasien und Gesamtschulen konnte dagegen kein bedeutender Zusammenhang zwischen dem CK, PCK und PK der Lehrkräfte und dem Lernerfolg der Schülerinnen und Schüler gemessen werden (Cauet et al., 2013). Die Ergebnisse zum Lernerfolg beruhen dabei auf einer Messung der Schülerleistung nach einer Unterrichtseinheit zur Mechanik, sowie der Videographierung einer Unterrichtssequenz. Das Professionswissen der Lehrkräfte wurde durch einen Paper-und-Pencil-Test erhoben. Cauet et al. (2013) interpretieren das Ergebnis dahingehend, dass schriftlich abprüfbares Wissen möglicherweise keinen Einfluss auf die Schülerleistung habe. Auch Vogelsang et al. (2013) argumentieren, dass bei der schriftlichen Erfassung des Professionswissens nicht sichergestellt sei, dass zwingend dasjenige Wissen erfasst werde, das zum kompetenten Unterrichten notwendig sei. Die Autoren schlagen daher gemeinsame Analysen von Professionswissen und Unterrichtshandeln vor, um der Forschungslücke im Feld der Wirksamkeit des Professionswissens zu begegnen. Ein weiteres Problem mag darin bestehen, dass der Schritt vom Professionswissen zu Unterrichtsmerkmalen zu groß ist. Neuere Arbeiten fokussieren daher stärker auf den Planungsprozess, um zunächst Hinweise auf die Wirksamkeit des Professionswissens auf professionelles Handeln zu ermöglichen (Stender, 2014).

Zur Wirksamkeit des Professionswissens existieren damit bisher noch wenig konsistente Befunde. Es zeigt sich jedoch, dass Operationalisierungen und Untersuchungsmethoden offenbar einen deutlichen Einfluss auf die generierten Befunde haben. Man kann also resümieren, dass sowohl bei der Entwicklung valider Testinstrumente, als auch bei der Herstellung von Zusammenhängen zu Merkmalen der Unterrichtsprozesse weiterer Forschungsbedarf besteht, wobei das Ziel der vorliegenden Arbeit darin besteht, einen Beitrag zu Ersterem zu leisten.

(c) Entwicklung

Um die Wirksamkeit der Lehrerinnen- und Lehrerausbildung im Hinblick auf das vermittelte Professionswissen untersuchen zu können, muss identifiziert werden, in welchen Ausbildungsabschnitten sich dieses Wissen maßgeblich entwickelt. Für eine Optimierung der Lernabläufe ist darüber hinaus die Abbildung der zeitlichen Entwicklung des Professionswissens erforderlich. Für eine optimale Anordnung von Lehr- und Lerninhalten ist es ferner notwendig, die Inhalte im Sinne eines kumulativen Lernens sinnvoll aufeinander aufzubauen.

Hinsichtlich der strukturellen Entwicklung liefern verschiedene Studien Hinweise dafür, dass Fachwissen eine notwendige Voraussetzung für die Entwicklung fachdidaktischen Wissens darstellt (Brunner et al., 2006; 2011; Krauss et al., 2008; Riese, 2009; Kirschner, 2013). Dies äußert sich darin, dass die Stichproben zwar Probanden mit hohem Fachwissen und niedrigem fachdidaktischen Wissen enthalten, jedoch keine mit niedrigem Fachwissen und hohem fachdidaktischen Wissen.

Bezüglich der Entwicklung des Professionswissens in Relation zu Studienzeit bzw. Lerngelegenheiten ist die Befundlage für Fachwissen und fachdidaktisches Wissen nicht eindeutig. Während in der MT21-Studie gezeigt werden konnte, dass für angehende Mathematiklehrkräfte der Umfang fachlicher Lerngelegenheiten signifikant sowohl mit fachlichem als auch mit fachdidaktischem Professionswissen korreliert, konnte dagegen keine signifikante Vorhersagekraft fachdidaktischer Lerngelegenheiten für den Erwerb fachdidaktischen Wissens festgestellt werden (Blömeke et al., 2010, S.45/46). Die COACTIV-Studie liefert insoweit ähnliche Befunde, als dass das größere Ausmaß der Lerngelegenheit von Gymnasiallehrkräften im Studium gegenüber Nichtgymnasiallehrkräften zu einem deutlich stärker ausgeprägten Fachwissen führt. Beim fachdidaktischen Wissen jedoch verfügen Gymnasiallehrkräfte trotz geringem Ausmaß an Lerngelegenheit über ein stärker ausgeprägtes fachdidaktisches Wissen. Kleickmann und Anders (2011) führen als Erklärung hierfür die Bedeutung des Fachwissens für das fachdidaktische Wissen und Selektionsmechanismen zwischen Gymnasial- und Nichtgymnasiallehramt aufgrund von Personenmerkmalen an. Beim pädagogischen Wissen führte das größere Ausmaß an Lerngelegenheiten der Nichtgymnasiallehrkräfte zu einem höher ausgeprägten pädagogischen Wissen. Die Studie liefert zudem den Befund, dass Quereinsteiger ohne fachdidaktische und pädagogische Ausbildung in diesen Wissensbereichen schwächer abschneiden. Dieser Befund wird durch Riese et al. (2011) gestützt, die das schlechtere Abschneiden von Quereinsteigern in PCK-Tests durch fehlende Lerngelegenheiten im Studium erklären. Für das Fach Physik wurde im Rahmen des ProWiN-Projekts das Fachwissen von angehenden Lehrkräften, Referendaren und Lehrkräften untersucht und verglichen (Borowski et al., 2011). In dem eingesetzten Test erreichten Studierende durchschnittlich 37,9% der maximalen Punktzahl, Referendare 48,1% und Lehrkräfte 59,1%, womit sich die drei Probandengruppen jeweils signifikant voneinander unterscheiden. Für die im ProWiN-Projekt erhobenen Daten für Referendare deuten ebenfalls auf eine Abhängigkeit des fachdidaktischen Wissens von der unterrichtspraktischen Erfahrung hin (Borowski et al., 2010).

Ergebnisse zur Entwicklung des Professionswissens angehender Physiklehrkräfte innerhalb des Studiums liefern lediglich die Arbeiten der Paderborner Gruppe (Riese, 2009; Riese und Reinholt, 2012), die auf einer querschnittlichen Erfassung des Professionswissens von Studierenden für das Gymnasial- und für das Haupt- und Realschullehramt beruht. Die Testergebnisse von Studierenden im Grundstudium wurden mit denen im Hauptstudium verglichen. Über sämtliche Wissensbereiche und angestrebte Schulformen hinweg zeigte sich ein Vorsprung der Studierenden des Hauptstudiums, wobei fast alle beobachteten Messungen signifikante Unterschiede lieferten. Die Unterschiede zwischen

Grund- und Hauptstudium waren dabei bei den Studierenden des Gymnasiallehramts am größten. Die Effektstärken lagen nur im mittleren Bereichen (CK: Cohens $d = .4$; PCK: $d = .04$; PK: $d = .5$), was die Autoren damit erklären, dass sich die Studierenden der beiden verglichenen Gruppen teilweise hinsichtlich ihres Studienfortschritts nur gering voneinander unterscheiden könnten (Riese und Reinhold, 2012).

Die Befundlage zur Entwicklung des Professionswissens kann damit insofern zusammengefasst werden, als dass in allen Wissensbereichen des Professionswissens deutliche Zuwächse im Studium – sofern untersucht - zu verzeichnen sind. Dies zeigt sich besonders an den Vergleichen von Studierenden des Grundstudiums und Hauptstudiums (Riese, 2009; Riese und Reinhold, 2012) bzw. an Studierenden und Referendaren (Borowski et al., 2011) zeigt. Differenziertere Erkenntnisse zur Entwicklung innerhalb des Lehramtsstudiums im deutschen Bildungssystem liegen jedoch bislang nicht vor. Zudem wurde bei den vorliegenden Studien, die das Fach Physik betreffen, das Fachwissen wiederum lediglich anhand einer Fokussierung auf den Bereich der klassischen Mechanik (Borowski et al., 2010, Riese, 2009) vorgenommen, während das fachdidaktische Wissen bezogen auf die Inhaltsbereiche Mechanik und Elektrizitätslehre (Borowski et al., 2011) bzw. auf das Experimentieren (Riese, 2009) erhoben wurde. Befunde zur Entwicklung des Professionswissens, die auf einer umfassenderen Operationalisierung des physikalischen Fachwissens und des fachdidaktischen Wissens beruhen, liegen auch hier nicht vor. Die Frage nach der Entwicklung des Fachwissens und des fachdidaktischen Wissens erfordert daher die Operationalisierung einzelner Facetten des Professionswissens.

2.2 Facetten des Professionswissens im Fach Physik

Um einer umfassenden Operationalisierung der Wissensbereiche gerecht zu werden, müssen Facetten des Professionswissens zum einen den jeweiligen Wissensbereich sinnvoll strukturieren und zum anderen die Gesamtheit des Wissensbereiches abdecken. Der Fokus der vorliegenden Arbeit liegt dabei auf der Konzeptualisierung der fachbezogenen Wissensbereiche, also auf dem Fachwissen und dem fachdidaktischen Wissen.

2.2.1 Fachwissen

Die vorab genannten Projekte fokussieren bei der Modellierung des Fachwissens im Fach Physik auf einzelne Inhaltsbereiche, beispielsweise das Fachwissen im Bereich Mechanik (Riese, 2009; Woitkowski, 2011; Kirschner, 2012, Woitkowski, 2015). Dabei wird häufig darauf verwiesen, dass laut Friege und Lind (2004) Wissen im Bereich Mechanik einen guten Indikator für allgemeines Physikwissen darstelle (z.B. bei Riese und Reinhold, 2012, S.119). Die Ergebnisse der Studie von Friege und Lind zeigen jedoch lediglich, dass Faktenwissen in den Bereichen Mechanik und Elektrizitätslehre eng mit der Problemlösefähigkeit im Bereich Mechanik zusammenhängt ($r = .73$ bzw. $r = .74$). Ein Beleg, dass Wissen im Bereich Mechanik eine bessere oder schlechtere Prädiktivität für die allgemeine Ausprägung des Physikwissens bietet, ergibt sich daraus nicht. Diese können nur durch Vergleiche der Wissensentwicklung in verschiedenen Inhaltsbereichen gewonnen werden.

Durch eine Beschränkung auf den Einsatz von Aufgaben zur Mechanik wird nur ein Bruchteil des schulrelevanten Fachwissens (KMK, 2005) und der Lerngelegenheiten angehender Physiklehrkräfte im Studium abgebildet (KMK, 2008). Zudem wird in der Professionswissensforschung angenommen, dass das Fachwissen vom jeweiligen Inhaltsbereich abhängig ist (Shulman, 1986; Baumert und Kunter, 2006). Beispielsweise können Studierende durch die Auswahl einer Vertiefungsrichtung im Stu-

dium umfangreiches Wissen zu einer einzelnen Facette des Fachwissens, z.B. zur Quantenmechanik, erwerben. Bisherige Untersuchungen bilden solches Vertiefungswissen nicht adäquat ab. Hierfür ist die explizite Differenzierung zwischen Facetten des Fachwissens nötig. Eine Analogie beim fachdiaktischen Wissen ergibt sich durch die Wahlfreiheit der Studierenden zwischen einzelnen fachdidaktischen Seminaren. Die im Studium angebotenen Seminare decken üblicherweise distinkte fachdidaktische Inhalte ab.

Wenn das Ziel einer Untersuchung der Binnenstruktur des Professionswissens die Identifikation von Abhängigkeiten zwischen einzelnen Facetten beinhaltet, ist eine Erfassung sämtlicher relevanter Inhaltsbereiche ohnehin unumgänglich. Diese sind nach Messick (1995) so zu modellieren, dass einerseits kein verwandtes Konstrukt, wie z.B. Intelligenz oder Lesefähigkeit, erfasst wird, andererseits jedoch auch keine relevanten Teilbereiche des zu erfassenden Konstrukts ausgelassen werden. Bei den für angehende Physiklehrkräfte relevanten Teilbereichen der Physik handelt es sich primär um solche, die im Schulunterricht behandelt werden. Diese ergeben sich aus den Bildungsstandards der Kultusministerkonferenz (Ständigen Konferenz der Kultusminister der Länder). Ferner sind diese Inhalte in den daran angelehnten, aktuellen Lehrbüchern für den schulischen Physikunterricht enthalten. Neben dem an Schülerinnen und Schüler vermittelten Fachwissen sollten Lehrkräfte jedoch auch über ein umfassenderes Überblickswissen verfügen. Die Vermittlung dieses Wissens ist Ziel der universitären Lehramtsausbildung. Die an der Universität vermittelten Inhalte sind durch die Hochschulcurricula der ausbildenden Universitäten festgelegt. Die für angehende Physiklehrkräfte relevanten Inhaltsbereiche ergeben sich damit maßgeblich aus den Bildungsstandards der Kultusministerkonferenz, den Modulhandbüchern deutscher Universitäten sowie Lehrbüchern und Standardwerken für das Physikstudium und den Physikunterricht. Die in diesen Quellen enthaltenen Inhalte haben sich, im Vergleich zu aktuellen fachdidaktischen Theorien, über lange Zeiträume entwickelt und gelten daher als etabliert.

(a) Bildungsstandards

Die Kultusministerkonferenz hat zwei Arten von Bildungsstandards verabschiedet, die für die Frage, über welches Fachwissen Physiklehrkräfte verfügen müssen, von Bedeutung sind. Zum einen die Bildungsstandards für den Mittleren Schulabschluss im Fach Physik (KMK, 2005), zum anderen die ländergemeinsamen inhaltlichen Anforderungen für die Fachwissenschaften und Fachdidaktiken in der Lehrerbildung (KMK, 2008). Die Bildungsstandards für den Mittleren Schulabschluss geben, zusammen mit den Einheitlichen Prüfungsanforderungen für die Abiturprüfung (EPA) (KMK, 1979; Leisen, 2004) an, welche Kompetenzen Schülerinnen und Schüler im Unterricht mindestens erreichen sollen. Hieran müssen Physiklehrkräfte ihren Unterricht ausrichten. Die ländergemeinsamen inhaltlichen Anforderungen für die Fachwissenschaften und Fachdidaktiken geben an welche Anforderungen an die Lehrerinnen- und Lehrerausbildung gestellt werden. Im Folgenden werden die Inhalte dieser Verordnungen, bezogen auf das Fach Physik, dargestellt.

Bildungsstandards für den Mittleren Schulabschluss

Die Bildungsstandards im Fach Physik für den mittleren Schulabschluss (KMK, 2005) nennen Beiträge des Faches Physik zur Bildung und welche Kompetenzbereiche das Fach Physik umfasst. Die Bedeutung der Physik für Naturwissenschaft und Technik wird hier betont, da wesentliche Bereiche unserer

modernen Gesellschaft, wie z.B. die Energieversorgung und die Medizin, auf hochtechnisierten Verfahren beruhen, deren Grundlagen aus der Physik stammen.

Die Verordnung fordert, dass Schülerinnen und Schüler mit dem Erwerb des mittleren Schulabschlusses über Kompetenzen in den folgenden vier Bereichen verfügen:

(1) Das *Fachwissen* der Schülerinnen und Schüler umfasst Wissen über physikalische Phänomene, Begriffe, Prinzipien, Fakten, Gesetze und Basiskonzepte (Materie, Wechselwirkung, Systeme und Energie). (2) Im Bereich *Erkenntnisgewinnung* verfügen die Schülerinnen und Schüler über Kenntnisse experimenteller Untersuchungsmethoden, sowie der Nutzung von Modellen. (3) Schülerinnen und Schüler sind fähig zu fachbezogener *Kommunikation*, d.h., sie sind in der Lage sich Informationen sach- und fachgerecht zu erschließen und sich darüber mit anderen auszutauschen. (4) Schülerinnen und Schüler sind fähig zur *Bewertung* physikalischer Sachverhalte in verschiedenen Kontexten. Zusätzlich zur Beschreibung dieser vier Kompetenzbereiche definieren die Bildungsstandards drei Anforderungsbereiche. So wird bei der Kompetenz im Bereich Fachwissen zwischen den Anforderungsbereichen *Wiedergabe*, *Anwendung* und *Transfer* unterschieden.

Um diese Kompetenzbereiche erfolgreich vermitteln zu können, muss die Lehrkraft notwendigerweise selbst über die geforderten Kompetenzen verfügen. Darüber hinaus benötigen Lehrkräfte die Kompetenz zu deren adäquater Vermittlung. Um dies sicherzustellen, hat die Kultusministerkonferenz ländergemeinsame inhaltliche Anforderungen in der Lehrerbildung verabschiedet, die im Folgenden dargestellt werden sollen.

Einheitlichen Prüfungsanforderungen in der Abiturprüfung

Zur Realisierung einer bundesweiten Vergleichbarkeit der Abiturprüfung über die Bundesländer hinweg, verabschiedete die Kultusministerkonferenz 1979 einheitliche Prüfungsanforderungen für die Abiturprüfung (KMK, 1979). Diese Prüfungsanforderungen nennen über die Kompetenzen und Anforderungsbereiche hinaus konkrete Inhalte, über die Abiturienten verfügen sollen. Genannt werden hier als grundlegende Inhalte: Wissen aus den Sachgebieten Felder, Wellen, Quanten und Materie. Als optionale, ergänzende Sachgebiete werden Astrophysik, Chaos, Dynamik, Elektronik, Festkörperphysik, Relativitätstheorie und Thermodynamik genannt (Leisen, 2004).

Bildungsstandards für die Lehrerbildung

Die ländergemeinsamen inhaltlichen Anforderungen für die Fachwissenschaften und Fachdidaktiken in der Lehrerbildung (KMK, 2008) sollen als Grundlage für die Evaluierung lehramtsbezogener Studiengänge dienen und sind damit vor dem Hintergrund dieser Dissertationsarbeit von immenser Bedeutung. Die Bildungsstandards für Lehrerbildung enthalten die Standards für insgesamt 20 Fächer. Sie legen zu Beginn fachunabhängig fest, in welchen Phasen der Ausbildung welche Kompetenzen von angehenden Lehrkräften erworben werden sollen. Die Ausbildung angehender Lehrerinnen und Lehrer kann demnach in drei Phasen unterteilt werden. Die erste Phase, deren Wirksamkeit auch im Rahmen dieser Dissertationsarbeit untersucht wird, ist das universitäre Studium. Im Studium sollen angehende Lehrkräfte grundlegende Kompetenzen im Fachwissen, zu Erkenntnis- und Arbeitsmethoden sowie zu fachdidaktischen Anforderungen aufbauen. Die zweite Phase der Lehrerinnen- und Lehrerausbildung ist der Vorbereitungsdienst und dient primär dem Erwerb unterrichtspraktischer Kompetenzen, deren theoretische Grundlagen im Studium gelegt bzw. angebahnt worden sind. Die

dritte Phase besteht in den laufenden Fort- und Weiterbildungen neben der aktiven Rolle als Lehrkraft (KMK, 2008, S.2f).

In der ersten, also der universitären Ausbildungsphase, liegt der Fokus der Ausbildung auf der Vermittlung anschlussfähigen Fachwissens, von Erkenntnis- und Arbeitsmethoden des Faches und fachdidaktischen Wissens (KMK, 2008, S.3). Das Fachwissen umfasst dabei die grundlegenden Fachgebiete, bezogen auf das Fach Physik, also die für den Schulunterricht relevanten Inhaltsbereiche. Zusätzlich zum Fachwissen in verschiedenen Inhaltsbereichen sollen Lehrkräfte über Orientierungswissen verfügen, das ihnen einen Überblick über aktuelle gesellschaftlich relevante Fragestellungen bezüglich ihres Faches erlaubt. Ferner werden die ideengeschichtliche Entwicklung des Fachwissens, sowie die Verknüpfung des Fachwissens mit fächerübergreifenden Konzepten gefordert. Das fachdidaktische Wissen angehender Lehrkräfte umfasst Wissen über Instruktionsstrategien, das es ihnen ermöglicht, die Bildungswirksamkeit verschiedener fachdidaktischer Konzepte bezogen auf fachwissenschaftliche Inhalte zu analysieren, curriculares Wissen über die Strukturierungsansätze von Fachwissen, die Kenntnis über Ergebnisse fachdidaktischer und lernpsychologischer Forschung sowie die Fähigkeit, diese Erkenntnisse zu nutzen. Des Weiteren umfasst die Fachdidaktik Wissen zum Assessment, also Wissen darüber, wie man die Leistungen von Schülerinnen und Schülern fach- und anforderungsgerecht beurteilt, sowie Kenntnisse über die besonderen kognitiven Merkmale von Schülerinnen und Schülern, also solche Merkmale, die das Lernen von Schülerinnen und Schülern betreffen.

Der fachspezifische Teil der Bildungsstandards für die Lehrerbildung (KMK, 2008, S.38f) konkretisiert darüber hinaus die obigen allgemeinen Anforderungen und nennt darüber hinaus explizite Fachinhaltsbereiche, die Gegenstand des Unterrichts in den Sekundarstufen I und II sind. Das Lehramtsstudium für die Sekundarstufe I umfasst demnach Wissen in Mechanik, Thermodynamik, Elektrodynamik, Optik, Atom- und Quantenphysik, sowie ein Überblickswissen in Festkörper-, Kern- und Elementarteilchenphysik und Kosmologie. Für die Sekundarstufe II bzw. das Gymnasium wird darüber hinaus ein höherer Vertiefungsgrad des Wissens in Festkörper-, Kern- und Elementarteilchenphysik, Kosmologie, Theoretischer Mechanik, Thermodynamik, Elektrodynamik und Quantentheorie gefordert. Experimentelle und mathematische Grundlagen der Inhaltsbereiche werden als Anforderungen an einen Studienabschluss ebenfalls explizit gefordert.

(b) Hochschulcurricula

Nach Artikel 5, Absatz 3, des Deutschen Grundgesetzes gilt: „Kunst, Wissenschaft, Forschung und Lehre sind frei“. In Deutschland gilt damit das Prinzip der „Freiheit der Lehre“, d.h., dass Universitäten und ihre Dozenten bezüglich des Angebots und Inhalts von Lehrveranstaltungen prinzipiell nicht an staatliche Vorgaben, wie z.B. die aus den Bildungsstandards gebunden sind. Dennoch unterscheiden sich die Lehrangebote von Hochschulen hinsichtlich der Lehramtsausbildung im Fach Physik nur unwesentlich – primär hinsichtlich der Ausstattung ihrer physikalischen Sammlungen und experimenteller Angebote (Dreisigacker, 1990), jedoch kaum hinsichtlich ihrer inhaltlichen Angebote. In der Praxis orientieren sich Hochschulen bei der Ausarbeitung ihrer Modulpläne für das Lehramtsstudium an den entsprechenden Standards und Lehrplänen. Im Fach Physik unterscheiden sich die Rahmenvorgaben bzw. Modulpläne einzelner Universitäten inhaltlich also kaum voneinander. Um dies zu verdeutlichen sollen im Folgenden beispielhaft die Modulpläne von sechs Universitäten aus sechs verschiedenen Bundesländern verglichen werden, in denen für die vorliegende Studie Daten erhoben worden sind. Für den Vergleich wird die Christian-Albrechts-Universität zu Kiel, die Freie Universität

Berlin, die Gottfried Wilhelm Leibniz Universität Hannover, die Universität Hamburg, die Ludwig-Maximilians-Universität München und die Universität Duisburg-Essen betrachtet. Der Vergleich hebt dabei auf die inhaltliche Ausgestaltung des Physikstudiums ab. Formale Studienkriterien wie allgemeine Prüfungsbestimmungen usw. werden nicht betrachtet, da sie für die inhaltliche Betrachtung in diesem Kapitel nicht von Belang sind.

Der Modulplan für ein Bachelorstudium für das Lehramt im Fach Physik an der Universität Kiel (FPO, 2007) gibt die folgenden Lehrveranstaltungen vor: Physik I zur Mechanik und Wärmelehre, Physik II zur Elektrizitätslehre und Optik, Physik III zur Atom- und Quantenmechanik, Physik IV zur Kern-, Teilchen-, Astrophysik und Kosmologie, Theoretische Mechanik und Theoretische Physik für Lehramtsstudierende. Für das Masterstudium sind zusätzlich die folgenden Lehrveranstaltungen vorgesehen: Physik V zur Festkörper- und Oberflächenphysik, Fortgeschrittenenpraktikum I und II für Lehramtsstudierende, Fachdidaktik I, II und III.

Für die Universität Hannover werden in der Prüfungsordnung für den fächerübergreifenden Bachelorstudiengang (POH, 2013) die Fachinhalte der Module für ein Lehramtsstudium im Fach Physik festgelegt. Die dort genannten Fachmodule, die verpflichtend zu belegen sind, umfassen: Modul 1 zur Mechanik und Relativität, Modul 2 zur Elektrizitätslehre, Modul 3 zu mathematischen Methoden der Physik und theoretischer Elektrodynamik, Modul 4 zur Experimentalphysik mit den Inhalten Optik, Atomphysik, Quantenphänomene, Moleküle, Kerne, Teilchen und Festkörper, sowie das physikalische Grundpraktikum, Modul 5 zur theoretischen Physik. Neben diesen Pflichtmodulen nennt die Prüfungsordnung noch Wahlpflichtmodule, von denen zwei aus den Bereichen Festkörperphysik, Kohärente Optik, Atom- und Molekülphysik und Strahlenschutz, sowie ein Modul zum Lehren und Lernen im Physikunterricht oder eines aus dem Hauptfach-Bachelor-Studiengang zu wählen sind.

Die fachspezifischen Bestimmungen für Physik als Fach eines Bachelor-Studiengangs der Universität Hamburg (FBP, 2013) sehen die folgenden Module vor, deren Noten zu jeweils unterschiedlicher Gewichtung in die Gesamtnote der Bachelorprüfung eingehen: Physik I zur Mechanik und Wärmelehre, Physik II zur Elektrodynamik und Optik, Physik III zur Quantenphysik und Statistischen Physik, Theoretische Physik I zur theoretischen Mechanik, Theoretische Physik II zur Elektrodynamik und Quantenmechanik, sowie Anfänger- und Fortgeschrittenenpraktika.

Das Modulhandbuch (MHDB, 2013) der Ludwig-Maximilians-Universität München nennt neben mathematischen Grundlagen an Inhaltsbereichen der Physik: Mechanik, Wärme und Elektromagnetismus, theoretische Mechanik, elektromagnetische Wellen und Optik, Atom- und Molekülphysik, Elektrodynamik, Kern- und Teilchenphysik, Festkörperphysik und statistische Physik.

Die Prüfungsordnung und das Modulhandbuch der Universität Duisburg-Essen (FPO, 2011; MHB, 2011) weist im Studienplan die folgenden Module für Lehramtsstudierende aus: Mechanik, Elektrodynamik, Quantenphysik, Vielteilchensysteme, Physik als Unterrichtsfach, Theoretische Physik und Physik im Kontext mit den Themen Energie und Energieversorgung, Physik der festen Erde, Wirtschaftsphysik, Nanotechnologie sowie Ozean und Atmosphäre.

Die Studien- und Prüfungsordnung der Freien Universität Berlin (SPO, 2015) nennt für das Lehramtsstudium die folgenden Module für Lehramtsstudierende: Grundlagen der Experimentalphysik, physikalische Grundkompetenzen, Einführung in die Struktur der Materie, physikalisches Grundpraktikum 1 & 2, Demonstrationspraktikum, Theoretische Physik 1 & 2 sowie Moderne Physik. Die Module zur Experimentalphysik enthalten dabei Vorlesungen und Übungen zu den folgenden Themen: klassische

Mechanik, Thermodynamik, Relativität, Elektrodynamik und Optik. Das Modul zu den physikalischen Grundkompetenzen behandelt Inhalte aus der Mechanik, Thermodynamik, Elektrodynamik Optik und Atomphysik. In den Praktika werden Experimente zu den Inhaltsbereichen Mechanik, Thermodynamik, Atom- und Kernphysik (Teil 1) sowie Elektrizitätslehre und Optik (Teil 2) behandelt.

Zusammenfassend lässt sich feststellen, dass Hochschulcurricula üblicherweise in den Anfängervorlesungen für Lehramtsstudierende die Inhaltsbereiche Mechanik, Elektromagnetismus, Optik und Wärmelehre thematisieren und dass spätere Vorlesungen Inhalte zur Festkörper-, Teilchen, Atom- und Kernphysik sowie Quantenmechanik und Relativitätstheorie umfassen.

(c) Lehrbücher

Lehrbücher und Nachschlagewerke der Physik spielen sowohl für den Schulunterricht, wie auch für die Ausbildung von Lehrkräften eine wichtige Rolle. Sie enthalten, systematisch und didaktisch aufbereitet die grundlegenden Inhalte der Schul- oder Universitätsphysik. Die Lehrbücher für den Schulunterricht in Deutschland orientieren sich größtenteils an den Lehrplänen der verschiedenen Bundesländer. Sie stellen damit ein implementiertes Curriculum dar (Härtig et al., 2012). Sie werden in der schulischen Praxis zumeist von der jeweiligen Schule beschafft und den Schülerinnen und Schülern zur Verfügung gestellt. Die unterrichtenden Lehrkräfte können Sie dadurch für ihre Unterrichtsvorbereitung nutzen und sollten ihren Unterricht so gestalten, dass die Schülerinnen und Schüler zum Nachschlagen auf das jeweils eingesetzte Lehrbuch zurückgreifen können. Die intensive Nutzung von Lehrbüchern im Physikunterricht konnte durch Studien von Merzyn (1994) und Härtig et al. (2012) belegt werden. Ähnlich verhält es sich mit den Lehrbüchern für die Universitätsphysik. Hier werden in der Praxis von den Fachdozenten der Universitäten meist Empfehlungen für ergänzende Literatur gegeben. Die Auswahl und Beschaffung der Literatur liegt dann im Verantwortungsbereich der Studierenden. Zwar existieren einzelne Lehrbücher, die sich speziell an Lehramtsstudierende richten wie zum Beispiel die Lehrbücher für Theoretische Physik von Schilcher (2010) und Schmüser (2012a; 2012b), in der Regel werden jedoch allgemeine Standardwerke für das Physikstudium wie z.B. von Tipler (2014), Demtröder (2010, 2012, 2013a, 2013b), Gerthsen (2006), Halliday (2013) herangezogen.

Die im Folgenden betrachteten Lehrbücher zur Physik lassen sich also je nach Zielgruppe unterscheiden. Eine Gruppe der Schulbücher richtet sich primär an Schülerinnen und Schüler und thematisiert Inhaltsbereiche der Mittel- und Oberstufe. Die andere Gruppe richtet sich primär an Studierende der Physik oder anderer Fachrichtungen, für die physikalische Kenntnisse vorausgesetzt werden. Die Lehrbücher der zweiten Gruppe beinhalten neben den Inhaltsbereichen der Schulphysik noch solche, die vor allem an Universitäten gelehrt werden. Im Folgenden werden die Inhalte bekannter Schul- und Universitätslehrbücher auf die enthaltenen Fachinhalte hin analysiert. Da die Nomenklatur bezüglich der Fachinhaltsbereiche in den verschiedenen Lehrbüchern nicht einheitlich ist, wird – soweit fachlich möglich und sinnvoll – im Rahmen dieser Arbeit eine Zuordnung von Inhalten zu den bereits im vorherigen Kapitel herausgearbeiteten Inhaltsbereichen Mechanik, Elektromagnetismus, Optik, Wärmelehre, Festkörperphysik, Atom- und Kernphysik, Quantenmechanik und Relativitätstheorie vorgenommen. Ein Beispiel ist die Zuordnung von Inhalten zur *Elektrizitätslehre* und zum *Magnetismus* zum übergreifenden Inhaltsbereich *Elektromagnetismus* bzw. von Inhalten zur *geometrischen Optik* und *Wellenoptik* zum übergreifenden Bereich *Optik*.

Physiklehrbücher für den Schulunterricht

Die zugelassenen Lernmittel und Schulbücher für den gymnasialen Physikunterricht in den einzelnen Bundesländern können dem Deutschen Bildungsserver (2016) entnommen werden. Zu den etabliertesten deutschsprachigen Lehrbüchern für Schulphysik gehören „Metzler Physik“ (Grehn und Krause, 2007) sowie das Lehrbuch von Dorn und Bader (2001). Das Lehrbuch Metzler Physik (Grehn und Krause, 2007) aus dem Schroedel-Verlag, welches in zahlreichen deutschen Bundesländern, zum Beispiel in Nordrhein-Westfalen, Schleswig-Holstein und Bayern, im Oberstufenunterricht eingesetzt wird (Deutscher Bildungsserver, 2016) enthält sämtliche Inhalte für den Oberstufenunterricht. Auf 577 Seiten werden 15 Kapitel zu spezifischen Inhaltsbereichen der Physik behandelt. Inhalte zur Mechanik sind dabei in den Kapiteln (1) bis (3) mit den Bereichen Mechanik, Gravitation und Mechanische Schwingungen und Wellen dargestellt. Kapitel (4) behandelt die Thermodynamik gefolgt von drei Kapiteln (5) bis (7) über elektrische Ladungen und Felder, bewegte Ladungsträger und magnetische Felder sowie elektromagnetische Schwingungen und Wellen. Kapitel (8) behandelt Chaotische Vorgänge. Es folgen Kapitel (9) über Spezielle Relativitätstheorie, (10) Einführung in die Quantenphysik, (11) Atomphysik, (12) Festkörperphysik und Elektronik, (13) Kernphysik, (14) Elementarteilchenphysik und (15) Astrophysik. Die Bereiche Mechanik und Elektrizitätslehre bilden damit, mit jeweils drei Kapiteln, inhaltliche Schwerpunkte des Buches. Kapitel (8) über Chaotische Vorgänge und Kapitel (14) über Elementarteilchenphysik sind mit jeweils 10 Seiten gegenüber den anderen Kapiteln sehr knapp gehalten.

Ein weiteres gebräuchliches Schulbuch ist der Gesamtband von Dorn und Bader (2001) bzw. die Einzelbände für die Sekundarstufen I und II (vgl. Dorn und Bader, 2000), die ebenfalls im Schroedel-Verlag erschienen sind. Der Einfachheit halber wird im Folgenden Bezug auf den Gesamtband genommen. Dieser behandelt auf 512 Seiten insgesamt 21 Kapitel zur Schulphysik. Kapitel (1) behandelt die geometrische Optik. Die Kapitel (2) bis (7) behandeln Teilgebiete, die der Mechanik zuzuordnen sind: (2) Kräfte und Zusammenwirken von Körpern, (3) Geradlinige Bewegungen und Kräfte, (4) Fall- und Wurfbewegungen, (5) Erhaltungssätze, (6) Kreisbewegungen, (7) Gravitationsbewegung. Die Kapitel (8) Teilchenmodell und Druck und (9) Temperatur und Wärme behandeln Inhalte der Thermodynamik. Inhalte zum Elektromagnetismus sind in den Kapiteln (10) Magnetismus, (11) Ladung – Strom – Spannung, (12) Halbleiter, (13) Das elektrische Feld, (14) Das Magnetfeld und Teilchen in Feldern, (15) Induktion und Wechselstrom organisiert. Es folgen nochmals zwei Kapitel mit Inhalten der Mechanik zu Schwingungen und Wellen (Kapitel 16 und 17). Inhalte zur Wellenoptik finden sich in Kapitel (18) zu Interferenzphänomenen. Die Inhaltsbereiche Relativitätstheorie und Quantenmechanik werden relativ knapp im Kapitel (19), Physik des 20. Jahrhunderts, behandelt. Abschließend folgen Kapitel (20) zur Radioaktivität und Kernphysik und (21) zu Mensch und Energie, in der die Rollen von Energiebereitstellung und Energienutzung thematisiert werden. Man kann resümieren, dass in dem Lehrbuch von Dorn und Bader die Mechanik mit 8 Kapiteln gefolgt vom Elektromagnetismus mit 6 Kapiteln die inhaltlichen Schwerpunkte bilden. Noch recht umfassend sind Inhalte zur Optik und Wärmelehre behandelt. Die Inhaltsbereiche Festkörperphysik (im Rahmen des Kapitels zu den Halbleitern), Relativitätstheorie und Quantenmechanik (im Kapitel zur Physik des 21. Jahrhunderts) sind hier eher knapp gehalten.

Als drittes Beispiel für etablierte Lehrbücher im Schulbereich wird die Reihe ‚Physik für Gymnasien‘ I und II (Boysen et al., 1990; Boysen et al., 1995) bzw. ‚Physik Oberstufe Gesamtband‘ (Diel et al., 2009) aus dem Cornelsen Verlag betrachtet. In Band I werden die Inhaltsbereiche Optik und Mechanik behandelt. In Band II folgen weitere Inhalte zur Mechanik, sowie Wärmelehre, Elektrizitätslehre

und Kernphysik. Der Oberstufenband behandelt schließlich weitere Inhalte der Mechanik, Felder, Schwingungen und Wellen, Quantenphysik und Struktur der Materie, Relativität und Astrophysik sowie Thermodynamik.

Zusammenfassend lässt sich feststellen, dass in den betrachteten Schulbüchern der inhaltliche Fokus jeweils auf den Inhaltsbereichen Mechanik und Elektromagnetismus liegt. Diese nehmen in den betrachteten Lehrbüchern jeweils den größten Raum ein. Die Inhaltsbereiche Optik und Wärmelehre werden ebenfalls stets ausführlich behandelt. Themen moderner Physik, wie zum Beispiel Atom- und Kernphysik, Relativitätstheorie, Quantenmechanik und Festkörperphysik werden thematisiert, sind jedoch im Vergleich zu den vorher genannten Inhaltsbereichen eher knapp gehalten.

Physiklehrbücher für die Universität

Wie auch bei den Lehrbüchern für den Schulunterricht sollen an dieser Stelle exemplarisch etablierte Standardwerke für das Physikstudium auf die in ihnen behandelten Inhalte untersucht werden. Da sich im Lehrbetrieb üblicherweise sowohl für das reine Physikstudium, wie auch für das Lehramtsstudium von Dozenten empfohlen und von Studierenden verwendet werden, kann man davon ausgehen, dass sie die Prüfungsrelevanten Inhalte im Sinne eines implementierten Curriculums enthalten. Im Vergleich zu den zuvor betrachteten Lehrbüchern für den Schulunterricht enthalten die Lehrbücher für die Universität ähnliche Inhaltsbereiche und setzen ähnliche Schwerpunkte. Die betrachteten Inhaltsbereiche sind jedoch meist breiter angelegt und Inhalte der modernen Physik, also Phänomene der Relativitätstheorie, Quantenmechanik, Festkörperphysik, Teilchenphysik und Kosmologie gewinnen gegenüber den Schulbüchern an Stellenwert.

Als erstes etabliertes Lehrbuch ist das international anerkannte Standardwerk *Fundamentals of Physics* in der aktuellen 10. Auflage (Halliday, Resnick und Walker, 2013) zu nennen. Die Gesamtausgabe umfasst insgesamt 44 Kapitel auf 1232 Seiten. Das erste Kapitel kann dabei nicht direkt einem Inhaltsbereich zugeordnet werden, da es sich inhaltsübergreifend mit Problemen physikalischer Messungen, insbesondere Genauigkeit und Fehlern, auseinandersetzt. Die Kapitel 2 bis 17 befassen sich mit verschiedenen Bereichen der Mechanik: Ein- und mehrdimensionale Bewegungen, Wirkungen von Kräften, kinetische Energie, Arbeit, Energieerhaltung, Rotationsbewegungen, Gleichgewichte, Elastizität, Gravitation, Strömungsbewegungen sowie Schwingungen und Wellen. Damit bildet dieser Themenbereich einen inhaltlichen Schwerpunkt des Buches. Es folgen drei Kapitel (18 bis 20) zur Thermodynamik und 13 Kapitel (21 bis 33) zum Elektromagnetismus, womit dieser, neben der Mechanik, den zweiten inhaltlichen Schwerpunkt bildet. Es folgen zwei Kapitel (34 und 35) zur geometrischen Optik und Wellenoptik, Kapitel 37 zur Relativitätstheorie, zwei Kapitel (38 und 39) zur Quantenmechanik, Kapitel 40 zur Atomphysik, Kapitel 41 zur Festkörperphysik, Kapitel 42 und 43 zur Kernphysik sowie abschließend Kapitel 44 zur Teilchenphysik und der Urknalltheorie.

Ein weiteres internationales Standardwerk sind die *Feynman Lectures on Physics* (Feynman et al., 2009). Das dreibändige Werk enthält insgesamt 52 Kapitel auf 1376 Seiten. Es startet mit einem Kapitel zum Aufbau der Materie aus Atomen sowie deren Veränderung durch chemische Reaktionen. Das zweite Kapitel behandelt den Stand der Physik um das Jahr 1920 und die Rolle der Quantenphysik in diesem physikalischen Kontext. Angeschlossen ist die Beziehung der Physik zu den Wissenschaften Chemie, Biologie, Astronomie, Geologie und Psychologie. Nach diesen einführenden Kapiteln startet ein großer Bereich mit 16 Kapiteln zur Mechanik, der im wesentlichen mechanische Energieformen und deren Erhaltung, Bewegungsvorgänge, Dynamik, Gravitation, Impuls, Kräfte, Arbeit und Potenti-

elle Energie, Trägheitsmomente, Oszillatoren und Überlagerung von Schwingungen umfasst. Eingebettet in diesen Themenkomplex zur Mechanik sind zwei Kapitel zu mathematischen Grundlagen der Algebra und Vektorrechnung, sowie drei Kapitel zur Relativitätstheorie und Raumzeit. Anschließend folgen ein zehn Kapitel umfassender Themenkomplex zur Optik. Mit dem Fermatschen Prinzip wird zunächst die geometrische Optik eingeleitet, gefolgt von Phänomenen der Wellenoptik. Eingebettet in diesen Themenkomplex ist auch hier ein Kapitel zu relativistischen Effekten der Optik, wie zum Beispiel der Aberration. An die Optik schließen sich zwei Kapitel zur Quantenmechanik an, die Welle-Teilchen-Dualismus behandeln und somit noch eng mit der Wellenoptik verknüpft sind. Danach beginnt ein acht Kapitel umfassender Themenkomplex zur kinetischen Gastheorie, Diffusion, den Gesetzen der Thermodynamik und Entropie. Es folgen fünf Kapitel zu Schwingungen und Wellen in der Akustik, die wiederum am ehesten der Mechanik zuzuordnen sind und ein abschließendes Kapitel zu Analogien in physikalischen Gesetzen aus unterschiedlichen Inhaltsbereichen. Zusammenfassend lässt sich sagen, dass auch bei den Feynman Lectures der Hauptschwerpunkt auf den Inhaltsbereichen Mechanik und Optik liegt, wobei den Themen der Elektrizitätslehre und des Magnetismus keine eigenständigen Kapitel gewidmet sind. Themen zu diesen Bereichen sind selbstverständlich in dem Werk ebenfalls enthalten, sind jedoch in den Kontext anderer Inhaltsbereiche eingegliedert wie z.B. Elektromagnetische Wellen in das Kapitel zur Interferenz.

Ein im deutschsprachigen Raum weitverbreitetes Standardwerk stellt die Reihe zur *Experimentalphysik* vom Demtröder (2012, 2013a, 2010, 2013b) dar. In vier Bänden sind die wichtigsten Inhalte der Experimentalphysik zu den Bereichen Mechanik und Thermodynamik (Band 1), Elektrizitätslehre und Optik (Band 2), Atome, Moleküle und Festkörper (Band 3) sowie Kern-, Teilchen- und Astrophysik (Band 4) enthalten.

Ebenfalls im deutschsprachigen Raum weit verbreitet ist das Lehrbuch von Tipler (2014) für Wissenschaftler und Ingenieure. Es behandelt die folgenden Inhaltebereiche: Mechanik, Schwingungen und Wellen, Thermodynamik, Elektrizität und Magnetismus, Optik, Relativitätstheorie, Quantenmechanik, Atome und Moleküle, Festkörperphysik sowie Kern- und Teilchenphysik.

Ein weiteres deutschsprachiges Standardwerk der Physik auf universitärem Niveau ist die Lehrbuchreihe *Grundkurs Theoretische Physik* von Nolting (Nolting, 2006a, 2006b, 2005a, 2009, 2006c, 2007, 2005b). Die Reihe behandelt in sieben Bänden Themen der theoretischen Physik, wobei sich die Bände 1 und 2 der klassischen bzw. analytischen Mechanik widmen, Band 3 der Elektrodynamik, Band 4 der speziellen Relativitätstheorie und Thermodynamik, Band 5 der Quantenmechanik, Band 6 der statistischen Physik und Band 7 Viel-Teilchen-Theorie. Die statistische Physik behandelt dabei hauptsächlich Themen, die der Thermodynamik zugeordnet werden können. Bei den Inhalten aus Band 7 zur Viel-Teilchen-Theorie handelt es sich hauptsächlich um Inhalte, die der Mechanik und Festkörperphysik zugeordnet werden können.

Ein bekannter Gesamtband zur Physik ist *Gerthsen Physik* (Gerthsen et al., 2006). Die Inhalte zur Physik sind auf 1176 Seiten auf insgesamt 20 Kapitel aufgeteilt. Wie die übrigen Standardwerke beginnt auch *Gerthsen Physik* mit Inhalten zur Mechanik in den ersten vier Kapiteln. In den Kapiteln 1 und 2 werden zunächst grundlegende Eigenschaften der Mechanik von Massenpunkten und starren Körpern behandelt. Das 3. Kapitel behandelt die Mechanik deformierbarer Festkörper sowie von Flüssigkeiten und Gasen. Es folgt ein Kapitel zu Schwingungen und Wellen. Der zweite Themenkomplex, dem sich das Werk widmet, ist die Wärmelehre, die in Kapitel 5 abgehandelt wird. Die Kapitel 6, 7 und 8 widmen sich dem Elektromagnetismus: Kapitel 6 enthält Inhalte zu Ladungen und Strömen mit

dazugehörigen Phänomenen, wie beispielsweise der Bildung von Magnetfeldern. Auf die Zusammenhänge zwischen Elektrizität und Magnetismus wird im anschließenden Kapitel 7 zur Elektrodynamik genauer eingegangen, ehe im Kapitel 8 der Elektromagnetismus mit Phänomenen wie dem Photoeffekt und Gasentladungen in Zusammenhang gebracht wird. Themen zum Elektromagnetismus nehmen damit wie in den übrigen Werken nach der Mechanik den zweitgrößten Raum in *Gerthsen Physik* ein. Es folgen zwei Kapitel zur Optik, die in Geometrische Optik (Kapitel 9) und Wellenoptik (Kapitel 10) aufgeteilt sind. Die bereits in Kapitel 8 und 10 angeschnittenen Phänomene zu elektromagnetischer Strahlung werden im Kapitel 11 aufgegriffen und weitergeführt, allerdings hier eher unter thermodynamischen Gesichtspunkten (z.B. Wärmestrahlung und Stefan-Boltzmann-Gesetz). In den folgenden Kapiteln werden die Relativitätstheorie (Kapitel 12), Teilchenphysik (Kapitel 13) und Atom- und Quantenphysik (Kapitel 14) abgehandelt. Zwei Besonderheiten finden sich in *Gerthsen Physik* insoweit, als dass Inhalten zur Laserphysik (Kapitel 15: Laserphysik) und zu Berührungspunkten zwischen Physik und Chemie (Kapitel 16: Die Elemente und die Chemie) jeweils ein eigenes Kapitel gewidmet sind, die sich in anderen Lehrbüchern so nicht wiederfinden. In Kapitel 17 wird schließlich die Kern- und Teilchenphysik behandelt. Abschließend folgen zwei Kapitel zur Statistischen Physik (Kapitel 19) und Nichtlinearen Dynamik bzw. Chaostheorie (Kapitel 20). In den wesentlichen Punkten reiht sich *Gerthsen Physik* damit bezüglich seiner inhaltlichen Schwerpunktsetzung in die schon betrachteten Werke ein. Astrophysikalische Inhalte sind in den Themenkomplex der Mechanik eingebettet. Damit gibt es vier Kapitel zur Mechanik, drei zum Elektromagnetismus, zwei zur Optik und ein umfangreiches Kapitel zur Wärmelehre. Inhalte zur Festkörperphysik, Atom- und Kernphysik, Quantenmechanik und Relativitätstheorie sind mit jeweils eigenständigen Kapiteln enthalten oder werden mit verwandten Bereichen benachbart abgehandelt, wie Quantenmechanik mit Phänomenen der Atomphysik.

Zusammenfassende Feststellungen

Die Bildungsstandards nennen für den Schulunterricht Mindeststandards die anhand von Basiskonzepten, Kompetenzbereichen und Anforderungsniveaus von sämtlichen Schülerinnen und Schülern erreicht werden sollen. Für die Ausbildung angehender Lehrkräfte werden Kompetenzen und Anforderungen in spezifischen Inhaltsbereichen definiert, über die Lehrkräften nach ihrer Ausbildung verfügen sollen.

Die Curricula deutscher Hochschulen nennen konkrete Inhaltsbereiche, die im Rahmen der Ausbildung, zum Teil verpflichtend, zum Teil wahlweise, behandelt werden sollen. Die Inhaltsbereiche der Hochschulcurricula für Lehramtsstudiengänge decken sich dabei größtenteils mit den in den Bildungsstandards für angehende Lehrkräfte genannten Inhaltsbereichen. Zum Teil gehen die Inhalte der Hochschulcurricula über die Inhalte der Bildungsstandards hinaus.

Die betrachteten Lehrbücher, sowohl zur Schulphysik als auch zur Universitätsphysik unterscheiden sich inhaltlich nur marginal voneinander. Alle nehmen bei der Präsentation der Lehrinhalte eine systematische Gliederung nach physikalischen Inhaltsbereichen vor. Die Bedeutung und Relevanz dieser Inhaltsbereiche ist zumeist deckungsgleich. Ferner enthalten alle betrachteten Lehrbücher einen gemeinsamen Kern von Inhaltsbereichen, der als wesentlich und damit relevant für die Ausbildung angehender Lehrkräfte im Fach Physik anzusehen ist. Der Umfang der Inhalte geht dabei bei den Lehrbüchern auf universitärem Niveau zum Teil deutlich sowohl über die Anforderungen der Bildungsstandards als auch der Hochschulcurricula hinaus. Der fachliche Einstieg erfolgt üblicherweise

mit einer umfassenden Einheit zum Inhaltsbereich Mechanik, der sowohl bezüglich der Seiten- und Kapitelanzahl den inhaltlichen Schwerpunkt bildet. Inhalte zur Astronomie und Astrophysik sind dabei entweder direkt der Mechanik zugeordnet oder werden in separaten Kapiteln behandelt. Die Anordnung der folgenden Inhalte ist zwar uneinheitlich, gemeinsam ist allen Lehrbüchern jedoch, dass Inhalte zur Elektrizitätslehre und zum Magnetismus – separat oder kombiniert – den zweitumfangreichsten Inhaltsbereich bilden. Als weitere Schwerpunkte nehmen je nach Lehrbuch noch die Inhaltsbereiche Optik und Thermodynamik besonders viel Raum ein. Inhalte zur Festkörperphysik, Atom- und Kernphysik, Relativitätstheorie und Quantenmechanik werden, jedoch in unterschiedlichem Umfang, behandelt. Spezielle Inhalte, wie z.B. Chaostheorie und Laserphysik, werden nur in einzelnen Lehrbüchern ausführlich, in den meisten nur am Rande behandelt.

2.2.2 Fachdidaktisches Wissen

Beim PCK im Sinne Shulmans (1986) handelt es sich um einen Wissensbereich, in dem sowohl Wissen über das Fach als auch über Pädagogik zum Tragen kommen. Damit bildet das fachdidaktische Wissen eine lehrerspezifische Wissensdomäne, denn Lehrkräfte grenzen sich durch den Einfluss des pädagogischen Wissens von Fachwissenschaftlern und durch den Einfluss von fachlichem Wissen von Erziehungswissenschaftlern ab. Ein Beispiel für solches Wissen wäre die Kenntnis typischer Schülerfehlvorstellungen zu einem bestimmten Fachinhaltsbereich. Fachwissenschaftlern fehlt hier üblicherweise die Kenntnis typischer Schülerfehlvorstellungen und Lernschwierigkeiten, denen entgegengewirkt werden sollte. Erziehungswissenschaftler verfügen dagegen nicht unbedingt über die Fähigkeit die fachliche Angemessenheit von Schülerfehlvorstellungen beurteilen zu können.

Beim PCK handelt es sich im Vergleich zum CK und PK um ein relativ junges Konstrukt, das noch verhältnismäßig wenig erforscht ist. PCK äußert sich aufgrund seiner Natur als handlungsnahes Wissen von Lehrkräften eher implizit und lässt sich nur schwer verbalisieren (Kind, 2009; Riese, 2009). Eine Folge davon ist, dass es zum PCK keine explizit formulierten, verbindlichen Standards gibt. National wie international liegt zwar eine Vielzahl von Ansätzen zur Beschreibung dessen, was PCK umfassen sollte, vor (vgl. Park und Oliver, 2008), es existiert jedoch bislang kein etablierter Konsens zur Konzeptualisierung von PCK (Abell, 2007; Gramzow, 2014). Dieser Dissens spiegelt sich auch in der Ausbildung angehender Lehrkräfte wieder. Während die Ausbildung im Fach Physik in Deutschland bundesweit relativ kanonisch aufgebaut ist, gibt es im fachdidaktischen Bereich erhebliche Unterschiede. Beispielsweise wurden in den Bildungsstandards der Kultusministerkonferenz Empfehlungen zur Einführung von Basiskonzepten erteilt, jedoch werden diese nicht in allen Hochschulen im Rahmen der fachdidaktischen Ausbildung gelehrt.

Bisherige Konzeptualisierungen fachdidaktischen Wissens unterschieden sich mehr oder weniger deutlich voneinander. Ein Modell des Lehrerwissens von Grossmann (1990) zum Beispiel fokussiert beim PCK auf fachbezogenes Wissen über Schülerkognition, Curricula und Instruktionsstrategien, während Modellierungen von Tamir (1988) und Hashweh (2005) zusätzlich Assessment, also Wissen über die Bewertung und Beurteilung von Schülerinnen und Schülern, mit einbeziehen. Ein Überblick über unterschiedliche Konzeptualisierungen von PCK liegt mit einem Review von Park und Oliver (2008) vor, das auf einer Analyse von zwölf Arbeiten zum fachdidaktischen Wissen beruht (Shulman, 1987; Tamir, 1988; Grossman, 1990; Smith und Neale, 1989; Cochran et al., 1993; Geddis et al., 1993; Fernandez-Balboa und Stiehl, 1995; Magnusson et al., 1999; Hashweh, 2005; Loughran et al., 2006). Dabei ergibt sich der Befund, dass die von Grossmann (1990) genannten Facetten (siehe Abbildung

2.5) *Schülerkognition, Curriculum und Instruktionsstrategien* die mit Abstand gebräuchlichsten Facetten darstellen, während Wissen über *Assessment* und *Kontexte*, wenn auch weniger häufig, genannt werden.

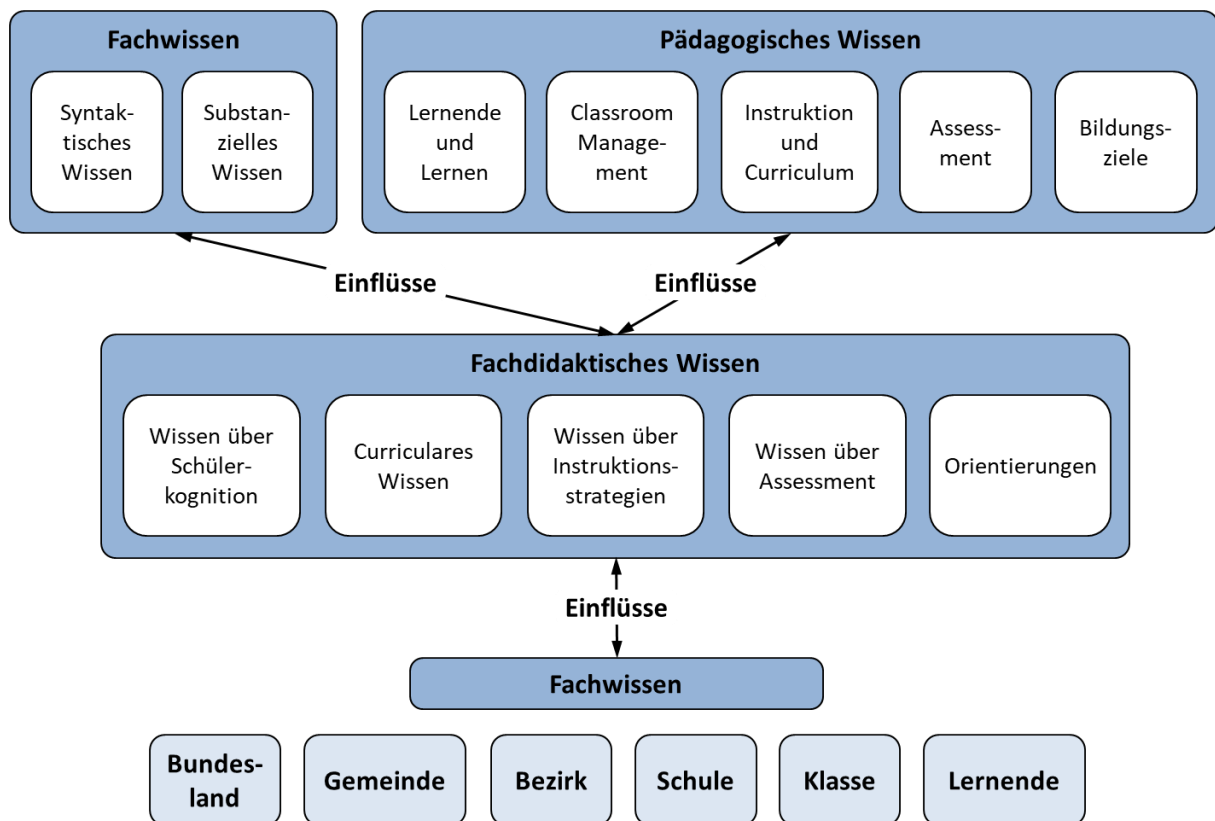


Abbildung 2.5: Komponenten des Professionswissens nach Grossman (1990).

Neuere entsprechende Übersichten, die auch die Arbeiten aus den Projekten COACTIV (Kunter et al., 2011a; 2011b), MT21 (Blömeke et al., 2008), *Quality of Instruction in Physics* (QuiP) (Olszewski, 2010), ProwiN (Tepner et al, 2012) sowie von Riese (2009) berücksichtigen, findet man bei Kirschner (2013, S.32) und Gramzow (2014, S.83). Da diese die obigen Befunde bezüglich der Häufigkeit der genannten Bereiche bestätigen, kann man davon ausgehen, dass das laut Gess-Newsome (2015) am häufigsten zitierte Modell für PCK von Magnusson et al. (1999), dargestellt in Abbildung 2.6, einen gebräuchlichen Kern von Facetten fachdidaktischen Wissens widerspiegelt.

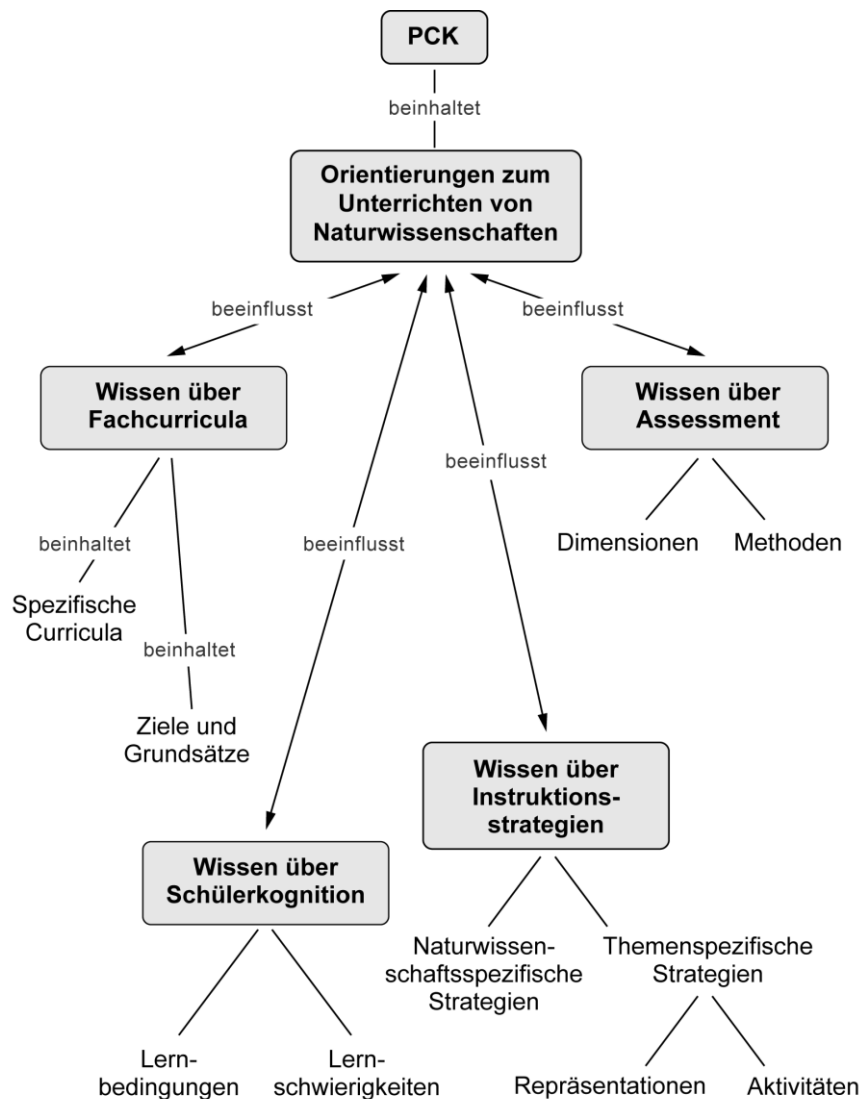


Abbildung 2.6: Komponenten fachdidaktischen Wissens für naturwissenschaftliches Unterrichten (Magnusson et al., 1999).

Magnusson et al. (1999) konkretisieren in ihrer Modellbeschreibung die Bedeutung der Facetten zum Wissen über Curricula, Schülerkognition, Instruktionsstrategien und Assessment. Sie bauen dabei auf den Arbeiten von Grossman (1990) und Tamir (1988) auf, die beim fachdidaktischen Wissen in den Naturwissenschaften die fünf Komponenten (a) *Orientierungen zum naturwissenschaftlichen Unterricht*, (b) *Wissen und Überzeugungen zu naturwissenschaftlichen Curricula*, (c) *Wissen und Überzeugungen zu inhaltspezifischen Schülerkognitionen*, (d) *Wissen und Überzeugungen zu Assessment* und (e) *Wissen und Überzeugungen zu Instruktionsstrategien naturwissenschaftlichen Unterrichts* unterscheiden. Bei den Orientierungen zum Unterrichten von Naturwissenschaften lassen sich bestimmte Schwerpunkte unterscheiden, die mit unterschiedlichen Zielsetzungen verbunden sind: Die Vermittlung naturwissenschaftlicher Arbeitsweisen, die Wichtigkeiten akkuraten und genauen Arbeitens, die Verwendung didaktisch sinnvoller Unterrichtsstrategien, die Anwendung von Umdeutungsstrategien, die Aktivierung der Lernenden im Unterrichtsgeschehen, der Einsatz schülerzentrierten Unterrichts im Sinne des forschend-entdeckenden sowie angeleiteten Lernens und die Einbindung von Projektarbeiten in den Unterricht. Die Orientierung der Lehrkraft zu diesen Schwerpunkten beeinflusst das Wissen über Curricula, Schülerkognition, Assessment und Instruktionsstrategien.

Wissen über Curricula umfasst laut Magnusson et al. (1999) zwei Kategorien: Kenntnisse über die Ziele naturwissenschaftlichen Unterrichts und über Lerneinheiten und Lernmaterial. Zum Wissen über Ziele naturwissenschaftlichen Unterrichts gehört das Wissen darüber, nach welchen Klassenstufen Schülerinnen und Schüler über bestimmtes Wissen und bestimmte Kompetenzen verfügen. Dieses Wissen ermöglicht es Lehrkräften im Unterricht optimal auf dem vorhandenen Vorwissen und den vorhandenen Kompetenzen der Lernenden aufzubauen. Unter dem Wissen über Lerneinheiten und Lernmaterial verstehen die Autoren Kenntnisse über Bildungsprogramme und vorhandene, etabliertes Unterrichtsmaterialien für bestimmte Inhalte und Klassenstufen.

Die Facette zum Wissen über Schülerkognition beinhaltet Kenntnisse, die Lehrkräfte benötigen, um ihre Schülerinnen und Schüler im Lernprozess zu unterstützen. Dazu gehört zum einen die angemessene Vermittlung naturwissenschaftlicher Konzepte und zum anderen Wissen über typische Lernschwierigkeiten und wie man diesen begegnen kann. Ersteres umfasst das Wissen der Lehrkraft über die Kompetenzen der Lernenden. Dieses Wissen ist nötig, um die zu bewältigende Hürde bei gestellten Aufgaben im Unterricht sinnvoll an die Vorkenntnisse und Fähigkeiten der Lernenden anzupassen. Das Wissen über typische Lernschwierigkeiten dagegen umfasst Wissen über Themen, die Lernenden typischerweise Schwierigkeiten bereiten. Die Ursache hierfür kann zum Beispiel die Abstraktheit des Lerngegenstandes sein. Ferner gehört zum Wissen über typische Lernschwierigkeiten die Kenntnis häufig auftretender Fehlvorstellungen, Kenntnisse über Strategien, wie man solchen Fehlvorstellungen im Unterricht begegnet und Kenntnisse darüber, welche Problemlösestrategien Lernende benötigen, um bestimmte Aufgaben bewältigen zu können.

Das Wissen über Assessment beinhaltet zum einen Wissen über die zu beurteilenden Dimensionen naturwissenschaftlicher Bildung. Das bedeutet Wissen darüber, welche Aspekte bestimmter Inhaltsbereiche für Lernende relevant sind und damit bewertet werden sollten. Zum anderen beinhaltet es Wissen darüber, welche Methoden der Bewertung es gibt und wie diese didaktisch sinnvoll umgesetzt werden können.

Wissen über Instrukionsstrategien kann aufgeteilt werden in Wissen über allgemeine, fachspezifische Instrukionsstrategien und Wissen über themenspezifische Instrukionsstrategien. Ersteres umfasst allgemeine Ansätze zur naturwissenschaftlichen Unterweisung, während letzteres auf Besonderheiten bestimmter Inhaltsbereiche abzielt. Beispiele für fachspezifische Instrukionsstrategien im Physikunterricht sind das Erkunden, Ausprobieren und Anwenden. Die Verwendung von Illustrationen, Beispielen und Analogien stellen Beispiele für themenspezifische Instrukionsstrategien dar.

Neuere Ansätze für eine gemeinsame und im Forschungsfeld anerkannte Modellierung fachdidaktischen Wissens greifen ebenfalls die von Magnusson et al. (1999) spezifizierten Kernfacetten auf und ergänzen *Wissen über Instrukionsstrategien und Darstellungsformen, Schülerkognition und naturwissenschaftliche Praxis und Einstellung* (Gess-Newsome, 2015). Bei diesen themenspezifischen Facetten handelt es sich um professionelles Wissen, bei dem man davon ausgeht, dass es kanonisch im Rahmen der Ausbildung vermittelt wird. Dieses Wissen kann eindeutig beschrieben werden und eignet sich daher für die Erfassung mittels Testverfahren für die Bewertung der Entwicklung des Professionswissens einer Lehrkraft (Gess-Newsome, 2015). Als Beispiel hierfür wird das Unterrichten zum Thema Kraft und Bewegung bei Drittklässlern angegeben. Hier sollte die Lehrkraft wissen, dass zunächst vereinfachend einzelne Kräfte und deren explizite Wirkung als Bewegungsänderung zu thematisieren sind, um die Lernenden nicht zu überfordern. Ferner ist es wichtig zu wissen, dass Drittklässler zum Verständnis dieses, im allgemeinen noch unbekanntes Konzept, Hilfen in Form von Visuali-

sierungen benötigen und eigene Erfahrungen, zum Beispiel durch Experimente oder Simulationen, machen müssen, um die Entwicklung von Fehlvorstellungen zu vermeiden (Gess-Newsome, 2015).

Von den themenspezifischen Facetten in der zweiten Ebene des Modells wird angenommen, dass sie, moderiert durch Verstärker und Filter wie Orientierungen und Überzeugungen der Lehrkraft, dem Vorwissen und dem jeweiligen Kontext, die im Modell als dritte Ebene dargestellt sind, direkt auf die Unterrichtspraxis wirken. Die Unterrichtspraxis ist im Modell als vierte Ebene dargestellt. Sie umfasst zum einen das fachdidaktische Wissen einer Lehrkraft, sowie den Unterrichtskontext. Unter dem fachdidaktischen Wissen einer Lehrkraft wird dabei zum einen das persönliche fachdidaktische *Wissen*, also das Wissen über Inhalte sowie über deren Relevanz und über die zugehörige Unterrichtsplanung verstanden. Zum anderen beinhaltet das fachdidaktische Wissen im Modell die *Fähigkeit* ein bestimmtes Thema auf eine bestimmte Art für bestimmte Lernende mit einem bestimmten Lernziel aufzubereiten (Gess-Newsome, 2015). Die Fähigkeit bzw. das Können beinhaltet es, während des eigenen Unterrichts Probleme zu erkennen und flexibel darauf reagieren zu können, also zum Beispiel spontane Änderungen der Unterrichtsplanung zu erwägen und umsetzen zu können. Dies kann zum Beispiel bei Verständnisschwierigkeiten einzelner Lernender das Anbieten zusätzlicher Repräsentationsform oder alternativer Erklärungen beinhalten. Dieses zweiteilige fachdidaktische Wissen, also Wissen bzw. Kenntnisse und Fähigkeit bzw. Können, wirkt in der Unterrichtspraxis je nach Kontext auf den Lernerfolg. Zwischen der Unterrichtspraxis und dem Lernerfolg stehen als fünfte Ebene im Modell wieder Verstärker und Filter, die die Wechselwirkung zwischen Unterrichtspraxis und Lernerfolg moderieren. Bei diesen moderierenden Verstärkern und Filtern handelt es sich um die Einstellungen, das Vorwissen und das Verhalten der Lernenden. Der Lernerfolg stellt schließlich die sechste und unterste Ebene des Modells dar. Es wird davon ausgegangen, dass der Lernerfolg der Lernenden auf sämtliche zuvor genannte Ebenen des Modells zurückwirkt, insofern, als dass er eine Lerngelegenheit für die Lehrkraft darstellt, sich auf die Motivation und Einstellung der Lehrkraft, sowie der Lernenden auswirkt und sich somit auch wechselwirkend auf die Unterrichtspraxis auswirkt.

Über das PCK-Modell durch Magnusson et al. (1999) und das Modell professioneller Handlungskompetenz von Baumert und Kunter (2006) beschränkt sich dieses Modell also nicht auf die Nennung und Gliederung von Wissens- bzw. Kompetenzkomponenten, sondern stellt darüber hinaus die Wechselwirkung der Komponenten im Unterrichtskontext dar, wobei zwischen Planungs- und Handlungsrelevantem Wissen unterschieden wird.

Das im Rahmen des Summits entwickelte Konsens-Modell für PCK könnte zukünftig einen wichtigen Schritt hinsichtlich der Vereinheitlichung von Modellierungen fachdidaktischen Wissens darstellen, wenngleich es komplexe Annahmen enthält, zu denen bislang keine Befunde für eine empirische Überprüfung vorliegen.

Als Resümee lässt sich feststellen, dass eine Vielzahl unterschiedlicher Konzeptualisierungen fachdidaktischen Wissens existiert, die sich hinsichtlich ihrer Schwerpunktbildung und der Ausdifferenzierung einzelner PCK-Facetten unterscheiden. Es lassen sich jedoch gemeinsame Kernfacetten fachdidaktischen Wissens identifizieren, die in einer Vielzahl von Konzeptualisierungen berücksichtigt werden. Dazu gehören insbesondere *Wissen über Schülerkognition* und *Wissen über Instruktionsstrategien* (Park und Oliver, 2008; Depaepe et al., 2013) und die von Magnusson et al. (1999) verwendeten und im Konsens-Modell des Professionswissens (siehe Abbildung 2.3) enthaltenen Facetten *Wissen über Assessment* und *Wissen über Curricula*.

2.2.3 Pädagogisches Wissen

Facetten des pädagogischen Wissens werden meist auf den Unterrichtskontext im Klassenverband bezogen betrachtet (Shulman, 1987) und beinhalten fachunabhängige Kenntnisse über (a) die Natur des Wissens, (b) des Lernens bzw. des Wissenstransfers und (c) der Motivation (Collins et al., 1992). Baumert und Kunter (2006) nehmen bei der Konzeptualisierung des Professionswissens im Rahmen der COACTIV-Studie eine Gliederung in (a) *konzeptuelles bildungswissenschaftliches Grundlagenwissen*, (b) *allgemeindidaktisches Konzeptions- und Planungswissen*, (c) *Wissen über Unterrichtsführung und Orchestrierung von Lerngelegenheiten* und (d) *Wissen über fachübergreifende Prinzipien des Diagnostizierens, Prüfens und Bewertens* vor.

Die Facetten des Professionswissens im COACTIV-Projekt werden von Voss und Kunter (2011) genauer beschrieben. Sie unterscheiden zunächst zwischen (1) *Wissen über Klassenprozesse* und (2) *Wissen über Schüler und Quellen für Heterogenität*. Das Wissen über Klassenprozesse lässt sich in drei weitere Bereiche unterteilen: (a) *Wissen über effektive Klassenführung* bezieht sich auf die Steuerung des Klassengefüges, so dass die zum Unterrichten verfügbare Zeit optimal genutzt wird, (b) *Wissen über Unterrichtsmethoden* beinhaltet die allgemeine Kenntnis und Fähigkeit verschiedene Unterrichtsmethoden effizient einzusetzen und (c) *Wissen über Leistungsbeurteilung*, also einerseits der Diagnose von Leistung zum Zweck einer Rückmeldung des Lernstandes und Optimierung des folgenden Unterrichts, zum anderen Bewertung im Sinne der Benotung der Schülerinnen und Schüler. Das Wissen über die Schüler und Quellen für Heterogenität wird weiter unterteilt in (d) *Wissen über individuelle Lernprozesse*, die ein Verständnis der Psychologie der Lernenden einschließt und eine optimale Unterstützung der individuellen Lernprozesse ermöglichen soll sowie (e) *Wissen über individuelle Besonderheiten von Schülerinnen und Schülern*, d.h. das Wissen und die Fähigkeit mit leistungsheterogenen Lerngruppen und verhaltensauffälligen Schülerinnen und Schülern pädagogisch angemessen umzugehen.

2.3 Wissensarten

Aus der Expertiseforschung ist bekannt, dass sich beim Vergleich von Novizen und Experten in einer Domäne zwischen Wissen und wissensbasiertem Handeln unterscheiden lässt (Bromme & Rambow, 2000). Dabei zeigt sich, dass Faktenwissen zwar eine notwendige Voraussetzung für die Fähigkeit Probleme zu lösen darstellt, dass es sich beim Wissen in einer bestimmten Domäne und der Problemlösefähigkeit in dieser Domäne jedoch um unterscheidbare Konstrukte handelt (Friege & Lind, 2004). Nach Klieme (2004) schließt Kompetenz sowohl Wissen als auch Können, im Sinne von fähigem Handeln, ein. Man kann also davon ausgehen, dass Wissen zu einer Fachrichtung in unterschiedlichen Formen vorliegen kann, die zu unterschiedlichem Handeln führen. Im Folgenden werden daher verschiedene Möglichkeiten zur Kategorisierung von Arten des professionellen Wissens im Fach Physik betrachtet.

Arten professionellen Wissens lassen sich unter verschiedenen Gesichtspunkten betrachten. Eine Kategorisierung ist zum Beispiel nach kognitivem Anspruchsniveau möglich. Sie kann sich danach richten, in welcher Institution das jeweilige Wissen vermittelt wird. In dieser Hinsicht ist insbesondere Schulwissen von Universitätswissen zu unterscheiden. Eine weitere Möglichkeit ergibt sich durch eine Kategorisierung nach Komplexität und Vernetzungsgrad. Hier kann zwischen Wissen, welches auf einem Fakt beruht, Wissen, welches auf zwei Fakten beruht und Wissen, welches auf dem Zusammenhang mehrerer Fakten beruht, unterschieden werden (Woitkowski, 2015).

Eine weitere Möglichkeit der Kategorisierung zielt auf die Art der Tätigkeit ab bei der Professionswissen herangezogen wird. Bezogen auf Lesestrategien unterscheiden Paris et al. (1983) hier zunächst zwischen deklarativem und prozeduralem Wissen. Dabei meint deklaratives Wissen zum einen einfaches und nicht notwendigerweise zusammenhängendes Faktenwissen. Darunter fällt beispielsweise die Kenntnis einer Definition, Formel oder eines einfachen Gesetzes. Zum anderen umfasst deklaratives Wissen auch Einstellungen und Selbsteinschätzungen. Als Beispiele hierfür werden die generelle Einstellung zum Lesen, beispielsweise "Lesen ist langweilig", und die Einschätzung der eigenen Lesekompetenz, beispielsweise "Ich bin ein langsamer Leser", angegeben. Unter prozedurales Wissen fällt dagegen das Wissen über typische Vorgänge und Abläufe. Als Beispiel wird hierzu das Wissen über die Technik des Überfliegens oder Querlesens genannt. Da deklaratives und prozedurales Wissen allein für ein strategisches Überblickswissen bzw. Handeln unzureichend seien, führen Paris et al. (1983) als dritte Wissensart konditionales Wissen ein. Dieses beinhaltet Wissen darüber, wann, wie und unter welchen Voraussetzungen verschiedene Fakten gelten oder welcher Handlungsablauf wann zielführend ist und darüber hinaus, wie sich dies jeweils begründen lässt. Ein Beispiel hierfür wäre zu wissen, in welcher Situation das Überfliegen bzw. Querlesen eine angemessene Technik darstellt und dies auch begründen zu können. Das konditionale Wissen stellt somit eine Art Überblicks- und Begründungswissen dar, das bei der Lösung eines Problems zur Auswahl hilfreicher Fakten und Verfahrensweisen befähigt.

Die drei genannten Wissensarten wurden von Shavelson et al. (2005) aufgegriffen, die sie als geordnete Menge, die einem zunehmenden Grad der Professionalisierung bzw. der Fähigkeit einer Person entsprechen, aufgreifen. Die erste Stufe dieser Fähigkeitsskala stellt dabei das deklarative Wissen ("knowing that") dar. Innerhalb des deklarativen Wissens lassen sich noch Unterscheidungen bezüglich des Wissensumfanges und der Struktur treffen. Bezüglich der Struktur wird zwischen Fakten, Definitionen und Beschreibungen unterschieden. Die zweite Stufe der Fähigkeitsskala stellt das prozedurale Wissen ("knowing how") dar. Dieses beinhaltet bestimmte Regeln und Abläufe für typische Handlungen in einer Wissensdomäne. Beim konditionalen Wissen unterscheidet Shavelson et al. (2005) im Gegensatz zu Paris et al. (1983) zwischen schematischem Wissen über allgemeine Schemata und mentale Modelle ("knowing why"), welches die dritte Stufe der Fähigkeitsskala darstellt und strategischem Wissen über Strategien und domänenspezifischen Vorgehensweisen ("knowing when, where and why"), die schließlich die vierte und letzte Stufe der Personenfähigkeit darstellt.

2.4 Korrelate des Professionswissens

Nachdem in den bisherigen Abschnitten die theoretischen Grundlagen, sowie Befunde zur Struktur und Entwicklung des Professionswissens zusammengetragen wurden, soll nun betrachtet werden, inwieweit das Professionswissen von Lehrkräften und angehenden Lehrkräften mit anderen Konstrukten zusammenhängen kann.

Zu nennen sind hier zum einen die von Weinert (2001) sowie Baumert und Kunter (2006) identifizierten Aspekte, die neben dem Professionswissen zur professionellen Handlungskompetenz beitragen und weitere individuelle Persönlichkeitsmerkmale einer Lehrkraft. Dazu gehören (a) Überzeugungen und Werthaltungen, zum Beispiel in Bezug auf die eigene Fachdisziplin. Stehen diese im Widerspruch zu fachlich angemessenen Vorstellungen, so führt dies zu einem Konflikt, der die positive Entwicklung professionellen Wissens hemmen kann (Blömeke, 2008). Weitere Aspekte der professionellen Handlungskompetenz, die im Zusammenhang mit der Entwicklung des Professionswissens betrachtet wer-

den müssen, sind (b) motivationale, volitionale und soziale Orientierungen, (c) selbstregulative Fähigkeiten, wie zum Beispiel Selbstwirksamkeit und (d) kognitive Fähigkeiten (Baumert und Kunter, 2006).

Im Rahmen des COACTIV-Projekts wurden die Wirkungen von Persönlichkeitsmerkmalen wie kognitive Fähigkeiten, Überzeugungen, Motivation und Selbstregulation auf die Kompetenzentwicklung untersucht (Klusmann, 2011; Kunter et al., 2011a; 2011b; Voss et al., 2011a). Die Befunde zeigten, dass sich alle Bereiche positiv auf die Unterrichtsqualität auswirken, was als Bestätigung des Modells der professionellen Handlungskompetenz (siehe Abbildung 2.2) verstanden werden kann. Bezüglich der Motivation zeigten die Daten jedoch, dass motivationale Orientierungen nicht als zeitlich stabil gelten können, sondern mitunter sogar im Laufe eines Jahres Veränderungen unterworfen sind. Ferner konnte die COACTIV-Studie laut Kunter (2011b, S.64) hinsichtlich der individuellen Eingangsvoraussetzungen der Probanden nur begrenzte Erkenntnisse zur Entwicklung der professionellen Kompetenz liefern, da die Probanden erst nach Abschluss der beruflichen Ausbildung getestet wurden. Hier sei eine Ergänzung der Studie hilfreich, welche die Phase des Referendariats eingehender Lehrkräfte mit einschließt.

Der Einfluss der Personenmerkmale Geschlecht, Alter, Bundesland in dem unterrichtet wird, Unterrichtserfahrung, begleitendes Mathematikstudium und Lehramtszugang (gymnasial vs. nichtgymnasial) von Physiklehrkräften auf die Ausprägungen von Fachwissen und fachdidaktischem Wissen wurde im Rahmen des ProWiN-Projekts untersucht (Kirschner, 2013). Dabei konnte gezeigt werden, dass Lehrer im Mittel sowohl über ein signifikant (jeweils $p < .05$) höheres Fachwissen als auch über ein höheres fachdidaktisches Wissen verfügen als Lehrerinnen. Jüngere Lehrkräfte verfügen im Mittel über signifikant geringeres fachdidaktisches Wissen. Lehrkräfte in Bayern verfügen im Mittel, verglichen mit anderen Bundesländern, über ein signifikant höheres Fachwissen, wohingegen Lehrkräfte in Nordrhein-Westfalen im Mittel über ein signifikant geringeres Fachwissen verfügen als in anderen Bundesländern. Ebenso konnte gezeigt werden, dass Unterrichtserfahrung mit signifikant höherem Fachwissen zusammenhängt (Kirschner, 2013).

Empirische Befunde für Zusammenhänge zwischen Überzeugungen und dem Fachwissen, fachdidaktischem Wissen und pädagogischem Wissen angehender Physiklehrkräfte wurden von Riese und Reinhold (2012) untersucht. Hierbei zeigten sich besonders hohe und signifikante Korrelationen zwischen dem Professionswissen und wissenschaftstheoretischen Überzeugungen: Für die Skala *Beliefs zur Natur des Wissens* ergaben sich Korrelationen von $r = .32$ ($p < .001$) zum Fachwissen, $r = .31$ ($p < .001$) zum fachdidaktischen Wissen und $r = .28$ ($p < .01$) zum pädagogischen Wissen. Für die *Bedeutung des Experiments bei der Erkenntnisgewinnung* ein Zusammenhang von $r = .31$ ($p < .001$) zum Fachwissen, von $r = .32$ ($p < .001$) zum fachdidaktischen Wissen und von $r = .29$ ($p < .001$) zum pädagogischen Wissen. Beliefs über *Rezeptartiges Lernen* korrelierten dagegen erwartungskonform negativ mit $r = -.19$ ($p < .01$) zum Fachwissen, $r = -.28$ ($p < .001$) zum fachdidaktischen Wissen und $r = -.20$ ($p < .01$) zum pädagogischen Wissen. Die Autoren erklären dies damit, dass Probanden, die ein rezeptartiges Lernen im Unterricht bevorzugen, dies auch für ihr eigenes Lernen tun, was der Erlangung eines tiefen Verständnisses im Wege stünde. Für weitere Beliefs zu epistemologischen Vorstellungen, Freude an neuen Dingen, Einstellungen zum Experimentieren im Physikunterricht sowie Schülerorientierung konnten keine oder nur geringe Zusammenhänge zu einzelnen Wissensbereichen gefunden werden. Überraschenderweise lieferten auch Skalen zu allgemeinen Selbstwirksamkeitserwartungen durchweg keine signifikanten Zusammenhänge zum Professionswissen. Die Selbstwirksamkeit bezogen auf das Physik-Lernen korrelierte nur mit dem Fachwissen (mit $r = .35$,

$p < .001$), jedoch weder mit dem fachdidaktischen noch mit dem pädagogischen Wissen, was die Autoren auf die Fachbezogenheit der verwendeten Skala zurückführen. Bei Erklärungen solcher Korrelationen muss auch beachtet werden, dass zunächst lediglich auf einen Zusammenhang, jedoch nicht auf die Reihenfolge einer Kausalkette geschlossen werden kann (vgl. Riese und Reinhold, 2010). Anhand dieser Ergebnisse kann festgehalten werden, dass für angehende Physiklehrkräfte zumindest die Angemessenheit wissenschaftstheoretischer Vorstellungen mit einem durchweg hohen Professionswissen einhergeht. In der Arbeit wurde zudem ein signifikanter Einfluss des Geschlechts zugunsten männlicher Probanden gefunden. Ebenfalls sollten kognitive Fähigkeiten der Probanden berücksichtigt werden. Als einfachster Indikator hierfür bietet sich die Betrachtung der Abiturnote an. Für genauere Betrachtungen sollten kognitive Grundfähigkeiten differenziert gemessen werden (vgl. Woitkowski, 2015).

Zusammenfassung des Forschungsstandes

Die zentrale Rolle der Lehrkraft für die Qualität von Unterricht ist unstrittig. Darüber hinaus kann von einem positiven Einfluss der professionellen Kompetenz im Allgemeinen und des Professionswissens im Speziellen auf die Unterrichtsqualität ausgegangen werden.

Darüber, was das Professionswissen einer Lehrkraft genau umfassen sollte, besteht bislang noch kein umfassender Konsens. Etabliert hat sich jedoch eine strukturelle Einteilung des Professionswissens in die Bereiche Fachwissen, fachdidaktisches Wissen und pädagogisches Wissen (siehe Abschnitt 2.1.1). Ob sich diese drei Bereiche, insbesondere für angehende Lehrkräfte, getrennt voneinander erfassen lassen und inwieweit sie jeweils hinsichtlich der Unterrichtsqualität wirksam sind, ist empirisch noch nicht belegt (siehe Abschnitt 2.1.2). Darüber, was die drei Bereiche jeweils umfassen, besteht am ehesten beim Fachwissen ein Konsens. Bildungsstandards der Bundesländer, Lehrpläne der Hochschulen und gängige Lehrwerke sind bezüglich der enthaltenen Fachinhaltsbereiche weitestgehend deckungsgleich (siehe Abschnitt 2.2.1). Zum fachdidaktischen Wissen existiert eine Vielzahl unterschiedlicher Konzeptualisierungen, die jedoch vier gemeinsame Kernfacetten enthalten (siehe Abschnitt 2.2.2). Ähnliches gilt für das pädagogische Wissen. Hier werden fünf Facetten des pädagogischen Wissens als relevant betrachtet (siehe Abschnitt 2.2.3). Unabhängig von den drei Bereichen des Professionswissens wird von drei verschiedenen Wissensarten ausgegangen (siehe Abschnitt 2.3).

Zu Merkmalen, die die Entwicklung von Professionswissen begünstigen können, gibt es bislang nur wenige Erkenntnisse. Diese beziehen sich zumeist auf ausgebildete Lehrkräfte. Der Einfluss von Persönlichkeitsmerkmalen und Begleitumständen auf das Professionswissen, insbesondere in der Ausbildungsphase, ist in weiten Teilen noch nicht systematisch erforscht (siehe Abschnitt 2.4).

3. Forschungsfragen

Im Folgenden werden aus dem Forschungsstand resultierende Defizite abgeleitet. Zu den einzelnen Forschungslücken wird jeweils eine Forschungsfrage und eine Hypothese formuliert.

3.1 Erste Forschungsfrage

In den erwähnten Studien COACTIV zu Mathematiklehrkräften (Kunter et al., 2011), ProwiN zu Physiklehrkräften (Kirschner, 2013) und Riese (2009) zu angehenden Physiklehrkräften konnte jeweils eine empirische Trennbarkeit der Bereiche CK, PCK und PK festgestellt werden, was für eine Berücksichtigung dieser drei Bereiche bei der Modellierung des Professionswissens spricht. Dabei muss allerdings berücksichtigt werden, dass die betrachteten Untersuchungen, sowohl im COACTIV-Projekt als auch im ProwiN-Projekt, jeweils auf ausgebildete Lehrkräfte abzielten. Ferner wurden die Befunde im COACTIV-Projekt im Fach Mathematik generiert. Eine generelle Übertragbarkeit dieser Ergebnisse auf die naturwissenschaftlichen Fächer, und insbesondere auf das Fach Physik, ist damit nicht notwendigerweise gegeben. Im ProwiN-Projekt wurde zwar das Professionswissen im Fach Physik untersucht, allerdings nur für Lehrkräfte nach Ausbildungsabschluss. Die Studie von Riese (2009) für angehende Physiklehrkräfte fokussierte auf das Professionswissen während der universitären Ausbildungsphase. Diese Untersuchung des Fachwissens beschränkt sich jedoch auf den Inhaltsbereich Mechanik und beim fachdidaktischen Wissen auf das Experimentieren.

Es liegen also bislang keine Erkenntnisse darüber vor, ob sich die bisher gefundenen Strukturen des Professionswissens, also die Zusammenhänge von CK, PCK und PK, auf der Basis einer inhaltlich breit angelegten Operationalisierung für angehende Physiklehrkräfte bestätigen lassen. Die erste Forschungsfrage fokussiert daher auf die Struktur des Professionswissens.

Forschungsfrage 1 (FF1)

Inwieweit lässt sich die Strukturierung des Professionswissens in Fachwissen, fachdidaktisches Wissen und pädagogisches Wissen für angehende Physiklehrkräfte bestätigen?

Hypothese 1 (H1)

Eine empirische Trennbarkeit der drei Wissensbereiche Fachwissen, fachdidaktisches Wissen und pädagogisches Wissen wird auch für angehende Physiklehrkräfte erwartet.

3.2 Zweite Forschungsfrage

Neben der Struktur des Professionswissens stellt sich die Frage nach der Ausprägung des Fachwissens und des fachdidaktischen Wissens angehender Physiklehrkräfte zu verschiedenen Phasen der universitären Ausbildung. Die Ausprägung dieser Merkmale in Abhängigkeit von der Studiendauer soll daher im Zuge der zweiten Forschungsfrage untersucht werden. Da die Gewichtung von Fachstudium und fachdidaktischem Studium über die Studiendauer hinweg nicht homogen ist, sind deutlich unterschiedliche Lerngelegenheiten für Fachwissen und fachdidaktisches Wissen zu verschiedenen Studienphasen zu erwarten. Legt man zugrunde, dass im Grundstudium bzw. in den ersten zwei Studienjahren die Grundvorlesungen des Faches im Fokus stehen und Schulpraktika sowie fachdidakti-

sche Vorlesungen im Studienverlauf erst später in den Fokus rücken, dann könnte sich dies in der Entwicklung des Professionswissens messbar widerspiegeln.

Forschungsfrage 2 (FF2)

Wie ausgeprägt sind Fachwissen und fachdidaktisches Wissen angehender Physiklehrkräfte in verschiedenen Phasen des Lehramtsstudiums?

Hypothese 2 (H2)

Mit zunehmender Studiensemesterzahl wird ein stetiger Anstieg des Professionswissens erwartet. Im ersten Studienjahr wird der deutlichste Anstieg des Fachwissens erwartet, während im zweiten und dritten Studienjahr der deutlichste Anstieg des fachdidaktischen Wissens erwartet wird.

3.3 Dritte Forschungsfrage

Schließlich stellt sich die Frage, wie sich die Entwicklung des Professionswissens angehender Physiklehrkräfte positiv beeinflussen lässt. Im Rahmen der dritten Forschungsfrage soll deshalb untersucht werden, welche Faktoren die Ausprägung des Professionswissens positiv oder negativ beeinflussen. Neben dem in Forschungsfrage 2 genannten Studienfortschritt und den damit verbundenen Lerngelegenheiten sollten auch Persönlichkeitsmerkmale von Lehramtskandidaten einen Einfluss auf die Entwicklung des Professionswissens haben. Der Forschungsstand (siehe Abschnitt 2.4) legt einen allgemein hohen Einfluss von kognitiven Fähigkeiten, Motivation, Interesse und Selbstwirksamkeit nahe. Ebenfalls könnte der sozioökonomische Hintergrund eine Rolle spielen, da die Bereitstellung zeitlicher und finanzieller Ressourcen eine notwendige Bedingung für Studierfähigkeit darstellen. Ferner könnten auch Vorwissen und gegebenenfalls Unterrichtserfahrungen eine Rolle spielen, die sich im Studienverlauf abschwächt.

Forschungsfrage 3 (FF3)

Welche Faktoren beeinflussen die Ausprägung des Professionswissens angehender Physiklehrkräfte?

Hypothese 3 (H3)

Es wird erwartet, dass eine stärkere Ausprägung von kognitiven Fähigkeiten, Motivation, Interesse, Selbstwirksamkeit, Vorwissen und Unterrichtserfahrung bzw. ein solider sozioökonomischer Hintergrund sich jeweils positiv auf die Ausprägung des Professionswissens angehender Physiklehrkräfte auswirken.

4 Methode

Die Bearbeitung der im vorherigen Kapitel formulierten Forschungsfragen erfolgte im Rahmen eines übergeordneten Forschungsvorhabens, des Projekts „Messung professioneller Kompetenzen in mathematischen und naturwissenschaftlichen Lehramtsstudiengängen“ (KiL). In diesem Projekt wurden Testinstrumente für das Professionswissen in den Naturwissenschaften und Mathematik entwickelt, pilotiert und normiert. Über die Testentwicklung sowie die Pilotierungs- und Normierungsstudie hinaus wurden im Rahmen der vorliegenden Arbeit, zusätzlich zum KiL-Projekt, verschiedene Validierungsstudien durchgeführt, um eine hohe Testgüte sicherzustellen.

Im folgenden Abschnitt 4.1 wird zunächst das KiL-Projekt vorgestellt. Danach wird das Design der vorliegenden Untersuchung beschrieben (Abschnitt 4.2), sowie die vorgenommene Modellierung des Professionswissens dargestellt (Abschnitt 4.3). Im Abschnitt 4.4 wird schließlich detailliert die Konstruktion und Validierung der neuentwickelten Messinstrumente erläutert, die Schwerpunkte der vorliegenden Arbeit sind.

4.1 Das KiL-Projekt

Ziel des von der *Wissenschaftsgemeinschaft Gottfried Wilhelm Leibniz e.V.* (Leibniz-Gemeinschaft) geförderten KiL-Projekts war die Bereitstellung und Kalibrierung von Testinstrumenten zur Messung des Professionswissens angehender Lehrkräften in den Fächern Biologie, Chemie, Mathematik und Physik. Dazu wurden Testaufgaben in den genannten Fächern konstruiert, in einer Pilotierungsstudie erprobt und überarbeitet. Schließlich wurden die überarbeiteten Aufgaben in einer Hauptstudie normiert.

Den ersten Schritt bildete dabei die Sichtung von Modulplänen und Prüfungsordnungen von Universitäten, die Lehramtsstudiengänge in den naturwissenschaftlichen Fächern und Mathematik anbieten. Hieraus wurden, unter Berücksichtigung der Bildungsstandards für Lehrerbildung (KMK, 2008), Übersichten erstellt, die die wesentlichen Inhalte des Studiums darstellen.

Der zweite Schritt sah eine entsprechende Konzeptualisierung und Operationalisierung der identifizierten Inhalte des fachwissenschaftlichen, fachdidaktischen und pädagogisch-bildungswissenschaftlichen Professionswissens durch Testaufgaben vor. Die Konstruktion der erforderlichen Testaufgaben erfolgte dabei durch die fachdidaktischen Abteilungen des Leibniz-Instituts für die Pädagogik der Naturwissenschaften und Mathematik (IPN) in Kooperation mit der Erziehungswissenschaftlichen Abteilung und Arbeitsgruppen der Pädagogischen Psychologie der Christian-Albrechts-Universität zu Kiel (CAU). Erstere brachten ihre Expertise besonders bei der Konstruktion der fachlichen und fachdidaktischen Messinstrumente ein, letztere bei der Konstruktion der bildungswissenschaftlichen Messinstrumente. Die projektintern konstruierten Aufgaben wurden anschließend durch Aufgabenkonstruktionen von externen Experten ergänzt. Um sicherzustellen, dass die Experten bei der Aufgabenkonstruktion dieselben Formate und Gütekriterien einhalten, wurde ein detaillierter, projektinterner Leitfaden für die Aufgabenkonstruktion erstellt. Die intern und extern konstruierten Aufgaben wurden ihren Inhaltsbereichen entsprechend zu Aufgabenpools zusammengefasst.

Nach dieser Operationalisierung der Konstrukte durch umfangreiche Aufgabenpools bestand der dritte Schritt in der Erprobung der Aufgabenpools durch eine bundesweite Pilotierungsstudie mit

einer anvisierten Stichprobe von etwa $n = 150$ Studierenden pro Aufgabe. Nach einer anschließenden Überarbeitungsphase wurden aus den Aufgabenpools Aufgaben zur Normierung in der Hauptstudie ausgewählt. Das resultierende Testinstrument sollte im Rahmen der Hauptstudie einer anvisierten Stichprobe von etwa $N = 250$ Physikstudierenden vorgelegt werden.

Neben der Sicherstellung allgemeiner Testgütekriterien wurden im Rahmen der querschnittlichen Hauptstudie erste Erkenntnisse zur Struktur und Entwicklung des Professionswissens gewonnen. Darüber hinaus wurden erste Hinweise zu Voraussetzungen, welche die Aneignung von Professionswissen im Lehramtsstudium begünstigen, generiert.

4.2 Design

Die vorliegende Arbeit bedient sich des im Rahmen des KiL-Projekts entwickelten Instruments und der damit erhobenen Daten. Daher werden im Folgenden die Schritte, die zur Entwicklung eines reliablen und validen Testinstruments nötig sind, ausgeführt.

Der erste Schritt besteht in der Definition und Modellierung des Professionswissens angehender Physiklehrkräfte. Um Vergleichbarkeit und Anschlussfähigkeit an bestehende Studien herzustellen, wurde dabei auf die im Abschnitt 2.2 beschriebenen Facetten zurückgegriffen. Die Modellierung diente anschließend dazu, das Konstrukt Professionswissen durch hochqualitative Pools von Testaufgaben zu operationalisieren.

Der zweite Schritt bestand in der anschließenden Auswahl, Adaption, Entwicklung und Präpilotierung erster Aufgaben. Durch zahlreiche interne Projektmeetings, in denen die Aufgaben innerhalb der Arbeitsgruppen vorgestellt und diskutiert wurden, wurde die inhaltliche Passung und generelle Qualität der Aufgaben weiter verbessert. Die Präpilotierung beinhaltet kleinere Studien zur Sicherstellung von Testgütekriterien. Als wichtigste Testgütekriterien werden hierbei *Reliabilität*, *Validität* und *Objektivität* angesehen (Bühner, 2006; Moosbrugger und Kelava, 2008).

Die *Reliabilität* wird dabei als die Zuverlässigkeit angesehen, mit der der Test ein bestimmtes Konstrukt erfasst. Das heißt, ein Test ist genau dann reliabel, wenn er ein Merkmal exakt und reproduzierbar erfasst (Wirtz und Caspar, 2002). Damit ist jedoch noch nicht sichergestellt, dass der Test auch das richtige, intendierte Merkmal erfasst. Dafür muss der eingesetzte Test noch das Gütekriterium der *Validität* erfüllen (Wirtz und Caspar, 2002; Bühner, 2006; Moosbrugger und Kelava, 2008).

Hinsichtlich der *Validität* der Testergebnisse wurde auf den Validitätsbegriff von Messick (1995) zurückgegriffen, da dieser eine besonders differenzierte Betrachtung und systematische Untersuchung verschiedener Aspekte der Validität ermöglicht. Messick (1995) unterscheidet zwischen den folgenden Validitätsaspekten: (1) Der Aspekt der Inhaltsvalidität umfasst die Relevanz, Repräsentationsfähigkeit und Qualität der durch das Instrument erfassten Inhalte. (2) Die kognitive Validität bezieht sich auf die Passung zwischen den tatsächlich beim Bearbeiten des Testinstruments stattfindenden kognitiven Prozessen mit den intendierten Prozessen. Eine Aufgabe, die von Probanden durch nicht intendiertes, fachfremdes Wissen gelöst werden kann, würde beispielsweise die kognitive Validität der Testergebnisse herabsetzen. (3) Der strukturelle Aspekt der Validität bezieht sich darauf, wie genau die Testwerte die Ausprägung des zu erfassenden Konstrukts repräsentieren. (4) Der Aspekt der Generalisierbarkeit beinhaltet, inwieweit die durch das Messinstrument für eine Anzahl n an Probanden erhaltenen Ergebnisse auf die Gesamtheit einer Zielgruppe verallgemeinert werden kön-

nen. (5) Der externe Aspekt der Validität ist dann gegeben, wenn das Testinstrument nur das intendierte Konstrukt, also in diesem Fall das Professionswissen, erfasst und sich von der Erfassung anderer Konstrukte, wie zum Beispiel allgemeiner kognitiver Fähigkeiten, abgrenzt. (6) Der Aspekt der Konsequenzen bezieht sich auf die Schlussfolgerungen, die aus den Testergebnissen abgeleitet werden können und welche, beispielsweise politischen oder gesellschaftlichen Konsequenzen aus den Testergebnissen folgen könnten. Hierfür wurden verschiedene Aufgaben jeweils kleineren Stichproben von studentischen Hilfskräften zur testweisen Bearbeitung vorgelegt. Das Hauptaugenmerk dieser Testphase lag auf der Anpassung von Testzeiten und möglichen Verständnisschwierigkeiten bei der Aufgabenformulierung, also der Wahrung der kognitiven Validität bei der Konstruktion der Aufgaben. In den nächsten Schritten wurde die Validität der Ergebnisse, die das Messinstrument liefert, durch verschiedene Studien zusätzlich empirisch überprüft und anschließend optimiert.

Ferner sollte ein Test unbedingt auch *objektiv* sein, das heißt, das Testergebnis soll nicht von der Person abhängig sein, die den Test leitet, auswertet oder das Ergebnis interpretiert (Moosbrugger und Kelava, 2008, S.8). Weitere wichtige Gütekriterien, die bei der Konstruktion zu beachten sind, sind (d) die *Skalierung* und *Normierung* des Tests, also die Passung des Tests für die anvisierte Zielgruppe und die Fähigkeit, zwischen den Merkmalsausprägungen verschiedener Probanden differenzieren zu können, (e) die *Testökonomie*, also die Angemessenheit von Zeit und Geld, die für die Durchführung des Tests aufgewendet werden müssen, (f) die *Zumutbarkeit* der Testdurchführung, also insbesondere die Dauer des Tests für die Probanden und (g) die *Fairness* des Tests, also die Eigenschaft des Tests, keine bestimmten Probandengruppen in unangemessener Weise zu bevorzugen oder zu benachteiligen (Moosbrugger und Kelava, 2008).

Die Sicherstellung von Reliabilität, Validität und Normierung ist ein aufwändiger Prozess, der – wie nachfolgend beschrieben wird – eine ausführliche Prüfung der eingesetzten Testaufgaben erfordert. Die Objektivität des Tests sicherzustellen ist verhältnismäßig einfach. Dies geschieht durch die Verwendung von geschlossenen Aufgaben und der Erstellung exakter Anweisungen zum Kodieren der halboffenen und offenen Aufgaben. Die Sicherstellung von Ökonomie und Zumutbarkeit des entwickelten Tests erfordert den Einsatz von Aufgaben, deren Bearbeitung nur kurze Zeit, etwa 1-2 Minuten, benötigt. Die Fairness wird schließlich durch die Einbindung der Aufgaben in neutrale, breitgestreute Kontexte gewährleistet.

Als dritter Schritt im Gesamtdesign wurde ein größerer Aufgabenpool zur Generierung von Aufgabenkennwerten in einer Pilotierungsstudie eingesetzt. Dabei wurden neben dem Professionswissen zum Zweck der Überprüfung der externen und strukturellen Validität auch Personenmerkmale, wie zum Beispiel kognitive Grundfähigkeiten, erhoben. Für die Pilotierungsstudie wurde eine Probandenzahl von $n \approx 150$ anvisiert.

Den vierten Schritt stellten Expertenratings dar. Dabei wurden wesentliche Aufgabenkennwerte aus der Pilotierungsstufe Personen mit besonderer Expertise im Bereich Professionswissen und Testentwicklung vorgelegt, um die Aufgabengüte zu diskutieren. Ein weiteres Ziel des Expertenratings lag darin, die inhaltliche Passung der Aufgaben durch unabhängige Experten zu prüfen wodurch die inhaltliche Validität der Testergebnisse sichergestellt werden sollte.

Als fünfter Schritt schloss sich an das Expertenrating noch eine Fragebogenstudie an, in der die Aufgaben auf Merkmale hin untersucht wurden, welche die kognitive Validität der erhobenen Daten

beeinträchtigen könnten. Anhand der Erkenntnisse aus Pilotierungsstudie, Expertenratings und Fragebogenstudie wurde der Aufgabenpool dann noch einmal überarbeitet.

Diese Überarbeitung und Neukonstruktion von Aufgaben stellte den sechsten Schritt dar. In diesem Rahmen wurden sämtliche Aufgaben intensiv überarbeitet. Einzelne Aufgaben mussten verworfen werden. Dies machte eine Neukonstruktion einzelner Aufgaben zu bestimmten Modellbereichen erforderlich.

Im siebten Schritt wurde der finale Aufgabenpool für die Hauptstudie in eine Studie zum Lauten Denken auf seine kognitive Validität hin untersucht.

Den achten und letzten Schritt stellt abschließend die Hauptstudie dar, in dem schließlich ausgewählte Aufgaben einer größeren Stichprobe vorgelegt wurden. Anvisiert waren hier $n \approx 250$ Probanden pro Item. Hierfür wurden ausschließlich Items eingesetzt, die sich im Zuge der bisherigen Entwicklung als besonders geeignet erwiesen haben. Mit den in der Hauptstudie erhobenen Daten wurden anschließend Analysen zu Struktur, Entwicklung und Wirksamkeit des Professionswissens durchgeführt. Bei der Pilotierungs- und Hauptstudie handelte es sich jeweils um Querschnittstudien innerhalb der universitären Phase der Lehramtsausbildung. Dabei wurde also einmalig die Ausprägung des Professionswissens für verschiedene Studienphasen erfasst. Teilnehmen durften demnach Studierende aller Semester, die für einen Lehramtsstudiengang im Fach Physik eingeschrieben waren.

Die Anlage der vorliegenden Studie ist in Abbildung 4.1 dargestellt. Die einzelnen dargestellten Schritte und die daraus gewonnenen Erkenntnisse werden in den folgenden Abschnitten detailliert erläutert. Den Validierungsstudien im Studiendesign, also den Schritten drei, vier, fünf und sieben kommt dabei eine besondere Bedeutung vor, da sie für die Sicherstellung der wichtigsten Testgütekriterien Reliabilität und Validität (vgl. Moosbrugger und Kelava, 2008) vonnöten sind.

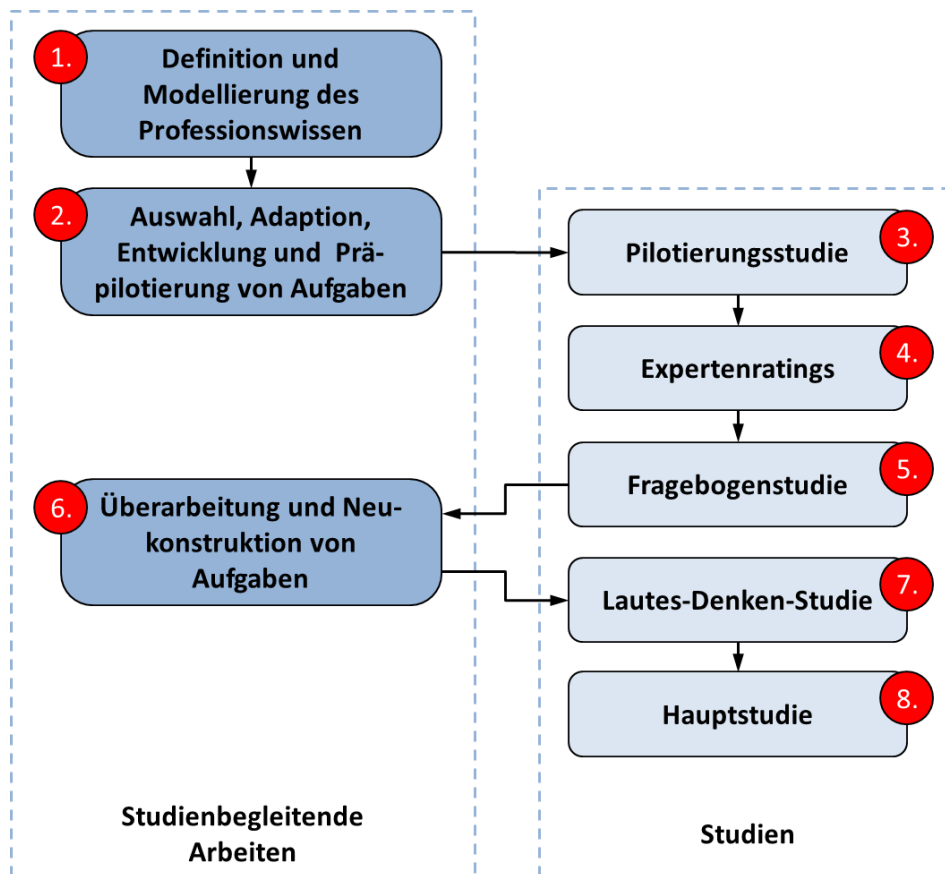


Abbildung 4.1: Anlage der vorliegenden Untersuchung.

4.3 Definition und Modellierung des Professionswissens

Um ein bestimmtes Konstrukt, hier das Professionswissen angehender Lehrkräfte, reliabel und valide messen zu können, ist es nötig, dieses Konstrukt möglichst genau zu definieren (Messick, 1995). Zu diesem Zweck wird ein Modell des Konstrukts erstellt. Neben der Definition des Konstrukts soll das Modell bei der Konstruktion von Testaufgaben die Herstellung einer Passung zwischen Aufgaben und Konstrukt erleichtern. Das Modell soll dabei sicherstellen, dass das verwendete Messinstrument das zu erfassende Konstrukt repräsentiert, ohne dass Bereiche des Konstrukts unterrepräsentiert werden oder dass irrelevante Bereiche außerhalb des Konstrukts erfasst werden. Der Schwerpunkt liegt dabei auf der Erfassung des Fachwissens und des fachdidaktischen Wissens. Das pädagogische Wissen wird in der vorliegenden Arbeit nur am Rande behandelt, da dies Gegenstand eigenständiger Untersuchungen innerhalb des KiL-Projekts darstellt (vgl. Kleickmann et al., 2014).

Bei der Modellierung werden, entsprechend der im theoretischen Hintergrund behandelten Konzeptualisierungen und Facetten des Professionswissens, zunächst Modellierungen für einzelne Facetten entwickelt. Diese werden schließlich zu einem Gesamtmodell für das Fachwissen und das fachdidaktische Wissen angehender Physiklehrkräfte zusammengeführt. Da für die einzelnen Facetten auch zwischen den Arten des vorliegenden Wissens unterschieden werden kann, wird anschließend an die Facetten für Fachwissen, pädagogisches Wissen und fachdidaktisches Wissen eine Kategorisierung bezüglich der Wissensarten entwickelt.

4.3.1 Inhaltsbereiche des Fachwissens

Wie in Abschnitt 4.2 zum Design der Studie erwähnt, besteht eine Voraussetzung für das Erreichen der Ziele der vorliegenden Arbeit in einer breiten Erfassung des Professionswissens, die sämtliche schulrelevanten Inhaltsbereiche umfasst. Daher sollen sämtliche im Theorieteil identifizierten unterrichtsrelevanten Inhaltsbereiche des Fachwissens (siehe Abschnitt 2.2.1) berücksichtigt werden. Dies erscheint nicht nur insofern sinnvoll, als dass es eine gute Repräsentation des Konstrukts liefert, sondern auch weil dadurch einer potentiellen Schwäche querschnittlicher Untersuchungen begegnet werden kann. Würde die Erfassung des Fachwissens beispielsweise lediglich auf den Inhaltsbereich der Optik fokussieren, so würden Probanden des zweiten Semesters, die diesen Inhaltsbereich kürzlich behandelt haben mutmaßlich ein besseres Ergebnis erzielen als solche, für die die Behandlung des Inhaltsbereichs deutlich länger zurückliegt. Die Streuung über viele Inhaltsbereiche hilft, solche Verzerrungen der Ergebnisse zu minimieren. Im vorliegenden Projekt werden damit alle als relevant betrachteten Inhaltsbereiche in das Modell des Professionswissens aufgenommen. Diese sind (a) klassische Mechanik, (b) Elektrizitätslehre und Magnetismus, (c) Optik (geometrische- und Wellenoptik), (d) Thermodynamik, (e) Festkörperphysik, (f) Atom- und Kernphysik, (g) Spezielle Relativitätstheorie und (h) und Quantenmechanik.

Nachfolgend wird für diese acht Inhaltsbereiche des Fachwissens jeweils eine kurze Charakterisierung angegeben, um darzustellen welche Fachinhalte dem jeweiligen Bereich innewohnen.

Klassische Mechanik

Der Inhaltsbereich zur klassischen Mechanik umfasst statische Zustände sowie Translations- und Rotationsbewegungen von Festkörpern unter dem Einfluss von Kräften. Dabei wird auf die in der Schulphysik und in den Grundvorlegungen zur Mechanik etablierte Newtonsche Beschreibung zurückgegriffen. Auf eine Beschreibung durch die Lagrange- oder Hamilton-Mechanik wird verzichtet, da diese Formulierungen der klassischen Mechanik nicht in allen Grundvorlegungen zur Mechanik verwendet bzw. gelehrt werden. Sie werden an der Universität zumeist in Vorlesungen zur theoretischen Physik behandelt, die nicht von allen Lehramtsstudierenden besucht werden. Die Erfassung von Wissen zu diesen Inhalten könnte daher die inhaltliche Validität der Testergebnisse gefährden. Um ferner eine Abgrenzung zur Quantenmechanik und Relativitätstheorie zu gewährleisten, beschränkt sich die Betrachtung zu diesem Inhaltsbereich auf makroskopische Körper und auf kleine Geschwindigkeiten.

Elektrizitätslehre und Magnetismus

Dieser Inhaltsbereich umfasst die klassische Elektrostatik, Elektrodynamik, Magnetismus und elektrodynamische Wechselwirkung. Auf die Einbindung von expliziten Inhalten zu den Maxwell-Gleichungen, insbesondere für bewegte Systeme, wird aus zwei Gründen verzichtet: Erstens, da diese Inhalte an der Universität oftmals in Vorlesungen zur theoretischen Physik gelehrt werden und somit nicht von allen Lehramtsstudierenden besucht werden und zweitens um eine Abgrenzung zur speziellen Relativitätstheorie zu gewährleisten.

Optik

Der Inhaltsbereich zur Optik umfasst die geometrische Optik, also die Lichtausbreitung im Strahlenmodell und die Wellenoptik, also die Betrachtung von Licht als elektromagnetische Welle. Zur geometrischen Optik gehören Reflexions-, Brechungs- und Abbildungsgesetze sowie die Funktionsweise einfacher optischer Instrumente. Zur Wellenoptik gehören die Phänomene der Beugung, Interferenz, Polarisation und Streuung.

Thermodynamik

Die Thermodynamik umfasst Prozesse, die unter dem Einfluss von Umwandlung oder Transport von thermischer Energie stattfinden. Neben Aggregatzuständen, Zusammenhängen zwischen Druck, Volumen und Temperatur, sind hier die Hauptsätze der Thermodynamik von Bedeutung, also Energieumwandlungen in abgeschlossenen Systemen, der Entropiebegriff sowie der Nullpunkt der Temperatur.

Festkörperphysik

Zur Festkörperphysik gehören Eigenschaften von Kristall- und Gitterstrukturen fester Materie. In diesem Zusammenhang werden auch Materialeigenschaften thematisiert, die die elektrische Leitfähigkeit beeinflussen. Damit wird in diesem Inhaltsbereich auch die Halbleiterphysik und Bereiche der Elektronik behandelt, sofern sie auf Materialeigenschaften fester Körper fokussieren.

Atom- und Kernphysik

In diesem Inhaltsbereich werden Inhalte zur Atomhülle und zum Atomkern zusammengefasst. Zur Atomphysik gehört dabei die Physik der Atomhülle und insbesondere das Thema Atommodelle, welches in der Schulphysik spiralcurricular behandelt wird. Die Kernphysik umfasst dagegen Kernverändernde Prozesse wie Kernspaltung, Kernfusion und die damit zusammenhängenden Phänomene. In diesem Zusammenhang ist besonders die Energieumwandlung in Kernreaktoren als gesellschaftsrelevantes Thema von Bedeutung.

Spezielle Relativitätstheorie

Die spezielle Relativitätstheorie beschreibt Phänomene die bei der Bewegung von Körpern bei Großen Geschwindigkeiten, also in der Größenordnung der Lichtgeschwindigkeit, auftreten. Hierzu zählen Längenkontraktion, Zeitdilatation und relativistische Geschwindigkeitsaddition und optische Aberration. Phänomene die mit relativistischen Betrachtungen zur Gravitation zusammenhängen, werden in diesem Bereich nicht behandelt, da sie zur allgemeinen Relativitätstheorie gehören, die wiederum nicht in allen Lehrveranstaltungen für Lehramtsstudierende behandelt werden.

Quantenmechanik

Die Quantenmechanik umfasst das Verhalten von Körpern in kleinen Bezugssystemen, das heißt unterhalb der atomaren Größenordnung. Wichtige Inhalte zu diesem Bereich sind der Tunneleffekt, die Unschärferelation und das Standardmodell der Elementarteilchenphysik.

4.3.2 Inhaltsbereiche des fachdidaktischen Wissens

Wie schon beim Fachwissen, sollen auch für das fachdidaktische Wissen Facetten definiert werden, die für eine angemessene Operationalisierung des Konstrukts hilfreich sind. Beim fachdidaktischen Wissen fällt eine Konzeptualisierung deutlich schwerer als beim Fachwissen, da für das fachdidaktische Wissen bisher keine allgemeingültige und anerkannte Definition existiert (Park und Oliver, 2008; Gramzow, 2013; Berry et al., 2015). Eine einheitliche Modellierung fachdidaktischen Wissens wäre wünschenswert, um die Vergleichbarkeit von Studienergebnissen in diesem Bereich zu verbessern (vgl. auch Gramzow, 2013). Zum derzeitigen Forschungsstand hat sich jedoch, wie im theoretischen Hintergrund beschrieben, noch keine Konzeptualisierung vollständig durchsetzen können (siehe Abschnitt 2.2.2). Es ist daher sinnvoll auf möglichst etablierte Konstrukte zurückzugreifen. Den im Theorierteil identifizierten Kernbereichen des fachdidaktischen Wissens folgend werden daher die von Magnusson et al. (1999) in Anlehnung an Grossmann (1990) unterschiedenen Facetten zur Modellierung des fachdidaktischen Wissens herangezogen. Diese wurden laut Park und Oliver (2008) in zahlreichen Konzeptualisierungen verwendet, erlauben eine breite Abdeckung des fachdidaktischen Wissens und versprechen eine bestmögliche Anschlussfähigkeit an bestehende Forschungsergebnisse. Diese Kernfacetten erscheinen unter den mannigfaltigen Konzeptualisierungsmöglichkeiten am etabliertesten und finden sich auch im Konsensmodell des fachdidaktischen Wissens wider (siehe Abbildung 2.3) (Gess-Newsome, 2015). Diese Kernfacetten sind (a) Wissen über Schülerkognition, (b) Wissen über Instruktionsstrategien, (c) Curriculares Wissen und (d) und Wissen über Assessment und naturwissenschaftliche Grundbildung. Im Folgenden wird die Bedeutung dieser Kernbereiche für die Modellierung im KiL-Projekt näher beschrieben.

Nachfolgend wird für die vier Kernbereiche des fachdidaktischen Wissens jeweils eine kurze Charakterisierung angegeben, um darzustellen welche fachdidaktischen Inhalte dem jeweiligen Kernbereich innewohnen.

Schülerkognition

Der Bereich Schülerkognition bezieht sich auf das Wissen über den typischen Verlauf von Lernprozessen und über Faktoren, die den Lernprozess beeinflussen. Ein Beispiel dafür wäre Wissen über die Interessantheit von Kontexten für Schülerinnen und Schüler. Dies umfasst die Kenntnis des typischen oder konkreten Vorwissens, der Fähigkeiten und des Interesses von Lerngruppen, Kenntnisse über besondere Schwierigkeiten beim Lernen physikalischer Inhalte, Kenntnisse über die Anwendung physikalischer Denkweisen (wie zum Beispiel der Modellierung abstrakter Inhalte), Kenntnisse typischer Fehlvorstellungen und Verständnisschwierigkeiten von Lerngruppen, Kenntnisse über deren Interesse und Motivation sowie deren Rollen für das Verständnis von Physik. Die Zuordnung von Aufgaben zum fachdidaktischen Inhaltsbereich Schülerkognition ist oftmals nicht eindeutig möglich, da es Grenzfälle zu anderen fachdidaktischen Inhaltsbereichen gibt. Grenzfälle die sowohl Aspekte des Bereichs Schülerkognition als auch Aspekte des Bereichs Assessment enthalten, stellen zum Beispiel Kenntnisse über Diagnose von Vorwissen, Fähigkeiten und Fehlvorstellungen von Schülerinnen und Schülern dar. Liegt der Fokus einer Aufgabe eher auf der Diagnose, so wäre die Aufgabe eher dem Bereich Assessment zuzuordnen. Liegt er dagegen eher auf dem Vorwissen, den Fähigkeiten oder den Fehlvorstellungen, so wäre die Aufgabe eher dem Bereich Schülerkognition zuzuordnen. Zahlreiche Beispiele für solche Grenzfälle sind in einem Leitfaden für die Aufgabenentwicklung enthalten, der den Experten für die Aufgabenentwicklung zur Verfügung gestellt wurde. Die Zuordnung der Auf-

gaben zu den Inhaltsbereichen wird damit vorerst den entwickelnden Experten überlassen. Später wird die Zuordnung im Rahmen des Expertenmeetings jedoch noch einmal kritisch diskutiert und geprüft.

Instruktionsstrategien

Der Bereich Instruktionsstrategien umfasst Wissen über Strategien zur inhalts- und lerngruppenabhängigen Unterrichtsgestaltung. Dazu gehören Kenntnisse über allgemeingültige Strategien, wie zum Beispiel das Auslösen eines kognitiven Konflikts zur Anbahnung eines Konzeptwechsels, die Fähigkeit zu deren bewusster Anwendung, Benennung, Erläuterung und Begründung, Kenntnisse über spezielle Strategien für konkrete Fachinhalte, Kenntnisse darüber, welche Methoden, Medien und Arbeitsweisen für bestimmte Inhalte angemessen sind und Kenntnisse über geeignete Modelle, Analogien und Veranschaulichungen zur Darstellung eines Sachverhalts.

Curriculum

Der Bereich Curriculum bezieht sich auf das Wissen über Lehrziele und Anordnungen von Inhalten im Physikunterricht. Dazu gehören Kenntnisse über naturwissenschaftliche Grund- und Allgemeinbildung, über Kompetenzen, die Schülerinnen und Schüler im Schulunterricht erwerben sollen, über Lehrplänen und Bildungsstandards.

4.3.3 Inhaltsbereiche des pädagogischen Wissens

Modellierung und Operationalisierung des pädagogisch-psychologischen Wissens wurde im Projekt KiL von der Abteilung Erziehungswissenschaften am Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik und der Arbeitsgruppe Psychologie der Christian-Albrechts-Universität zu Kiel um Prof. Dr. Jens Möller übernommen (Hohenstein et al., 2014; Hohenstein et al., 2017). Die Inhaltsbereiche pädagogischen Wissens werden als wichtige Facetten des Professionswissens nur der Vollständigkeit halber genannt.

Bestehende Ansätze aus verfügbaren Testverfahren erfassen in der Regel unterrichtsbezogenes pädagogisch-psychologisches Wissen, wie Wissen über Klassenführung, Unterrichtsmethoden oder Lernvoraussetzungen von Schülerinnen und Schülern (König et al., 2011; Voss et al., 2011b). Bei der universitären Ausbildung von Lehrkräften wird pädagogisches und psychologisches Wissen jedoch, gemäß den Empfehlungen der Kultusministerkonferenz (KMK) im Bereich Bildungswissenschaften, breiter aufgefasst (KMK, 2004; Kunina-Habenicht et al., 2013). Es umfasst beispielsweise auch Wissen über Schulentwicklung, persönliche Weiterbildung und Beratung. Die Modellierung des pädagogischen Wissens im vorliegenden Projekt baut auf diesen Ansätzen auf, geht jedoch darüber hinaus, indem die fünf Kernfacetten Facetten (a) Lehren, (b) Lernen und Entwicklung, (c) Klassenführung, (d) Leistungsbeurteilung und (e) Individuelle Besonderheiten einbezogen und fünf weitere Facetten aus den KMK-Standards übernommen werden: (f) Bildungssystem und Schulorganisation, (g) Methoden und Ergebnisse bildungswissenschaftlicher Forschung, (h) Reflexion und Weiterentwicklung der eigenen Professionalität, (i) Schulentwicklung und Werte, (j) Normen und soziale Konflikte. Eine ausführliche Beschreibung dieser Facetten wird von Hohenstein et al. (2017) geliefert.

4.3.4 Wissensarten

Für die Kategorisierung von Wissensarten existieren unterschiedliche Konzeptualisierungen, die in den vergangenen Jahren in verschiedenen Studien für die Modellierung des Professionswissens genutzt wurden. Diese unterscheiden sich hinsichtlich ihrer Systematik und sind daher nicht unmittelbar miteinander vergleichbar. Da sich zudem noch keine Kategorisierung im Forschungsfeld durchgesetzt hat, macht es Sinn auf möglichst etablierte Kategorisierungen zurückzugreifen. Diese sollten zum einen die Herstellung einer guten Passung zwischen Testinstrument und modelliertem Konstrukt erleichtern und zum anderen eine möglichst hohe Anschlussfähigkeit und Vergleichbarkeit mit bisherigen Studien erlauben. Betrachtet man das Forschungsfeld zum Professionswissen von Lehrkräften und angehenden Lehrkräften in den Naturwissenschaften und Mathematik, so erscheint eine Kategorisierung nach kognitivem Anspruchsniveau oder nach Komplexität und Vernetzungsgrad (ein Fakt, zwei Fakten, ein Zusammenhang etc.) wie sie von Woitkowski (2015) vorgenommen wurde, für eine Erfassung des Professionswissens weniger anschlussfähig zu sein als eine Kategorisierung nach kognitionspsychologischen Gesichtspunkten, wie sie in den Projekten COACTIV (Baumert und Kunter, 2006) und ProwiN (Kirschner, 2013) verwendet wurde. Auch Riese (2009) unterscheidet zwischen deklarativen Bereichen professionellen Wissens und explizierbarem prozeduralen Wissen, wengleich er, wie auch Vogelsang (2014), auf die Kluft zwischen Wissen und Handeln hinweist (vgl. auch Vogelsang und Reinhold, 2013). Damit stellt sich die generelle Frage zum Zusammenhang zwischen Wissen und Handeln. Obgleich diese Frage in der vorliegenden Arbeit nicht beantwortet werden kann, so soll doch möglichst Handlungs- und Prozessorientiertes Wissen erfasst werden. Zu diesem Zweck soll bei dem vorliegenden Modell neben einer Kategorie des deklarativen Faktenwissens auch eine zweite Kategorie prozeduralen Wissens über typische Handlungsabläufe erfasst werden. Zusammen mit einer dritten Kategorie über stärker vernetztes Wissen, die für eine Differenzierung zwischen Novizen und Experten in einer Domäne nötig ist, ergibt sich damit eine prinzipielle Unterscheidung zwischen drei Wissensarten. Dies erlaubt ferner eine Anschlussfähigkeit und Vergleichbarkeit mit den Projekten COACTIV und ProwiN herzustellen. Daher wird in der vorliegenden Arbeit also auf eine dreistufige Aufgliederung der Arten professionellen Wissens zurückgegriffen, die sich bei der Erfassung des Professionswissens in den genannten Projekten bewährt hat (Baumert und Kunter, 2006; Kirschner, 2013). Um eine möglichst gute Trennbarkeit zwischen den Wissensarten zu gewährleisten, werden für die Modellierung der dritten Kategorie, also des vernetzten Wissens, das strategische und schematische Wissen in dieser Kategorie zusammengeführt. Damit lassen sich bei der Modellierung der Wissensarten die folgenden Wissensarten unterscheiden: (a) deklaratives Wissen, (b) prozedurales Wissen und (c) schematisches und strategisches Wissen. Diese Wissensarten werden im Folgenden kurz erläutert:

- (a) *Deklaratives Wissen* bezieht sich auf die Kenntnis von Fakten, Definitionen und Gesetzmäßigkeiten („*knowing that*“); zum Beispiel, dass Kraft gleich Masse mal Beschleunigung ist oder das Prinzip der Energieerhaltung.
- (b) *Prozedurales Wissen* bezieht sich auf die Kenntnis von Regeln und Abläufen („*knowing how*“); zum Beispiel, wie Experimente durchzuführen sind oder wie man die Differentialgleichung für das mathematische Pendel löst. Die Anwendung prozeduralen Wissens kann sowohl bewusst als auch unbewusst erfolgen.

(c) *Schematisches und strategisches Wissen* bezeichnet die Kenntnis von Schemata, Modellen und Strategien („*knowing why, when and where*“), beispielsweise die Rolle eines harmonischen Oszillators im Gedankengebäude der Physik oder die Implikationen, die der Übergang von einer laminaren zu einer turbulenten Strömung hat.

4.3.5 Modellierung des Fachwissens und fachdidaktischen Wissens

Insofern, als dass sich – wie im Theorieteil aufgezeigt – die Aufgliederung des Konstrukts Professionswissen in die Wissensbereiche Fachwissen, fachdidaktisches Wissen und pädagogisches Wissen in den bisherigen Arbeiten bewährt hat, gibt es hinreichenden Grund, weiter daran festzuhalten. Nicht zuletzt, um eine wissenschaftliche Anschlussfähigkeit der vorliegenden Forschungsarbeit zu gewährleisten und Resultate zu liefern, die sich vergleichbar in die Forschungslandschaft eingliedern, wird diese Konzeptualisierung des Professionswissens auch für die Modellierung im KiL-Projekt herangezogen, d.h. eine Einteilung in die Bereiche PK, CK und PCK wird auch hier zur Anwendung kommen. Da sich die vorliegende Arbeit aus genannten Gründen im Folgenden auf das Fachwissen und das fachdidaktische Wissen konzentriert, wird an dieser Stelle auf eine Modellierung des pädagogischen Wissens verzichtet; es folgt also die Modellierung des Fachwissens und des fachdidaktischen Wissens.

Für das Fachwissen ergibt sich aus den Dimensionen *Wissensart* und *Fachinhalt* ein zweidimensionales Modell. Da sich fachdidaktische Inhalte sowohl nach Fachinhalt wie auch nach fachdidaktischem Inhalt differenzieren lassen, ergibt sich für das fachdidaktische Wissen ein dreidimensionales Modell mit der zusätzlichen Dimension *fachdidaktischer Inhalt*. Dabei ist noch zu berücksichtigen, dass Aufgaben zum fachdidaktischen Wissen nicht zwangsläufig einem spezifischen Fachinhalt angehören müssen. Manche Aufgaben, zum Beispiel solche, die das Experimentieren im Physikunterricht thematisieren ohne auf ein konkretes Experiment abzielen, können keinem konkreten Fachinhalt zugeordnet werden. Sie werden daher inhaltsübergreifend bezeichnet, da sie sich theoretisch auf sämtliche Inhaltsbereiche beziehen könnten. Da die Dimensionen der Wissensarten und des Fachinhaltes sowohl beim Fachwissen als auch beim fachdidaktischen Wissen auftauchen und jeweils identisch sind, lassen sich die beiden Modelle wie in der folgenden Abbildung 4.2 gemeinsam darstellen.

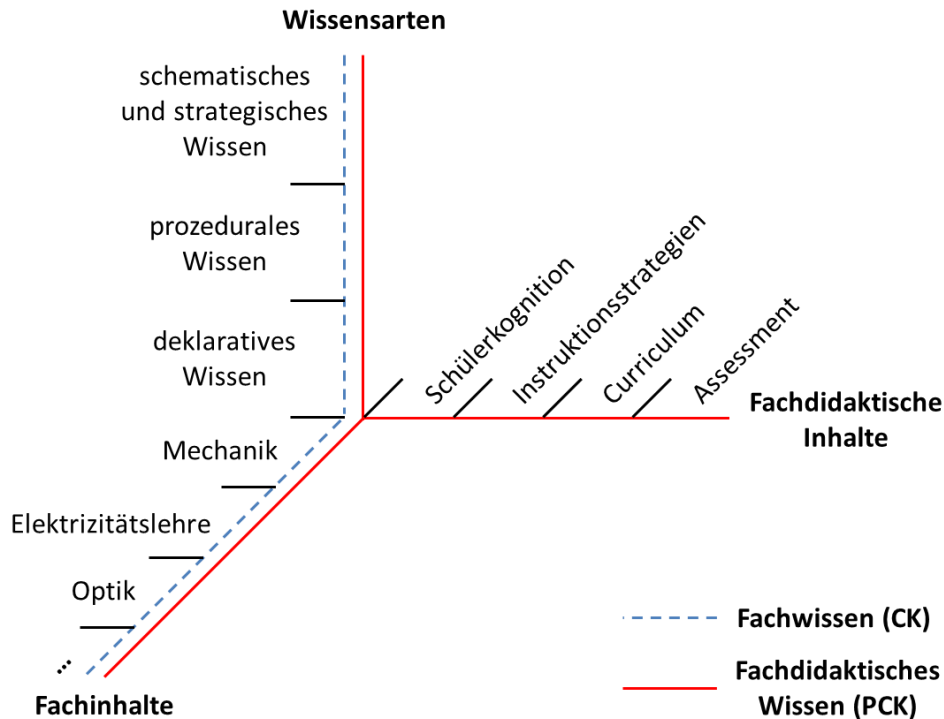


Abbildung 4.2: Modell für das Fachwissen und das fachdidaktische Wissen.

Der von den Dimensionen dieses Modells aufgespannte zwei- bzw. dreidimensionale Raum soll im Folgenden mit Testaufgaben gefüllt werden, so dass zwei Aufgabenpools entstehen, die jeweils das Konstrukt des Fachwissens bzw. des fachdidaktischen Wissens repräsentieren. Da die Fachinhalte für das Fachwissen als besonders bedeutsam angesehen werden, wird beim Fachwissen auf eine homogene Abdeckung der Fachinhaltsbereiche durch die Aufgaben fokussiert. Analog dazu werden für das fachdidaktische Wissen die fachdidaktischen Inhalte als besonders bedeutsam angesehen und daher hier Wert auf eine homogene Abdeckung der fachdidaktischen Inhalte gelegt. Das bedeutet, dass für das Fachwissen bzw. für das fachdidaktische Wissen für jeden Inhaltsbereich eine Mindestanzahl von Aufgaben vorliegt. Die Aufgabenanzahl pro Bereich stellt dabei einen Kompromiss dar, einerseits bezüglich der Testökonomie und andererseits bezüglich einer reliablen Abdeckung des jeweiligen Bereichs. Da, aufgrund des von Gramzow (2015, S. 157) beschriebenen Reliabilitäts-Validitäts-Dilemmas die breite, umfassende Abdeckung eines Bereichs zwar die Validität erhöht, es jedoch zu einer niedrigeren inneren Konsistenz des erfassten Konstrukts und damit zu einer verringerten Reliabilität kommen kann, muss bei der Abdeckung der einzelnen Inhaltsbereiche eine Entscheidung zwischen einer Fokussierung auf inhaltliche Breite und damit hoher Validität oder auf inhaltliche Konsistenz und damit hohe Reliabilität getroffen werden. Da die breite Abdeckung der zu erfassenden Inhalte ein wesentliches Ziel der vorliegenden Arbeit darstellt, wird der Fokus bei der Aufgabenkonstruktion auf ersterem liegen. Das heißt, es wird Wert darauf gelegt einen Inhaltsbereich durch Aufgaben zu möglichst unterschiedlichen Themen breit zu repräsentieren, auch wenn dies auf Kosten der inneren Konsistenz geschieht, also mit einer geringeren Reliabilität einhergeht.

Weiterreichende Erläuterungen zu den Testgütekriterien, sowie die einzelnen Schritte zur Konstruktion der Testinstrumente zum Fachwissen und zum fachdidaktischen Wissen werden in den folgenden Abschnitten berichtet.

4.4 Konstruktion und Validierung der Messinstrumente

Für die Erhebung des Professionswissens von Lehrkräften bieten sich verschiedene Arten von Messinstrumenten an: Ein *Interview*, bei dem das Wissen der Probanden individuell abgefragt wird; ein Vignettentest, bei dem videographierte Unterrichtssituationen gezeigt werden zu denen die Probanden einen Reflexionsbogen ausfüllen und ein *Paper-und-Pencil-Test*, bei dem den Probanden schriftliche Fragen zur Beantwortung vorgelegt werden. Für die vorliegende Untersuchung ist die individuelle Interviewmethode aufgrund der angestrebten hohen Probandenzahlen ungeeignet. Der Vignettentest hingegen ist wegen der angestrebten inhaltlichen Breite, die eine Vielzahl von Unterrichtsvignetten und damit viel Testzeit verlangt, nicht geeignet. Es wird also aus Gründen der Testökonomie ein Paper-und-Pencil-Test, also ein schriftlicher Fragebogen, eingesetzt.

Im Folgenden werden die Schritte 2 bis 7 aus der Darstellung des Studiendesigns (siehe Abbildung 4.1), die zur Konstruktion und Validierung des finalen Messinstruments für die Hauptstudie durchgeführt wurden, ausgeführt.

4.4.1 Auswahl, Adaption, Entwicklung und Präpilotierung von Aufgaben

Der nächste Schritt nach dem Abschluss der Modellierung besteht in der Auswahl und Entwicklung entsprechender Testaufgaben. Hierfür wurden zunächst vorhandene Testinstrumente zum fachlichen bzw. fachdidaktischen Wissen im Fach Physik gesichtet und auf ihre Modellpassung hin geprüft. Es existieren zahlreiche erprobte Instrumente aus dem englischsprachigen Raum, die im College eingesetzt werden. Das bekannteste Beispiel aus dieser Gruppe ist der *Force Concept Inventory* (FCI), der in 29 Fragen zum Kraftkonzept wesentliche Kenntnisse und Anwendungen der Newtonschen Axiome abfragt (Hestenes et al., 1992). Ein weiteres Testinstrument zum Inhaltsbereich Mechanik ist der *Test of Understanding Graphs-Kinematics* (TUG-K) der in 21 Fragen jeweils das Verständnis von Graphen zu Bewegungsabläufen erfasst (Beichner, 1994). Ein weiteres Testinstrumente zur Mechanik, welches aber auch Fragen zur Thermodynamik enthält, ist der 35 Fragen umfassende *Energy Concept Inventory* (ECI), der auf etwas niedrigerem Niveau das Energiekonzept von Schülerinnen- und Schülern testet (Swackhamer und Hestenes, 2005). Zum Inhaltsbereich Elektrizitätslehre gibt es den *Brief Electricity and Magnetism Assessment* (BEMA), das die Themen Elektrizitätslehre, Magnetismus, sowie deren Wechselwirkungen in 22 Fragen erfasst (Maloney et al., 2001; Ding et al., 2006). Zum Thema Quantenmechanik existiert der 25 Fragen umfassende *Quantum Physics Conceptual Survey* (QPCS), der Wissen zum Photoeffekt, zum Standardmodell der Elementarteilchenphysik, zu Beugungseffekte von Elektronen und zur Heisenbergschen Unschärferelation testet (Wutti-prom et al., 2009). Zum Thema Teilchenphysik existiert zudem noch der 12 Fragen umfassende *Physical Changes Concepts Test* (PCCT) (Haidar und Abraham, 1991). Eine Übersicht der betrachteten Testinstrumente zum Fachwissen ist in Tabelle 4.1 dargestellt.

Tabelle 4.1: Übersicht veröffentlichter Testinstrumente zum Fachwissen.

Testinstrument	Inhalt	Aufgabenanzahl
FCI	Kraftkonzept	29
TUG-K	Kinematik	21
ECI	Energiekonzept	35
BEMA	Elektrizitätslehre	22
QPCS	Quantenmechanik	25
PCCT	Teilchenphysik	12
Gesamt		144

Durch Adaptierung und Übersetzung von Aufgaben aus diesen veröffentlichten Instrumenten lassen sich einzelne Bereiche des Modells für das Fachwissen abdecken (siehe Abbildung 4.2). Da die Instrumente jedoch bestimmte Themengebiete fokussieren, sind für eine breite Abdeckung des Fachwissen keinesfalls hinreichend. So liegen zum Inhaltsbereich Mechanik mit dem FCI und dem TUG-K zwar eine Vielzahl von Aufgaben zur Statik und zur translativen Kinetik vor, beispielsweise fehlen jedoch gänzlich Aufgaben zu Rotationsbewegungen. Zudem ist der Inhaltsbereich Thermodynamik mit den vorliegenden Aufgaben unterrepräsentiert. Die die Inhaltsbereiche Optik, Festkörperphysik, Atom- und Kernphysik und Relativitätstheorie werden durch vorliegende Aufgaben abgedeckt. Diese Lücken bei der Modellabdeckung müssen wurden mittels gezielter Neukonstruktionen durch Fachexperten geschlossen werden.

Bei der Konstruktion fachdidaktischer Aufgaben musste mangels geeigneter publizierter Erhebungsinstrumente ausschließlich auf neukonstruierte Aufgaben zurückgegriffen werden. Dabei wurden besonders relevante Kontexte fachdidaktischen Wissens wie z.B. das Experimentieren im Unterricht oder typische Schülerfehlvorstellungen zu bestimmten Themen (vgl. Kircher et al., 2009; Wiesner et al., 2011), identifiziert und in die neukonstruierten Aufgaben eingebunden.

Für die Neukonstruktion von Aufgaben, sowohl zum Fachwissen als auch zum fachdidaktischen Wissen, wurden Experten (Doktoranden, Post-Docs und Professoren) mit besonderer Expertise im Bereich der Aufgabenkonstruktion herangezogen. Zur Sicherstellung der genannten Testgütekriterien wurde zudem ein projektinterner Leitfaden für die Aufgabenentwicklung formuliert, der das zugrundeliegende Modell, eine Beschreibung der Wissensarten und Inhaltsbereiche, die folgende Beschreibung der Aufgabenformate sowie allgemeine Hinweise für die Aufgabenkonstruktion enthält. Als Hilfe für die Entwickler und zur Standardisierung der Aufgabenentwicklungen erhielten die Aufgabenentwickler ein Word-Dokument als Vorlage. In diese Vorlage konnten die zugehörigen fachlichen- und fachdidaktischen Inhaltsbereiche, Aufgabenkennwerte, die Aufgabe selbst, die Musterlösung zur Aufgabe, sowie weitere Hinweise und Kommentare eingefügt werden. Die folgenden Tabellen 4.2 und 4.3 enthalten jeweils eine Übersicht, über die Aufgaben zum Fachwissen und zum fachdidaktischen Wissen, die von den einzelnen Entwicklern auf diese Weise konstruiert wurden. Die einzelnen Entwickler für die Fach- und fachdidaktischen Aufgaben werden dabei durch eine zugewiesene, fortlaufende Nummer anonymisiert.

Tabelle 4.2: Übersicht über neukonstruierte Aufgaben zum Fachwissen (blau).

Nr.	Inhaltsbereich	Anzahl
1	Optik	15
2	verschiedene - überwiegend Festkörperphysik und Atom- und Kernphysik	30
3	spezielle Relativitätstheorie	20
Gesamt		65

Tabelle 4.3: Übersicht über neukonstruierte Aufgaben zum fachdidaktischen Wissen (rot).

Nr.	Inhaltsbereich	Anzahl
4	verschiedene - überwiegend Schülerkognition und Instrukionsstrategien	30
5	verschiedene - überwiegend Schülerkognition und Instrukionsstrategien	20
6	verschiedene - überwiegend Schülerkognition und Instrukionsstrategien	30
7	verschiedene - überwiegend Schülerkognition und Instrukionsstrategien	30
8	verschiedene - überwiegend Schülerkognition und Instrukionsstrategien	30
9	verschiedene - überwiegend Assessment	30
10	Curriculum	10
Gesamt		180

Da die Inhaltsbereiche für das Fachwissen relativ trennscharf sind, und da hier bereits einige Inhalte durch die Adaption der Aufgaben aus den veröffentlichten Testinstrumenten recht gut abgedeckt waren, wurden für das Fachwissen von den Entwicklern gezielt Aufgaben zu bestimmten Inhaltsbereichen konstruiert (siehe Entwickler 1 bis 3 in Tabelle 4.2). Da die Inhaltsbereiche zum fachdidaktischen Wissen hingegen weniger Trennscharf sind und da hierzu noch keine Aufgaben vorlagen wurden es hier zunächst den Entwicklern überlassen, Aufgaben zu verschiedenen fachdidaktischen Inhalten zu konstruieren. Dabei stellte sich heraus, dass überwiegend Aufgaben zur Schülerkognition und zu Instrukionsstrategien entwickelt wurden (siehe Entwickler 4 bis 8 in Tabelle 4.3). Daher wurden anschließend durch zwei Entwickler noch gezielt Aufgaben zu den fachdidaktischen Inhaltsbereichen Assessment und Curriculum konstruiert (siehe Entwickler 9 und 10 in Tabelle 4.3). Die adaptierten und von den Experten konstruierten Aufgaben wurden anschließend an studentischen Hilfskräften überprüft und im Rahmen regelmäßiger, projektinterner Entwicklungstreffen diskutiert und bewertet. Für die weiteren Schritte zur Entwicklung eines ersten Testinstruments wurden solche Aufgaben ausgewählt, die für die studentischen Hilfskräfte verständlich, weder zu einfach noch zu anspruchsvoll waren und deren Bearbeitung eine möglichst geringe Zeit in Anspruch nahm. Aufgaben, die diesen Kriterien nicht entsprachen wurden verworfen. Von den ursprünglichen 209 adaptierten und konstruierten Aufgaben zum Fachwissen (siehe Tabellen 4.2) blieb ein Pool von 171 Aufgaben übrig.

Von den 180 Aufgaben zum fachdidaktischen Wissen (siehe Tabelle 4.3) blieb ein Pool von 166 Aufgaben übrig.

Die Aufgaben aus den veröffentlichten Testinstrumenten erfassen überwiegend Schulwissen im geschlossenen Multiple-Choice-Format. Die in der vorliegenden Arbeit angestrebte breite Erfassung des Professionswissens angehender Physiklehrkräfte verlangt jedoch aufgrund des zum Teil höheren Anspruchs auch die Erfassung komplexerer Inhalte. Diese erfordern neben dem Multiple-Choice-Format weitere Antwortformate. Beispiele hierfür sind frei formulierte Antworten oder Antworten durch die Anfertigung einer Skizze oder Zeichnung. Hierzu werden in der vorliegenden Studie verschiedene Aufgabenformate eingesetzt, die im Folgenden erläutert werden.

Aufgabenformate

Eine Aufgabe setzt sich wie folgt aus mehreren Bestandteilen zusammen (siehe Abbildung 4.3). Sie beginnt in der Regel mit einem Aufgabenstamm, auf den ein oder mehrere Items folgen. Der Aufgabenstamm beschreibt eine Situation oder einen Sachverhalt. Jedes Item wiederum besteht aus einem Itemstamm, der eine Frage oder eine Anweisung enthält, und einem Antwortteil. Der Antwortteil besteht je nach Itemformat aus mehreren Antwortoptionen (geschlossenes Format), einem ergänzenden Text (halboffenes Format) oder einem Feld zum Eintragen einer Antwort (offenes Format). Neben Aufgaben, die aus einem Aufgabenstamm und Item(s) bestehen, können auch solche Aufgaben entwickelt werden, die keinen Aufgabenstamm benötigen, da die Aufgabe bereits durch den Itemstamm eindeutig formuliert ist.

Generell werden in der vorliegenden Studie drei Aufgabenformate eingesetzt: (1) geschlossene Formate, bei denen aus einer vorgegebenen Auswahl an Antwortoptionen die richtigen bzw. passenden auszuwählen sind, (2) halboffene Aufgabenformate, bei denen von der Testperson eine kurze, zum Beispiel nur aus einem Wort bestehende, Antwort zu geben ist und (3) offene Aufgabenformate, die ebenfalls eine von der Testperson zu erstellende Antwort verlangen, die aber das freie Verfassen eines kurzen Textes, das Erstellen einer Skizze oder Ähnliches erfordert.

Aufgabenstamm	
Eine Kugel der Masse m wird mit einem dünnen Faden an eine Kugel der Masse $2m$ gehängt. Anschließend werden die Kugeln losgelassen und fallen reibungsfrei im konstanten Schwerfeld der Erde.	
Item	
Wie groß ist die Zugkraft im Faden?	} Itemstamm
x 0	
<input type="checkbox"/> $\frac{1}{2} m g$	} Antwortteil
<input type="checkbox"/> $m g$	
<input type="checkbox"/> $2 m g$	
<input type="checkbox"/> $3 m g$	
<input type="checkbox"/> $3 m g$	

Abbildung 4.3: Beispiel zur Verdeutlichung der Bestandteile einer Aufgabe.

Zu den geschlossenen Formaten gehören drei wichtige Gruppen: Multiple-Choice-Items, Richtig-Falsch-Items und Zuordnungs-Items.

Multiple-Choice-Items

Ein Multiple-Choice-Item besteht aus einem Aufgabenstamm, der eine Situation oder einen Sachverhalt schildert, einem Itemstamm, der eine Frage oder Arbeitsanweisung formuliert und mehreren Antwortoptionen. Erfordert ein Item keine Beschreibung der Situation oder des Sachverhalts, kann sie auch nur aus einem Itemstamm und dazugehörigen Antwortoptionen bestehen (siehe Abbildung 4.4). Üblicherweise umfassen die Antwortoptionen genau eine richtige Antwort, den so genannten Attraktor, und drei oder vier falsche, aber plausible Antwortoptionen, die sogenannten Distraktoren.

Multiple-Choice-Item

Welche der folgenden Aussagen gilt für jeden unelastischen Stoß?

- Die kinetische Energie des Gesamtsystems bleibt erhalten.
- Der Gesamtimpuls bleibt erhalten.
- Die Geschwindigkeit der einzelnen Stoßpartner bleibt erhalten.
- Die Bewegungsrichtung jedes Stoßpartners bleibt erhalten.

Abbildung 4.4: Beispiel für ein Multiple-Choice-Item.

Richtig-Falsch-Items

Ein Richtig-Falsch-Item besteht aus einem Itemstamm, der als Frage oder als Aussage formuliert sein kann, und aus zwei Antwortoptionen, von denen eine richtig und die andere falsch ist (siehe Abbildung 4.5). Richtig-Falsch-Items werden vorrangig verwendet, wenn es gilt, isolierte Fakten wie Namen, Daten oder Definitionen abzufragen. Um die Ratewahrscheinlichkeit zu minimieren, sollten Richtig-Falsch-Items möglichst blockweise verwendet werden. Ansonsten läge die Wahrscheinlichkeit für das Erraten der richtigen Antwort für ein Einzelitem bei 50%. Das bedeutet, dass 4 bis maximal 10 Einzelfragen oder -aussagen in einer Aufgabe zusammengefasst werden. Ein solches Item gilt nur dann als vollständig richtig gelöst, wenn alle Einzelfragen oder -aussagen korrekt beantwortet wurden. Damit die Punktvergabe der Kompetenz des Testkandidaten entspricht, sollte auch erwogen werden, eine Aufgabe als teilweise richtig gelöst zu bewerten, sobald ein bestimmter Prozentsatz der Einzelfragen korrekt bearbeitet wird.

Richtig-Falsch-Item

Obwohl selbst kein Stern, emittiert der Planet Jupiter pro Zeiteinheit eine bestimmte Energiemenge. Geben Sie für die folgenden Prozesse jeweils an, ob sie eine nennenswerte Energiequelle des Planeten darstellen.

	Ja	Nein
Absorption	x	<input type="checkbox"/>
Kontraktion	x	<input type="checkbox"/>
Oxidation	<input type="checkbox"/>	x
Kernfusion	<input type="checkbox"/>	x
Kernzerfall	<input type="checkbox"/>	x

Abbildung 4.5: Beispiel für eine Aufgabe mit Richtig-Falsch-Items.

Zuordnungs-Items

Zuordnungs-Items bestehen aus einer Liste von Prämissen und einer Liste von Antwortoptionen (siehe Abbildung 4.6). Jeder Prämisse ist dabei eine Antwortoption zuzuordnen.

Zuordnungs-Item

Welche der folgenden Begriffe sind für Translationen und Rotationen jeweils analog zueinander? Tragen Sie den entsprechenden Buchstaben in die Klammer hinter dem Namen des Begriffs ein. Jede Antwortalternative kann einmal, mehrmals oder gar nicht richtig sein.

Weg (E)	A: Drehimpuls
Masse (D)	B: Drehmoment
Kraft (B)	C: Rotationsenergie
Impuls (A)	D: Trägheitsmoment
	E: Winkel

Abbildung 4.6: Beispiel für ein Zuordnungsitem.

Kurzantwort-Items

Bei den halboffenen Items handelt es sich um Kurzantwort-Items, bei denen keine Antwortoptionen vorgegeben werden. Sie verlangen eine kurze, eigenständig generierte Antwort. Diese kann beispielsweise aus einzelnen Wörtern, Zahlen, Symbolen bestehen. Beispiele sind die Beschriftung einer Skizze oder die Vervollständigung eines Satzes. Ein Beispiel ist in Abbildung 4.7 gegeben.

Kurzantwort-Item

Ein Geiger-Müller-Zählrohr ist ein Gerät zur Messung von

ionisierender Strahlung

Abbildung 4.7: Beispiel für ein Kurzantwort-Item.

Offene Items

Items mit offenem Antwortformat erlauben der Testperson, auf ein bestimmtes Item frei zu antworten. Der Lösungsweg wird dabei nicht vorgegeben (siehe Abbildung 4.8).

Item

Ein Gummiball wird von einem hohen Gebäude fallen gelassen. Der Ball springt, nachdem er auf dem Boden aufgetroffen ist, wieder hoch, erreicht jedoch nicht seine ursprüngliche Höhe. Beschreiben Sie die Geschwindigkeit und Beschleunigung des Balls aus physikalischer Sicht vom Beginn des Falls bis zum Erreichen seiner maximalen Höhe nach dem Aufprall. Unterteilen Sie die Bewegung dabei in geeignete Phasen.

Lösung des Items

Phase 1: Die Starthöhe des Gummiballs sei h_1 . Er hat die Anfangsgeschwindigkeit $v_0 = 0$ und wird mit der Erdbeschleunigung g gleichmäßig in Richtung des Erdbodens beschleunigt. Der Ball erreicht den Erdboden mit einer Geschwindigkeit $v_1 = \sqrt{2gh_1}$.

Phase 2: Beim Aufprall auf dem Erdboden kehrt sich die Bewegungsrichtung des Gummiballs um. Ein Teil der kinetischen Energie des Balls wird bei dem Stoß mit dem Boden in innere Energie umgewandelt.

Phase 3: Der Ball prallt mit einer Geschwindigkeit v_2 (mit $|v_2| < |v_1|$) vom Erdboden weg in die Höhe. Dabei wirkt die Erdbeschleunigung g entgegen seiner Bewegungsrichtung.

Phase 4: Sobald die Geschwindigkeit des Balls auf den Wert Null abgesunken ist, hat er seine maximale Höhe h_2 (mit $h_2 < h_1$) erreicht.

Abbildung 4.8: Beispiel für ein offenes Item.

Um die Einhaltung des Testgütekriteriums der Objektivität sicherzustellen, ist neben einer Aufgabe selbst vor allem die Beurteilung der Probandenantworten von Bedeutung. Durch die Formulierung möglichst genauer Kodierschemata wird sichergestellt, dass die Beurteilung möglichst unabhängig von der bewertenden Person ist. Dazu werden im Kodierschema genaue Kriterien für die Vergabe der vollen Punktzahl (Full Credit), gegebenenfalls für teilweise richtige Antworten für die Teilpunkte vergeben werden können (Partial Credit) und für nicht richtige Antworten (No Credit) gegeben. Die Kri-

terien werden durch eine Liste typischer Beispielantworten ergänzt und teilweise um Grenzfälle und Anmerkungen erweitert.

Präpilotierung

Im Rahmen dieser Präpilotierung wurde auch eine Machbarkeitsstudie an den Universitäten in Flensburg und Kiel durchgeführt. Dabei konnten Studierende aller Studiengänge, also nicht nur Lehramtsstudierende, und sämtlicher Studiensemester an dieser Machbarkeitsstudie teilnehmen. Zielsetzung dieser Machbarkeitsstudie war (1) die Gewinnung von Erkenntnissen über die Bereitschaft von Studierenden an Testungen zum professionellen Wissen teilzunehmen, (2) von Erkenntnissen zur Selektivität bzw. möglicher Auswahleffekte der Rekrutierung und (3) von ersten Erkenntnissen über das fachliche und pädagogische Wissen der Studierenden und deren Beanspruchung durch einen mehrstündigen Test. Im Rahmen dieser Machbarkeitsstudie wurden den Probanden Testinstrumente zu kognitiven Fähigkeiten, zum pädagogischen Wissen und für das Fachwissen eingesetzt. Die Testzeit inklusive Pausen betrug dabei etwa 4 Stunden. Für das Fachwissen im Fach Physik wurde ein Test eingesetzt, der ins Deutsche übersetzte Aufgaben aus den veröffentlichten Testinstrumenten *Brief Electricity and Magnetism Assessment* (BEMA) (Ding et al., 2006) und *Force Concept Inventory* (FCI) (Hestenes et al., 1992) enthielt. Für die Studie konnten über alle naturwissenschaftlichen Fächer und Mathematik hinweg $n = 294$ Studierende gewonnen werden. 70% der Teilnehmer waren weiblich, das Alter betrug durchschnittlich 23 Jahre ($SD = 3,4$) und der Anteil der Lehramtsstudierenden an den Teilnehmern betrug 79%. Die Studierenden erhielten für ihre Teilnahme jeweils eine Aufwandsentschädigung von 50 Euro. Bei der Durchführung konnte festgestellt werden, dass die Teilnehmenden die Testaufgaben trotz der hohen Testdauer von 4 Stunden diszipliniert und ernsthaft bearbeitet haben. Die Häufigkeit und Länge der Pausen wurde als optimal wahrgenommen. Die Ergebnisse der Machbarkeitsstudie können insofern zusammengefasst werden, als dass (1) die Rekrutierung von Studierenden in ausreichender Quantität bei Bezahlung einer entsprechenden Aufwandsentschädigung kein Problem darstellte, (2) keine offensichtlichen Auswahleffekte zu beobachten waren und (3) die eingesetzten Aufgaben aus BEMA und FCI gut zum Fähigkeitslevel der Studierenden passten. Hierbei muss allerdings beachtet werden, dass die Testaufgaben aus BEMA und FCI nur Teile der Inhaltsbereiche Elektrizitätslehre und Mechanik abdecken, also nicht die Forderung nach einer breit angelegten Operationalisierung des Professionswissens erfüllen.

4.4.2 Pilotierungsstudie

Die Pilotierungsstudie dient primär der Untersuchung der Reliabilität entwickelter Aufgabenpools sowie des strukturellen und externen Aspekts der Konstruktvalidität nach Messick (1995). Nachfolgend werden Design, eingesetztes Instrument, Stichprobe und Ergebnisse der Pilotierungsstudie vorgestellt.

Design der Pilotierungsstudie

Eine hohe Bedeutung für den Erfolg der Pilotierungsstudie kam der Rekrutierung ausreichend vieler Probanden zu. Dies stellt insofern eine Herausforderung dar, da im Vergleich zu den Fächern Biologie, Chemie und Mathematik die Anzahl eingeschriebener Studierender im Fach Physik relativ gering ist. Darüber hinaus ist der Anteil von sogenannten Parkstudierenden, also eingeschriebenen Personen, die jedoch nicht real studieren, im Fach Physik besonders hoch (Matzdorf und Düchs, 2013). Um

trotzdem die anvisierte Anzahl von etwa 150 Studierenden zu erreichen, wurde die querschnittliche Pilotierungsstudie bundesweit an zwölf Standorten durchgeführt. Hauptkriterium für die Wahl der Standorte war dabei die Anzahl von Lehramtsstudierenden in den naturwissenschaftlichen Fächern und Mathematik. Die Auswahl für die Durchführung der Pilotierungsstudie fiel auf die Universitäten in Hamburg, Kassel, Potsdam, Oldenburg, Bremen, Regensburg, Berlin, Kiel, Freiburg, Rostock, Bielefeld und München (siehe Abbildung 4.9).

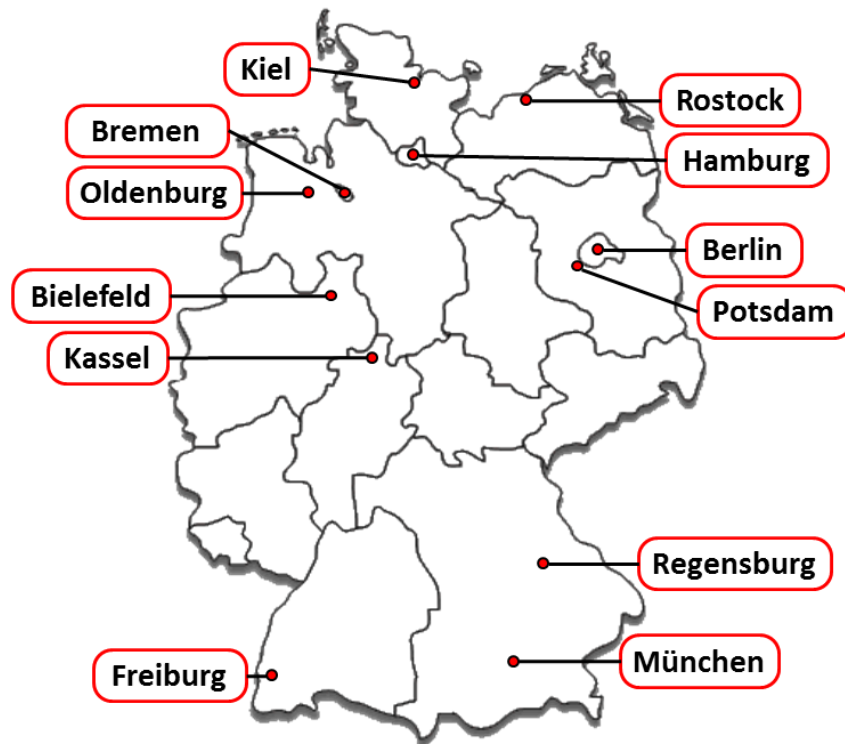


Abbildung 4.9: Erhebungsstandorte in der Pilotierungsstudie.

Instrument der Pilotierungsstudie

Aus der Phase der Aufgabenentwicklung und Präpilotierung entstanden Pools mit insgesamt 171 Aufgaben für die Erfassung des Fachwissens und 166 Aufgaben für die Erfassung des fachdidaktischen Wissens. Dabei wurde darauf geachtet, dass für jeden Fachinhaltsbereich und für jeden fachdidaktischen Inhaltsbereich eine hinreichende Anzahl von Aufgaben für die Pilotierungsphase zur Verfügung steht. Ungefähr 75% der Aufgaben hatten ein geschlossenes, 15% ein offenes und 10% ein halboffenes Format. Im Rahmen von internen Projektmeetings und Diskussionsrunden wurden davon 80 Aufgaben mit 84 Items zum Fachwissen und 70 Aufgaben mit 78 Items zum fachdidaktischen Wissen ausgewählt. Diese wurden im Rahmen einer bundesweiten Pilotierungsstudie getestet. Um erste Informationen über Zusammenhänge zwischen Fachwissen, fachdidaktischem Wissen und pädagogischem Wissen generieren zu können, wurden daneben noch 20 Aufgaben zum pädagogischen Wissen eingesetzt. Sämtliche Aufgaben wurden dazu, gruppiert nach ihren fachlichen bzw. fachdidaktischen Inhaltsbereichen, in Aufgabenblöcke aufgeteilt, deren Bearbeitungsdauer etwa 20 Minuten darstellen sollten. Da die Bearbeitungsdauer mit durchschnittlich etwa 2 Minuten pro Aufgabe veranschlagt war, enthielt ein solcher Aufgabenblock also etwa 10 Aufgaben.

Aufgrund der unterschiedlichen Anzahl von Inhaltsbereichen zum Fachwissen und zum fachdidaktischen Wissen wurden die Aufgabenblöcke zum Fachwissen auf zwei Testhefte aufgeteilt. Durch die zwei Testhefte konnte, auf Kosten der Probandenzahl pro Aufgabe, eine größere Aufgabenanzahl getestet werden. Die Aufgabenblöcke wurden dabei wie folgt auf die zwei Testhefte aufgeteilt: Testheft I enthielt vier Blöcke zum Fachwissen zu den Bereichen Mechanik, Optik, Festkörperphysik und spezielle Relativitätstheorie sowie vier Blöcke zum fachdidaktischen Wissen in den Bereichen Schülerkognition, Instruktionsstrategien, Curriculum und Assessment. Testheft II enthielt ebenfalls vier Blöcke zum Fachwissen, allerdings zu den Bereichen Elektrizitätslehre, Thermodynamik, Atom- und Kernphysik und Quantenmechanik, drei Blöcke zum fachdidaktischen Wissen zu den Bereichen Schülerkognition, Instruktionsstrategien und Curriculum sowie einen Block zum pädagogischen Wissen. Da Aufgaben- und Itemzahlen in den beiden Testheften voneinander abweichen, ist die genaue Anzahl der Aufgaben und Testitems zu den verschiedenen Inhaltsbereichen in Tabelle 4.4 dargestellt.

Tabelle 4.4: Testheftdesign für die Pilotierungsstudie: Aufteilung der Testitems auf zwei Testhefte (blau unterlegt: CK, rot unterlegt: PCK, gelb unterlegt: PK).

Inhalt von Testheft I	Itemzahl	Inhalt von Testheft II	Itemzahl
Mechanik	13	Elektrizitätslehre	10
Optik	10	Thermodynamik	12
Festkörperphysik	10	Atom- und Kernphysik	10
Relativitätstheorie	8	Quantenmechanik	11
Schülerkognition	15	Schülerkognition II	10
Instruktionsstrategien I	8	Instruktionsstrategien II	15
Curriculum I	9	Curriculum I	10
Assessment	11	Pädagogisches Wissen (PK)	20

Bei den Aufgaben zur Erfassung des Fachwissens hatten etwa 80% der Aufgaben ein geschlossenes Format. Bei den Aufgaben zur Erfassung des fachdidaktischen Wissens lag dieser Anteil bei etwa 60%. Da jedes Testheft acht Aufgabenblöcke mit einer Bearbeitungszeit von jeweils 20 Minuten enthielt, lag die Bearbeitungszeit sämtlicher Aufgabenblöcke eines Testhefts bei 160 Minuten. Die Aufgaben bestanden meist aus einem Item, jedoch enthielten die Testhefte auch Aufgaben, die aus mehreren Items bestehen.

Neben den Merkmalen zum Professionswissen wurden noch Begleitvariablen zu Persönlichkeitsmerkmalen, wie kognitive Fähigkeiten, Semesterzahl und Schulnoten erhoben. Um einer Ermüdung der Probanden entgegenzuwirken, wurden zudem Pausen in den Ablauf der Testung aufgenommen. Die gesamte Testungsdauer betrug damit für jede Testperson 240 Minuten inklusive der Pausenzeiten.

Die Testhefte der Probanden wurden zentral am Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik elektronisch erfasst. Danach wurden die Probandenantworten anhand der Kodieranleitungen von geschulten Hilfskräften und Doktoranden kodiert. Bei den offenen Aufgaben wurde zweifache Kodierung durchgeführt, um die Objektivität der Bewertung zu erhöhen. Aufgrund der genauen Kodieranleitung ergab sich dabei eine Interraterübereinstimmung von etwa 98%.

Stichprobe der Pilotierungsstudie

Zur Rekrutierung der Stichprobe wurden Werbeposter, Flyer und Bekanntmachungen in Lehrveranstaltungen eingesetzt. Den teilnehmenden Probanden wurde als Anreiz eine Aufwandsentschädigung von 40 Euro gezahlt. Insgesamt konnten $N = 166$ Studierende des Lehramts mit Hauptfach Physik aus allen Studiensemestern für die Befragung gewonnen werden. Davon bearbeiteten 81 Probanden das Testheft I und 85 Probanden das Testheft II.

Die Verteilung der Stichprobe anhand der einzelnen Standorte und Studiensemester ist in Tabelle 4.5 dargestellt:

Tabelle 4.5: Stichprobe der Pilotierungsstudie.

Standorte	Testheft I		Testheft II	
	Probandenzahl N	mittlere Semesterzahl	Probandenzahl N	mittlere Semesterzahl
Berlin	8	6,9	8	7,4
Bielefeld	5	5,2	6	4,0
Bremen	6	5,0	9	5,1
Freiburg	8	4,3	9	5,4
Hamburg	6	6,2	3	6,0
Kassel	7	4,1	9	3,1
Kiel	11	4,4	11	4,0
München	2	8,0	3	2,0
Oldenburg	5	5,8	3	3,3
Potsdam	8	4,3	9	3,6
Regensburg	8	5,7	8	6,6
Rostock	7	8,9	7	7,7
Gesamt	81	5,5	85	5,0

Da die durchschnittliche Studiendauer für ein Physikstudium im Jahr 2010 bei 11 Semestern lag (Nienhaus, 2010) und da über sämtliche Semester hinweg Daten von Probanden vorliegen, kann anhand der mittleren Studiendauer der Probanden von 5,5 bzw. 5,0 Semestern von einer Repräsentativität der Stichprobe für die Grundgesamtheit der Physikstudierenden ausgegangen werden.

Ergebnisse der Pilotierungsstudie

Zur Auswertung der Pilotierungsstudie wurden klassische Analysen getrennt für beide Testhefte durchgeführt. Als Maß für die Reliabilität der einzelnen Skalen wurde Cronbachs α berechnet. Die Auswertung ergab ein Cronbachs $\alpha = .71$ für Testheft I beziehungsweise $\alpha = .77$ für Testheft II im Bereich Fachwissen und $\alpha = .73$ für Testheft I bzw. $\alpha = .57$ für Testheft II im Bereich fachdidaktisches Wissen (siehe Tabelle 4.6).

Tabelle 4.6: Reliabilitäten und Itemzahlen für Fachwissen und fachdidaktisches Wissen in der Pilotierungsstudie.

	Testheft I CK	Testheft II CK	Testheft I PCK	Testheft II PCK
Cronbachs α	.71	.77	.73	.57
Itemzahl n	41	43	43	35

Als ausreichendes Maß für Reliabilität wird für Cronbachs α zumeist ein Wert von mindestens .6 gefordert (Wirtz und Caspar, 2002; Moosbrugger und Kelava, 2008; George, 2011). Die Reliabilitäten der Gesamtskalen sind damit für das Fachwissen insgesamt ausreichend, beim Fachdidaktischen Wissen für Testheft II jedoch noch optimierungsbedürftig. Die Werte der einzelnen Subskalen, lagen beim Fachwissen zwischen $\alpha = .46$ und $\alpha = .62$ (mit einem Ausreißer von $\alpha = .15$) und beim fachdidaktischen Wissen zwischen $\alpha = .29$ und $\alpha = .59$ (siehe Tabellen 4.7 und 4.8).

Tabelle 4.7: Skalenskennwerte für das Fachwissen aus der Pilotierungsstudie (Itemanzahl N , mittlere Lösungswahrscheinlichkeit p , Standardabweichung s , minimale Lösungswahrscheinlichkeit p_{min} , maximale Lösungswahrscheinlichkeit p_{max} , Cronbachs α).

Fachwissen (CK)	N	p	s	p_{Min}	p_{Max}	α
Mechanik	13	.50	.23	.05	.84	.62
Elektrizitätslehre	10	.33	.19	.03	.65	.48
Optik	10	.32	.18	.10	.56	.46
Thermodynamik	12	.41	.15	.19	.63	.50
Festkörperphysik	10	.22	.20	.06	.73	.48
Atom- und Kernphysik	10	.32	.26	.03	.80	.50
Relativitätstheorie	8	.15	.09	.01	.26	.15
Quantenmechanik	11	.37	.20	.01	.72	.56

Um das Professionswissen der Zielgruppe möglichst gut messen zu können, wird eine mittlere Lösungswahrscheinlichkeit von $p \approx .50$ angestrebt. Dabei sollten zu allen Inhaltsbereichen auch einzelne einfache und schwierige Items vorkommen, um zwischen Probanden mit besonders niedrigem oder hohem Professionswissen noch differenzieren zu können. Dieses Kriterium wird bei allen Inhaltsbereichen erfüllt. Die Items zum Fachwissen haben sich anhand der mittleren Lösungswahrscheinlichkeiten, abgesehen vom Inhaltsbereich Mechanik, jedoch als etwas zu schwierig für die Zielgruppe erwiesen. Bei den Inhaltsbereichen Relativitätstheorie und Festkörperphysik waren die Items sogar deutlich zu schwierig. Die entsprechenden Kennwerte für das fachdidaktische Wissen sind in der folgenden Tabelle 4.8 dargestellt.

Tabelle 4.8: Skalenkennwerte für das fachdidaktische Wissen aus der Pilotierungsstudie (Itemanzahl N , mittlere Lösungswahrscheinlichkeit p , Standardabweichung s , minimale Lösungswahrscheinlichkeit p_{min} , maximale Lösungswahrscheinlichkeit p_{max} , Cronbachs α).

Fachdidaktisches Wissen	N	p	s	p_{Min}	p_{Max}	α
Schülerkognition I	15	.60	.27	.12	.95	.29
Instruktionsstrategien I	8	.27	.12	.06	.37	.51
Curriculum I	9	.32	.23	.01	.60	.46
Assessment	11	.40	.20	.10	.70	.44
Schülerkognition II	10	.18	.22	.00	.70	.40
Instruktionsstrategien II	15	.47	.22	.05	.88	.59
Curriculum II	11	.23	.17	.01	.50	.34

Die mittleren Lösungswahrscheinlichkeiten für das fachdidaktische Wissen liegen bis auf wenige Ausnahmen im akzeptablen Bereich, wenngleich die Aufgaben auch hier im Durchschnitt etwas zu schwierig waren. Anhand der genannten Kriterien für Cronbachs α sind die Reliabilitäten der Subskalen damit bis auf die Ausnahme der Mechanikskala optimierungsbedürftig, da Cronbachs α in diesen Fällen kleiner als .60 ist.

Zusammenfassend kann festgestellt werden, dass die durchschnittliche Lösungswahrscheinlichkeit der Items für das Fachwissen 32% betrug. Die durchschnittliche Lösungswahrscheinlichkeit für die Items zum fachdidaktischen Wissen betrug entsprechend 42%. Damit hatten sich die Aufgaben, insbesondere die fachwissenschaftlichen, als zu schwer für die Probanden erwiesen. Da man eine Normalverteilung beim Professionswissen annimmt, sollten die Verteilung der Aufgabenschwierigkeiten dieser Verteilung entsprechen, um optimal zwischen den Probanden differenzieren zu können. Die durchschnittliche Lösungsrate sollte demnach optimalerweise etwa bei 50% liegen, wobei viele Aufgaben eine mittlere Schwierigkeit aufweisen sollten und wenige Aufgaben eine niedrige bzw. hohe Schwierigkeit. Die durchschnittlichen Lösungsraten werden in den folgenden Konstruktionsphasen herangezogen, um die Aufgabenschwierigkeiten an diese Anforderungen anzupassen. Betrachtet man die Lösungsraten bezogen auf die in der Modellierung verwendeten Wissensarten, so ergibt sich beim Fachwissen eine leichte Tendenz dazu derart, dass Aufgaben, die die Anwendung deklarativen Wissens erfordern, durchschnittlich einfacher sind. Aufgaben, die schematisch/strategisches Wissen erfordern, scheinen dagegen im Vergleich am schwierigsten zu sein (siehe Tabelle 4.9). Für das fachdidaktische Wissen ist keine solche Tendenz erkennbar gewesen.

Tabelle 4.9: Aufgabenschwierigkeiten beim Fachwissen in Abhängigkeit von der Wissensart.

	Lösungsrate Testheft I CK	Lösungsrate Testheft II CK
deklarativ	.31	.41
prozedural	.32	.36
schematisch/strategisch	.28	.22

Im Folgenden werden für Validitätsanalysen weitere erhobene Merkmale herangezogen: Zur Untersuchung der externen und strukturellen Validität des Tests wurden Zusammenhänge mit der Abiturnote und der Semesterzahl der Probanden berechnet. Um ferner sicherzustellen, dass die Skalen für das Fachwissen und das fachdidaktische Wissen nicht nur die Intelligenz bzw. allgemeine kognitive

Fähigkeiten der Probanden messen, wurden die kognitiven Fähigkeiten der Probanden durch drei zusätzliche Tests erhoben: Ein verbaler Test zu Wortklassifikationen (V-Test), ein nonverbaler Test zu Figurenalogien (N-Test) und ein quantitativer Test zu Zahlenfolgen (Q-Test). Die Korrelationen der Testscores mit den Begleitvariablen sind in Tabelle 4.10 dargestellt.

Tabelle 4.10: Korrelationen zwischen Testscores und Begleitvariablen für die Pilotierungsstudie.

	Testheft I CK	Testheft II CK	Testheft I PCK	Testheft II PCK
Abiturnote	-.22 ^{**}	-.20 ⁺	-.22 [*]	-.15 ^{**}
Semesterzahl	.36 ^{**}	.32 ^{**}	.28 [*]	.20 [*]
V-Test	.30 ^{**}	.24 ^{**}	.30 ^{**}	.25 [*]
N-Test	.33 ^{**}	.28 ^{**}	.17	.38 ^{**}
Q-Test	.33 ^{**}	-.05	.17 [*]	.16

Die Abiturnote wurde dabei als Schulnote auf einer Skala von 1 (= sehr gut) bis 6 (= ungenügend) kodiert. Die negative Korrelation zwischen Note und Testscore wird damit plausibel. Die teils unterschiedlichen Werte zwischen den Testheften können durch die unterschiedlichen Aufgaben in den Testheften erklärt werden. Auffällig ist dabei insbesondere der relativ starke Zusammenhang zwischen der Semesterzahl und dem Fachwissen der Probanden im Vergleich zum fachdidaktischen Wissen. Dies deutet auf einen stärkeren Zuwachs beim Fachwissen im Vergleich zum fachdidaktischen Wissen hin, was etwa dem Anteil der Lerngelegenheiten im Studium entspricht, zumal in den Anfangsvorlesungen im 1. und 2. Semester meist noch keine oder nur wenige Lerngelegenheiten für fachdidaktisches Wissen im Studium bestehen.

Bezüglich der kognitiven Fähigkeiten konnte Riese (2009) einen korrelativen Zusammenhang von .28 zwischen kognitiven Fähigkeiten, gemessen anhand der Abiturnote von Studierenden, und dem Fachwissen messen. Zwischen kognitiven Fähigkeiten und fachdidaktischen Wissen konnte er einen etwas geringeren Zusammenhang von .14 messen (Riese, 2009). Bezogen auf die Ergebnisse der vorliegenden Studie ergeben sich insoweit, bis auf eine Ausnahme beim Q-Test im Testheft II, plausible Korrelationen zwischen den kognitiven Fähigkeiten der Probanden und deren Testscores. Zwischen den Summenscores beim Fachwissen und fachdidaktischen Wissen ergab sich für beide Testhefte jeweils eine Korrelation von .40 ($p < .001$). Dies liegt zwar etwas unterhalb, jedoch in derselben Größenordnung wie die Korrelationen, die im Projekt ProwiN (Kirschner, 2015) für Lehrkräfte ($r = .45^{***}$) und in der Arbeit von Riese (2009) für angehende Lehrkräfte ($r = .45^{***}$) ermittelt wurden (siehe Abschnitt 2.1.2).

Um einerseits der festgestellten Problematik zu vieler schwieriger Aufgaben gerecht zu werden und andererseits die inhaltliche Validität der verwendeten Items, bzw. der mithilfe der Items gemessenen Ergebnisse, sicherzustellen, wurden die eingesetzten Aufgaben anschließend Experten zur Begutachtung vorgelegt.

4.4.3 Expertenratings

Die Begutachtung der in der Pilotierungsstudie eingesetzten Aufgaben dient der anschließenden Überarbeitung und Verbesserung der Testinstrumente für den Einsatz im Rahmen der Hauptstudie der vorliegenden Arbeit. Konkrete Ziele dabei waren erstens die Reduktion der Items um etwa die

Hälfte, um in der Hauptstudie anstelle von zwei Testheften nur noch ein Testheft einzusetzen. Die Anzahl der Probanden pro Testaufgabe sollte so erhöht werden. Zweitens wurden inhaltliche Verbesserungen der bestehenden Aufgaben und die Sicherstellung der Repräsentation des Konstrukts durch die Gesamtheit der Aufgaben im Hinblick auf die Hauptstudie angestrebt. Die Begutachtung der Aufgaben durch Experten soll damit sicherstellen, dass das entwickelte Testinstrument dem inhaltlichen Aspekt der Konstruktvalidität nach Messick (1995) genügt. Drittens sollte die mittlere Aufgabenschwierigkeit gesenkt werden, um sicherzustellen, dass das finale Testinstrument möglichst gut zwischen den unterschiedlichen Niveaustufen der angestrebten Zielgruppe differenzieren kann.

Expertenrating für die Aufgaben zum Fachwissen

Die Aufgaben zum Fachwissen wurden von einem Fachwissenschaftlern mit besonderer Expertise in der Lehrerinnen- und Lehrerausbildung anhand eines Fragebogens begutachtet. Der Fragebogen sollte dabei sicherstellen, dass besonders wichtige Qualitätsaspekte der Aufgaben nicht unberücksichtigt bleiben. So wurde unter anderem gefragt, ob die einzelnen Items einer Aufgabe zum intendierten Inhaltsbereich passt, ob der Inhalt der Items einen Vorlesungsgegenstand für die Zielgruppe darstellt, ob die Schwierigkeit der Items als angemessen beurteilt wird und ob sie verständlich und eindeutig formuliert sind. Der Gutachter konnte zusätzlich weitere Kommentare zu den Aufgaben abgeben (siehe Abbildung 4.10).

Block 7: Bewertung der Items (nicht vom Aufgabenersteller auszufüllen)							
Bitte schätzen Sie hier die Items anhand der folgenden Aspekte durch Eintragen einer Ziffer von 1 bis 4 (1=trifft nicht zu, 2=trifft eher nicht zu, 3=trifft eher zu, 4=trifft voll zu) ein und kommentieren Sie ihre Einschätzung gegebenenfalls.							
Kriterium	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Kommentar
Die Zuordnung des Items zum Inhaltsbereich passt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Der Fachinhalt des Items wird in den Vorlesungen für Lehramtsstudierende an meinem Standort gelehrt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Das Item kann von Studierenden, die die entsprechenden Lehrveranstaltungen besucht haben, gelöst werden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Das Item erfasst ein wichtiges Prinzip des jeweiligen Inhaltsbereichs.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Die Zuordnung des Items ist zur Wissensart passt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Der Aufgabenstamm ist für Studierende verständlich formuliert.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Das Item ist eindeutig formuliert.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Das Item kann im angegebenen Zeitrahmen von Studierenden gelöst werden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Die Musterlösung für das Item ist nachvollziehbar.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Es gibt neben der Musterlösung noch einfachere Lösungen für das Item.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Das Item ist per Ausschlussverfahren oder andere (auch nichtfachliche) Ansätze lösbar.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Für das Item sind typische falsche Antwortoptionen angegeben.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Die vorgeschlagene Punktvorgabe, (insbesondere für partial credits) ist für das Item angemessen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Die Lösung des Items erfordert physikalisches Fachwissen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Abbildung 4.10: Auszug aus dem Fragebogen für die Bewertung der Items.

Anschließend wurde der Gutachter für jede Aufgabe zu einer zusammenfassenden Bewertung aufgefordert, wobei wiederum die Möglichkeit bestand zu einzelnen Aspekten der Aufgabe freie Kommentare abzugeben (siehe Abbildung 4.11).

Block 8: Bewertung der Aufgabe (nicht vom Aufgabenersteller auszufüllen)*Bitte schätzen Sie hier die Aufgabe anhand der folgenden Aspekte ein und kommentieren Sie ihre Einschätzung gegebenenfalls.*

Kriterium	Bewertung				Kommentar
Für wie geeignet halten Sie die Aufgabe, um das Professionswissen von Lehramtsstudierenden zu erfassen?	nicht geeignet <input type="checkbox"/>	eher nicht geeignet <input type="checkbox"/>	eher geeignet <input type="checkbox"/>	sehr geeignet <input type="checkbox"/>	<input type="checkbox"/>
Wie schätzen Sie die Aufgabenschwierigkeit im Hinblick auf die Zielgruppe (Lehramt) ein?	sehr einfach <input type="checkbox"/>	eher einfach <input type="checkbox"/>	eher schwer <input type="checkbox"/>	sehr schwer <input type="checkbox"/>	<input type="checkbox"/>
Die einzelnen Items der Aufgabe sind unabhängig voneinander zu lösen.	trifft nicht zu <input type="checkbox"/>	trifft eher nicht zu <input type="checkbox"/>	trifft eher zu <input type="checkbox"/>	trifft zu <input type="checkbox"/>	<input type="checkbox"/>

Abbildung 4.11: Fragebogen für die zusammenfassende Aufgabenbewertung.

Nach der zusammenfassenden Bewertung einer Aufgabe wurde der Gutachter aufgefordert entsprechend seiner Bewertung mögliche Adaptionmöglichkeiten zu der Aufgabe anzugeben (siehe Abbildung 4.12).

Block 9: Adaption (nicht vom Aufgabenersteller auszufüllen)*Bitte geben Sie hier an, was Sie – entsprechend ihren obigen Einschätzungen – an der Aufgabe oder einzelnen Items verändern würden.*

Abbildung 4.12: Fragebogen für Adaptionmöglichkeiten zu einer Aufgabe.

Expertenrating für die Aufgaben zum fachdidaktischen Wissen

Bei den Aufgaben für die Erfassung des fachdidaktischen Wissens handelte es sich, im Gegensatz zu den Aufgaben für die Erfassung des Fachwissens, gänzlich um Neuentwicklungen. Daher wurden diese im Rahmen des Expertenratings, verglichen mit den Aufgaben zum Fachwissen, einer aufwändigeren Begutachtung unterzogen. Hierzu wurden die einzelnen Aufgabenblöcke zunächst jeweils einem von insgesamt acht Experten mit besonderer Expertise in der Professionswissensforschung und von unterschiedlichen Standorten zur Begutachtung anhand des oben beschriebenen Fragebogens vorgelegt. Bei den begutachtenden Experten handelte es sich um drei Post-Docs und fünf Professoren. Zusätzlich zu diesen Einzelbegutachtungen wurden alle fachdidaktischen Aufgaben im Rahmen eines zweitägigen Expertenmeetings gemeinsam mit allen Gutachtern diskutiert. Aufgaben- und Skalennennwerte aus der Pilotierungsstudie, wie Lösungsraten, Trennschärfen und Reliabilitäten wurden den Experten als zusätzliche Information mitgeteilt. Die Lösungsrate eines Items ist ein Maß für die Aufgabenschwierigkeit. Sie gibt an, mit welcher Wahrscheinlichkeit eine Aufgabe im Durchschnitt erfolgreich bearbeitet wurde. Eine niedrige Lösungsrate ist also ein Hinweis auf eine hohe Aufgabenschwierigkeit. Die Trennschärfe eines Items gibt die Korrelation zwischen den Probandenergebnissen des jeweiligen Items und der jeweiligen Skala an. Hat beispielsweise ein Item zum Inhaltsbereich Schülerkognition eine niedrige Trennschärfe, so ist dies ein Hinweis darauf, dass das Item überwiegend von denjenigen Probanden erfolgreich bearbeitet wurde, die insgesamt nur wenige Items zum Inhaltsbereich Schülerkognition erfolgreich bearbeiten konnten. Eine niedrige Trennschärfe weist also darauf hin, dass das Item ein anderes Konstrukt misst als die Skala.

Im Rahmen der Diskussion wurde für jede einzelne Aufgabe festgelegt, ob diese entweder verworfen, überarbeitet oder unverändert eingesetzt werden sollte. Um sicherzustellen, dass genügend Diskussionszeit für die einzelnen Aufgaben zur Verfügung stand, wurde ein Zeitplan für das Expertenrating erstellt. Dieser sah am Nachmittag des ersten Tages die Vorstellung des Projekts und des Modells für die Aufgabenkonstruktion vor (siehe Abschnitt 4.3.5). Anschließend standen jeweils zwei Stunden Diskussionszeit für die Aufgabenblöcke zu den Bereichen Schülerkognition und Instruktionsstrategien

vor. Am zweiten Tag standen vormittags noch 45 Minuten für den Abschluss der Diskussion vom Vortag zur Verfügung. Anschließend wurden wieder jeweils zwei Stunden Diskussionszeit für die Aufgabenblöcke zu den Bereichen Curriculum und Assessment veranschlagt. Eine Schlussdiskussion von 45 Minuten bildete den Abschluss des Expertenmeetings.

Im Rahmen der Diskussionen wurde vor allem auf die fachliche und fachdidaktische Relevanz der Aufgabeninhalte, auf die Formulierung der Aufgaben sowie die Aufgabenschwierigkeiten fokussiert. Änderungsvorschläge der Experten wurden, soweit möglich, während der Diskussion in die Aufgaben eingearbeitet und anderenfalls dokumentiert. Hierzu wurden die Diskussionsbeiträge für die späteren Auswertungen und Umsetzungen audiographiert und gespeichert.

Ergebnisse der Expertenratings

Als Ergebnis des Expertenmeetings sind in alle Aufgaben Verbesserungen hinsichtlich einer einfacheren und präziseren Aufgabenformulierung eingeflossen. Etwa bei der Hälfte der Aufgaben konnte im Rahmen des Expertenmeetings ein Konsens erzielt werden. Das bedeutet, dass diese Aufgaben von allen Experten als geeignet für die Erfassung des fachdidaktischen Wissens beurteilt wurden. Die andere Hälfte der Aufgaben wurde verworfen bzw. es wurden weitreichende Änderungsvorschläge seitens der Experten genannt. In diesen Fällen wurden die Aufgaben aus inhaltlichen Gründen von einigen Experten abgelehnt. Dies waren Aufgaben, für deren fachdidaktische Inhalte an einigen, aber nicht an allen Standorten der bundesweit angelegten KiL-Studie Lerngelegenheiten für die Studierenden existieren oder deren Inhalte nicht von allen Experten als relevant erachtet wurden. Zudem wurden einige Aufgaben von den Experten als zu schwierig kritisiert. Für einen Teil dieser Aufgaben wurden seitens der Experten Überarbeitungsvorschläge erteilt, die eine Optimierung und den weiteren Einsatz der Aufgaben ermöglichen sollten.

Des Weiteren wurde von den Experten beurteilt, ob die einzelnen Bereiche des Fachwissens bzw. des fachdidaktischen Wissens durch die Gesamtheit der Aufgaben zu den einzelnen Bereichen angemessen repräsentiert werden. Hierzu wurde insbesondere zu den Aufgaben zum Fachwissen Ergänzungsvorschläge gemacht, sodass hier für die acht Aufgabenblöcke jeweils 1-2 Aufgaben zum Aufgabenpool hinzugefügt werden konnten.

Als weiteres Ergebnis, welches sich sowohl auf den CK-, wie auch auf den PCK-Begutachtungsprozess bezieht, wurden einige Aufgaben von den Experten für geeignet befunden, obwohl diese in der Pilotierungsstudie problematische Kennwerte lieferten. Da hierfür zunächst keine Erklärung vorlag, sollte dieser Effekt einer näheren Untersuchung unterzogen werden. Die Untersuchung dieser speziellen Aufgaben ist Gegenstand der im Folgenden vorgestellten Studie.

Eine solche Beispielaufgabe mit einer relativ niedrigen Lösungswahrscheinlichkeit von 39,2% in der Pilotierungsstudie. Die Aufgabe hatte darüber hinaus mit einer Item-Skala-Korrelation von .13 und einer Item-Subskala-Korrelation von .12 in der Pilotierungsstudie schwache Aufgabenkennwerte geliefert. Sie wurde jedoch von den Experten inhaltlich als geeignet eingestuft wurde. Die ursprüngliche Aufgabe ist mit den Lösungswahrscheinlichkeiten aus der Pilotierungsstudie in Abbildung 4.13 dargestellt.

Aufgabenstamm		
In einer Studie wurde erhoben, wie das Interesse am Physikunterricht mit anderen Variablen zusammenhängt. Dabei zeigte sich einerseits, dass das Interesse am Fach Physik immer dann groß ist, wenn die Schülerinnen und Schüler überzeugt sind, im Physikunterricht etwas leisten zu können. Andererseits ergab sich, dass das Fachinteresse fast gar nicht vom Sachinteresse an der Physik oder vom Fasziniertsein an Technik und Naturphänomenen abhängt.		
Item	Richtig	Falsch
Welche der folgenden Aussagen dazu sind richtig, welche falsch?		
Im Physikunterricht sollte man nicht zu viel Rücksicht auf die Interessen der Schülerinnen und Schüler nehmen.	11,1%	88,9%
Der vorausgegangene Physikunterricht geht offenbar weitgehend an den Interessen der Schülerinnen und Schüler vorbei.	75,9%	24,1%
Die Mädchen sind besonders benachteiligt, weil sie nachweislich im Vergleich zu Jungen weniger überzeugt sind, im Physikunterricht etwas leisten zu können.	71,3%	28,7%
Aus der Studie lässt sich nichts schließen, denn jede Klasse ist anders.	29,1%	70,9%

Abbildung 4.13: Aufgabenbeispiel aus dem Expertenrating, richtige Antwortanteile in grün, falsche Antwortanteile in rot.

Die Richtigkeit der einzelnen Aussagen wurde von den Probanden zwar überwiegend korrekt eingeschätzt, jedoch haben nur 39,2% der Probanden alle vier Aussagen richtig bewertet. Von den Experten wurden die Länge des Aufgabenstammes und die Formulierung der vierten Aussage kritisiert. Die durch die Experten geänderte Fassung der Aufgabe ist in Abbildung 4.14 dargestellt.

Aufgabenstamm		
In einer Studie wurde erhoben, wie das Interesse am Physikunterricht mit anderen Variablen zusammenhängt.		
Item	Richtig	Falsch
Welche der folgenden Aussagen dazu sind richtig, welche falsch?		
Im Physikunterricht sollte man nicht zu viel Rücksicht auf die Interessen der Schülerinnen und Schüler nehmen.	<input type="checkbox"/>	<input type="checkbox"/>
Der vorausgegangene Physikunterricht geht offenbar weitgehend an den Interessen der Schülerinnen und Schüler vorbei.	<input type="checkbox"/>	<input type="checkbox"/>
Die Mädchen sind besonders benachteiligt, weil sie nachweislich im Vergleich zu Jungen weniger überzeugt sind, im Physikunterricht etwas leisten zu können.	<input type="checkbox"/>	<input type="checkbox"/>
Aus der Studie lässt sich nichts schließen, denn jede Klasse ist anders.	<input type="checkbox"/>	<input type="checkbox"/>

Abbildung 4.14: Aufgabenbeispiel aus dem Expertenrating, richtige Antwortanteile in grün, falsche Antwortanteile in rot.

4.4.4 Fragebogenstudie

Ziel der Fragebogenstudie, die im Rahmen einer Masterarbeit (Hübner, 2013) durchgeführt wurde, war es einerseits Gründe für die zu niedrigen Lösungsraten und Trennschärfen einzelner Aufgaben in der Pilotierungsstudie zu identifizieren und andererseits Optimierungsmöglichkeiten für solche Aufgaben zu finden. Die Mängel dieser einzelnen Aufgaben wurden aufgrund der positiven Begutachtung durch die Experten nicht im inhaltlichen Bereich vermutet. Naheliegender war es deshalb, zu untersuchen, ob es durch einen Mangel kognitiver Validität der Aufgaben zu Verständnisschwierigkeiten seitens der Probanden kommen konnte. In der Literatur finden sich verschiedene Verfahren zur Untersuchung der kognitiven Validität, die sich hinsichtlich Testökonomie und Informationsgewinn stark voneinander unterscheiden (siehe Tabelle 4.11). Für die Überprüfung von Items aus Fragebögen sind die Formate des Lauten Denkens, des Interviews und des Fragebogens besonders etabliert. (Van Someren et al., 1994; Thelk, 2006; Ifenthaler, 2006; Beggow, 2013). Eine Übersicht dieser Verfahren ist in Tabelle 4.11 dargestellt.

Beim Lauten Denken wird der Testperson der zu untersuchende Test vorgelegt, der bei gleichzeitiger Verbalisierung von Gedanken, Antworten, Handlungs- und Arbeitsschritten „lautdenkend“ zu bearbeiten ist. So können die stattfindenden kognitiven Prozesse dokumentiert und in einer Auswertung mit den zu erwartenden Prozessen verglichen werden (vgl. Hadenfeldt und Neumann, 2014). Dabei werden in der Regel nicht die Ursachen für unerwartete kognitive Prozesse erfasst. Für einen Informationsgewinn zur Verbesserung der Aufgaben wirken sich die wenig anregende Denkkumgebung, der kognitive Anspruch all seine Gedanken zu äußern, die begrenzten Rahmen von Testaufgaben und die begrenzte Anzahl möglicher Testteilnehmer aufgrund des hohen zeitlichen Aufwands bei der Auswertung, sowie die nicht standardisierten Datensätze negativ aus. Vorteile sind hingegen die offene Arbeitsumgebung und das mögliche Beobachten anderer Prozesse während der Bearbeitung.

Tabelle 4.11: Stärken und Schwächen der Testverfahren zur Erfassung kognitiver Validität hinsichtlich verschiedener Aspekte (+ : Methode gut geeignet; o : Methode mittelmäßig geeignet; - : Methode wenig geeignet).

Aspekt	Lautes Denken	Interview	Fragebogen
Erhebungsaufwand	o	-	+
Auswertungsaufwand	-	o	+
Erhebungsobjektivität	+	o	+
Auswertungsobjektivität	o	o	o
Individualität der Ergebnisse	+	+	-
Authentizität der Testsituation	+	-	-

Die durch Interviews generierten Daten können durch einen standardisierten Fragenkatalog homogener als beim Lauten Denken sein. Beiden Methoden gemein ist der vergleichsweise hohe Aufwand der Durchführung und Auswertung. Während beim Lauten Denken der hohe Aufwand bei der Analyse in der Natur qualitativer Daten und deren Auswertung liegt, kann der Aufwand der Durchführung und Auswertung beim Interview durch eine Standardisierung des Fragenkatalogs verringert werden. Die Tatsache, dass ein Interview mit hoch standardisiertem Fragenkatalog nur noch wenig Vorteil gegenüber einer Befragung mittels Fragebogen bietet, legt eben diese Methode als Alternative nahe. Durch eine Befragung mittels Fragebogen können vom Testleiter unabhängige Ergebnisse erzielt

werden. Gleichzeitig ist der Durchführungsaufwand durch die mögliche parallele Testung vieler Probanden deutlich reduziert. Nicht zuletzt ist eine standardisierte Auswertung möglich. Dabei sollen Items des Fragebogens so gestaltet sein, dass so wenig Informationsverlust wie möglich auftritt, d.h., dass die zur Beurteilung der kognitiven Validität notwendigen Informationen möglichst umfassend erhoben werden. Die Methode einer Fragebogenbefragung bietet also Möglichkeiten, einer effizienten und objektiven Untersuchung der kognitiven Validität. Aus diesen Gründen wurde für die vorliegende Studie die Methode des Fragebogens gewählt. Im Folgenden wird, ausgehend von einem in den Vereinigten Staaten entwickelten Fragenkatalog zur Verbesserung von Testaufgaben, der für diese Studie entwickelte und eingesetzte Fragebogen vorgestellt.

Entwicklung des Erhebungsinstruments

Die American Association for the Advancement of Science (AAAS) veröffentlichte im Rahmen des Project 2061 einen Fragenkatalog (AAAS, 2007), der der Entwicklung von Leistungstestaufgaben dient und die Grundlage für den in der Masterarbeit (Hübner, 2013) verwendeten Fragebogen, den Cognitive Lab-Questionnaires (CLQ), darstellt. Der Fragenkatalog der AAAS wurde dazu ins Deutsche übersetzt, für die Erfassung der kognitiven Validität im Sinne der Aufgabenentwicklung für die vorliegende Arbeit neu gegliedert und erweitert. Das Hauptaugenmerk lag dabei auf der Gewinnung von Informationen über die kognitive Validität von Testaufgaben des Fachwissens und fachdidaktischen Wissens. Der CLQ stellt damit einen standardisierten, universell einsetzbaren Fragebogen zur Untersuchung der kognitiven Validität solcher Testaufgaben dar. Insbesondere sollten durch ihn Informationen über Aufgabenmerkmale generiert werden, die zu einer Herabsetzung der kognitiven Validität führen können. Solche Merkmale können beispielsweise Unklarheiten oder unbekannte Begriffe in der Aufgabenstellung sein.

Insgesamt umfasst der CLQ zwölf Fragen: Die Fragen 1 und 2 befassen sich mit der Erfassung der Aufgabe durch den Probanden. Hierbei sollen unklare Begriffe und Formulierungen identifiziert werden. Die Fragen 3 bis 7 befassen sich mit der gewählten Lösung und dem zugehörigen Lösungsweg und sollen die eventuelle Anwendung unangemessener Lösungsstrategien erfassen. Die Fragen 8 und 9 erfassen Eigenschaften von ggf. in der Aufgabe enthaltenen Grafiken und Formeln. Das Ziel hierbei ist zu erfahren, ob eventuell beigefügte grafische Elemente ihren Zweck erfüllen. Die Fragen 10 und 11 erfassen Informationen zum Inhalt der untersuchten Aufgabe. Hier liegt das Ziel darin, zu erfahren, ob der Inhalt der Aufgabe von den Probanden als valide angesehen wird. Die Frage 12 gibt den Probanden schließlich Gelegenheit, Verbesserungsvorschläge zur untersuchten Aufgabe zu geben.

Studiendesign

Im Rahmen der Masterarbeit wurde der CLQ zur Untersuchung der Validität von 20 Aufgaben für die Erfassung des Fach- und des fachdidaktischen Wissens eingesetzt. Die hier untersuchten 20 Aufgaben wurden anhand von Kennwerten, die in der Pilotierungsstudie des Projekts KiL gewonnen wurden, ausgewählt, weil sie unerwartet schlechte psychometrische Eigenschaften hinsichtlich ihrer Schwierigkeit oder ihrer Trennschärfe gezeigt haben. Im Rahmen des Expertenratings wurden diese Aufgaben jedoch als inhaltlich prinzipiell geeignet eingeschätzt. Jeweils 10 dieser Aufgaben entstammten aus den Bereichen des Fachwissens und des fachdidaktischen Wissens.

Stichprobe

Für die Untersuchung wurden die 20 Aufgaben, jeweils zusammen mit dem CLQ, einer Stichprobe von $N = 13$ Probanden vorgelegt. Die Probanden waren Lehramtsstudierende unterschiedlicher Semester mit Hauptfach Physik an der Christian-Albrechts-Universität zu Kiel. Da zunächst lediglich Informationen über die Ursachen für die schlechten psychometrischen Eigenschaften gewonnen werden sollen, wurde eine Stichprobengröße von mindestens 10 Probanden anvisiert, jedoch erlaubt der CLQ eine testökonomische Erweiterung auf größere Stichproben. Für die Bearbeitung der 20 Aufgaben einschließlich des Ausfüllens des CLQ haben sich aufgrund vorheriger Pilotierung des Fragebogens 90 Minuten Zeit als ausreichend erwiesen.

Auswertung der Fragebögen

Die Suche nach kritischen Aufgabenmerkmalen wurde individuell für jede untersuchte Aufgabe vorgenommen. Dabei wurde zunächst nach möglichen Auffälligkeiten in den zur Aufgabe generierten Daten, d.h. den Probandenantworten zur jeweiligen Aufgabe, gesucht. Als Auffälligkeit gilt zum Beispiel das mehrfache Auftreten unbekannter Begriffe im Aufgabentext, niedrige Lösungswahrscheinlichkeiten, insbesondere unterhalb der Ratewahrscheinlichkeit bei Multiple-Choice-Items, mehrheitlich geratene Lösungen bei geschlossenen Aufgaben oder Distraktoren bei Multiple-Choice-Items, die von der Mehrzahl der Probanden ausgewählt wurden. Anhand der einzelnen Fragen des CLQ konnten damit Ursachen für Auffälligkeiten gefunden oder zumindest eingegrenzt werden. Unbekannte Begriffe, zum Beispiel, können anhand von Frage 1 direkt identifiziert werden. Werden gehäuft bestimmte Begriffe, wie zum Beispiel im unten aufgeführten Aufgabenbeispiel (siehe Abbildung 4.15), als unbekannt genannt, so kann anhand der Ergebnisse ein Optimierungsvorschlag entworfen werden. Der Aufgabentext des Optimierungsvorschlags verzichtet dann auf die genannten Begriffe, erläutert sie oder ersetzt sie. Weitere Ursachen für niedrige Lösungswahrscheinlichkeiten, die anhand der Fragen 2 bis 7 gefunden wurden, sind unnötig komplizierte Beschreibungen von Sachverhalten oder Fragestellungen im Aufgabentext. Die von den Probanden genannten Punkte wurden in einer Liste gesammelt und nach Häufigkeit sortiert. Mehrfach genannte Aspekte wurden wiederum als kritische Aufgabenmerkmale aufgefasst.

Ergebnisse der Fragebogenstudie

Bei den als kritisch identifizierten Aufgabenmerkmalen lässt sich zwischen formalen und inhaltlichen Merkmalen unterscheiden. Dabei versteht man unter formalen Aufgabenmerkmalen das Aufgabenformat und die Formulierung der Aufgabe. Beispiele sind Verneinungen im Aufgabentext, die von den Probanden leicht überlesen werden können oder im Aufgabentext verwendete Begriffe, die einigen Probanden unbekannt sind. Inhaltliche Aufgabenmerkmale sind solche, die sich auf das zum Lösen der Aufgabe notwendige Fachwissen oder fachdidaktische Wissen beziehen.

In den meisten Fällen handelte es sich bei den identifizierten kritischen Aufgabenmerkmalen um formale Aufgabenmerkmale. Für die Testung professionellen Wissens sind solche Aufgaben besonders problematisch, da sie dazu führen können, dass die Aufgabe anstelle von Fachwissen oder fachdidaktischem Wissen eher das Text- oder Sprachverständnis des jeweiligen Probanden erfassen. Ein Beispiel für solche Aufgabenmerkmale ist in Abbildung 4.15 dargestellt.

Bei den formalen Aufgabenmerkmalen zeigte sich ferner, dass Verneinungen und Schlüsselworte, wie „nicht“, „immer“, „nur“, „muss“, in den Aufgabentexten weniger oft überlesen werden, wenn diese durch Unterstreichung hervorgehoben werden. Ebenso erwies es sich als kognitiv entlastend für die Probanden, wenn erläuternde grafische Elemente wie Abbildungen oder Diagramme zu einer Aufgabe beigelegt werden und im Aufgabentext Bezug auf diese genommen wird. Bezogen auf inhaltliche Aufgabeneigenschaften zeigte sich, dass unbekannte Fachbegriffe, entbehrliche Informationen und die Formulierung des Aufgabenkontextes bei Probanden zu kognitiven Schwierigkeiten bei der Lösung einer Aufgabe führen können. Zur Verbesserung der Aufgabenqualität des Testinstruments wurden diese Erkenntnisse auf alle bestehenden Aufgaben angewendet, indem die Aufgaben nochmals auf ihre formalen Aufgabenmerkmale hin überprüft wurden und ggf. überarbeitet wurden.

Wie eine Aufgabe anhand der Ergebnisse der Fragebogenstudie optimiert werden konnte, soll im folgenden Beispiel einer Aufgabe zum Fachwissen im Bereich Relativitätstheorie gezeigt werden. Abbildung 4.15 zeigt dabei die Aufgabe, so wie sie in der Pilotierungsstudie eingesetzt wurde.

Aufgabenstamm
Daniel fliegt mit $v = 0,8 c$ in Richtung Polarstern. Elton bleibt auf der Erde zurück und betrachtet am Nachthimmel den Gürtel des Orion, welcher auf dem Himmelsäquator liegt und folglich einen Winkel von $\pi/2$ zum Polarstern besetzt.
Item
Unter welchem Winkel sieht Daniel diese Konstellation?
<input type="checkbox"/> mehr vorne, also $< \pi/2$ <input type="checkbox"/> mehr vorne, also $\pi/2$ <input type="checkbox"/> mehr hinten, also $> \pi/2$ <input type="checkbox"/> kann den Gürtel nicht sehen

Abbildung 4.15: Problematische Aufgabe zum Fachwissen aus der Pilotierungsstudie.

Die richtige Lösung der Aufgabe ergibt sich durch die Berücksichtigung der scheinbaren Ortsveränderung des Gürtels des Orion aufgrund der Endlichkeit der Lichtgeschwindigkeit (Aberration). Richtig ist bei der ersten Version der Aufgabe deshalb die erste Antwortoption: „mehr vorne, also $< \pi/2$ “.

Die Aufgabe wurde den befragten Probanden als eine von zehn Aufgaben aus dem Bereich des Fachwissens vorgelegt. Die Lösungsrate lag für diese Aufgabe bei 23%, also etwa im Bereich der Rätewahrscheinlichkeit. Tatsächlich gaben 62% der Probanden an, geraten zu haben. Nahezu alle Probanden waren sich bei ihrer Lösung nicht sicher. Dieser Befund legte eine detailliertere Untersuchung der qualitativen Daten zur Verständlichkeit der Aufgabe nahe. Dabei konnten mehrere Punkte identifiziert werden, die die Validität der Testaufgabe mindern:

- (a) Die Testaufgabe enthält für viele Probanden unbekannte Begriffe. Hierzu zählen: Polarstern, Himmelsäquator, Gürtel des Orion. Dies deutet darauf hin, dass die Probanden nicht mit dem Kontext der Aufgabe vertraut sind.
- (b) Die Variablen „ v “ und die Konstante „ c “ werden in der Aufgabe nicht eingeführt. Dies führt zu Verständnisschwierigkeiten einiger Probanden.

- (c) Die Aufgabenstellung wird durch die Einführung von Namen unnötig lang, was eine höhere Lesekompetenz seitens der Probanden erfordert und gleichzeitig womöglich die Motivation, sich intensiver mit der Aufgabe zu beschäftigen, senkt.
- (d) Der Sachverhalt ist offenbar zu kompliziert dargestellt und daher schwer zu erfassen. 62% der Studienteilnehmer wünschten sich eine Skizze zur Verdeutlichung.

Auf Basis der Befunde wurde die Aufgabe überarbeitet, wobei die entsprechenden Mängel beseitigt wurden (vgl. Abbildung 4.16).

Aufgabenstamm

Ein Beobachter fliegt mit 80% der Lichtgeschwindigkeit, wie in der Abbildung gezeigt, an einem Stern vorbei.



Item

In welcher Richtung sieht er den Stern wenn dieser sich, wie oben dargestellt, im rechten Winkel zu seiner Flugbahn befindet?

- Er sieht ihn genau im rechten Winkel.
- Er sieht ihn nicht im rechten Winkel, sondern etwas vor sich.
- Er sieht ihn nicht im rechten Winkel, sondern etwas hinter sich.
- Er kann ihn gar nicht sehen.

Abbildung 4.16: Überarbeitete Aufgabe zum Fachwissen.

Aufbauend auf den Erkenntnissen der Fragebogenstudie, aber auch der Pilotierung und des Expertenratings, wurde der Aufgabenpool einer umfassenden Überarbeitung unterzogen.

Zusammenfassung der Erkenntnisse aus der Fragebogenstudie

Die Ergebnisse der Fragebogenstudie zeigen anhand des eingesetzten CLQ Aufgabeneigenschaften auf, die Ursachen für eine Beeinträchtigung der kognitiven Validität untersuchter Testaufgaben sein können. Besonders Aufgabeneigenschaften, wie die Formulierung, verwendete Begriffe, die Darstellung von Sachverhalten und die Verwendung von Fachbegriffen führten bei den Probanden der Studie zu Problemen, die eine Bearbeitung der Aufgaben erschwerten. Weitere Erprobungen des eingesetzten CLQ mit größeren Stichproben wären an dieser Stelle wünschenswert. Die weitere Validierung des CLQ ist jedoch nicht Ziel der vorliegenden Studie. Stattdessen sollte der CLQ hier als Hilfsmittel zur Aufklärung geringer kognitiver Validität ausgewählter, problembehafteter Testaufgaben dienen. Die hierzu vorgestellten Ergebnisse konnten trotz der geringen Stichprobengröße konkrete Anhaltspunkte für problematische Aufgabenmerkmale und ferner Anhaltspunkte für die Optimierung der untersuchten Aufgaben liefern. Dabei waren Erhebungs- und Kodieraufwand deutlich geringer als beim Lauten Denken und traditionellem Interview. Mit dem CLQ liegt damit insofern ein alternatives Instrument zu den etablierten Verfahren Lautes Denken und Interview für die Untersuchung kogniti-

ver Validität vor. Ob die Qualität der gefundenen Ergebnisse an die der traditionellen Methoden heranreicht, bleibt jedoch zunächst unklar. Um eine Vergleichbarkeit der Ergebnisse verschiedener Methoden zu untersuchen, bietet sich eine Vergleichsstudie der Verfahren CLQ, Lautes Denken und Interview an. Da dies jedoch nicht Ziel der vorliegenden Arbeit ist, wird eine solche Vergleichsstudie hier nicht in Angriff genommen.

4.4.5 Überarbeitung und Neukonstruktion von Aufgaben

In der folgenden Phase der Überarbeitung und Neukonstruktion von Aufgaben wurden die Erkenntnisse der Pilotierungsstudie, des Expertenratings und der Interviewstudie auf den vorliegenden Aufgabenpool angewendet. Dazu wurden sämtliche Aufgaben auf die verwendeten Begriffe, Klarheit der Aufgabenformulierung, ggf. mögliche visuelle Hilfen und die gesonderte Kenntlichmachung verneinender Formulierungen begutachtet und an die gewonnen Erkenntnisse angepasst. Dazu wurde der Aufgabenpool der Testaufgaben zum Fachwissen und fachdidaktischen Wissen, dargestellt in Tabelle 4.12, nochmals verkleinert.

Tabelle 4.12: Aufgabenpool für CK und PKC vor der Pilotierungsstudie (blau: CK; rot: PCK).

Inhalt von Testheft I	Itemzahl	Inhalt von Testheft II	Itemzahl
Mechanik	13	Elektrizitätslehre	10
Optik	10	Thermodynamik	12
Festkörperphysik	10	Atom- und Kernphysik	10
Relativitätstheorie	8	Quantenmechanik	11
Schülerkognition I	15	Schülerkognition II	10
Instruktionsstrategien I	8	Instruktionsstrategien II	15
Curriculum I	9	Curriculum II	10
Assessment	11		

Die Gesamtheit der verbliebenen Aufgaben sollte einerseits den jeweiligen Inhaltsbereich laut der befragten Experten noch angemessen repräsentieren und andererseits die Anzahl der Aufgaben so reduzieren, dass die Aufgaben in einer angemessenen Zeitspanne durch die Probanden zu bewältigen sind. Als maximale Testzeit wurde hier ein Zeitrahmen von 4 Stunden, inklusive Pausen, anvisiert. Eine solche Zeitspanne intensiver schriftlicher Arbeit wird auch in Prüfungssituationen mitunter von Studierenden abverlangt und kann daher als zumutbar gelten. Der verkleinerte Aufgabenpool ist in Tabelle 4.13 dargestellt.

Tabelle 4.13: Aufgabenpool für CK und PKC nach der Überarbeitungsphase (blau: CK; rot: PCK).

Block	Inhaltsbereich	Aufgabenanzahl
1	Mechanik	9
2	Elektrizitätslehre	8
3	Optik	8
4	Thermodynamik	8
5	Festkörperphysik	7
6	Atom- und Kernphysik	7
7	Relativitätstheorie	5
8	Quantenmechanik	6
9	Schülerkognition	10
10	Instruktionsstrategien	10
11	Curriculum	7
12	Assessment	7

4.4.6 Lautes-Denken-Studie

Nach der umfassenden Überarbeitung und Verkleinerung des Aufgabenpools, aufgrund der Erkenntnisse aus Pilotierungsstudie, Expertenratings und Fragebogenstudie erschien es angemessen, den Optimierungserfolg, insbesondere bezogen auf die kognitive Validität der Testaufgaben, anhand einer weiteren Studie zu prüfen. Das Ziel in der im Folgenden vorgestellten Untersuchung war es daher, Hinweise zur kognitiven Validität des überarbeiteten Testinstruments zur Erfassung des Fachwissens und des fachdidaktischen Wissens angehender Physiklehrkräfte zu generieren und zu bewerten. Die Untersuchung wurde wieder im Rahmen einer Masterarbeit durchgeführt (Kurdinat, 2013). Zentrale Frage war dabei, inwieweit bei der Bearbeitung der untersuchten Aufgaben intendierte, konstruktrelevante Lösungsstrategien Anwendung finden. Zu den intendierten und konstruktrelevanten Lösungsstrategien zählen solche, die inhaltliche Berührung zu den im Kodiermanual genannten Musterlösungen haben, die also zielführend für die Lösung der jeweiligen Aufgabe sind und die dem Inhaltsbereich der jeweiligen Aufgabe zugeordnet werden können. Ob eine Aufgabe mit der gewählten Lösungsstrategie richtig oder falsch gelöst wird, spielt in diesem Zusammenhang zunächst keine Rolle.

Studiendesign

Um eine repräsentative Auswahl von Aufgaben aus dem zu untersuchenden Testinstrument auszuwählen, wurden per Zufall 24 Aufgaben aus dem nach der Überarbeitungsphase verbliebenen Aufgabenpool gezogen. Die zufällige Auswahl wurde jedoch dadurch eingeschränkt, dass aus jedem Inhaltsbereich zwei Aufgaben gezogen wurden. Dies entspricht einem Anteil von etwa einem Viertel des Aufgabenpools der im Rahmen der Hauptstudie der vorliegenden Arbeit eingesetzt wurde. Die Stichprobe der Studie zum Lauten Denken besteht analog dazu aus Studierenden, die sich an der Universität Kiel in einem Lehramtsstudiengang mit dem Studienziel Bachelor, Master oder Staatsexamen befinden. Aus dieser Zielgruppe wurden $N = 10$ Probanden getestet, die sich in unterschiedlichen Phasen des Studiums befinden, also sowohl Studienanfänger als auch angehende Absolventen umfassen. Das Laute Denken fordert vom Probanden die permanente Verbalisierung seiner Aktivitä-

ten und Gedanken während er sich in einer Testsituation befindet. Da diese Anforderung über die tatsächliche Testsituation hinausgeht und diese Form der Kommunikation ungewohnt ist, wurde für jeden Probanden eine kurze Übungs- und Eingewöhnungsphase in die Technik des Lauten Denkens der eigentlichen Testsituation vorangestellt. So hatten die Probanden jeweils Gelegenheit, sich vor dem Beginn der Testung an das Testformat zu gewöhnen. Ferner wurde die Testdauer auf eine Stunde beschränkt, um Ermüdungseffekte bei den Probanden zu minimieren. Um darüber hinaus die Objektivität der Durchführung zu gewährleisten, handelte der Testleiter nach eigens für die Untersuchung erstellten Richtlinien. Diese legen fest, wie der Testleiter in bestimmten Situationen und auf bestimmte Fragen der Probanden reagiert. Während sich der Proband in der Testsituation befindet und „laut denkt“, finden keine Interventionen durch den Testleiter statt. Die einzige nötige Intervention kann die Aufforderung zum lauten Denken darstellen, falls der Proband für eine bestimmte Zeitspanne, in der vorliegenden Studie für zehn Sekunden, keine Gedanken verbalisiert. Im Interventionsfall gab der Testleiter neutrale Anweisungen an den Probanden, wie zum Beispiel „Sprich weiter!“, „Bitte lies die Aufgabe laut vor!“ oder „Bitte bearbeite die Aufgabe selbstständig!“. Diese Einwurfe waren als direkte Aufforderung formuliert, so dass eine Antwort durch den Probanden nicht erforderlich war. Der Testleiter hielt sich zudem während der Studie nicht im Sichtfeld des Probanden aufhalten. Durch diese Maßnahmen wurde die Beeinflussung des Probanden minimiert und die Authentizität der Testsituation blieb erhalten. Die Handlungen des Probanden wurden während der Durchführung der Studie zwecks späterer Auswertung in Ton und Bild aufgezeichnet.

Auswertung der Studie zum Lauten Denken

Zur Auswertung wurden die Rohdaten der Studie zum Lauten Denken nach der Durchführung der Erhebung in drei Schritten bearbeitet. In einem ersten Schritt werden die aufgezeichneten Audiodaten transkribiert. Dazu werden die aufgenommenen Lautäußerungen der Probanden für die weiteren Analysen schriftlich festgehalten. Unverständliche und abgebrochene Worte und Sätze sowie Pausen, je nach Länge, werden dabei gekennzeichnet. Interventionen des Testleiters und die Zeitpunkte der Lautäußerungen werden dabei ebenfalls erfasst und dokumentiert. Im zweiten Schritt werden die transkribierten Rohdaten in Segmente aufgeteilt, die zunächst allgemeinen Prozessen entsprechen. Diese Segmentierung der Rohdaten hat großen Einfluss auf die Auswertung, da ihr Ergebnis der Identifikation logischer kognitiver Prozesse zugrunde liegt. Prinzipiell existieren drei mögliche Verfahren der Segmentierung. Hierfür hat sich die Unterteilung der einzelnen Segmente durch festgelegte Zeitintervalle, durch die Ideenstruktur der Probanden oder durch Satzenden etabliert (Ericsson und Simon, 1980; 1993). Da eine Unterteilung in zeitliche Intervalle für diese Untersuchung nicht zielführend ist und eine Unterteilung nach Ideen das Problem der Objektivierbarkeit mit sich bringt, wurde für die vorliegende Untersuchung die Segmentierung nach Sätzen gewählt. Dabei werden die Segmente zunächst entsprechend ihrer sprachlichen Struktur in Sätze und Teilsätze unterteilt und anschließend bei Fällen von Sinnzusammenhängen zu den Segmenten zusammengesetzt (Hadenfeldt et al., 2014). Den dritten Schritt bei der Bearbeitung der Rohdaten stellt die Kategorisierung der Segmente dar. Das hier verwendete Kategoriensystem lehnt sich an einen Vorschlag von Deffner (1984) an. Hiernach werden zunächst vier Kategorien von Aussagen unterschieden: (1) Aussagen mit Bezug zur Aufgabenstellung, z.B. „Welche Größe sollte berechnet werden?“, (2) Strategieaussagen, z.B. „Ich muss hier ein Kräftegleichgewicht aufstellen.“, (3) Aussagen zur Testsituation, z.B. „Wieviel Zeit habe ich noch?“ und (4) Aussagen, die keiner der anderen Kategorien zugeordnet werden können. Im Kontext der kognitiven Validierung sind besonders die Strategieaussagen von Interesse. Sie werden in

konstruktrelevante Aussagen, nicht-konstruktrelevante Aussagen und Aussagen, die nicht eindeutig zugeordnet werden können, differenziert.

Ergebnisse der Studie zum Lauten Denken

Die anhand des dargelegten dreischrittigen Verfahrens identifizierten Anteile an konstruktrelevanten, nicht-konstruktrelevanten, bzw. nicht entscheidbaren Aussagen der Probanden sind für die zufällig gezogenen Items in Tabelle 4.14 gezeigt. Die Kodierung der Aufgaben entspricht dabei der Einteilung in die Testblöcke entsprechend den Inhaltsbereichen.

Tabelle 4.14: Lösungsraten und Kategorisierungen für die untersuchten Aufgaben.

Item	Inhaltsbereich	Lösungsrate p	Konstrukt-relevant	Nicht-Konstrukt-relevant	Nicht entscheidbar
1	Mechanik	0,06	0,83	0,04	0,13
2	Mechanik	0,60	0,76	0,16	0,08
3	Elektrizitätslehre	0,73	0,73	0,18	0,09
4	Elektrizitätslehre	0,38	0,58	0,04	0,38
5	Optik	0,35	0,48	0,11	0,41
6	Optik	0,55	0,58	0,15	0,27
7	Thermodynamik	0,55	0,72	0,18	0,10
8	Thermodynamik	0,28	0,56	0,00	0,44
9	Festkörperphysik	0,83	0,63	0,18	0,19
10	Festkörperphysik	0,19	0,56	0,00	0,44
11	Atom- und Kernphysik	0,71	0,92	0,04	0,04
12	Atom- und Kernphysik	0,32	0,78	0,04	0,18
13	Relativitätstheorie	0,63	0,79	0,00	0,21
14	Relativitätstheorie	0,42	0,89	0,02	0,09
15	Quantenmechanik	0,48	0,76	0,02	0,22
16	Quantenmechanik	0,38	0,55	0,08	0,37
17	Schülerkognition	0,31	0,64	0,00	0,36
18	Schülerkognition	0,92	0,96	0,01	0,03
19	Instruktionsstrategien	0,36	0,58	0,40	0,02
20	Instruktionsstrategien	0,38	0,75	0,03	0,22
21	Curriculum	0,86	0,86	0,06	0,08
22	Curriculum	0,62	0,86	0,05	0,09
23	Assessment	0,21	0,74	0,24	0,02
24	Assessment	0,42	0,91	0,06	0,03

Das Verhältnis k zwischen der Anzahl der als konstruktrelevant identifizierten Lösungsstrategien n_k und der Gesamtanzahl der Lösungsstrategien n liefert den Anteil der konstruktrelevanten Lösungsstrategien. Es gelte also: $k = \frac{n_k}{n}$.

Da es sich bei den konstruktrelevanten Lösungsstrategien um die intendierten, inhaltlichen und fachbezogenen Lösungsprozesse handelt und bei den nicht-konstruktrelevanten Lösungsstrategien um nicht fachbezogene, lässt sich durch k insofern ein Maß für die kognitive Validität untersuchter Aufgaben, Skalen oder Testinstrumente definieren. Der Anteil der konstruktrelevanten, nicht-konstruktrelevanten und der nicht entscheidbaren Lösungsstrategien an der Gesamtzahl der Lösungsstrategien für die untersuchten Aufgaben, ist in Abbildung 4.17 dargestellt.

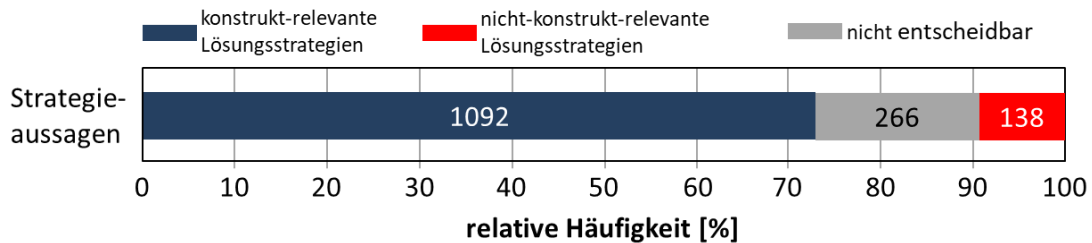


Abbildung 4.17: Gesamtanteil konstruktrelevanter, nicht-konstruktrelevanter und nicht entscheidbarer Anteile der verwendeten Lösungsstrategien (Kurdinat, 2013).

Die erste Frage nach der Anwendung konstruktrelevanter Lösungsstrategien lässt sich somit quantitativ beantworten. Am vorliegenden Beispiel des überarbeiteten Aufgabenpools (siehe Tabelle 4.13) erhält man $n_k = 1092$ und $n = 1496$ und damit: $\kappa = \frac{1092}{1496} \approx 0,73$. Damit erweist sich der konstruktrelevante Anteil der Lösungsstrategien als der Größte, der Anteil der nicht-konstruktrelevanten als der Kleinste, was als Indiz für die kognitive Validität des verwendeten Instruments gewertet werden kann.

Trägt man die relativen Häufigkeiten konstruktrelevanter, nicht entscheidbarer und nicht-konstruktrelevanter Lösungsstrategien für die einzelnen getesteten Aufgaben auf, so ergibt sich das in Abbildung 4.18 dargestellte Diagramm. Dabei zeigt sich, dass bis auf eine Aufgabe jeweils der Anteil konstruktrelevanter Lösungsstrategien oberhalb von 50% liegt. Für die eine Aufgabe mit einem Anteil konstruktrelevanter Lösungsstrategien unterhalb von 50% ist der Anteil der nicht entscheidbaren Lösungsstrategien besonders hoch; der Anteil der nicht-konstruktrelevanten Lösungsstrategien mithin klein. Der Durchschnitt konstruktrelevanter Lösungsstrategien liegt oberhalb von 73%, wohingegen der relative Anteil nicht-konstruktrelevanter Lösungsstrategien durchschnittlich unterhalb von 10% liegt. Für einzelne Aufgaben, zum Beispiel die Aufgaben mit den Nummern 19 und 23, sind jedoch deutliche Abweichungen von diesen Durchschnittswerten zu erkennen.

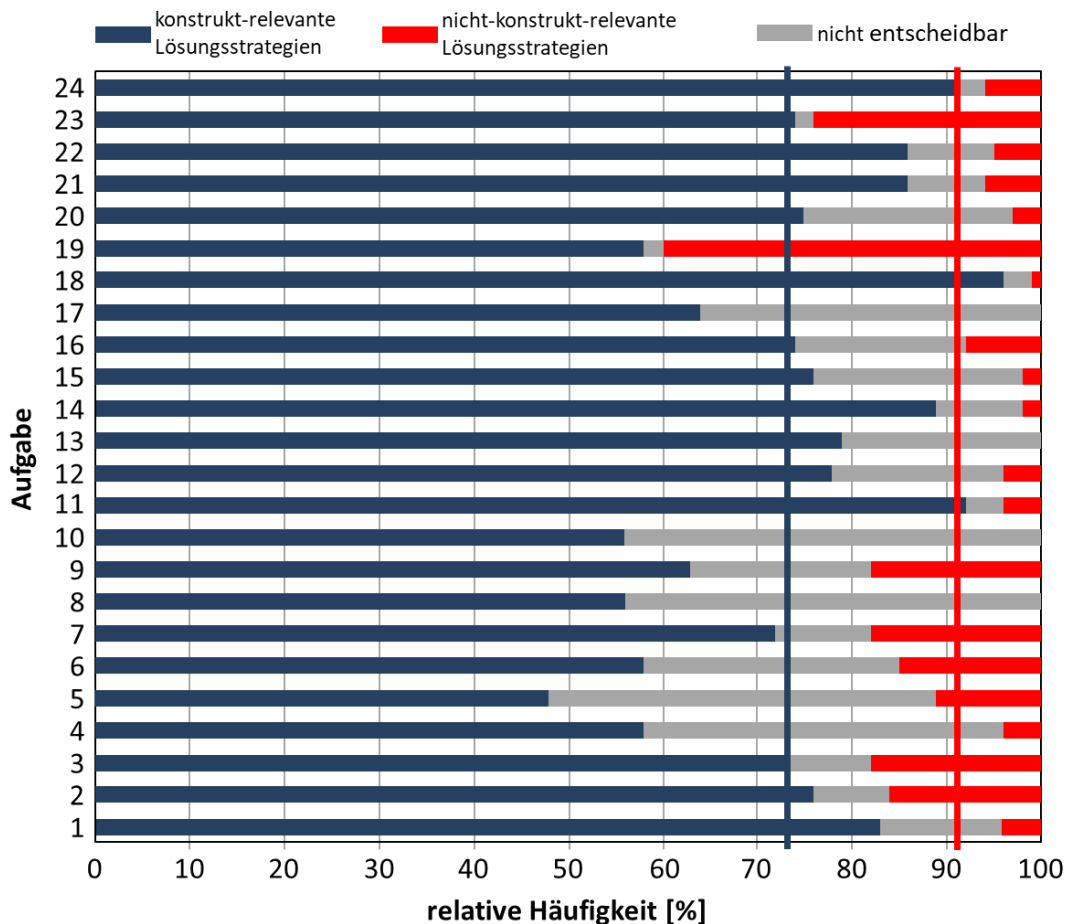


Abbildung 4.18: Anteile konstruktrelevanter, nicht-konstruktrelevanter und nicht entscheidbarer Anteile der verwendeten Lösungsstrategien bei der Lösung der Aufgaben (Kurdiat, 2013).

Bezüglich der Verteilung der Lösungsstrategien können zwei Feststellungen getroffen werden: Erstens, dass sich der Anteil nicht entscheidbarer Lösungsstrategien bei offenen Aufgaben als geringer erwies als bei geschlossenen Aufgaben. Zweitens, dass der Anteil nicht-konstruktrelevanter Lösungsstrategien sich im vorliegenden Fall bei den Inhaltsbereichen *Optik*, *Thermodynamik* und *Festkörperphysik* als überdurchschnittlich hoch erwies.

Berücksichtigt man noch die tatsächlich Lösungsraten der Probanden auf den Items, also die durchschnittliche Wahrscheinlichkeit mit der ein Item richtig gelöst wurde, so erhält man einen signifikanten Zusammenhang zwischen der Lösungsrate und dem Anteil konstruktrelevanter Lösungsstrategien. Die Korrelation beträgt $r = .406^*$ ($p < .05$). Dies ist insofern erwartungskonform, als dass man annehmen kann, dass die Anwendung konstruktrelevanter Lösungsstrategien eine höhere Wahrscheinlichkeit mit sich bringt, eine Aufgabe korrekt zu lösen.

Zusammenfassung zur Lautes-Denken-Studie

In der dargelegten Studie wurde die kognitive Validierung eines Paper-und-Pencil-Testinstruments zur Erfassung des Fachwissens und des fachdidaktischen Wissens angehender Physiklehrkräfte vorgestellt. Dazu wurde die Testmethode des Lauten Denkens gewählt. Mittels des gewählten Auswertverfahrens wurde gezeigt, dass diese Methode es erlaubt, quantitative Werte für die kognitive

Validität zu liefern. Die Bestimmung solcher Werte für einen Teilaspekt der Validität erleichtert die Sicherstellung dieses Testgütekriteriums und ermöglicht gegebenenfalls einen Vergleich verschiedener Testinstrumente.

Die Ergebnisse zeigen, dass die Probanden beim Bearbeiten der Testaufgaben überwiegend konstruktrelevante, also fachbezogene Lösungsstrategien verwenden. Die konstruktrelevanten Lösungsstrategien überwiegen dabei mit einem Anteil von 73% gegenüber 18% nicht entscheidbarer Anteile und 9% nicht-konstruktrelevanter Anteile. Für eine wertende Beurteilung dieser Ergebnisse mangelt es an einer Referenz. Die Beurteilung der erhaltenen Ergebnisse bleibt daher problematisch, solange keine Referenzen zu anderen, inhaltlich ähnlich angelegten Testinstrumenten vorliegen. Man kann jedoch feststellen, dass die Anwendung konstruktrelevanter Lösungsstrategien sowohl im Durchschnitt als auch für jede einzelne der getesteten Aufgaben deutlich überwiegt. Zur Validierung des vorgestellten Verfahrens sollten daher weitere Untersuchungen vorgenommen werden, um dessen Reproduzierbarkeit und Objektivität zu prüfen und um Referenzen für die Interpretation der Ergebnisse zu generieren.

4.4.7 Hauptstudie

Studiendesign

Das Studiendesign für die Hauptstudie der vorliegenden Untersuchung ähnelt dem der Pilotierungsstudie. Auch für die Hauptstudie wurde eine bundesweite Erhebung des Professionswissens an zwölf Standorten mittels Paper-und-Pencil-Test durchgeführt. Bei der Auswahl der Standorte handelte es sich wieder um solche, denen für die Ausbildung angehender Lehrkräfte im bundesvergleich eine bedeutende Rolle zukommt. Die Hauptstudie wurde dabei, wie schon die Pilotierungsstudie als Querschnittuntersuchung durchgeführt. Dementsprechend konnten Lehramtsstudierende mit dem Fach Physik, unabhängig von ihrem jeweiligen Studiensemester, an der Hauptstudie teilnehmen.

Unterschiede zur Pilotierungsstudie gibt es im Hinblick auf die anvisierte Probandenzahl und – daraus resultierend – auf das Testheftformats. Ziel der Hauptstudie waren die Rekrutierung von $n \approx 250$ Probanden pro Item. Um dieser Zahl möglichst nahe zu kommen, wurde die Aufgaben des entwickelten und validierten Aufgabenpools nicht mehr, wie noch in der Pilotierungsstudie, auf zwei Testhefte aufgeteilt. Stattdessen bearbeiteten alle teilnehmenden Probanden alle überarbeiteten Aufgaben. Die Verkleinerung des Aufgabenpools im Rahmen des Validierungsprozesses erlaubt diese Vorgehensweise, ohne dabei Testqualitätsmerkmale wie Testdauer und Zumutbarkeit für die Probanden zu verletzen.

Testinstrument

Nach der Zusammenstellung und Überarbeitung des Aufgabenpools wurden, anhand der Resultate der in den Abschnitt 4.4.1 bis 4.4.6 dargelegten Entwicklungsschritte, die geeignetsten Aufgaben zu einem Testinstrument für die Datenerhebung der Hauptstudie zusammengestellt.

Das Testheft erfasst dabei zunächst Angaben zur Person und Studiensituation der Probanden. Als Indikator für kognitive Fähigkeiten wird beispielsweise die Abiturnote erfasst, aber auch die letzten Noten im Fach Physik als Indikator für schulisches Vorwissen. Daneben werden noch Variablen für die Lehr- und Lerngelegenheiten der Probanden im Studium, fachspezifische Skalen zu Selbstwirksam-

keit, Motivation, Interesse, Studienzufriedenheit und Wissen über Nature of Science erfasst. Es folgen drei Abschnitte zur Erfassung des Professionswissens, die von den im Theorieteil beschriebenen Bereichen Fachwissen, fachdidaktisches Wissen und pädagogisches Wissen gebildet werden. Die Bereiche zum Fachwissen und zum fachdidaktischen Wissen sind wiederum, analog zur Vorgehensweise bei der Pilotierungsstudie (Abschnitt 4.4.2), in Aufgabenblöcke zu spezifischen Inhaltsbereichen gegliedert. Tabelle 4.15 zeigt die Organisation der Aufgaben zur Erfassung des Fachwissens und des fachdidaktischen Wissens.

Tabelle 4.15: Testaufgaben zur Erfassung des Fachwissens und des fachdidaktischen Wissens für die Hauptstudie (blau: CK; rot: PCK).

Block	Inhaltsbereich	Aufgabenanzahl
1	Mechanik	9
2	Elektrizitätslehre	8
3	Optik	8
4	Thermodynamik	8
5	Festkörperphysik	7
6	Atom- und Kernphysik	7
7	Relativitätstheorie	5
8	Quantenmechanik	6
9	Schülerkognition	10
10	Instruktionsstrategien	10
11	Curriculum	7
12	Assessment	7

Die Aufgaben sind innerhalb der Blöcke jeweils so gewählt, dass der jeweilige Inhaltsbereich möglichst breit erfasst wird. Am Beispiel des Inhaltsbereichs Mechanik bedeutet dies, dass die neun eingesetzten Aufgaben das Wissen der Probanden zu gänzlich unterschiedlichen Aspekten der Mechanik erfassen. Um dies zu verdeutlichen, sind die unterschiedlichen Aufgabeninhalte innerhalb des Bereiches Mechanik in der nachfolgenden Tabelle 4.16 dargestellt.

Tabelle 4.16: Testaufgaben zur Erfassung des Fachwissens zum Inhaltsbereich Mechanik.

Aufgabe	Thema
1	Gravitationsgesetz
2	3. Keplersches Gesetz
3	Trägheitsmoment
4	1. Newtonsches Axiom
5	Kinetische Energie
6	Energieerhaltung/Umwandlung
7	Geschwindigkeits-Zeit-Beziehung
8	Drehmoment
9	Schwingung/Pendel

Bezüglich des von Gramzow (2015) beschriebenen Reliabilitäts-Validitäts-Dilemmas liegt also der Fokus auf einer validen Erfassung des jeweiligen Bereichs, was – im Hinblick auf die Subskalen – zu einer eingeschränkten Reliabilität führen kann.

Im Folgenden wird die erzielte Stichprobe, bei der dieses Testinstruments im Rahmen der Hauptstudie der vorliegenden Arbeit zum Einsatz kam, vorgestellt.

Stichprobe

Im Rahmen der Hauptstudie wurden die Daten von 201 Probanden an insgesamt zwölf Universitätsstandorten erhoben: Bielefeld, Bremen, Frankfurt a.M., Freiburg, Hamburg, Hannover, Kassel, Kiel, München, Oldenburg, Regensburg und Rostock (siehe Abbildung 4.19).

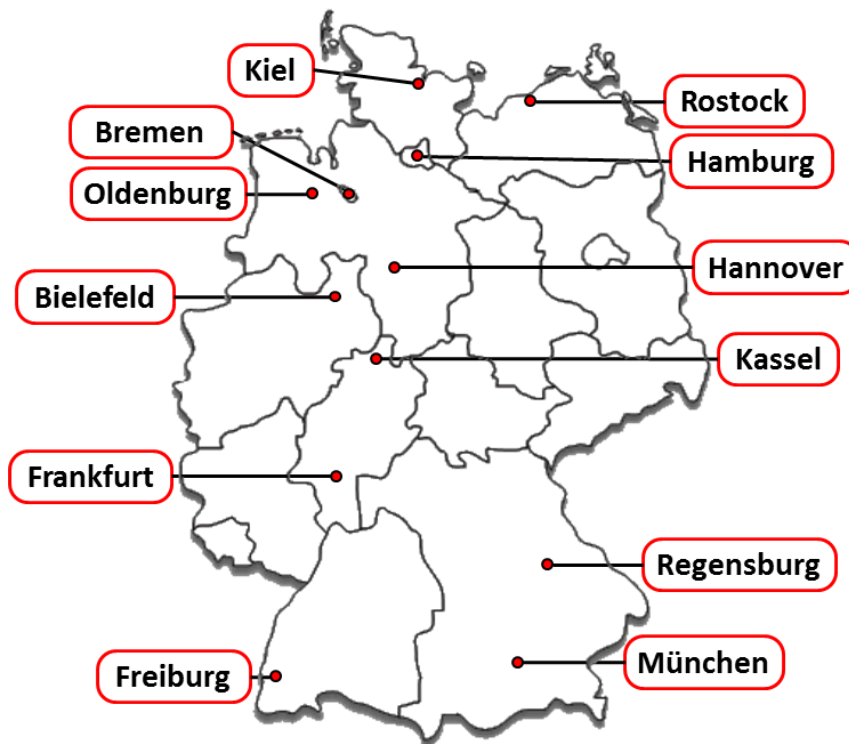


Abbildung 4.19: Erhebungsstandorte für die Hauptstudie.

Die Verteilung der Probanden auf die einzelnen Standorte ist in der folgenden Tabelle 4.17 dargestellt.

Tabelle 4.17: Verteilung der Probanden auf die einzelnen Erhebungsstandorte.

Standort	Probandenzahl
Bielefeld	19
Bremen	10
Freiburg	13
Frankfurt a.M.	8
Hamburg	10
Hannover	13
Kassel	21
Kiel	29
München	2
Oldenburg	4
Regensburg	41
Rostock	31
Gesamt	201

Bei den Probanden handelt es sich um Lehramtsstudierende mit Hauptfach Physik. Das Studienziel war bei 58,2% der Probanden das Staatsexamen, bei 28,9% der Bachelor of Education und bei 11,4% der Master of Education. 40,8 % der Probanden waren weiblich. Das durchschnittliche Alter der Probanden betrug 23,5 Jahre bei einer Standardabweichung von 3,3 Jahren, die durchschnittliche Semesterzahl lag bei 5,6 bei einer Standardabweichung von 2,7. Für die Betrachtung der Studienphasen der Probanden erscheint die Betrachtung der Semesterzahl jedoch wenig sinnvoll, da deutlich mehr Studierende ein Studium zu Beginn eines Wintersemesters aufnehmen. Daraus resultiert ein deutliches Ungleichgewicht zwischen den Probandenzahlen gerader und ungerader Studiensemester. Für die Analysen wird daher die in Abbildung 4.20 dargestellte die Verteilung der Probanden nach Studienjahr verwendet. Hier ergibt sich in der Stichprobe ein Mittelwert von 2,87 bei einer Standardabweichung von 1,3.

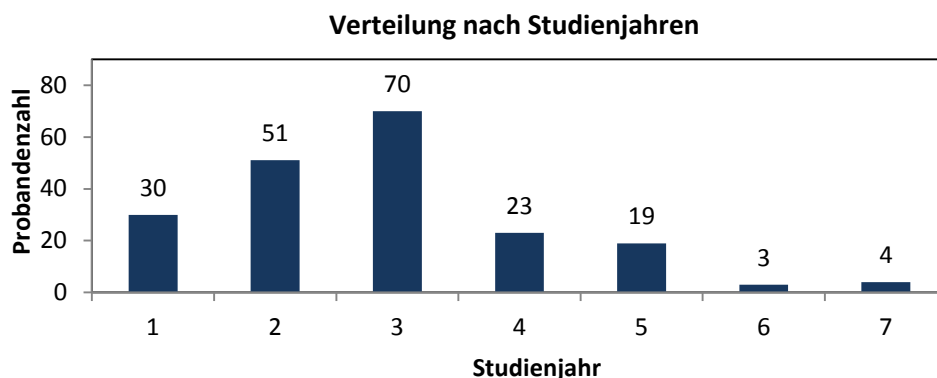


Abbildung 4.20: Verteilung der Probanden der Hauptstudie nach Studienjahren.

Bei 187 von 201 Probanden der Hauptstudie konnte erfasst werden, ob diese bereits an der Pilotierungsstudie des KiL-Projekts (siehe Abschnitt 4.4.2) teilgenommen hatten. Dies traf auf 42 von 187 zu, was einem Anteil von 22,5% entspricht.

Analyse

Für die Analysen wurde die Statistik- und Analysesoftware SPSS von IBM (IBM, 2011) und die Programmiersprache R für statistisches Rechnen (Hain, 2011; R Core Team, 2015) verwendet. Für die Bestimmung von Effektstärken wurde das Online-Instrument Psychometrika genutzt (Lenhard und Lenhard, 2015). Zudem wurden auch raschskalierte Daten genutzt, die mit der Software Winsteps (Linacre, 2015) generiert wurden.

Als erster Schritt der Datenanalyse wurde untersucht, ob die neuentwickelten Messinstrumente zur Erfassung des Fachwissens und des fachdidaktischen Wissens erwartungskonform funktioniert haben. Dazu wurden klassische Reliabilitätsanalysen der Haupt- und Subskalen angestellt. Als Maß für die Reliabilität der einzelnen Skalen wurde wiederum Cronbachs α herangezogen (Cronbach, 1951; Moosbrugger und Kelava, 2008). Auf der Ebene einzelner Items wurde die Lösungsrate, also der Anteil korrekter Lösungswege von der Gesamtzahl der Bearbeitungen und die Trennschärfe, also der korrelative Zusammenhang zwischen den Testwerten eines Items mit den Testwerten der zugehörigen Gesamtskala, bestimmt. Die Trennschärfen wurden dabei auf ihre Testgüte hin interpretiert. Allgemein gelten Items mit einer Trennschärfe von .4 bis .7 als „gute“ Trennschärfen (Moosbrugger und Keleva, 2008). Aufgrund des zuvor beschriebenen Reliabilitäts-Validitäts-Dilemmas (siehe Abschnitt 4.3.5) und der damit einhergehenden relativ niedrigen Inter-Item-Korrelationen müssen für den vorliegenden Test weniger strenge Maßstäbe angelegt werden. Items mit einer niedrigen, positiven Trennschärfe, z.B. nahe bei Null, erfassen Merkmalsausprägungen, die nahezu unabhängig von der übrigen Subskala sind. Jedoch kann dies nicht zwangsläufig als kritisches Itemmerkmal aufgefasst werden, da die Items bewusst so konstruiert wurden, dass sie Wissen über disjunkte Bereiche innerhalb eines Fachinhaltsbereichs erfassen. Wer beispielsweise die Keplerschen Gesetze kennt, muss nicht zwangsläufig auch über Wissen zu Trägheitsmomenten verfügen, auch wenn beide Inhalte grundsätzlich der Subskala Mechanik zugeordnet werden. Aus diesem Grunde wurden nachfolgend nur Items mit einer negativen Trennschärfe aus der weiteren Analyse ausgeschlossen. Dies kann nach Moosbrugger und Kelava (2008) durch Mängel in der Instruktion oder der Item-Formulierung bedingt sein. Solche Items sollten nicht genutzt werden, sofern eine keine sinnvolle Umkodierung möglich ist, was im vorliegenden Fall zutrifft. Nach diesem Kriterium wurde ein Item zum Fachwissen aus der Subskala „Relativitätstheorie“ aufgrund negativer Trennschärfe aus der Skala eliminiert. Eine negative Trennschärfe eines Items drückt aus, dass sich für die Lösungsrate des Items und die Gesamtskala eine negative Korrelation ergibt. Das heißt, dass ein Proband das Item mit höherer Wahrscheinlichkeit lösen kann, wenn er schlechter im Gesamttest abgeschnitten hat. Da die Ergebnisse für dieses Item damit im Sinne der Validität als unplausibel angenommen werden können, wird das Item für die weiteren Analysen nicht mehr berücksichtigt. Neben den Reliabilitätsanalysen werden durch die Untersuchung von Korrelationen des Testscores mit Begleitvariablen Hinweise auf die Validität der Messergebnisse analysiert.

Zur Beantwortung der ersten Forschungsfrage wird die Struktur des Professionswissens untersucht. Hierfür werden Strukturgleichungsmodelle mittels der Programmiersprache R (Hain, 2011) verglichen. Benutzt wurde hierfür das R-Paket Lavaan (Rosseel et al., 2013). Im Zuge der Auswertung wird zunächst die am besten passende Anzahl der Dimensionen für das Konstrukt des Professionswissens gesucht. Da theoretische Annahmen über die Struktur des Professionswissens existieren (siehe Abschnitt 2.1), wird hier die Methode der konfirmatorischen Faktorenanalyse angewendet. Der Unterschied zu einer explorativen Faktorenanalyse besteht in einer vorausgehenden Annahme über die Anzahl der Faktoren. Das Faktorenmodell wird also vor der Analyse spezifiziert und anschließend auf

seine Passung hin überprüft (Bühner, 2006). Wegen der geringen Stichprobengröße ($N = 201$) wird dabei auf eine Modellierung auf Basis von Einzelitems verzichtet. Stattdessen werden Scores der Subskalen gebildet und mit diesen die Modellierungen gerechnet. Die Subskalen für die einzelnen Wissensbereiche des Professionswissens, die für die Strukturprüfungen herangezogen werden, sind in Tabelle 4.18 dargestellt.

Tabelle 4.18: Subskalen der Wissensbereiche des Professionswissens.

Fachwissen	fachdidaktisches Wissen	pädagogisches Wissen
Mechanik	Schülerkognition	Lehren und Lernen
Elektrizitätslehre	Instruktionsstrategien	Lernen und Entwicklung
Optik	Curriculum	Klassenführung
Thermodynamik	Assessment	Leistungsbeurteilung
Festkörperphysik		
Atom- und Kernphysik		
Relativitätstheorie		
Quantenmechanik		

Als Parameter für die Modellpassung werden Comparative-Fit-Index (CFI), Tucker-Lewis-Index (TLI), Standardized Root Mean Square of Residuals (SRMR), Root Mean Square Error of Approximation (RMSEA), Akaike Information Criterion (AIC), und Bayesian Information Criterion (BIC) berechnet, verglichen und interpretiert.

Zur Beantwortung der zweiten Forschungsfrage wird die Entwicklung des Professionswissens anhand von Testscores und raschskalierten Testwerten analysiert. Die Analysen der Testscores wurden mit der Statistiksoftware SPSS (IBM, 2011) und die Raschanalyse mittels Winsteps (Linacre, 2015) durchgeführt. Dazu wird die durchschnittliche Ausprägung der Scores für Fachwissen und fachdidaktisches Wissen jeweils über die ersten fünf Studienjahre hinweg aufgetragen. Die Ausprägungen des Fachwissens und des fachdidaktischen Wissens wurden für die Raschanalyse von 200 bis 1200 skaliert. Mit dieser Skalierung soll sichergestellt werden, dass bei die Probandenwerte nicht den Wert Null erreichen können, woraus der falsche Schluss gezogen werden könnte, es gäbe Probanden die über kein Fachwissen oder fachdidaktisches Wissen verfügen. Auf diese Weise sollen, dem Validitätsaspekt der Konsequenzen nach Messick (1995) folgend, eventuelle individuelle negative Konsequenzen für Studienteilnehmer vermieden werden.

Die Ergebnisse der Probanden aus den Studienjahren 6 und 7 werden aufgrund der geringen Stichprobengröße für diesen Studienabschnitt nicht mehr für die Auswertung berücksichtigt. Die Analyse nach Rasch (1960) gehört Item-Response-Theorie (IRT) und ist als Ergänzung der klassischen Analysen anzusehen (Moosbrugger und Kelava, 2008). Sie ermöglicht es, sowohl die Fähigkeit von Probanden unterschiedlicher Populationen als auch Itemschwierigkeiten anhand derselben Skalierung direkt miteinander zu vergleichen (Neumann, Neumann und Nehm, 2011). Mit dem Online-Instrument Psychometrika (Lenhard und Lenhard, 2015), werden anschließend Effektstärken für den Zuwachs des Fachwissens und des fachdidaktischen Wissens zwischen Anfangs- und Endphase des Lehramtsstudiums berechnet.

Für die Beantwortung der dritten Forschungsfrage werden potentielle Prädiktoren des Professionswissens untersucht. Hierfür werden im Rahmen der Testung für sämtliche Probanden per Fragebogen Indikatoren wie die Lehr- und Lerngelegenheiten, kognitive Fähigkeiten, Schulwissen im Fach Physik, Wissen über Nature of Science, fachbezogene Selbstwirksamkeit und Selbsteinschätzung, Motivation, Interesse und Studienzufriedenheit erhoben. Für diese Personenmerkmale werden mittels R Korrelationen nach Spearman und Regressionen mit den Testscores berechnet, um Hinweise für mögliche Zusammenhänge zu generieren.

5 Ergebnisse

Nachfolgend wird anhand von Ergebnissen zu den wichtigsten Testgütekriterien, Reliabilität und Validität, geprüft, ob das eingesetzte Testinstrument den gestellten Anforderungen an eine hohe Testgüte gerecht wird. Erst anschließend werden Ergebnisse zur Struktur- und Entwicklung des Professionswissens, sowie zu Zusammenhängen des Professionswissens mit erhobenen Personenmerkmalen berichtet, um die in Kapitel 4 aufgeworfenen Forschungsfragen beantworten zu können.

5.1 Kennwerte des Messinstruments

5.1.1 Reliabilität

Im Folgenden werden zunächst die Reliabilitäten der neuentwickelten Hauptskalen für Fachwissen und fachdidaktisches Wissen anhand der Bemaßung durch Cronbachs α betrachtet. Diese sind in Tabelle 5.1 dargestellt.

Tabelle 5.1: Reliabilitäten für Fachwissen und fachdidaktisches Wissen in der Hauptstudie (blau: CK; rot: PCK).

	CK	PCK
Cronbachs α	.85	.74
mittlere Lösungswahrscheinlichkeit p	.50	.51
Itemzahl n	58	36

Die neuentwickelten Hauptskalen zeigen demnach gegenüber der Pilotierungsstudie anhand der Bemaßung durch Cronbachs α deutlich bessere Werte (vgl. Abschnitt 4.4.2). Die Reliabilität der Fachwissensskala kann als gut, die für das fachdidaktische Wissen als akzeptabel betrachtet werden (Bühner, 2006; Moosbrugger und Kelava, 2008). Die Aufgabenschwierigkeiten sind mit einer mittleren Lösungswahrscheinlichkeit von 50% beim Fachwissen und 51% beim fachdidaktischen Wissen sehr gut an die Fähigkeiten der Zielgruppe angepasst. Die Reliabilitäten der Hauptskalen sind für die angestrebten Analysen hinreichend hoch.

Da eine breite Erfassung über verschiedene Inhaltsbereiche hinweg das Ziel der Untersuchung war, werden im Folgenden die Skalenkennwerte für die Subskalen des Fachwissens und des fachdidaktischen Wissens betrachtet. In den Tabellen 5.2 und 5.3 sind die Itemzahl N , die mittlere Lösungswahrscheinlichkeit p , deren Standardabweichung s , die minimale Lösungswahrscheinlichkeit p_{min} auf den Subskalen, die maximale Lösungswahrscheinlichkeit p_{max} und Cronbachs α für die jeweilige Subskala für das Fachwissen (Tabelle 5.2) und das fachdidaktische Wissen (Tabelle 5.3) dargestellt.

Tabelle 5.2: Skalenkennwerte für das Fachwissen aus der Hauptstudie.

Fachwissen (CK)	N	p	s	p_{Min}	p_{Max}	α
Mechanik	9	.53	.27	.07	.86	.37
Elektrizitätslehre	8	.50	.24	.20	.86	.37
Optik	8	.42	.17	.15	.67	.59
Thermodynamik	8	.52	.21	.23	.85	.53
Festkörperphysik	7	.50	.22	.19	.83	.50
Atom- und Kernphysik	7	.50	.18	.23	.71	.43
Relativitätstheorie	4	.48	.22	.20	.67	.31
Quantenmechanik	6	.40	.12	.23	.62	.55

Die Werte zeigen, dass die Lösungswahrscheinlichkeit als Maß für die Aufgabenschwierigkeit für das Fachwissen durchweg als angemessen angesehen werden kann, da p jeweils zwischen 40% und 60% liegt, wobei 50% als optimale Lösungswahrscheinlichkeit angestrebt waren. Dabei zeigt sich, dass sich die schwierigsten Aufgaben zum Fachwissen im Bereich Quantenmechanik mit einer mittleren Lösungswahrscheinlichkeit von 40% befanden. Die maximale Lösungswahrscheinlichkeit ist in diesem Bereich mit 62% besonders gering, was sich auch darin widerspiegelt, dass selbst die leichteste Aufgabe in diesem Bereich von 38% der Probanden nicht gelöst werden konnte. Die übrigen maximalen Lösungsraten liegen im Bereich von 80%, was angemessen erscheint, um auch innerhalb der Subpopulation der Probanden mit schwachem Testergebnis differenzieren zu können. Die leichtesten Items lagen mit einer Lösungswahrscheinlichkeit von 53% im Bereich Mechanik, obwohl sich hier das schwerste Item mit einer minimalen Lösungswahrscheinlichkeit von 7% befand. Da die übrigen minimalen Lösungswahrscheinlichkeiten über alle Inhaltsbereiche hinweg im Bereich von 20% liegen, kann davon ausgegangen werden, dass – analog zu den Probanden mit schwachem Testergebnis – genügend schwierige Items vorhanden sind, so dass auch innerhalb der Subpopulation der Probanden mit starkem Testergebnis noch differenziert werden kann.

Die Reliabilitäten der Subskalen liegen durchweg unterhalb von .60. Dies kann durch das vorab beschriebene Reliabilitäts-Validitäts-Dilemma (siehe Abschnitt 4.3.5) begründet werden. Dennoch sind die Reliabilitäten bei den Subskalen zum Fachwissen damit nicht hinreichend hoch, um auf deren Basis belastbare Einzelanalysen anstellen zu können. Aus diesem Grund wird die Untersuchung über die Entwicklung des Professionswissens im Studienverlauf auf die Ergebnisse der Hauptskalen fokussieren, wenngleich die Ergebnisse auf den Subskalen ebenfalls berichtet werden. Auf tiefergehende Analysen der Subskalen wird aufgrund der geringen Reliabilität jedoch verzichtet.

Tabelle 5.3: Skalenkennwerte für das fachdidaktische Wissen aus der Hauptstudie.

Fachdidaktisches Wissen	N	p	s	p_{Min}	p_{Max}	α
Schülerkognition	10	.60	.21	.26	.92	.45
Instruktionsstrategien	10	.46	.14	.23	.65	.51
Curriculum	7	.45	.27	.10	.86	.39
Assessment	7	.37	.14	.21	.63	.45

Beim fachdidaktischen Wissen ergeben sich ähnliche Skalenkennwerte wie beim Fachwissen. Die Aufgabenschwierigkeiten sind über die Subskalen hinweg etwas weniger homogen, aber noch akzeptabel, da sie im Bereich der angestrebten 50% liegen. Die höchste mittlere Lösungswahrscheinlichkeit ergab sich hier im Aufgabenblock zum Assessment mit einer Lösungswahrscheinlichkeit von 37%, die leichtesten im Aufgabenblock zur Schülerkognition mit einer Lösungswahrscheinlichkeit von 60%. Die minimalen Lösungsraten der einzelnen fachdidaktischen Inhaltsbereiche sind mit Werten von 10% bis 27% im akzeptablen Bereich. Die maximalen Lösungsraten bei den Inhaltsbereichen Instruktionsstrategien bzw. Assessment liegen mit 65% bzw. 63% etwas niedriger als angestrebt, was auf eine schlechte Differenzierungsfähigkeit des Testinstruments für leistungsschwache Testpersonen hindeutet. Die geringe Standardabweichung von jeweils .14 in den beiden Bereichen kann ebenfalls als Hinweis auf eine schlechte Differenzierungsfähigkeit des Tests im oberen und unteren Leistungsspektrum aufgefasst werden.

Die Reliabilitäten der Subskalen zum fachdidaktischen Wissen sind, wie beim Fachwissen, für belastbare Einzelanalysen zu gering, da die Werte auch hier jeweils unterhalb von .60 liegen. Daher gilt auch hier, dass die Untersuchung über die Entwicklung des Professionswissens im Studienverlauf auf die Ergebnisse der Hauptskalen fokussieren wird, wenngleich die Ergebnisse auf den Subskalen ebenfalls berichtet werden. Unter der Annahme, dass die mittlere Inter-Item-Korrelation \bar{r} bei der Veränderung der Itemzahl N konstant bleibt, lässt sich der Wert für Cronbachs α (zur Erinnerung: Cronbachs $\alpha = \frac{N \cdot \bar{r}}{1 + (N-1) \cdot \bar{r}}$) für beliebige Itemzahlen auf einer Skala vorhersagen. Zur Verdeutlichung wird in der folgenden Tabelle 5.4 in dieser Weise gezeigt, dass etwa 20 Items pro Subskala bei der inhaltlichen Breite der Erfassung in der vorliegenden Arbeit für das Erreichen einer akzeptablen Reliabilität ausreichen.

Tabelle 5.4: Hochgerechnete Reliabilitäten bei der Verwendung von $N = 20$ Items pro Subskala für CK und PCK (blau: CK; rot: PCK).

Fachwissen (CK)	α_{20}
Mechanik	.56
Elektrizitätslehre	.59
Optik	.79
Thermodynamik	.73
Festkörperphysik	.75
Atom- und Kernphysik	.68
Relativitätstheorie	.68
Quantenmechanik	.78
Fachdidaktisches Wissen (PCK)	
Schülerkognition	.61
Instruktionsstrategien	.68
Curriculum	.71
Assessment	.69

5.1.2 Validität

Wie im vorangegangenen Abschnitt gezeigt, kann von einer guten Reliabilität für die Fachwissensskala und einer akzeptablen Reliabilität der Skala für das fachdidaktische Wissen ausgegangen werden. Eine hinreichende Reliabilität der eingesetzten Skalen ist ein notwendiges Kriterium für die Validität der Testergebnisse, insofern, als dass man davon ausgehen kann, dass die eingesetzten Hauptskalen zwei in sich konsistente Konstrukte erfassen. Ob es sich bei diesen Konstrukten um die intendierten Konstrukte handelt, also das Fachwissen und das fachdidaktische Wissen angehender Physiklehrkräfte, ist eine Frage der Validität der Testergebnisse, die im Folgenden behandelt wird.

Die Validierung der Messungen mittels der neuentwickelten Messinstrumente für das Fachwissen und das fachdidaktische Wissen stellt einen Schwerpunkt der vorliegenden Arbeit dar. Daher wird der besonders stark differenzierende Begriff der Konstruktvalidität von Messick (1995; vgl. Moosbrugger und Kelava, 2008) herangezogen, bei dem die Konstruktvalidität anhand von sechs Aspekten betrachtet wird. Diese sechs Aspekte werden im Folgenden diskutiert und anhand von Testergebnissen bzw. Begleitumständen der Testdurchführung beurteilt.

(a) Inhaltsvalidität

Der Begriff der Inhaltsvalidität bezieht sich auf die inhaltliche Passung zwischen den Items des Testinstruments und dem zu erfassenden Konstrukt (Messick, 1995; Moosbrugger und Kelava, 2008). Die inhaltliche Validität des Testinstruments sollte durch den Prozess der Aufgabenkonstruktion anhand des im Abschnitt 4.3 entwickelten Modells des Professionswissens und durch das im Abschnitt 4.4.3 dargestellte Expertenrating sichergestellt werden. Dabei repräsentiert das Aufgabenentwicklungsmodell die im Abschnitt 2.2.1 für das Fachwissen und im Abschnitt 2.2.2 für das fachdidaktische Wissen recherchierten Lernziele für die Ausbildung angehender Physiklehrkräfte. Durch die Entwicklung anhand des Modells wurde so eine Passung zwischen den Inhalten der Messinstrumente (siehe Abschnitt 4.4.7) und den für angehende Lehrkräfte relevanten Inhalten (siehe Abschnitt 2.2) hergestellt. Für die in der Hauptstudie eingesetzten Items wurde diese Passung durch das Expertenrating bestätigt.

(b) Kognitive Validität

Für die Herstellung und Überprüfung der kognitiven Validität wird an dieser Stelle auf die Ergebnisse der Fragebogenstudie (Abschnitt 4.4.4) und die Ergebnisse der Studie zum Lauten-Denken (Abschnitt 4.4.6) verwiesen. Erstere lieferte Anhaltspunkte zur Verbesserung der kognitiven Validität. Letztere zeigt die überwiegende Anwendung konstruktrelevanter Denkprozesse, die bei der Bearbeitung der Items in der Hauptstudie durch die Probanden zur Anwendung kommen.

(c) Strukturelle Validität

Zur Prüfung der strukturellen Validität muss untersucht werden, ob die Testergebnisse die Ausprägung des zu messenden Konstrukts repräsentieren. Hierfür muss das Testinstrument in der Lage sein, zwischen Probanden unterschiedlicher Wissensausprägung zu differenzieren. Aus den im vorangegangenen Abschnitt berichteten Lösungsraten (siehe Tabellen 5.2 und 5.3) ergeben sich eine mittlere Lösungsrate von jeweils 48% beim Fachwissen und beim fachdidaktischen Wissen. Dies sind erste Indikatoren für eine angemessene Aufgabenschwierigkeit in den Bereichen CK und PCK. Sie stellen

eine notwendige Voraussetzung für die Fähigkeit zur Differenzierung zwischen den Probanden dar. Da für die Verteilung der Probandenfähigkeit bzw. der Wissensausprägung eine Normalverteilung angenommen werden kann, sollten die Items der Testinstrumente für Fachwissen und fachdidaktisches Wissen bezüglich ihrer Schwierigkeit diese Normalverteilung im besten Fall widerspiegeln. Die einfache Betrachtung der mittleren Aufgabenschwierigkeiten ist hierfür nicht hinreichend, denn es muss gefordert werden, dass viele Items mit mittlerer Aufgabenschwierigkeit und wenige Items mit hoher und niedriger Aufgabenschwierigkeit innerhalb der Probandenfähigkeit existieren. Um zu prüfen, ob Items mit hoher und niedriger Aufgabenschwierigkeit existieren können innerhalb der Aufgabenblöcke die minimalen und maximalen Lösungsraten betrachtet werden.

Hier ergibt sich für den Aufgabenblock zur Mechanik eine minimale Lösungsrate von .07, d.h., dass die schwierigste Aufgabe zur Mechanik mit einer Rate von 7% von den Probanden gelöst wurde. Die einfachste Aufgabe wurde mit einer Rate von 86% von den Probanden gelöst. Die Forderung nach der Existenz von hohen und niedrigen Aufgabenschwierigkeiten ist damit erfüllt. Betrachtet man zusätzlich die Standardabweichung von .27 bei einem Mittelwert von 53%, so kann man davon ausgehen, dass auch die Forderung nach einer besonders guten Auflösung im mittleren Schwierigkeitsbereich erfüllt ist.

Einschränkungen bei den Aufgabenblöcken zum Fachwissen müssen bei den Inhaltsbereichen Relativitätstheorie und Quantenmechanik gemacht werden. Bei den Aufgaben zur Relativitätstheorie war die schwerste Aufgabe mit einer Lösungsrate von 20% zwar noch angemessen, die einfachste Aufgabe mit einer Lösungsrate von 67% etwas zu schwierig. Dazu kommt noch die geringe Anzahl von insgesamt nur 4 Items auf dieser Skala, was der Optimierung der Testökonomie geschuldet ist. Immerhin liegt auch hier der Mittelwert der Lösungsraten angemessenen 48%. Bei den Aufgaben zur Quantenmechanik ist die Spannweite der Aufgabenschwierigkeiten am geringsten. Die schwierigste Aufgabe hat hier eine Lösungsrate von 23%, die einfachste Aufgabe war dagegen mit einer Lösungsrate von nur 62% noch zu einfach. Hier zeigt sich auch die geringste mittlere Lösungsrate von 40%.

Ein ähnliches Bild ergibt sich auch bei den fachdidaktischen Inhaltsbereichen Instruktionsstrategien und Assessment. Bei den Aufgaben zu Instruktionsstrategien war die schwierigste Aufgabe mit einer Lösungsrate von 23% noch angemessen, während die einfachste Aufgabe auch hier mit einer Lösungsrate von 65% noch zu schwer ist. Die mittlere Lösungsrate liegt hier bei angemessenen 46%. Bei den Aufgaben zum Assessment hat die schwierigste Aufgabe eine Lösungsrate von 21%. Auch hier ist die leichteste Aufgabe mit einer Lösungsrate von 63% noch zu schwierig, was sich auch in der zu geringen mittleren Lösungsrate von 37% zeigt.

Von den dargelegten Einschränkungen abgesehen zeigt sich damit jedoch, dass die Verteilung der Aufgaben insgesamt gut mit der Verteilung der Probandenfähigkeit übereinstimmt, was für die Differenzierungsfähigkeit des Instruments innerhalb der Spannweite der Probandenfähigkeit spricht und als Argument für die strukturelle Validität des Testinstruments gewertet werden kann.

(d) Generalisierbarkeit

Der Aspekt der Generalisierbarkeit des Instruments ist dann gegeben, wenn die Stichprobe der Erhebung als repräsentativ für die Gesamtheit der Zielgruppe angenommen werden kann. Dies ist genau dann der Fall, wenn die Stichprobe hinsichtlich ihrer Merkmale eine ähnliche Struktur wie die Grundgesamtheit der Zielgruppe aufweist (Messick, 1995; Moosbrugger und Kelava, 2008). Für die

Generalisierbarkeit der Stichprobe spricht die Verteilung auf zwölf verschiedene Standorte, die über neun Bundesländer verteilt sind (Nordrhein-Westfalen, Bremen, Baden-Württemberg, Hessen, Hamburg, Niedersachsen, Schleswig-Holstein, Bayern und Mecklenburg-Vorpommern). Die Erfüllung des Kriteriums der Generalisierbarkeit steigt mit zunehmender Stichprobengröße bzw. sinkt für geringere Stichprobengrößen. Um Auswahleffekte auszuschließen, ist die Erfüllung zudem grundsätzlich nur durch zufällige Ziehung der Stichprobe zu erreichen. Da jedoch die Stichprobengröße in der vorliegenden Studie einen Kompromiss zwischen der Genauigkeit der Repräsentation der Grundgesamtheit und der Testökonomie darstellt, kann dieser Aspekt bei einer Stichprobengröße von 201 nur eingeschränkt als erfüllt gelten. Zudem handelt es sich bei den Probanden der vorliegenden Studie um Freiwillige, so dass ein gewisser Auswahleffekt angenommen werden muss. Beispielsweise kann angenommen werden, dass die Probanden der vorliegenden Stichprobe sich durch eine besonders hohe Motivation auszeichnen, wenn Sie freiwillig an einer vierstündigen Testung teilnehmen. Durch die Zahlung einer Aufwandsentschädigung von 40 € könnte die Motivation jedoch auch monetäre und damit fachfremde Gründe haben.

(e) Externe Validität

Um Hinweise für die externe Validität zu generieren, werden – wie schon bei der Pilotierung (siehe Abschnitt 4.4.2) – Korrelationen der Testscores mit Begleitvariablen berechnet, für die als Maß externer Konstrukte signifikante Zusammenhänge zum Professionswissen erwartet werden können. Die allgemeine kognitive Fähigkeit der Probanden wird durch die Abiturnote in Form der Schulnote (1 = sehr gut; ...; 6 = ungenügend) operationalisiert. Für die Schulleistung im Fach Physik vor Beginn des Studiums wird als Maß die letzte Schulnote im Fach Physik erfasst. Dazu wird der Studienfortschritt anhand der Semesterzahl, sowie der Lehramtszugang betrachtet. Beim Lehramtszugang wird zwischen gymnasial und nichtgymnasial unterschieden. Zum Vergleich werden die Korrelationen zu den Testscores für das pädagogische Wissen mit angegeben (siehe Tabelle 5.5).

Tabelle 5.5: Zusammenhänge zwischen Fachwissen, fachdidaktischem Wissen und pädagogischem Wissen mit externen Variablen (Kodierung des Lehramtszugangs: 1 = gymnasial, 0 = nichtgymnasial) (**: $p < .001$; *: $p < .05$; +: $p < .10$).

	Fachwissen	fachdidaktisches Wissen	pädagogisches Wissen
Abiturnote	-.46**	-.35**	-.30**
Letzte Schulnote im Fach Physik	-.26**	-.15*	-.18*
Fachsemester	.34**	.23**	.20*
Lehramtszugang	.39**	.18*	.12*

Die beobachteten negativen Korrelationen zwischen den Testscores und der Abiturnote bzw. der letzten Physiknote lassen sich durch die Skalierung des Notensystems der Schule erklären, da hier die niedrigste Note für die besten Leistungen vergeben wird, während bei den Testscores ein hoher Score auf ein hohes Wissen schließen lässt. Demnach sind die betrachteten Zusammenhänge durchaus plausibel, wobei zu beachten ist, dass die Zusammenhänge mit der Abiturnote als Indikator für kognitive Fähigkeiten gegenüber der letzten Physiknote überwiegt. Hohe Zusammenhänge zeigen sich ebenfalls bei der Semesterzahl, wobei der höchste Zusammenhang mit dem Fachwissen sich durch den hohen Anteil an Lerngelegenheiten für das Fachwissen im Studium erklären lässt.

Beim Lehramtszugang zeigt sich, dass angehende Lehrkräfte für das Gymnasium solchen mit angestrebtem nichtgymnasialen Lehramtszugang beim Fachwissen deutlich überlegen sind, was sich wiederum durch mehr Lerngelegenheiten im Studium erklären lässt. Bemerkenswert ist, dass angehende Gymnasiallehrkräfte sowohl beim fachdidaktischen Wissen als auch beim pädagogischen Wissen besser abschneiden, obwohl hier im Allgemeinen kein höherer Anteil an Lerngelegenheiten vorliegt. Es kann daher vermutet werden, dass das bessere Abschneiden hier durch die kognitiven Fähigkeiten moderiert wird. Insgesamt sprechen die berechneten Korrelationen aufgrund ihrer Plausibilität und ihrer Werte für die externe Validität des Instruments.

(f) Konsequenzen

Unter dem Aspekt der Konsequenzen des Testergebnisses versteht man sowohl die individuellen Konsequenzen für die Probanden als auch die politischen und gesellschaftlichen Konsequenzen, die aus der Interpretation und Veröffentlichung von Testergebnissen resultieren können (Messick, 1995). Zur Abschätzung der individuellen Konsequenzen muss diskutiert werden, inwieweit das Testverfahren in den Alltag der Probanden eingreift. Da die Teilnahme an der Testung erstens freiwillig und zweitens mit einer Gesamtdauer von vier Stunden als zumutbar beurteilt werden muss, ist hier keine nachhaltige Beeinflussung der Probanden zu erwarten.

Die Validität im Sinne der politischen und gesellschaftlichen Konsequenzen wäre im Falle der vorliegenden Arbeit gefährdet, wenn davon ausgegangen werden müsste, dass der Einsatz des Messinstruments oder die Veröffentlichung von Ergebnissen Lehren und Lernen von Professionswissen angehender Physiklehrkräfte in ungewünschter Weise verändern könnte (Moosbrugger und Kelava, 2008). Dies kann der Fall sein, wenn Dozenten ihre Probanden gezielt auf den Test vorbereiten und das Testverfahren so Einfluss auf die Lehre nimmt. Da die Testinstrumente im Vorfeld der Testung nicht veröffentlicht wurden, ist jedoch von einer derartigen Beeinflussung nicht auszugehen.

Zusammenfassung zur Testgüte

Das in der Hauptstudie eingesetzte Testinstrument erfasst das Fachwissen mit hoher Reliabilität und das fachdidaktische Wissen mit akzeptabler Reliabilität. Die Ergebnisse zur Konstruktvalidität zeigen, dass das Professionswissen angehender Physiklehrkräfte durch das Messinstrument besonders gut hinsichtlich des inhaltlichen und kognitiven Aspekts erfasst wird.

5.2 Struktur des Professionswissens

Um erste Erkenntnisse über die Struktur des Professionswissens zu erhalten, wird zunächst anhand von Modellvergleichen mittels einer konfirmatorischen Faktorenanalyse geprüft, ob die Passung eines dreidimensionalen Modells für die Struktur des Professionswissens anhand der erhobenen Daten bestätigt werden kann. Danach wird die innere Struktur des am besten passenden Modells weiter untersucht.

5.2.1 Modellprüfung

Aufgrund der bisherigen Befunde zum Professionswissen (siehe Abschnitt 2.1.2) ergibt sich die theoretische Annahme, dass sich das Professionswissen in die drei Bereiche Fachwissen bzw. Content Knowledge (CK), fachdidaktisches Wissen bzw. Pedagogical Content Knowledge (PCK) und pädagogisches Wissen bzw. Pedagogical Knowledge (PK) aufteilen lässt. Daher bietet sich die Prüfung eines dreidimensionalen Modells mittels explorativer Faktorenanalyse (siehe Abschnitt 4.4.7) an, in denen die Subskalen zum CK, PCK und PK jeweils auf einen von drei Faktoren laden. Die Alternative zur Trennung von CK, PCK und PK wäre die Untrennbarkeit von CK, PCK und PK, also deren Zusammenfallen zu einem empirisch nicht trennbaren Konstrukt. Für ein zweidimensionales Modell des Professionswissen existieren keine theoretischen Annahmen oder Befunde. Daher wird die Passung eines dreidimensionalen Modells mit einem eindimensionalen Modell verglichen, bei dem sämtliche Subskalen auf einen Faktor laden. Die Fit-Werte des Vergleichs der beiden Modellpassungen ist in Tabelle 5.6 dargestellt.

Tabelle 5.6: Modellvergleich für 1- und 3-dimensionales Modell des Professionswissens.

Modell	CFI	TLI	SRMR	RMSEA	AIC	BIC
1-dim	.786	.753	.085	.095	13763.142	13868.848
3-dim	.898	.879	.065	.066	13667.193	13782.809

Für die Interpretation der Fit-Werte existieren in der Literatur verschiedene Richtwerte: Für den Comparative-Fit Index (CFI) und den Tucker-Lewis-Index (TLI) liegen die Fit-Werte stets zwischen 0 und 1 wobei Werte oberhalb von .90 (Bentler, 1990; 1992; Baumgartner und Homburg, 1996) beziehungsweise .95 (Carlson und Mulaik, 1993; Hu und Bentler, 1998) als guter Fit gewertet werden können. Je höher also die Werte für CFI und TLI, desto besser die Modellpassung. Demnach liegen die Werte von CFI und TLI für das dreidimensionale Modell knapp außerhalb des guten Bereichs. Für das eindimensionale Modell liegen die Werte für CFI und TLI deutlicher außerhalb des guten Bereichs.

Beim Standardized Root Mean Square of Residuals (SRMR) und Root Mean Square Error of Approximation (RMSEA) bedeutet 0 einen perfekten Modellfit. Beim SRMR werden Werte unterhalb von .08 (Hu und Bentler, 1998) oder .10 (Homburg, Herrmann und Pflesser, 2008) als guter Fit interpretiert. Beim RMSEA werden Werte unterhalb von .06 (Hu und Bentler, 1992) oder .05 als guter Fit und Werte unterhalb von .08 als noch akzeptabler Fit interpretiert (Browne und Cudeck, 1993). Beim SRMR liegen die Fit-Werte des dreidimensionalen Modells damit nach beiden Interpretationen im guten Bereich. Die Fit-Werte des eindimensionalen Modells liegen beim SRMR nur bezüglich der Interpretation von Hu und Bentler (1998) innerhalb des guten Bereichs. Beim RMSEA zeigt nur das dreidimensionale Modell einen noch akzeptablen Fit, während das eindimensionale Modell außerhalb des akzeptablen Bereichs liegt.

Für das Akaike Information Criterion (AIC) und das Bayesian Information Criterion (BIC) ist der Modellfit jeweils umso besser, je niedriger die jeweiligen Werte sind. Vergleicht man also das drei- und eindimensionale Modell bezüglich AIC und BIC, so passt auch hier das dreidimensionale Modell jeweils besser zu den erhobenen Daten. Damit weisen sämtliche untersuchten Fit-Parameter darauf hin, dass das dreidimensionale Modell gegenüber dem eindimensionalen Modell zu bevorzugen ist.

5.2.2 Binnenstruktur des Professionswissens

Da nach den vorangegangenen Analysen ein dreidimensionales Modell des Professionswissens bestätigt wurde, werden nun die Zusammenhänge der drei Wissensbereiche betrachtet. Hierzu werden nächst manifeste Korrelationen zwischen den drei Bereichen CK, PCK und PK bestimmt (siehe Abbildung 5.1).

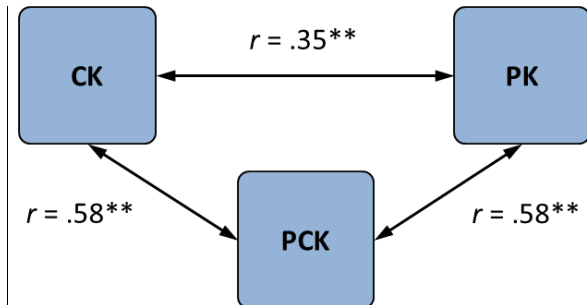


Abbildung 5.1: Manifeste Korrelationen zwischen Bereichen des Professionswissens.

Hier ergeben sich starke Zusammenhänge zwischen fachdidaktischem Wissen und Fachwissen bzw. fachdidaktischem Wissen und pädagogischem Wissen und ein mittlerer Zusammenhang zwischen Fachwissen und pädagogischem Wissen.

Betrachtet man die latenten Korrelationen zwischen den Testscores der drei Bereiche, so ergibt sich ein ähnliches Bild, wobei die Werte der Korrelationen höher ausfallen (siehe Abbildung 5.2).

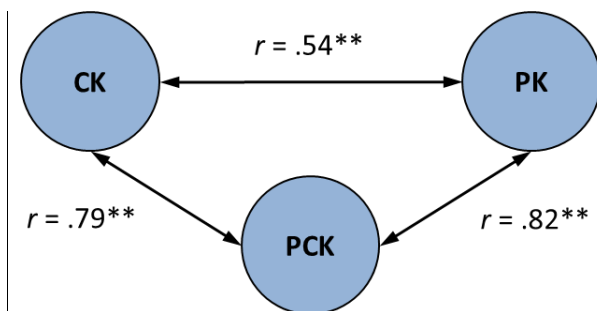


Abbildung 5.2: Latente Korrelationen zwischen Bereichen des Professionswissens.

Auch hier erkennt man einerseits einen starken und hochsignifikanten Zusammenhang zwischen fachdidaktischem Wissen und Fachwissen ($r = .79$) und andererseits zwischen fachdidaktischem Wissen und pädagogischem Wissen ($r = .82$). Ferner existiert ein mittlerer, aber immer noch hochsignifikanter Zusammenhang zwischen Fachwissen und pädagogischem Wissen.

Um die Binnenstruktur des Professionswissens tiefergehend zu untersuchen, wird nun analysiert, wie die drei Bereiche des Professionswissens durch ihre jeweiligen Subskalen repräsentiert werden. Hierzu werden die einzelnen Faktorladungen der Subskalen auf den jeweiligen Bereich des Professionswissens, ihre Standardfehler und die Korrelation mit dem jeweiligen Bereich des Professionswissens betrachtet (siehe Tabellen 5.7 bis 5.9). Die Faktorladung wird dabei jeweils für die erste Subskala auf den Wert 1,0 normiert. Für die Signifikanzniveaus der Korrelationen in den Tabellen 5.7 bis 5.9 gilt jeweils: $p < .001$.

Der Bereich Fachwissen wird durch die acht Facetten Mechanik, Elektrizitätslehre, Optik, Thermodynamik, Festkörperphysik, Atom- und Kernphysik, Relativitätstheorie und Quantenmechanik repräsentiert.

Tabelle 5.7: Repräsentation des Fachwissens durch Subskalen.

Subskala	Standardisierte Faktorladung	Standardfehler bez. Mechanik
Mechanik	0.61	.00
Elektrizitätslehre	0.57	.14
Optik	0.58	.16
Thermodynamik	0.69	.17
Festkörperphysik	0.67	.15
Atom- und Kernphysik	0.59	.14
Relativitätstheorie	0.47	.10
Quantenmechanik	0.60	.16

Das Fachwissen wird demnach besonders gut durch die Facetten Thermodynamik, Festkörperphysik, Mechanik und Quantenmechanik repräsentiert. Die Subskala zur Relativitätstheorie stellt hier einen Ausreißer dar, was auf die ungenügende Reliabilität dieser Subskala zurückgeführt werden kann.

Das fachdidaktische Wissen wird durch die Subskalen Schülerkognition, Instruktionsstrategien, Curriculum und Assessment repräsentiert.

Tabelle 5.8: Repräsentation des fachdidaktischen Wissens durch Subskalen.

Subskala	Standardisierte Faktorladung	Standardfehler bez. Schülerkognition
Schülerkognition	0.60	.00
Instruktionsstrategien	0.68	.18
Curriculum	0.59	.11
Assessment	0.51	.14

Anhand der Faktorladungen erkennt man, dass das fachdidaktische Wissen in besonderem Maße von den Subskalen Instruktionsstrategien und Schülerkognition repräsentiert wird. Die Faktorladungen der Subskalen Assessment und Curriculum sind mit .59 und .51 etwas geringer.

Das pädagogische Wissen wird durch die Subskalen Lehren, Lernen, Klassenführung und Leistungsbeurteilung repräsentiert.

Tabelle 5.9: Repräsentation des pädagogischen Wissens durch Subskalen.

Subskala	Standardisierte Faktorladung	Standardfehler bez. Lehren
Lehren	0.59	.00
Lernen	0.78	.19
Klassenführung	0.56	.19
Leistungsbeurteilung	0.47	.23

Das pädagogische Wissen wird nach Tabelle 5.9 mit Abstand am besten durch die Subskala zum Lernen repräsentiert. Die Subskala zur Leistungsbeurteilung stellt – wie schon die Subskala zur speziellen Relativitätstheorie beim Fachwissen – mit .47 einen Ausreißer nach unten dar.

Zusammenfassung

Die Ergebnisse der konfirmatorischen Faktorenanalyse zeigen, dass die empirische Trennbarkeit der drei Wissensbereiche des Professionswissens voneinander, also des Fachwissens, des fachdidaktischen Wissens und des pädagogischen Wissens, für angehende Physiklehrkräfte bestätigt werden kann. Trotz der Trennbarkeit hängen die einzelnen Bereiche stark miteinander zusammen, wobei sich besonders starke Zusammenhänge zwischen dem fachdidaktischen Wissen und dem Fachwissen bzw. dem pädagogischen Wissen ergeben.

5.3 Entwicklung des Fachwissens und fachdidaktischen Wissens

Nachdem im vorangegangenen Abschnitt die empirische Trennbarkeit der Bereiche des Professionswissens, Fachwissen, fachdidaktisches Wissen und pädagogisches Wissen bestätigt wurde, wird nun die Entwicklung des Professionswissens in den fachbezogenen Bereichen, Fachwissen und fachdidaktisches Wissen, betrachtet. Für die Untersuchung der Entwicklung des Fachwissens und des fachdidaktischen Wissens werden zunächst beide Konstrukte, operationalisiert durch die von den Probanden erzielten Punktzahlen der Probanden auf den verschiedenen Skalen, zu verschiedenen Phasen des Lehramtsstudiums betrachtet. Als Phasen werden die Studienjahre, gemessen vom 1. Bis zum 5. Jahr, herangezogen. Darüber hinausgehende Studienphasen werden nicht berücksichtigt da zum einen ab dem 6. Studienjahr üblicherweise die Regelstudienzeit überschritten wird und zum anderen die Studienjahre 6 und 7 nur mit Stichprobengrößen unter 10 Probanden im zugrundeliegenden Datensatz enthalten sind, was keine belastbaren Analysen zuließe. Die Tabelle 5.10 zeigt die Mittelwerte und Standardabweichungen der Ausprägung von Fachwissen und fachdidaktischem Wissen. Eine zusätzliche Varianzanalyse zeigt die Signifikanz der Unterschiede ($p < .01$) zwischen den universitären Jahrgangsstufen 1 bis 5 für das Fachwissen und das fachdidaktische Wissen.

Tabelle 5.10: Erzielte Scores beim Fachwissen und fachdidaktischen Wissen mit Standardabweichungen im Laufe des Lehramtsstudiums (blau: CK; rot: PCK).

Studienjahr	Fachwissen	SD	Differenz zum Vorjahr	fachdidaktisches Wissen	SD	Differenz zum Vorjahr
1	21,3	5,3		14,4	4,0	
2	27,6	4,9	+6,3	16,6	4,9	+2,2
3	29,3	8,6	+1,7	17,6	5,0	+1,0
4	34,4	6,9	+5,1	20,4	3,4	+2,8
5	32,1	5,1	-2,3	18,5	4,0	-1,9

Dividiert man die obigen Summenscores für Fachwissen und fachdidaktisches Wissen jeweils durch die zugehörigen Itemzahlen, so erhält man die mittleren Lösungsraten als Indikator für die Merkmalsausprägung. Diese, aufgetragen über die Studienjahre, ist in Abbildung 5.3 für das Fachwissen dargestellt.

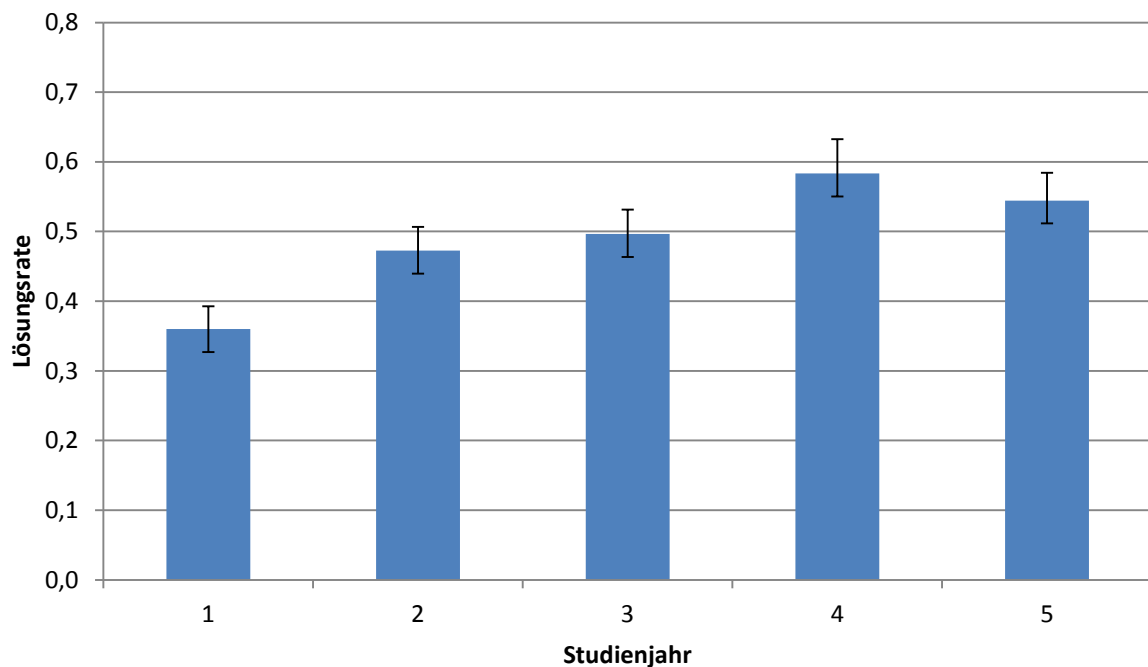


Abbildung 5.3: Entwicklung des Fachwissens während der ersten fünf Studienjahre.

Beim Fachwissen zeigt sich der stärkste Anstieg zwischen dem ersten und zweiten Studienjahr. Im weiteren Verlauf steigt das Fachwissen der untersuchten Probanden bis zum vierten Studienjahr weiter an, bevor sich ein leichter Rückgang zum fünften Studienjahr hin ergibt.

Die Lösungsraten als Indikator für die Merkmalsausprägung des fachdidaktischen Wissens sind, in gleicher Art wie beim Fachwissen, in Abbildung 5.4 über die Studienjahre hinweg aufgetragen.

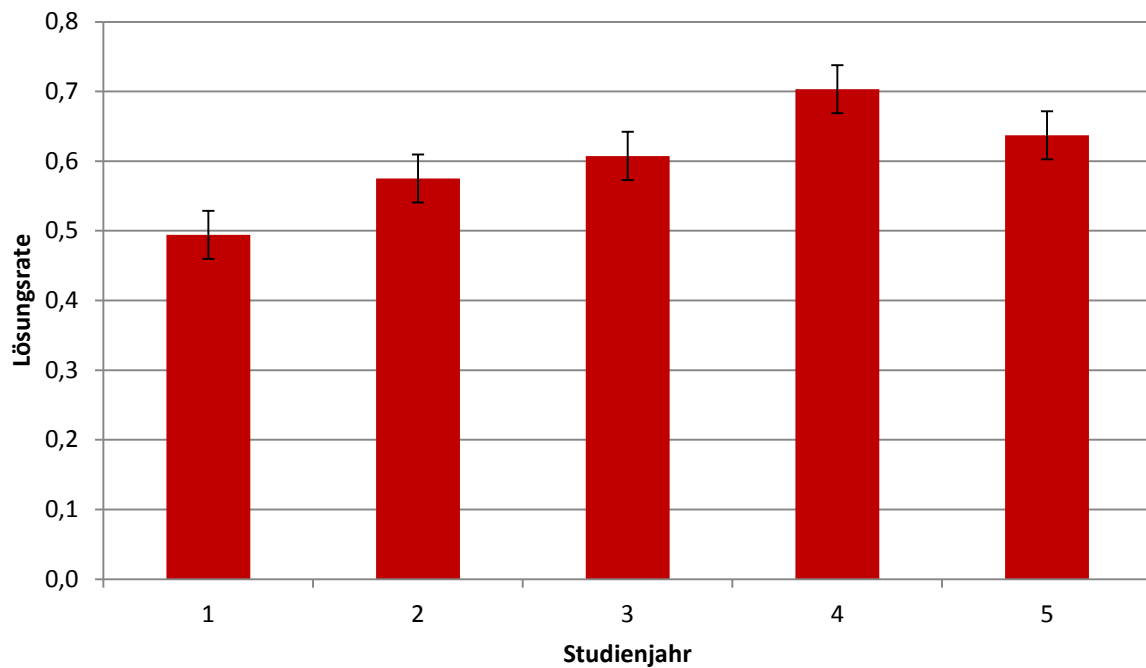


Abbildung 5.4: Entwicklung des fachdidaktischen Wissens während der ersten fünf Studienjahre.

Im Vergleich zum Fachwissen zeigen die Lösungsraten für das fachdidaktische Wissen eine strukturelle Gemeinsamkeit, nämlich einen Anstieg bis zum vierten Studienjahr und einen leichten Rückgang vom vierten zum fünften Studienjahr. Es zeigen sich jedoch auch zwei Unterschiede im Vergleich zum Fachwissen: Der stärkste Anstieg der Lösungsraten ist beim fachdidaktischen Wissen erst zwischen dem dritten und vierten Studienjahr zu verzeichnen. Zudem ist der Anstieg über die gesamte Studiendauer hinweg weniger stark ausgeprägt.

Entwicklung des Professionswissens auf Subskalenebene

Über die Betrachtung der Entwicklung des Professionswissens auf den Hauptskalen für CK und PCK hinaus erlaubt der Einsatz des vorliegenden, inhaltlich breit aufgestellten Messinstruments eine differenziertere Betrachtung des Professionswissens auf Subskalenebene. Hierfür wird, wie schon im vorangegangenen Abschnitt, die mittlere Lösungsrate für jede Subskala als Indikator für die Merkmalsausprägung über die Studienjahre hinweg aufgetragen. Die zugehörigen Daten sind mit Standardabweichung in Tabelle 5.11 dargestellt.

Tabelle 5.11: Lösungsraten der Subskalen zum Fachwissen mit Standardabweichungen im Laufe des Lehramtsstudiums (Studienjahr SJ; Standardabweichung SD).

SJ	Mechanik (SD)	Elektrizitätslehre (SD)	Optik (SD)	Thermodynamik (SD)	Festkörperphysik (SD)	Atom- u. Kernphysik (SD)	Relativitätstheorie (SD)	Quantenmechanik (SD)
1	.43 (.16)	.43 (.15)	.28 (.21)	.34 (.18)	.34 (.17)	.41 (.19)	.45 (.29)	.32 (.25)
2	.54 (.17)	.50 (.19)	.37 (.21)	.57 (.20)	.47 (.21)	.46 (.21)	.56 (.32)	.45 (.30)
3	.54 (.17)	.52 (.21)	.43 (.19)	.52 (.22)	.52 (.24)	.52 (.24)	.55 (.29)	.53 (.31)
4	.60 (.18)	.52 (.17)	.58 (.20)	.60 (.23)	.66 (.19)	.63 (.20)	.66 (.25)	.61 (.22)
5	.56 (.13)	.51 (.13)	.59 (.19)	.53 (.20)	.58 (.18)	.57 (.14)	.67 (.26)	.55 (.23)

Damit ergeben sich für die acht Subskalen zum Fachwissen die in Abbildung 5.5 dargestellten Entwicklungsverläufe. Dabei ist jedoch zu beachten, dass die Standardabweichungen zum Teil recht groß sind, die Messwerte also mit einer relativ hohen Ungenauigkeit behaftet sind. Die erste betrachtete Subskala ist die zum Inhaltsbereich Mechanik, der nach Friege und Lind (2003; 2004) als Prädiktor für physikalisches Fachwissen bzw. Problemlösekompetenz angesehen wird.

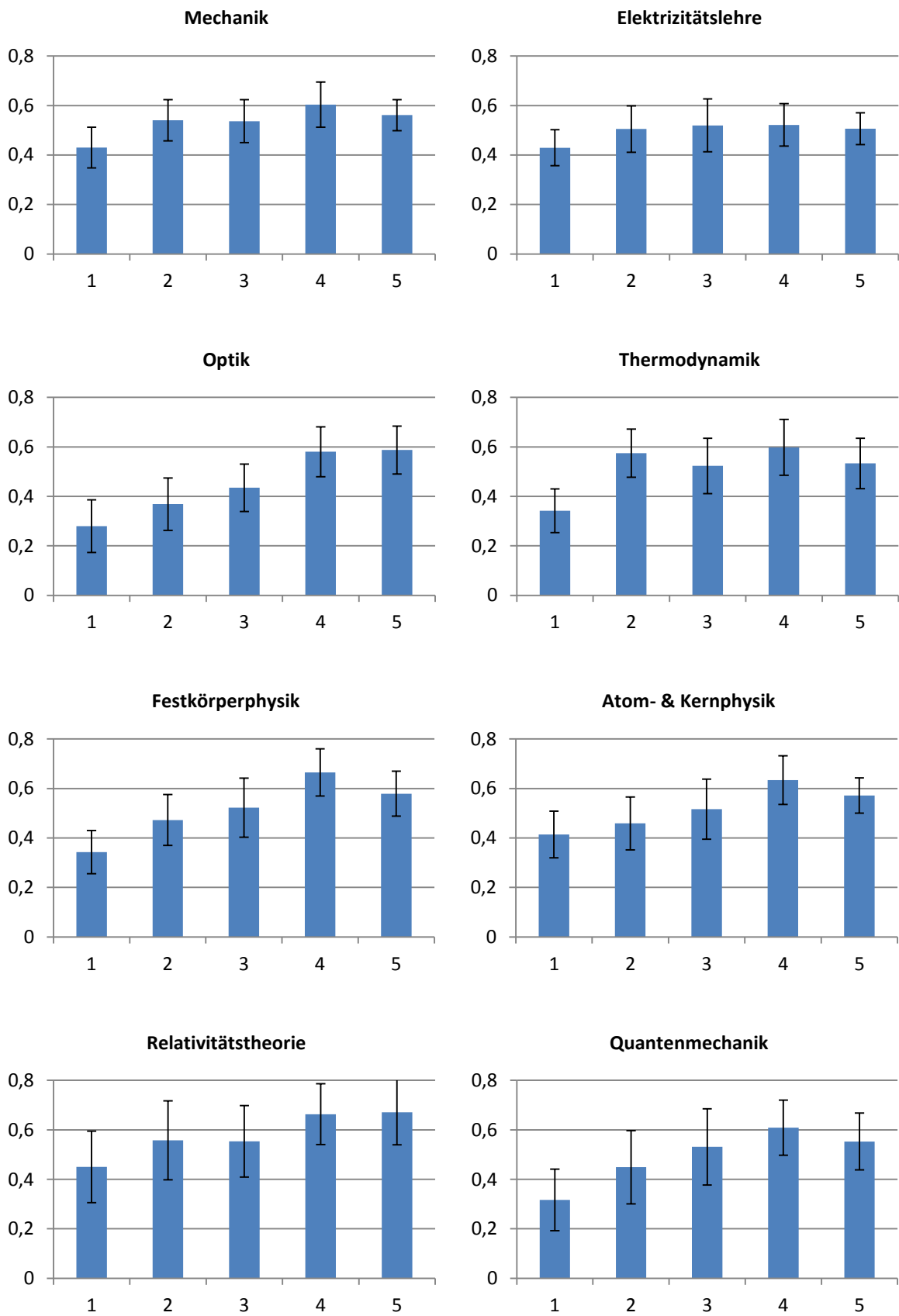


Abbildung 5.5: Entwicklung in den Fachinhaltsbereichen.

Ein Vergleich der Entwicklungen in den einzelnen Fachinhaltsbereichen in der Abbildung 5.5 zeigt, dass die meisten Subskalen dem Skalenverlauf der Hauptskala für das Fachwissen im Wesentlichen folgen, jedoch weniger glatt verlaufen. Trends, die für die generelle Entwicklung des Fachwissens gelten, also größter Anstieg von ersten zum zweiten Studienjahr und ein Abfall vom vierten zum fünften Studienjahr, sind auch für die Subskalen zu den Inhaltsbereichen Mechanik, Elektrizitätslehre und Quantenmechanik sichtbar. Die Subskala zum Inhaltsbereich Optik zeigt im Gegensatz zur generellen Entwicklung des Fachwissens den stärksten Anstieg zwischen dem dritten und vierten Studienjahr, sowie einen leichten Anstieg des Fachwissens vom vierten zum fünften Studienjahr. Studienfortschritte lassen sich jedoch auch an dieser Skala gut ablesen, da der Fortschritt bis zum vierten Studienjahr relativ kontinuierlich verläuft. Ferner sei daran erinnert, dass diese Subskala mit einem Cronbachs α von .59 (siehe Tabelle 5.2) die größte Reliabilität unter allen Subskalen zum Fachwissen aufweist.

Ebenso wie beim Fachwissen lassen sich auch für das fachdidaktische Wissen die Entwicklungsverläufe der Subskalen messen. Hierfür wird, wie zuvor, die mittlere Lösungsrate für jede der vier Subskalen als Indikator für die Merkmalsausprägung über die Studienjahre hinweg aufgetragen. Die zugehörigen Daten mit Standardabweichung sind in Tabelle 5.12 dargestellt.

Tabelle 5.12: Lösungsraten der Subskalen zum fachdidaktisches Wissen mit Standardabweichungen im Laufe des Lehramtsstudiums (Studienjahr SJ; Standardabweichung SD).

SJ	Schülerkognition (SD)	Instruktionsstrategien (SD)	Curriculum (SD)	Assessment (SD)
1	.52 (.18)	.37 (.17)	.46 (.15)	.35 (.20)
2	.60 (.15)	.45 (.20)	.46 (.19)	.41 (.25)
3	.61 (.14)	.47 (.21)	.52 (.19)	.41 (.23)
4	.63 (.12)	.55 (.19)	.61 (.17)	.57 (.25)
5	.59 (.13)	.55 (.16)	.57 (.15)	.40 (.24)

Damit ergeben sich für die Subskalen zum fachdidaktischen Wissen die in der Abbildung 5.6 dargestellten Entwicklungsverläufe.

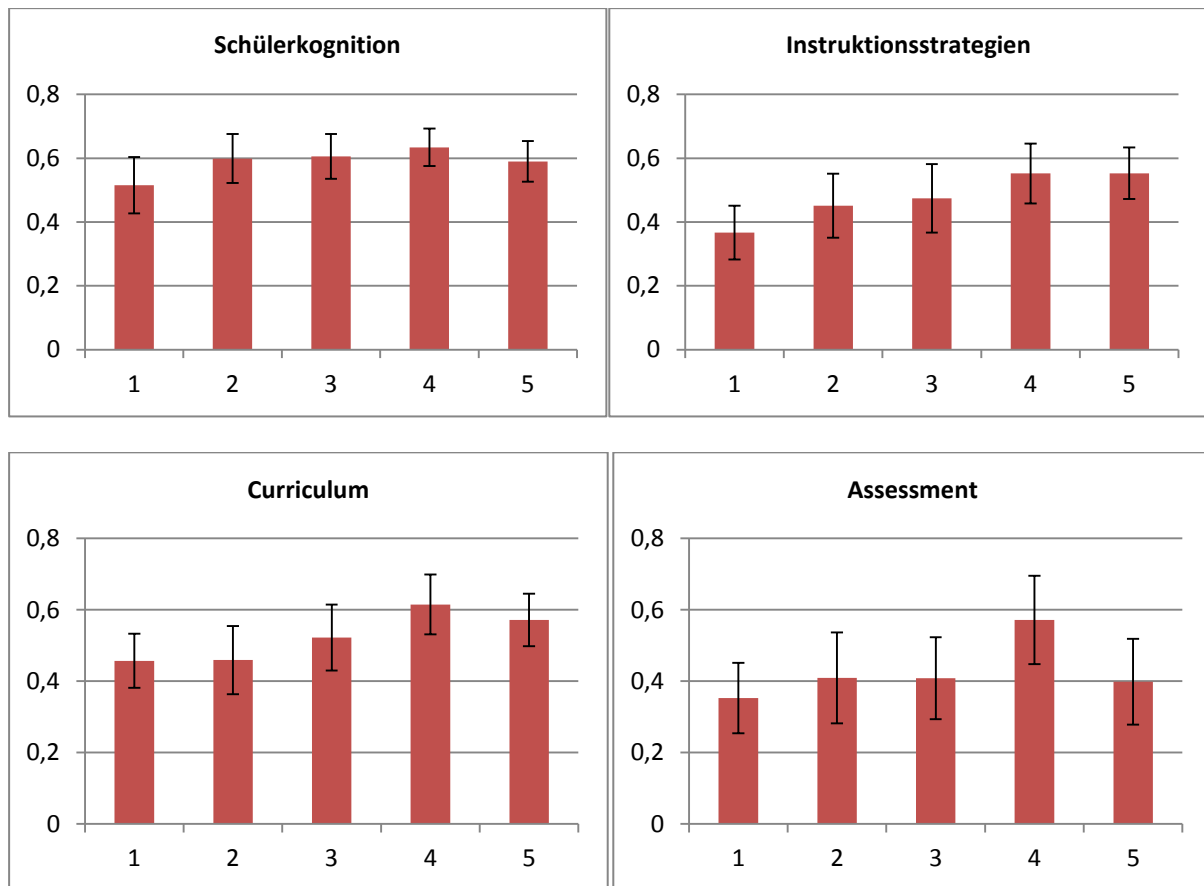


Abbildung 5.6: Entwicklung in den fachdidaktischen Inhaltsbereichen.

Hier zeigen sich, im Vergleich mit den Subskalen zum Fachwissen, deutlichere Unterschiede. Auf den Subskalen für Schülerkognition und Assessment lässt zwischen dem zweiten und dritten Studienjahr nur ein geringer Studienfortschritt ablesen. Die Subskala zum fachdidaktischen Inhaltsbereich Curriculum zeigt dagegen keinen nennenswerten Anstieg vom ersten zum zweiten Studienjahr. Lediglich die auf der Subskala zu Instruktionsstrategien lassen sich deutliche Unterschiede zwischen den einzelnen Jahrgangsstufen messen. Allerdings muss auch hier wieder berücksichtigt werden, dass die den Skalen zugrunde liegenden Reliabilitäten, gemessen durch Cronbachs α , zwischen .39 und .51 liegen. Die Daten lassen damit keine belastbare Analyse auf Basis der Subskalen zu. Die Subskala zu Instruktionsstrategien, die am besten zwischen den Jahrgangsstufen differenziert, erreicht hier mit .51 allerdings den Maximalwert im Bereich PCK, ist also noch am verlässlichsten.

Raschskalierung

Zur Überprüfung der Befunde für die Entwicklung des Professionswissen auf den Hauptskalen CK und PCK wurde zusätzlich zu den Betrachtungen der Summenscores eine Raschskalierung der Daten vorgenommen. Dazu wurden die Merkmalsausprägungen der für Fachwissen und fachdidaktisches auf einer Skala von 200 bis 1200 (siehe Abschnitt 4.4.7, *Analyse*) gemessen und wie oben über die ersten fünf Studienjahre hinweg aufgetragen (siehe Abbildung 5.7).

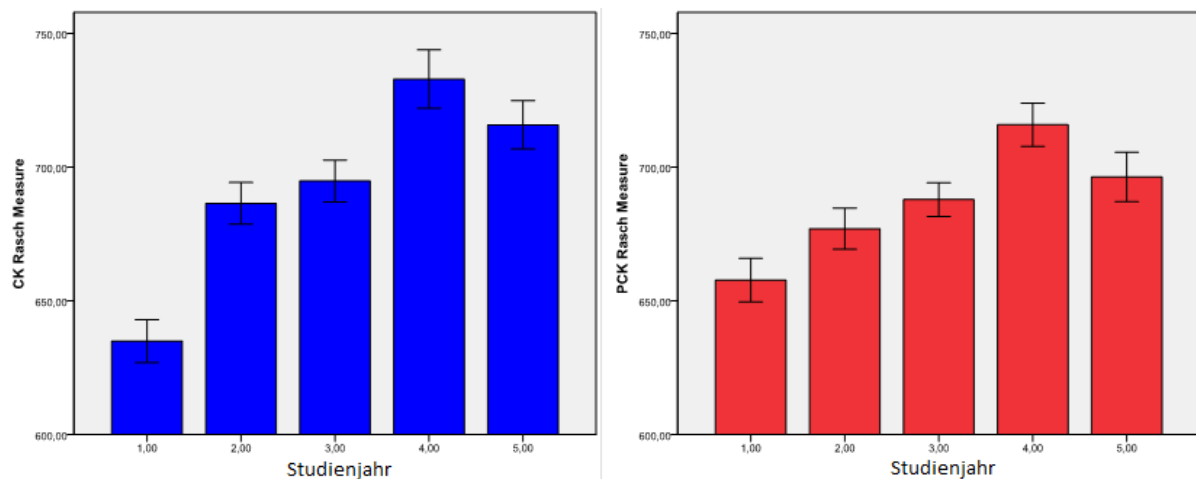


Abbildung 5.7: Raschskalierung der Entwicklung des Fachwissens und des fachdidaktischen Wissens.

Hier treten die oben erwähnten Gemeinsamkeiten und Unterschiede zwischen Fachwissen und fachdidaktischem Wissen noch deutlicher als in den Abbildungen 5.3 und 5.4 zu Tage: Die Ausprägung beider Werte steigt über den Studienverlauf hin bis zum vierten Studienjahr an und fällt dann leicht ab.

Für das Fachwissen im Fach Physik ergibt sich auf der Raschskala ein durchschnittlicher Zuwachs von $\tau = 14.30$ ($p < .001$) Raschpunkten pro Jahr. Der Gesamtzuwachs im Studienverlauf, also vom ersten bis zum fünften Studienjahr, beträgt 80,8 Raschpunkte (von 635,0 auf 715,8). Dabei lässt sich eine Effektstärke, gemessen durch Cohens d , von $d = 1,9$ beobachten, was einen sehr starken Effekt darstellt. Für das fachdidaktische Wissen ergibt sich ein durchschnittlicher Zuwachs von $\tau = 8.03$ ($p < .05$) Raschpunkten pro Jahr und ein Gesamtzuwachs von 38,6 (von 657,7 auf 696,3) bei einer Effektstärke von Cohens $d = 0,9$, was noch einen starken Effekt darstellt. Die Effekte in Schulvergleichsstudien legen nahe, dass pro Jahr ein Zuwachs zwischen mindestens $d = 0.15$ und $d = 0.40$ zu erwarten ist (Hattie, 2009, S.20; Geller et al., 2014), was von Studierenden offenbar leicht übertroffen wird.

Zusammenfassung

Die eingesetzten Instrumente zur Messung des Fachwissens und des fachdidaktischen Wissens sind in der Lage starke Effekte hinsichtlich der Entwicklung von Merkmalsausprägungen in verschiedenen Studienphasen zu messen. Sowohl die Ausprägung des Fachwissens als auch des fachdidaktischen Wissens steigt vom ersten zum vierten Studienjahr an und fällt zum fünften Studienjahr hin leicht ab. Dabei zeigt sich, dass die Ausprägung des Fachwissens besonders zu Beginn des Lehramtsstudiums stark ansteigt. Die Ausprägung des fachdidaktischen Wissens steigt im Studienverlauf später und insgesamt vergleichsweise weniger stark an.

5.4 Korrelate und Determinanten des Professionswissens

Um herauszufinden, welche Faktoren die Entwicklung des Professionswissens angehender Physiklehrkräfte beeinflussen können, werden Einflüsse möglicher Prädiktoren des Professionswissens auf die Ausprägungen des Fachwissens und des fachdidaktischen Wissens untersucht. Die Operationalisierungen der Personenmerkmale auf der Basis von Einzelitems werden im Folgenden dargestellt. Für die Operationalisierungen auf der Basis von Skalen wird für die Skala jeweils die Reliabilität, gemessen durch Cronbachs α , berichtet. Die Personenmerkmale bzw. Skalen sind dabei in folgenden Kate-

gorien gruppiert: (a) Kognitive Fähigkeiten und Wissen, (b) Motivation und Interesse, (c) Selbstwirksamkeit und Selbsteinschätzung, (d) Zufriedenheit und (e) Lehr- und Lerngelegenheiten.

(a) Kognitive Fähigkeiten und Wissen

Als Maß für allgemeine kognitive Fähigkeiten dient hier die Note der allgemeinen Hochschulreife. Das schulische Vorwissen im Fach Physik wird über die letzte Physiknote erfasst, sofern das Fach Physik vom Probanden in der Oberstufe belegt wurde. Daneben wird über eine 24 Items umfassende Skala das Wissen über die Natur der Naturwissenschaften bzw. Nature of Science (NOS) erfasst. Diese Operationalisierungen sind in Tabelle 5.13 dargestellt.

Tabelle 5.13: Operationalisierungen der Abiturnote und der letzten Physiknote.

Variable	Operationalisierung
Abiturnote	Welche Gesamtnote (z.B. 2,5) haben Sie im Zeugnis der Hochschulreife (Abitur) erzielt?
Physiknote	Falls Sie Physik in der Oberstufe belegt haben: Was war Ihre letzte Zeugnisnote?

Neben der Abiturnote als etabliertem Maß für allgemeine kognitive Fähigkeiten (vgl. Riese und Reinhold, 2012; Gramzow, 2014) und der letzten Physiknote in der Schule als Maß für Vorwissen über Schulphysik (vgl. Riese, 2009, S.140) wurde noch das Wissen der Probanden über *Nature of Science (NOS)* mittels einer Skala von 24 Items erfasst. Die Ergebnisse der Regression sind in Tabelle 5.14 dargestellt.

Tabelle 5.14: Regression von Fachwissen und fachdidaktischem Wissen durch Abiturnote, Physiknote und Wissen über Nature of Science (NOS) (**: $p < .001$; *: $p < .05$; +: $p < .10$).

	Itemzahl	Cronbachs α	Fachwissen	fachdidaktisches Wissen
Abiturnote	1		-.46**	-.35**
Physiknote	1		-.26**	-.15*
Wissen über NOS	24	.98	.12 ⁺	.09

Erwartungskonform hängen kognitive Fähigkeiten, operationalisiert durch die Abiturnote, höher mit dem Fachwissen (vgl. Woitkowski, 2015, S.217) als mit dem fachdidaktischen Wissen im Fach Physik zusammen. Für das Fachwissen stellen sie gleichermaßen den stärksten in der vorliegenden Arbeit identifizierten Indikator dar. Vorangegangene Arbeiten im Forschungsfeld zeigten sowohl vergleichsweise geringere Zusammenhänge als in der vorliegenden Arbeit, sowie eine gewisse Schwankung der empirischen Befunde, die auf die Operationalisierung der Konstrukte des Fachwissens und des fachdidaktischen Wissens durch die verwendeten Aufgaben zurückzuführen sein dürfte. Beispielsweise ergab sich in der Arbeit von Riese (2009, S.141/142) für angehende Physiklehrkräfte eine Korrelation von $r = -.28^{**}$ zwischen Abiturnote und Fachwissen und von $r = -.14^{**}$ zwischen Abiturnote und fachdidaktischem Wissen, in der Arbeit von Woitkowski (2015, S.218) eine Korrelation von $r = -.39^{**}$ für das Fachwissen und bei Gramzow (2014, S.145) $r = -.23^+$ für das fachdidaktische Wissen.

Die letzte Schulnote im Fach Physik hat dagegen nur einen mittleren Effekt auf das Fachwissen und einen geringen auf das fachdidaktischen Wissen. Wie schon die Korrelationen zwischen Testscores des Kognitiven Fähigkeitstests (KFT) mit Fachwissen und fachdidaktischem Wissen bei der Pilotierungsstudie (siehe Abschnitt 4.4.2) zeigten, scheint kognitive Leistungsfähigkeit einen größeren Einfluss auf die Ausprägung des Professionswissens zu haben, als schulisches Vorwissen im Fach.

Das Wissen über *Nature of Science*, als fachwissensnahe Domäne (Lederman, 1992), hat einen lediglich am 10%-Niveau signifikanten Effekt auf das Fachwissen der Probanden, wobei die Effektstärke mit .12 im Vergleich zu anderen Personenmerkmalen eher gering ausfällt.

(b) Motivation und Interesse

In diesem Bereich wird die Studienmotivation der Probanden, sowie ihr Interesse am Fach Physik untersucht. Tabelle 5.15 zeigt die Effekte der Skalen auf Fachwissen (CK) und fachdidaktisches Wissen (PCK).

Tabelle 5.15: Regression von Motivation und Interesse mit den Scores für Fachwissen und fachdidaktisches Wissen (**: $p < .001$; *: $p < .05$; +: $p < .10$).

	Itemzahl	Cronbachs α	Fachwissen	fachdidaktisches Wissen
Interesse an Physik	5	.82	.31**	.21*
Motivation im Studium	11	.73	.10	.12 ⁺

Bezüglich der Motivation zeigen sich nur geringe Effekte auf Fachwissen und fachdidaktisches Wissen. Dieser Befund ist jedoch vergleichbar mit dem von Riese (2009, S.157) wo lediglich ein schwacher signifikanter Zusammenhang ($r = .14^*$) zwischen Leistungsmotivation und Fachwissen gefunden wurde.

Einen deutlich besseren Indikator, besonders für das Fachwissen, stellt das Interesse am Fach Physik dar, das sowohl für Fachwissen als auch fachdidaktisches Wissen hochsignifikante mittlere Effekte liefert.

(c) Selbstwirksamkeit und Selbsteinschätzung

Auf Einzelitemebene sollen die Probanden ihre bisherigen Leistungen im Studium im Vergleich zu Ihren Kommilitoninnen und Kommilitonen auf einer fünfstufigen Likert-Skala einschätzen (Skalenstufen der Selbsteinschätzung: viel besser, etwas besser, genauso gut, etwas schlechter, viel schlechter). Diese Operationalisierungen sind in Tabelle 5.16 dargestellt.

Tabelle 5.16: Operationalisierungen der Selbsteinschätzung.

Variable	Operationalisierung
	Wie bewerten Sie in diesem Studienbereich ihre bisherigen Leistungen im Vergleich zu Ihren Kommilitoninnen und Kommilitonen? Meine Studienleistungen sind insgesamt gesehen...
Selbsteinschätzung Physik	...im Fach Physik...
Selbsteinschätzung Physikdidaktik	...in Fachdidaktik Physik...
Selbsteinschätzung Erziehungswissenschaft	...in Bildungs- bzw. Erziehungswissenschaften...

Zusätzlich werden in diesem Bereich Skalen zur allgemeinen Selbstwirksamkeit der Probanden bezogen auf ihren Studienerfolg, bezogen auf das Unterrichten von Physik, sowie auf die Planung und Durchführung von Physikexperimenten im Unterricht erfasst. Die Effekte auf Fachwissen und fachdidaktisches Wissen sind in Tabelle 5.17 aufgelistet.

Tabelle 5.17: Regressionen zur Selbstwirksamkeit mit den Scores für Fachwissen und fachdidaktischem Wissen (**: $p < .001$; *: $p < .05$; +: $p < .10$).

	Itemzahl	Cronbachs α	Fachwissen	fachdidaktisches Wissen
Selbstwirksamkeit bezogen auf				
... Studienerfolg	4	.64	.25**	.25**
... Physikunterricht	4	.74	-.05	.02
... Planung von Experimenten	7	.86	.00	-.00
... Durchführung von Experimenten	7	.86	-.02	.08
Selbsteinschätzung				
... im Fach Physik	1	-	-.29**	-.21*
... in Physikdidaktik	1	-	-.05	-.22*
... in Erziehungswissenschaften	1	-	-.05	-.16*

Bezüglich der Selbstwirksamkeit der untersuchten Stichprobe zeigt sich ein uneinheitliches Bild. Die Skalen der Selbstwirksamkeit bezogen auf den Studienerfolg und die Selbsteinschätzungen bezüglich Physik, Physikdidaktik und Erziehungswissenschaft zeigen signifikante Effekte auf den untersuchten Wissensbereichen des Professionswissens. Bei den Skalen zur Selbsteinschätzung muss beachtet werden, dass die Likert-Skala, auf der die Probanden ihre Antwort angegeben haben, höhere Werte für negative Selbsteinschätzung vorzieht. Insofern sind hier jeweils negative Regressionswerte zum Fachwissen, zum fachdidaktischen Wissen und zum pädagogischen Wissen zu erwarten. Die Selbsteinschätzung bezüglich Physikdidaktik und Erziehungswissenschaften, hängen lediglich signifikant mit dem fachdidaktischen Wissen zusammen. Dieses Resultat scheint aufgrund der in Abschnitt 5.4 berichteten substantiellen Zusammenhänge zwischen fachdidaktischem- und pädagogischem Wissen durchaus plausibel.

Fraglich bleibt, warum sich für die Skalen zur Selbstwirksamkeit bezogen auf Unterrichten und Experimentieren weder signifikante Effekte auf das Fachwissen noch auf das fachdidaktische Wissen zeigen, zumal gezeigt werden kann, dass die Skalen der Selbstwirksamkeit bezogen auf Studienerfolg einen deutlichen Effekt haben. Ein ähnliches Bild bezüglich der Selbstwirksamkeit findet sich in der Arbeit von Riese (2009, S.157) in der ebenfalls keine signifikanten Zusammenhänge für Fachwissen oder fachdidaktisches Wissen mit der Selbstwirksamkeit der Probanden gefunden wurde.

(d) Zufriedenheit

Zuletzt sollen noch Variablen für die Zufriedenheit im Studium und für die Freude am Unterrichten untersucht werden. Dabei werden die Items zur Zufriedenheit im Studium und Begeisterung und Freude beim Unterrichten auf einer vierstufigen Likert-Skala erfasst (trifft gar nicht zu, trifft eher nicht zu, trifft eher zu, trifft völlig zu). Bei den beiden Items zu Begeisterung und Zufriedenheit im Studium gilt die Einschränkung, dass diese nur bei existierender Unterrichtserfahrung erfasst wurden. Tabelle 5.18 zeigt die Operationalisierungen dieser Items.

Tabelle 5.18: Operationalisierungen der Variablen für Zufriedenheit.

Variable	Operationalisierung
Begeisterung beim Unterrichten	Ich unterrichte mit Begeisterung im Fach Physik
Freude am Unterrichten	Mir macht Unterrichten im Fach Physik großen Spaß

Die Effekte dieser Variablen auf das Fachwissen und das fachdidaktische Wissen sind in Tabelle 5.19 dargestellt.

Tabelle 5.19: Regression von Variablen zur Zufriedenheit mit Fachwissen und fachdidaktischem Wissen (**: $p < .001$; *: $p < .05$; +: $p < .10$).

	Itemzahl	Cronbachs α	Fachwissen	fachdidaktisches Wissen
Studienzufriedenheit	4	.81	-.01	.05
Begeisterung in Bezug auf Unterrichten von Physik	1	-	.18*	.20*
Freude in Bezug auf Unterrichten von Physik	1	-	.16*	.15 ⁺

Für die Studienzufriedenheit zeigen sich hier keine signifikanten Effekte auf das Fachwissen oder das fachdidaktische Wissen. Geringe bis mittlere, signifikante Effekte dagegen für die Begeisterung und Freude bezogen auf das Unterrichten, sofern Unterrichtserfahrungen vorliegen.

(e) Lehr- und Lerngelegenheiten

Als Indikator für die allgemeine Quantität von Lerngelegenheiten wird zunächst die Semesterzahl herangezogen, da davon ausgegangen werden kann, dass mit zunehmendem Studienfortschritt auch die Anzahl an wahrgenommenen Lerngelegenheiten steigt. Daneben werden die Quantitäten von Un-

terrichtshospitationen und Unterrichtsplanung, sowie die Gelegenheit eigener Lehrerfahrung erfasst (siehe Tabelle 5.20).

Tabelle 5.20: Operationalisierungen der Variablen für Lehr- und Lerngelegenheiten.

Variable	Operationalisierung
Semesterzahl	In welchem Fachsemester studieren Sie das Fach Physik?
Hospitation	In wie vielen Unterrichtsstunden haben Sie im Verlauf Ihrer bisherigen Lehramtsausbildung hospitiert?
Unterrichtsplanung	Wie viele Unterrichtsstunden haben Sie im Verlauf ihrer bisherigen Lehramtsausbildung selbst geplant und gehalten?
Lehrerfahrung	Haben Sie bereits Erfahrungen mit eigenem Unterricht gemacht (bspw. in Praktika)?

Die Effekte dieser Variablen auf die Ausprägungen von Fachwissen und fachdidaktischem Wissen ist in Tabelle 5.21 aufgelistet.

Tabelle 5.21: Regression der Variablen für Lehr- und Lerngelegenheiten im Fach Physik mit den Scores für Fachwissen und fachdidaktischem Wissen (* : $p < .001$; * : $p < .05$; + : $p < .10$).

	Itemzahl	Fachwissen	fachdidaktisches Wissen
Semesterzahl	1	.34**	.23*
Hospitation	1	.43**	.43**
Unterrichtsplanung	1	.33**	.34**
Lehrerfahrung	1	.18*	.25**

Wie bereits in der Pilotierung der eingesetzten Instrumente dargestellt, zeigen sich stärkere Effekte der Semesterzahl auf Fachwissen als auf fachdidaktisches Wissen. Dies kann durch den höheren Anteil an Lerngelegenheiten des Fachwissens gegenüber dem fachdidaktischen Wissen erklärt werden. Betrachtet man hingegen die absolute Quantität von Hospitationen, Unterrichtsplanung und Lehrerfahrung, so ergeben sich praktisch keine Unterschiede mehr. Dies spricht dafür, dass im Studium allgemein mehr Fachwissen als fachdidaktisches Wissen vermittelt wird. Eine stärkere Ausprägung von Erfahrungen aus schulpraktischen Studien, zum Beispiel aus studienbegleitenden Schulpraktika, geht jedoch gleichermaßen mit einem höheren Fachwissen und fachdidaktischen Wissen einher. Praktische Lehrerfahrung führt offenbar, sogar verglichen mit dem Fachwissenszuwachs, zu einem stärkeren Zuwachs an fachdidaktischem Wissen.

Insgesamt stellen die Variablen für Lehr- und Lerngelegenheiten die stärksten Prädiktoren für fachdidaktisches Wissen dar. Beim Fachwissen werden sie lediglich von der allgemeinen kognitiven Leistungsfähigkeit, Operationalisiert durch die Abiturnote (siehe Tabelle 5.14) übertroffen.

Zusammenfassung

Die stärksten und Indikatoren für Fachwissen und fachdidaktisches Wissen werden, gemessen an der Effektstärke, in der folgenden Tabellen 5.22 und 5.23 dargestellt:

Tabelle 5.22: Liste der stärksten Prädiktoren für das Fachwissen im Fach Physik.

Personenmerkmal	Stärke des Effekts auf Fachwissen
Abiturnote	.46
Anzahl bisheriger Unterrichtshospitationen	.43
Semesterzahl	.34
Anzahl bisher geplanter Unterrichtsstunden	.33
Interesse an Physik	.31
Selbsteinschätzung des Studienerfolgs im Fach Physik	.29
Letzte Physiknote an der Schule	.26
Selbstwirksamkeit bezogen auf Studienerfolg	.25
Begeisterung für das Unterrichten von Physik	.18
Lehrerfahrung	.18
Freunde in Bezug auf das Unterrichten von Physik	.16

Tabelle 5.23: Liste der stärksten Prädiktoren für das fachdidaktische Wissen im Fach Physik.

Personenmerkmal	Stärke des Effekts auf fachdidaktisches Wissen
Anzahl bisheriger Unterrichtshospitationen	.43
Abiturnote	.35
Anzahl bisher geplanter Unterrichtsstunden	.34
Selbstwirksamkeit bezogen auf Studienerfolg	.25
Lehrerfahrung	.25
Semesterzahl	.23
Selbsteinschätzung des Studienerfolgs in Physikdidaktik	.22
Selbsteinschätzung des Studienerfolgs im Fach Physik	.21
Interesse an Physik	.21
Begeisterung für das Unterrichten von Physik	.20
Selbsteinschätzung des Studienerfolgs in Erziehungswissenschaften	.16
Letzte Physiknote an der Schule	.15

Geringe Effekte für beide Wissensbereiche liefern die Motivation im Studium und die Begeisterung für das Unterrichten. Für Selbstwirksamkeitserwartungen bezogen auf den Unterricht und das Experimentieren im Unterricht, sowie für die Zufriedenheit im Studium konnten keine signifikanten Effekte auf das Fachwissen oder das fachdidaktische Wissen gefunden werden.

6 Diskussion

Im Folgenden werden die Güte der gewonnenen Daten und die Ergebnisse der Untersuchung im Hinblick auf die Forschungsfragen diskutiert.

6.1 Güte der gewonnenen Daten

Reliabilität

Mit den im Rahmen der vorliegenden Arbeit neu entwickelten Testinstrumenten für das Fachwissen und das fachdidaktische Wissen wurden erstmals Instrumente vorgestellt, mit denen das Professionswissen im Fach Physik über sämtliche schulelevante Inhaltsbereiche hinweg erfasst werden kann. Die eingesetzten Aufgaben zeigten dabei für die untersuchte Stichprobe eine gute Passung zwischen Schwierigkeitsgrad und der Probandenfähigkeit. Trotz seiner inhaltlichen Breite und des damit verbundenen Reliabilitäts-Validitäts-Dilemmas (siehe Abschnitt 4.3.5; vgl. auch Gramzow, 2014, S.157) zeigen die Skalen für Fachwissen und fachdidaktisches Wissen hinsichtlich der gemessenen Werte für Cronbachs α gute ($\alpha_{CK} = .85$) bzw. zufriedenstellende ($\alpha_{PCK} = .74$) Reliabilitäten. Zur besseren Einschätzung dieser Bemaßung der Reliabilität kann ein Vergleich zu ähnlichen Testinstrumenten dienen: In der Untersuchung von Riese (2009) zeigte das Testinstrument für das Fachwissen eine Reliabilität von $\alpha = .81$ bei 28 Items und für das fachdidaktische Wissen von $.74$ bei 39 Items. In der COACTIV-Studie erreichte die Skala des fachdidaktischen Wissen mit 22 Items eine Reliabilität von $\alpha = .78$, die für das Fachwissen mit 13 Items $\alpha = .83$ (Kunter et al., 2011). Die Instrumente zur Erfassung des Professionswissens im Fach Biologie zeigten mit 19 Items $\alpha = .71$ für das Fachwissen und mit 9 Items von $\alpha = .86$ für das fachdidaktische Wissen (Großschedl et al., 2014). Damit liegt das im Rahmen der vorliegenden Arbeit neu entwickelte Instrument bezüglich der Reliabilität auf Augenhöhe mit anderen Instrumenten im Forschungsfeld. Die getroffene Kategorisierung der eingesetzten Aufgaben erlaubt zudem die gezielte Erfassung des Professionswissens unter Einbeziehung oder Ausschluss bestimmter Inhaltsbereiche. Für die Erfassung einzelner Inhaltsbereiche muss jedoch berücksichtigt werden, dass die Einzelskalen für sich genommen keine reliabel Messung erlauben. Diesem Umstand könnte jedoch durch Einbeziehung zusätzlicher Items in die Subskalen abgeholfen werden (siehe Abschnitt 5.1.1).

Itemschwierigkeit

Neben der Möglichkeit des gezielten Einsatzes von Items für bestimmte Inhalte lässt die im Abschnitt 5.3 dargestellte Raschskalierung der Items auf einer Skala mit den Probanden auch den gezielten Einsatz von Items für bestimmte Schwierigkeitsgrade zu. Hiermit kann verlässlich eine im Schwierigkeitsgrad angepasste Teilmenge von Items für die Testung einer bestimmten Stichprobe, beispielsweise schwierigere Items für die Testung des Fachwissens von Referendaren nach dem Abschluss der universitären Lehramtsausbildung, herangezogen werden. Auch hier muss jedoch berücksichtigt werden, dass eine Verringerung der Itemzahlen mit einer Verschlechterung der Reliabilität einhergehen kann, so dass auch für eine solche Testung die Einhaltung einer Mindestzahl von 20 Items anzuraten wäre, um eine akzeptable Verlässlichkeit der Ergebnisse zu gewährleisten.

Validität

Neben der Reliabilität der eingesetzten Testinstrumente konnten durch die im Abschnitt 4.4 dargestellten und im Abschnitt 5.1.2 diskutierten Schritte zahlreiche Hinweise für die Validität der Testergebnisse nach Messick (1995) generiert werden. Danach sind sechs Aspekte der Konstruktvalidität zu unterscheiden. (1) Die *inhaltliche Validität* umfasst die Passung zwischen dem zu untersuchenden Konstrukt und dem Testinstrument. Um hinsichtlich dieses Aspekts der Validität eine hohe Testgüte sicherzustellen, wurden im Vorfeld der Untersuchung die Literatur zum Professionswissen sowie gängige Lehrpläne analysiert (Kapitel 2). Auf dieser Analyse basierend wurde ein Modell des Professionswissens erstellt (Abschnitt 4.3), welches als Grundlage für die Aufgabenkonstruktion diente. Die Passung zwischen dem Modell und den Testaufgaben wurde durch eine standardisierte Aufgabenkonstruktion mittels Handreichung und anschließende Expertenratings abgesichert. (2) Die *kognitive Validität* umfasst den Anteil intendierter Denkprozesse bei der Aufgabenbearbeitung durch die Probanden. Dies konnte durch eine entwicklungsbegleitende Fragebogenstudie (Abschnitt 4.4.4) abgesichert und durch den Nachweis der überwiegenden Anwendung konstruktrelevanter Lösungsstrategien bei der Bearbeitung der Testaufgaben im Rahmen einer Studie zum Lauten Denken nachgewiesen werden (Abschnitt 4.4.6). (3) Die *strukturelle Validität* der Testergebnisse ist gegeben, wenn die Ergebnisse des Tests das zu messende Konstrukt repräsentieren und. Dies wurde durch die Analyse von Korrelationen der Testwerte mit erhobenen Begleitvariablen sichergestellt (Abschnitt 5.1.2). Dabei zeigten sich erwartete Zusammenhänge, sowohl zum Fachwissen als auch zum fachdidaktischen Wissen. Ferner konnte durch die Überprüfung der Aufgabenschwierigkeit die Adressatengerechtigkeit der Aufgabenstellungen nachgewiesen werden. (4) Die Testergebnisse sind *generalisierbar*, wenn sie auf die Grundgesamtheit der Zielgruppe verallgemeinert werden können. Um dies sicherzustellen, wurde die Testung an einer möglichst großen und bundesweit verteilten Probandengruppe durchgeführt. Die Probandengruppe deckt zudem alle Fachsemester des Lehramtsstudiums ab. Hinsichtlich der Testprobanden müssen allerdings auch eine Einschränkung gemacht werden. Da die zur Untersuchung genutzten Daten in einer querschnittlichen Untersuchung erhoben wurden, können keine Entwicklungsverläufe für einzelne Probanden betrachtet werden. Zur Abgrenzung des fachdidaktischen Wissens vom Fachwissen wäre zudem ein zusätzlicher Validierungsschritt durch einen Einsatz des Testinstruments an Fachphysikern sinnvoll (Kirschner, 2013; Gramzow, 2014). Als Ergebnis einer solchen Untersuchung sollte man anhand der Lerngelegenheiten im Studium erwarten, dass Studierende, welche kein Lehramt anstreben, im Test zur Erfassung des fachdidaktischen Wissens schwächer, im fachwissenschaftlichen Test dagegen stärker als Lehramtsstudierende abschneiden. Ein solcher Schritt wurde in der vorliegenden Arbeit nicht vorgenommen. Dies sollte allerdings durch die ausführliche Diskussion der Gesamtheit der eingesetzten Aufgaben durch Experten (vgl. Abschnitt 4.4.3) mit besonderer Expertise im Bereich Fachdidaktik zumindest teilweise kompensiert werden. Eine zusätzliche mögliche Einschränkung der Validität ergibt sich durch den Umstand, dass ein knappes Viertel (22,5% bzw. 42 von 187 erfassten Fällen) der Probanden aus der Stichprobe der Hauptstudie bereits an der Pilotierungsstudie im Rahmen der vorliegenden Arbeit teilgenommen hatte. Dabei spielt der Umstand eine Rolle, dass man üblicherweise nach den Testdurchführungen einen Austausch zwischen den Probanden über die abgefragten fachlichen und fachdidaktischen Inhalte beobachten kann. Sowohl die Bearbeitung des Tests als auch der beobachtete Austausch der Probanden nach einer Testung können zusätzliche Lerngelegenheiten darstellen. Für die Probanden, die sowohl an der Pilotierungsstudie, als auch an der Hauptstudie teilgenommen haben, kann dies zu einem besseren Abschneiden in der Hauptstudie geführt haben. Dieser Effekt wird jedoch im vorliegenden Fall durch zwei Umstände relativiert: Erstens waren die Testaufgaben in der Pilotierungsstu-

die auf zwei Testhefte aufgeteilt, so dass Probanden durch die Bearbeitung des Tests jeweils nur einen Teil der Aufgaben der Hauptstudie kannten. Zweitens lag zwischen Pilotierung und Hauptstudie eine Phase der Überarbeitung der Items, so dass sich zahlreiche Aufgaben beim Einsatz in Pilotierung und Hauptstudie voneinander unterschieden. Da die Teilnahme an der Pilotierungsstudie als Variable in der Hauptstudie mit erfasst wurde, kann man zeigen, dass die Gruppe der Probanden, die an der Pilotierung teilgenommen haben, mit durchschnittlich 30,0 Punkten gegenüber 28,6 Punkten im Fachwissen und 18,1 Punkten gegenüber 17,4 Punkten im fachdidaktischen Wissen tatsächlich leicht besser abschneidet. Eine Varianzanalyse (ANOVA) zeigt jedoch, dass diese Unterscheide weder für das Fachwissen noch für das fachdidaktische Wissen signifikant sind (jeweils $p > .30$). (5) Die *externe* Validität der umfasst die Abgrenzung des erfassten Konstrukts von anderen, nicht intendierten Konstrukten. Dies wurde, wie bei der strukturellen Validität durch die Analyse von Korrelationen der Testwerte mit Begleitvariablen sichergestellt, wobei durchweg plausible Ergebnisse beobachtet wurden (Abschnitt 5.1.2). (6) Den letzten Aspekt der Konstruktvalidität stellen schließlich die *Konsequenzen* dar, die sich durch die Testergebnisse ergeben können. Dies umfasst sowohl individuelle, als auch gesamtgesellschaftliche Konsequenzen. Individuellen Konsequenzen wurden dabei durch eine sorgfältige Anonymisierung der beteiligten Personen und die Einhaltung gängiger Regeln des Datenschutzes sichergestellt. Hinsichtlich möglicher gesellschaftlicher oder politischer Konsequenzen wurde im Vorfeld der Untersuchung mit den teilnehmenden Institutionen vereinbart, dass keine Vergleiche von Studienstandorten angestellt werden. Aus diesem Grund wurde darauf in der vorliegenden Arbeit verzichtet.

6.2 Forschungsfragen

Nachfolgend werden die Ergebnisse der vorliegenden Arbeit im Hinblick auf die drei Forschungsfragen, wie sie im Kapitel 3 formuliert wurden, diskutiert.

Die erste Forschungsfrage der vorliegenden Arbeit zielte auf die Untersuchung der Struktur des Professionswissens angehender Physiklehrkräfte auf der Basis einer inhaltlich breiten Erfassung ab. Hierzu wurden als Referenz (im Abschnitt 2.1.2) Befunde aus vorangegangenen Arbeiten über die Struktur des Professionswissens berichtet. Dabei zeigte sich stets eine empirische Trennbarkeit dreier Wissensbereiche innerhalb des Professionswissens: (1) Fachwissen, (2) fachdidaktisches Wissen und (3) pädagogisches Wissen. Trotz der Trennbarkeit ergaben sich in diesen Untersuchungen stets substantielle Zusammenhänge zwischen diesen Bereichen. Beispielsweise wurde zwischen CK und PCK im ProwiN-Projekt (Kirschner, 2014) eine manifeste Korrelation von $r = .45$ für Physiklehrkräfte gemessen; In der Arbeit von Riese (2009) von $r = .49$ für angehende Physiklehrkräfte. Die Befunde dieser Arbeiten, und aus den im Abschnitt 2.1.2 zitierten, beruhen jedoch entweder auf fachfremden Untersuchungen, beispielsweise in der COACTIV-Studie für Mathematik bzw. auf der Untersuchung von Lehrkräften (COACTIV und ProwiN) oder auf einer Fokussierung auf einzelne Aspekte des Fachwissens bzw. fachdidaktischen Wissens (ProwiN und Riese). Die Ergebnisse der vorliegenden Arbeit (aus Abschnitt 5.2) bestätigen die vorliegenden Befunde jedoch. Sowohl die empirische Trennung der drei Wissensbereiche als auch die Qualität der Zusammenhänge zwischen diesen Bereichen konnte bestätigt werden. Dabei zeigt sich, dass – entsprechend der theoretischen Natur des Konstrukts des fachdidaktischen Wissens als Amalgam aus Wissen über das Fach und Wissen über Pädagogik (Shulman, 1986; 1987) – sowohl Fachwissen als auch pädagogisches Wissen besonders stark mit dem fachdidaktischen Wissen zusammenhängen. Zwischen Fachwissen und pädagogischem Wissen ergibt sich dagegen ein mittlerer Zusammenhang.

Die zweite Forschungsfrage zielte auf die Entwicklung der Ausprägung des Fachwissens und des fachdidaktischen Wissens zu unterschiedlichen Studienphasen ab. Die Studienphasen wurden hier durch die Jahrgangsstufe im Studium für die ersten fünf Jahre operationalisiert. Die berichteten Entwicklungsverläufe (Abschnitt 5.3) zeigen dabei stärkeren Zuwachs beim Fachwissen gegenüber dem Zuwachs beim fachdidaktischen Wissen. Die Effektstärke beim Zuwachs des Fachwissens ist mit Cohens $d = 1.9$ gegenüber dem fachdidaktischen Wissen mit $d = 0.9$ etwa doppelt so hoch. Ferner findet der stärkste Zuwachs bei Fachwissen offenbar zu Beginn des Studiums statt, wohingegen das fachdidaktische Wissen in der zweiten Hälfte des Studiums besonders stark ansteigt. Dieser Befund entspricht der üblichen Verteilung von Lerngelegenheiten im Lehramtsstudium. Zum einen liegt der Fokus der Wissensvermittlung im Lehramtsstudium hauptsächlich auf der Vermittlung von Fachwissen. Zum anderen liegen Lerngelegenheiten für das fachdidaktische Wissen eher in der zweiten Hälfte des Lehramtsstudiums, wohingegen angehende Lehrkräfte zum Studienbeginn üblicherweise vor allem Fachvorlesungen hören. Diese Passung zwischen der Organisation der Lerngelegenheiten im Studium und der Entwicklung des Professionswissens deutet auf eine generelle Effektivität, sowohl der fachwissenschaftlichen als auch der fachdidaktischen Ausbildung an den an der Hauptstudie beteiligten Standorten hin. Neben dieser generellen Wirksamkeit der Ausbildung ist jedoch auch ein unerwünschter Effekt zu beobachten, der der Diskussion bedarf: Sowohl bei der Entwicklung des Fachwissens als auch des fachdidaktischen Wissens zeigt sich ein Rückgang der Ausprägungen vom vierten zum fünften Studienjahr. Hierfür ist die Kombination zweier Erklärungsmodelle denkbar: Erstens kann vermutet werden, dass Studierende sich am Ende ihres Studiums für das Bestehen von Abschlussprüfungen hinsichtlich ihres verfügbaren Professionswissens stark spezialisieren. Vorher im Studium erworbenes Professionswissen, welches für das Bestehen von Abschlussprüfungen nicht benötigt wird, kann dabei mangels Anwendungsgelegenheit verloren gehen („*use it or lose it*“). Diese Erklärung stellt die Stabilität des im Studium vermittelten Wissens in Frage. Zweitens kann der Rückgang durch einen Auswahleffekt erklärt werden. Es ist anzunehmen, dass Studierende mit hohem Professionswissen ihr Studium durchschnittlich schneller abschließen als Studierende mit geringem Professionswissen. Insofern wäre bei einem Messzeitpunkt im 9. oder 10. Fachsemester mit einem geringeren Anteil von Probanden mit hohem Professionswissen zu rechnen. Dieser Umstand liegt in der Natur einer querschnittlichen Datenerhebung und könnte durch eine längsschnittliche Untersuchung auf ihre Gültigkeit hin überprüft werden.

Die dritte Forschungsfrage befasst sich mit Faktoren, die die Entwicklung des Professionswissens angehender Physiklehrkräfte beeinflussen. Dies zielt besonders darauf ab, welche Studienumstände eine erwünschte Entwicklung des Professionswissens verstärken können. Dazu wurden potentielle Faktoren durch verschiedene Personenmerkmale operationalisiert und im Rahmen der Hauptstudie miterhoben. Da die potentiell beeinflussenden Faktoren in der vorliegenden Arbeit (im Abschnitt 5.4) anhand von Korrelationen zwischen diesen Personenmerkmalen und den Ausprägungen des Fachwissens und fachdidaktischen Wissens untersucht wurden, können aus den Ergebnissen zwar Zusammenhänge, aber nur eingeschränkt Kausalitäten abgeleitet werden. Zudem betreffen einige der operationalisierenden Variablen generell unbeeinflussbare, aber wichtige Faktoren, wie zum Beispiel die kognitiven Grundfähigkeiten der Probanden, oder zumindest im Rahmen des Universitätsstudiums unbeeinflussbare Faktoren, wie zum Beispiel die schulische Vorbildung. Dennoch lassen sich aus den Befunden wichtige Erkenntnisse schließen. So konnte gezeigt werden, dass neben den Lerngelegenheiten im Studium auch Lehrgelegenheiten durch schulpraktische Studien mit einem höheren Professionswissen einhergehen. Ferner hängt auch eine positive Selbsteinschätzung im Hinblick auf Fähigkeiten und Studienerfolg sowohl mit dem Fachwissen als auch mit dem fachdidaktischen Wis-

sen signifikant zusammen. Zuletzt scheint besonders die Aufrechterhaltung des Interesses am Fach Physik, sowie die Zufriedenheit und die Motivation im Lehramtsstudium Potential für die Begünstigung einer positiven Entwicklung des Professionswissens zu liefern.

7 Zusammenfassung und Ausblick

Zusammenfassung der Ergebnisse

Die Ergebnisse der vorliegenden Studie bestätigen, dass es sich beim *Fachwissen*, beim *fachdidaktischen Wissen* und beim *pädagogischen Wissen* angehender Physiklehrkräfte um voneinander abgrenzbare Konstrukte handelt, deren Ausprägungen unabhängig voneinander gemessen werden können. Dabei zeigen sich stärkere Zusammenhänge zwischen Fachwissen und fachdidaktischem Wissen bzw. fachdidaktischem Wissen und pädagogischem Wissen und ein schwächerer Zusammenhang zwischen Fachwissen und pädagogischem Wissen. Die positive Entwicklung in diesen Bereichen konnte für die ersten vier Studienjahre ebenfalls nachgewiesen werden, was für die generelle Wirksamkeit des Lehramtsstudiums spricht. Personenmerkmale, die diese positive Entwicklung maßgeblich fördern können sind vor allem eine Einbindung schulpraktischer Studien in den Lernprozess, als auch eine positive Einstellung der angehenden Physik Lehrkraft im Hinblick auf den Studienerfolg und das Unterrichten.

Schlussfolgerungen für das Lehramtsstudium

Aus den Erkenntnissen über die Struktur und Entwicklung des Professionswissens, sowie auf die Untersuchung von Zusammenhängen mit Begleitfaktoren lassen sich Schlussfolgerungen auf eine sinnvolle Strukturierung der Lehrerinnen- und Lehrerausbildung machen. So kann aus der Bestätigung des Befunds über die theoretische Natur des fachdidaktischen Wissens als Konglomerat aus Fachwissen und pädagogischem Wissen geschlossen werden, dass sowohl eine hinreichende Ausprägung des Fachwissens als auch des pädagogischen Wissens notwendige Voraussetzungen für die Entwicklung fachdidaktischen Wissens darstellen. Die Kombination aus Fachwissen, pädagogischem Wissen und dem Vorhandensein von Lerngelegenheiten für das fachdidaktische Wissen stellt schließlich eine hinreichende Bedingung für dessen Entwicklung dar. Insofern erscheint die bestehende Organisation des Studiums, bei der zunächst auf die Vermittlung von Fachwissen und erst im späteren Studienverlauf verstärkt auf die aufbauende Vermittlung fachdidaktischen Wissens fokussiert wird, als sinnvoll.

Daneben deuten die Daten darauf hin, dass ein stärkerer Praxisbezug der Lehrerinnen- und Lehrerausbildung durch Einbindung von Lehrgelegenheiten in das Studium sinnvoll wäre. Wie gezeigt wurde, haben solche Lehrgelegenheiten das Potential sowohl die Ausprägung des fachdidaktischen Wissens aber auch des Fachwissens zu begünstigen. Hier muss allerdings einschränkend bedacht werden, dass der Zusammenhang zwischen der Quantität von Lehrgelegenheiten und der Ausprägung von Fachwissen und fachdidaktischem Wissen auch durch die dahinterliegende Variable des Studienfortschritts moderiert sein kann, da mit zunehmender Studiendauer – wie gezeigt – das Professionswissen bis einschließlich zum vierten Studienjahr kontinuierlich ansteigt und sich auch mehr Lehr- und Lerngelegenheiten bieten.

Ausblick

Befunde der Erforschung des Professionswissens von Referendaren und Lehrkräften zeigen, dass auch nach Abschluss der universitären Ausbildungsphase ein signifikanter Wissenserwerb, besonders im Bereich des fachdidaktischen Wissens, stattfindet (vgl. Kirschner, 2014, S.113). Die Tatsache, dass sich im Zuge der ProWiN-Studie gezeigt hat, dass Testinstrumente, die für die Erfassung des Professi-

onswissens von Lehrkräften in den Naturwissenschaften entwickelt wurden, auch für Studierende und Referendare geeignet sind (Borowski et al., 2010), kann man davon ausgehen, dass diese Kompatibilität umgekehrt in gleicher Weise gilt. Es wäre daher wünschenswert, das in dieser Arbeit entwickelte Testinstrument, über den Umfang der Erhebungen in der vorliegenden Studie hinaus, an weiteren Zielgruppen einzusetzen. Dies würde eine umfassendere Abbildung der Entwicklung professionellen Wissens über die Studienphase hinaus ermöglichen und könnte beispielweise die Frage klären, ob und ggf. wie sich die spezifischen Wissensbereiche Fachwissen, fachdidaktisches Wissen und pädagogisches Wissen im Laufe der gesamten Lehrerinnen- und Lehrerausbildung (also inklusive Referendariat) sowie im Laufe des Einsatzes im Schuldienst (also inklusive Lehrtätigkeit) weiterentwickeln. Einen weiteren Anschlusspunkt für Folgeuntersuchungen stellt ferner die Validierung der vorgestellten Testinstrumente durch einen Einsatz an fachfremden Studierenden für das Instrument zum Fachwissen und durch einen Einsatz an Fachwissenschaftlern im Fach Physik für das Instrument zum fachdidaktischen Wissen dar.

Zur weiteren Validierung des Testinstruments wäre es wünschenswert, Generalisierbarkeit der vorliegenden Ergebnisse, die auf einer querschnittlichen Untersuchung beruhen, weiter abzusichern. Zudem bleibt die Frage nach der Ursache für den gemessenen Rückgang bei der Ausprägung des Professionswissens zum Studienende hin offen. Diesen Umständen sollte durch eine echte längsschnittliche Untersuchung abgeholfen werden. Eine solche Untersuchung hat mit dem Start des IPN-Projekts *Kompetenzentwicklung in mathematischen und naturwissenschaftlichen Lehramtsstudiengängen* (KeiLa) begonnen. In diesem durch die Leibniz-Gemeinschaft geförderten Projekt werden die im vorliegenden Projekt entwickelten Messinstrumente weiter verwendet. Über einen Zeitraum von drei Jahre werden die Instrumente dabei in längsschnittlichen Erhebungen für zwei Kohorten eingesetzt. Bei der einen Kohorte handelt es sich zunächst um Studienanfänger im Bachelor-Studium, bei der anderen um fortgeschrittene Studierende im Master-Studiengang.

8 Literaturverzeichnis

- AAAS (2007). American Association for the Advancement of Science (Hrsg.): *Getting Assessment Right*. 2061 today, 17 (1).
- Abell, S.K. (2007). Research on Science Teacher Knowledge. In S.K.Abell (Ed.) *Handbook of research on science teacher education*, 1105-1149.
- Adams, P. E., & Krockover, G. H. (1997). Beginning science teacher cognition and its origins in the preservice secondary science teacher program. *Journal of research in Science Teaching*, 34(6), 633-653.
- Alonso, M., Finn, E. (1992). *Physics*. Addison-Wesley.
- Ball, D.L., Hill, H.C., Bass, H. (2005). Knowing mathematics for teaching. *American Educator*, Fall, 14-46.
- Baumert, J., Stanat, P., & Demmrich, A. (2001). PISA 2000: Untersuchungsgegenstand, theoretische Grundlagen und Durchführung der Studie. In *PISA 2000* (pp. 15-68). VS Verlag für Sozialwissenschaften.
- Baumert, J., Krauss, S., Brunner, M. (2003). COACTIV: Cognitive Activation in the Classroom: The Orchestration of Learning Opportunities for the Enhancement of Insightful Learning in Mathematics. Sonderdruck aus: *Max Planck Institute for Human Development (2003). Research Report 2001-2002*. Berlin, 100-105.
- Baumert, J., Blum, W., & Neubrand, M. (2004). Drawing the lessons from PISA 2000. *Zeitschrift für Erziehungswissenschaft*, 7(3), 143-157.
- Baumert, J. & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 9 (4), 469-520.
- Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., Klusmann, U., Krauss, S., Neubrand, M. & Tsai Y.-M. (2010). Teachers' Mathematical Knowledge, Cognitive Activation in the Classroom, and Student Progress. *American Educational Research Journal*, 47(1), 133–180.
- 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. 29-53).
- Baumert, J., & Kunter, M. (2011b). Das mathematikspezifische Wissen von Lehrkräften, kognitive Aktivierung im Unterricht und Lernfortschritte von Schülerinnen und Schülern 1. *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV*, 163.
- Baumgartner, H., & Homburg, C. (1996). Applications of structural equation modeling in marketing and consumer research: A review. *International journal of Research in Marketing*, 13(2), 139-161.
- Beggow, E.P.; Ha, M.; Nehm, R. H.; Pearl, D.; Boone, W. J. (2013). Assessing Scientific Practices Using Machine-Learning Methods: How Closely Do They Match Clinical Interview Performance. *Journal of Science Education*. Springer. New York.

- Beichner, R. J. (1994). Testing student interpretation of kinematics graphs. *American journal of Physics*, 62(8), 750-762.
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological bulletin*, 107(2), 238.
- Bentler, P. M. (1992). On the fit of models to covariances and methodology to the Bulletin.. *Psychological bulletin*, 112(3), 400.
- Berliner, D. C. (2001). Learning about and learning from expert teachers. *International journal of educational research*, 35(5), 463-482.
- Berry, A., Friedrichsen, P., & Loughran, J. (Eds.). (2015). *Re-examining Pedagogical Content Knowledge in Science Education*. Routledge.
- Black, P., & Wiliam, D. (2006). *Inside the black box: Raising standards through classroom assessment*. Granada Learning.
- Blömeke, S. (2008). *Professionelle Kompetenz angehender Lehrerinnen und Lehrer*. Waxmann Verlag.
- Blömeke, S., Felbrich, A., Müller, C., Kaiser, G., & Lehmann, R. (2008a). Effectiveness of teacher education. *Zeitschrift für Didaktik der Mathematik*, 40(5), 719-734.
- Blömeke, S., Felbrich, A., & Müller, C. (2008b). *Theoretischer Rahmen und Untersuchungsdesign*. In: Blömeke, S., Kaiser, G., Lehmann, R. (Hrsg.): *Professionelle Kompetenz angehender Lehrerinnen und Lehrer. Wissen, Überzeugungen und Lerngelegenheiten deutscher Mathematikstudierender und –referendare*. Waxmann, 15-48.
- Blömeke, S., Suhl, U., Kaiser, G., Felbrich, A., Schmotz, C., & Lehmann, R. (2010). Lerngelegenheiten und Kompetenzerwerb angehender Mathematiklehrkräfte im internationalen Vergleich. *Unterrichtswissenschaft*, 38(1), 29-50.
- Boone, W.J., Staver, J.R., Yale, M.S. (2014). *Rasch Analysis in the Human Sciences*. Springer.
- Borko, H., & Putnam, R. T. (1996). *Learning to teach*.
- Borowski, A., Neuhaus, B. J., Tepner, O., Wirth, J., Fischer, H.E., Leutner, D., Sandmann, A. & Sumfleth, E. (2010a). Professionswissen von Lehrkräften in den Naturwissenschaften (ProwiN) – Kurzdarstellung des BMBF-Projekts. *Zeitschrift für Didaktik der Naturwissenschaften*, 16, 341-349.
- Borowski, A., Olszewski, J., & Fischer, H. E. (2010b). Fachdidaktisches Wissen von Physikreferendaren. *Der mathematische und naturwissenschaftliche Unterricht*, 63(5), 260-263.
- Borowski, A., Kirschner, S., Liedtke, S., & Fischer, H. E. (2011). Vergleich des Fachwissens von Studierenden, Referendaren und Lehrenden in der Physik. *PhyDid A-Physik und Didaktik in Schule und Hochschule*, 1(10), 01-09.
- Boysen, G., Glunde, H., Heise, H., Muckenfuß, H., Schepers, H., Wiesmann, H.-J. (1990). *Physik für Gymnasien 1 – Länderausgabe C*. Schroedel-Verlag.
- Boysen, G., Glunde, H., Heise, H., Muckenfuß, H., Schepers, H., Schepers, H., Wiesmann, H.-J. (1995). *Physik für Gymnasien 2 – Länderausgabe C*. Schroedel-Verlag.

- Breit, S., Schwantner, U., & Grafendorfer, A. (2007). *PISA 2006: internationaler Vergleich von Schülerleistungen*. Leykam.
- Bromme, R. (1992). Der Lehrer als Experte. Zur Psychologie des professionellen Lehrerwissens.
- Bromme, R. (1994). Beyond subject matter: A psychological topology of teachers' professional knowledge. *Didactics of mathematics as a scientific discipline*, 73-88.
- Bromme, R. (1997). Kompetenzen, Funktionen und unterrichtliches Handeln des Lehrers. In F. E. Weinert (Hrsg.), *Psychologie des Unterrichts und der Schule. Enzyklopädie der Psychologie, Serie I*, Bd. 3, 177-212.
- Bromme, R., & Rambow, R. (2000). *Experten-Laien-Kommunikation als Gegenstand der Expertiseforschung: Für eine Erweiterung des psychologischen Bildes vom Experten*.
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. *Sage Focus Editions*, 154, 136-136.
- Bruckner, J., Günther, A. (2013). *Pisa-Schock hilft benachteiligten Schülern*. Süddeutsche Zeitung SZ.de. 3. Dezember 2013.
- Brunner, M., Kunter, M., Krauss, S., Klusmann, U., Baumert, J., Blum, W., Neubrand, M., Dubberke, T., Jordan, A., Löwen, K. & Tsai, Y.-M. (2006). *Die professionelle Kompetenz von Mathematiklehrkräften: Konzeptualisierung, Erfassung und Bedeutung für den Unterricht – Eine Zwischenbilanz des COACTIV-Projekts*. In M. Prenzen & L. Allolio-Näcke (Hrsg.), *Untersuchungen zur Bildungsqualität von Schule*, 54-82.
- Bühner, M. (2006). *Einführung in die Test- und Fragebogenkonstruktion*. Pearson Studium Verlag.
- Carlson, M., & Mulaik, S. A. (1993). Trait ratings from descriptions of behavior as mediated by components of meaning. *Multivariate Behavioral Research*, 28(1), 111-159.
- Carlson, J., Stokes, L., Helms, J., Gess-Newsome, J., & Gardner, A. (2015). THE PCK SUMMIT. *Re-examining Pedagogical Content Knowledge in Science Education*, 14.
- Cauet, E., Liepertz, S., Kirschner, S., Borowski, A., & Fischer, H. E. (2013). Professionswissen von Physiklehrkräften und Schülerleistung. *Naturwissenschaftliche Bildung zwischen Science- und Fachunterricht. Gesellschaft für Didaktik der Chemie und Physik, Jahrestagung in München*.
- Cochran, K.F., DeRuiter, J.A., King, R.A. (1993). Pedagogical content knowledge: An integrative model for teacher preparation. *Journal of Teacher Education*, 44, 263-272.
- Cochran, K. F. (1997). Pedagogical content knowledge: Teachers' integration of subject matter, pedagogy, students, and learning environments. *Research matters to the Science Teacher*, 14, 97-102.
- Collins, A., Greeno, J., Resnick, L. B., Berliner, B., & Calfee, R. (1992). Cognition and learning. B. Berliner & R. Calfee, *Handbook of Educational Psychology*, New York: Simon & Shuster MacMillan.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *psychometrika*, 16(3), 297-334.
- Deffner, G. (1984). *Lautes Denken – Untersuchung zur Qualität eines Datenerhebungsverfahrens*. Frankfurt a.M.: Verlag Peter Lang.
- Demtröder, W. (2012). *Experimentalphysik 1: Mechanik und Wärme*. Springer Lehrbuch, 6. Auflage.

- Demtröder, W. (2013a). *Experimentalphysik 2: Elektrizität und Optik*. Springer Lehrbuch, 6. Auflage.
- Demtröder, W. (2010). *Experimentalphysik 3: Atome, Moleküle und Festkörper*. Springer Lehrbuch, 4. Auflage.
- Demtröder, W. (2013b). *Experimentalphysik 4: Kern-, Teilchen- und Astrophysik*. Springer Lehrbuch, 4. Auflage.
- 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.
- Deutscher Bildungsserver (2016). Zugelassene Lernmittel und Schulbücher. URL: <http://www.bildungsserver.de/Zugelassene-Lernmittel-und-Schulbuecher-522.html>. Aufgerufen am 03.07.2016.
- Diel, B., Erb, R., Lindner, K., Schmalhofer, C., Schön, L.-H., Tillmanns, P., Winter, R. (2009). *Physik Oberstufe: Gesamtband*. Cornelsen, 2009.
- Ding, L., Chabay, R., Sherwood, B. & Beichner, R. (2006). Evaluating an electricity and magnetism assessment tool: Brief electricity and magnetism assessment. In: *Physics Education Research*, 2, 010105.
- Doll, J., & Prenzel, M. (2001). Das DFG-Schwerpunktprogramm „Bildungsqualität von Schule“ (BIQUA): Bedingungen mathematischer, naturwissenschaftlicher und überfachlicher Kompetenzen. *TIMSS–Impulse für schule und Unterricht*, 5, 99-103.
- Dorn, F. & Bader, F. (2000). *Physik Gymnasium Gesamtband SekII*. Schroedel GmbH.
- Dorn, F. & Bader, F. (2001). *Physik in einem Band SI + SII – allgemeine Ausgabe 2001*. Schroedel GmbH.
- Douvdevany, O., Dreyfus, A., & Jungwirth, E. (1997). Diagnostic instruments for determining junior high-school science teachers' understanding of functional relationships within the 'living cell'. *International Journal of Science Education*, 19(5), 593-606.
- Dreisigacker, E. (1990). SPIEGEL-Rangliste Physik: Fakten, Analyse, Meinungen. *Physikalische Blätter*, 46(5), 158-160.
- Einhaus, E. & Schecker, H (2005). Thermodynamics concept inventory - Development of an standardized test for thermodynamics. In: *Proceedings of the Fifth International ESERA Conference on Contributions of Research to Enhancing Students' Interest in Learning Science*, Barcelona.
- Elbaz, F. (1983). *Teacher Thinking. A Study of Practical Knowledge*. Croom Helm Curriculum Policy and Research Series. Nichols Publishing Company, 155 West 72nd Street, New York, NY 10023.
- Ergönenc, J., Neumann, K. & Fischer, H. E. (2014). The Impact of Pedagogical Content Knowledge on Cognitive Activation and Student Learning. In H. E. Fischer, P. Labudde, K. Neumann & J. Viiri (Hrsg.), *Quality of Instruction in Physics – Results from a trinational video study* (S. 145-160). Münster: Waxmann.
- Ericsson, K.A. & Simon, H.A. (1980). Verbal Reports as Data. *Psychological Review*, 87 (3), S. 215-251.

- Ericsson, K.A. & Simon, H.A. (1993). *Protocol Analysis: Verbal Reports as Data*, Revised Edition. Cambridge: The MIT Press.
- Fahrholz, B. (Ed.). (2002). *Nach dem PISA-Schock: Plädoyers für eine Bildungsreform*. Hoffmann und Campe.
- FBP (2013). *Neufassung der Fachspezifischen Bestimmungen für Physik als Fach eines Studienganges mit dem Abschluss "Bachelor of Science" (B.Sc.) vom 06.03.2013*. Herausgeber: Der Präsident der Universität Hamburg, Referat 31 – Qualität und Recht. Nr. 2 vom 8. Januar 2014.
- Fernandez-Balboa, J.M., Stiehl, J. (1995). The generic nature of pedagogical content knowledge among college professors. *Teaching and Teacher Education*, 11(3), 293-306.
- Feynman, R.P., Leighton, R.B., Sands, M., Gottlieb, M.A., Leighton, R. (2009). *Feynman Vorlesungen über Physik*. Oldenbourg Wissenschaftsverlag.
- Fischer, H. E., Borowski, A. (2009). Professionswissen und Fortbildung von Physiklehrern. In: Kircher, E., Girwidz, R., & Häußler, P.(Hrsg.). *Physikdidaktik* (pp. 689-707). Springer Berlin Heidelberg.
- Fischer, H. E., Borowski, A., Kauertz, A., & Neumann, K. (2010). Fachdidaktische Unterrichtsforschung–Unterrichtsmodelle und die Analyse von Physikunterricht. *Zeitschrift für Didaktik der Naturwissenschaften*, 16, 59-75.
- Fischer, H. E., Borowski, A., & Tepner, O. (2012). Professional knowledge of science teachers. In *Second international handbook of science education* (pp. 435-448). Springer Netherlands.
- FPO (2007). Fachprüfungsordnung (Satzung) der Mathematisch-Naturwissenschaftlichen-Fakultät der Christian-Albrechts-Universität zu Kiel für Studierende der Zwei-Fächer-Bachelor- und Masterstudiengänge Physik mit den Abschlüssen Bachelor of Arts (B.A.), Bachelor of Science (B.Sc.) und Master of Education (M.Ed.) (Fachprüfungsordnung Physik (2-Fächer)). 29.11.2007.
- FPO (2011). Fachprüfungsordnung für das Studienfach Physik im Bachelor Studiengang mit Lehramtsoption Gymnasien/Gesamtschulen an der Universität Duisburg-Essen. 30.08.2011.
- Friege, G., & Lind, G. (2003). Allgemeine und fachspezifische Problemlösekompetenz. *Zeitschrift für Didaktik der Naturwissenschaften*, 9, 63-74.
- Friege, G., & Lind, G. (2004). Aus Bildung und Wissenschaft-Leistungsmessung im Leistungskurs. *Mathematische und Naturwissenschaftliche Unterricht*, 57(5), 259-265.
- Gardner A, Gess-Newsome,J (2011). *A PCK Rubric to Measure Teachers' Knowledge of Inquiry-Based Instruction Using Three Data Sources*. Annual Meeting of the National Association for Research in Science Teaching, Orlando, April 6, 2011.
- Geddis, A.N., Onslow, B., Beynon, C., Oesch, J. (1993). Transforming content knowledge: Learning to teach about isotopes. *Science Education*, 77(6), 575-591.
- Geller, C., Neumann, K., Boone, W.J. & Fischer, H.E. (2014). What makes the finnish different in science? Assessing and comparing students' science learning in three countries. *International Journal of Science Education*, 36(18), 3042-3066.
- George, D. (2011). *SPSS for windows step by step: A simple study guide and reference, 17.0 update, 10/e*. Pearson Education India.

- Gerthsen, C., Meschede, D., & Vogel, H. (2006). *Gerthsen Physik*. Springer-Verlag, 23. Auflage.
- Gess-Newsome, J., & Carlson, J. (2013, September). The PCK summit consensus model and definition of pedagogical content knowledge. In *the Symposium "Reports from the Pedagogical Content Knowledge (PCK) Summit, ESERA Conference*.
- Gess-Newsome, J. (2015). A model of teacher professional knowledge and skill including PCK: Results of the thinking from the PCK Summit, *Re-examining Pedagogical Content Knowledge in Science Education*, 28-42.
- Getzels, J. W., & Jackson, P. W. (1963). The teacher's personality and characteristics. *Handbook of research on teaching*, 506-582.
- Gramzow, Y., Riese, J., Reinhold, P. (2013). Modellierung fachdidaktischen Wissens angehender Physiklehrkräfte, *Zeitschrift für Didaktik der Naturwissenschaften*.
- Gramzow, Y. (2015). *Fachdidaktisches Wissen von Lehramtsstudierenden im Fach Physik: Modellierung und Testkonstruktion*. Logos Berlin.
- Grehn, J. & Krause J. (2007). Metzler Physik Schülerband SII. Schroedel GmbH.
- Grossman, P.L. (1990). The making of a teacher: Teacher knowledge and teacher education. *Teachers College Press*.
- Großschedl, J., Mahler, D., Kleickmann, T., & Harms, U. (2014). Content-related knowledge of biology teachers from secondary schools: Structure and learning opportunities. *International Journal of Science Education*, 36(14), 2335-2366.
- Großschedl, J., Harms, U., Kleickmann, T., & Glowinski, I. (2015). Preservice Biology Teachers' Professional Knowledge: Structure and Learning Opportunities. *Journal of Science Teacher Education*, 26(3), 291-318.
- Hadenfeldt, J. C., Repenning, B., & Neumann, K. (2014). Die kognitive Validität von Ordered Multiple Choice Aufgaben zur Erfassung des Verständnisses von Materie. *Zeitschrift für Didaktik der Naturwissenschaften*, 20(1), 57-68.
- Haidar, A. H. & Abraham, M. R. (1991). A Comparison of Applied and Theoretical Knowledge of Concepts Based on the Particular Nature of Matter. In: *Journal of research in science teaching*. Vol 28, NO.10, 919-938.
- Hain, J. (2011). Statistik mit R-Grundlagen der Datenanalyse, RRZN-Handbuch, 1. Auflage, Mai.
- Halliday, D., Resnick, R., & Walker, J. (2013). *Fundamentals of physics extended*. 10. Auflage. John Wiley & Sons.
- Härtig, H., Kauertz, A., & Fischer, H. E. (2012). Das Schulbuch im Physikunterricht. *Mathematische und Naturwissenschaftliche Unterricht*, 65(4), 197.
- Hascher, T. (2011). Forschung zur Wirksamkeit der Lehrerbildung. *Handbuch der Forschung zum Lehrerberuf*, 418-440.
- Hashweh, M.Z. (1987). Effects of subject-matter knowledge in the teaching of biology and physics. *Teaching & Teacher Education*, Vol. 3, No.2, 109-120.

- Hashweh, M.Z. (2005). Teacher pedagogical constructions: a reconfiguration of pedagogical content knowledge. *Teachers and Teaching: Theory and Practice*, 11(3), 273-292.
- Hattie, J. (2003). *Teachers Make a Difference, What is the research evidence?*.
- Hattie, J. (2005). What is the nature of evidence that makes a difference to learning?. *2005-Using data to support learning*, 7.
- Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Abingdon: Routledge.
- Helmke, A. (2006). Was wissen wir über guten Unterricht. *Pädagogik*, 2, 42-45.
- Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force Concept Inventory. *The Physics Teacher*, 30, 141-158.
- Hill, H.C., Rowan, B., Ball, D. (2005). Effects of Teachers Mathematical Knowledge for Teaching on Student Achievement. *American Educational Research Journal*. Vol. 42, No. 2, 371-406.
- Hohenstein, D. P. F., Zimmermann, F., Kleickmann, T., Köller, O., & Möller, J. (2014). Sind die bildungswissenschaftlichen Standards für die Lehramtsausbildung in den Curricula der Hochschulen angekommen?. *Zeitschrift für Erziehungswissenschaft*, 17(3), 497-507.
- Hohenstein, F., Kleickmann, T., Zimmermann, F., Köller, O. & Möller, J. (2017). *Erfassung von pädagogischem und psychologischem Wissen in der Lehramtsausbildung: Entwicklung eines Messinstruments*. *Zeitschrift für Pädagogik*, 1, 91-113.
- Homburg, C., Herrmann, A., Pflesser, C., & Klarman, M. (2008). Methoden der Datenanalyse im Überblick. *Handbuch Marktforschung—Methoden, Anwendungen, Praxisbeispiele*, 3, 151-173.
- Hu, L. T., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological methods*, 3(4), 424.
- Hübner, A. (2013). *Untersuchung der kognitiven Validität von Aufgaben zur Testung des professionellen Wissens von Lehramtsstudierenden*. Unveröffentlichte Masterarbeit, Christian-Albrechts-Universität zu Kiel.
- IBM (2011). IBM SPSS STATISTIC program, version 19 statistical software packages. *IBM corporation, New York*.
- Ifenthaler, D. (2006). *Diagnose lernabhängiger Veränderung mentaler Modelle*. Dissertation. Freiburg.
- Jüttner, M., & Neuhaus, B. J. (2013). Validation of a paper-and-pencil test instrument measuring biology teachers' pedagogical content knowledge by using think-aloud interviews. *Journal of Education and Training Studies*, 1(2), 113-125.
- 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.
- Kennedy, M. M., Ahn, S. & Choi, J. (2008). The value added by teacher education. *Handbook of research on teacher education*, 3, 1249-1273.

- Kerstan, T. (2011). *Der heilsame Schock – Zehn Jahre nach der Veröffentlichung der ersten Pisa-Studie. Was bleibt?* Zeit Online, 02.12.2011.
- Kind, V. (2009). Pedagogical content knowledge in science education: perspectives and potential for progress. *Studies in science education*, 45(2), 169-204.
- Kircher, E., Girwidz, R., & Häußler, P. (2009). Professionswissen und Fortbildung von Physiklehrern. In *Physikdidaktik* (pp. 689-707). Springer Berlin Heidelberg.
- Kirschner, S. (2013). *Modellierung und Analyse des Professionswissens von Physiklehrkräften* (Vol. 161). Logos Verlag Berlin GmbH.
- Kleickmann, T., & Anders, Y. (2011). Lernen an der Universität. In *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV* (pp. 305-315). Waxmann.
- Kleickmann, T., Großschedl, J., Harms, Z., Heinze, A., Herzog, S., Hohenstein, F., Köller, O., Kröger, J., Lindmeier, A., Loch, C., Mahler, D., Möller, J., Neumann, K., Parchmann, I., Steffensky, M., Taskin, V. & Zimmermann, F. (2014). Professionswissen von Lehramtsstudierenden der mathematisch-naturwissenschaftlichen Fächer-Testentwicklung im Rahmen des Projekts KiL. *Unterrichtswissenschaft*, 42(3), 280-288.
- Klieme, E. (2004). *Was sind Kompetenzen und wie lassen sie sich messen?* (pp. 10-13).
- Klieme, E., Jude, N., Baumert, J., & Prenzel, M. (2010). PISA 2000–2009: Bilanz der Veränderungen im Schulsystem. In *PISA 2009. Bilanz nach einem Jahrzehnt* (pp. 277-300).
- Klusmann, U. (2011). Individuelle Voraussetzungen von Lehrkräften. In *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV* (pp. 298-304). Waxmann.
- KMK (1979). *Vereinbarung über Einheitliche Prüfungsanforderungen in der Abiturprüfung*. Beschluss der Kultusministerkonferenz vom 01.06.1979 in der Fassung vom 24.10.2008.
- KMK (2005). *Bildungsstandards im Fach Physik für den Mittleren Schulabschluss*. Beschluss der Kultusministerkonferenz vom 16.12.2004. München: Luchterhand.
- KMK (2008). *Ländergemeinsame inhaltliche Anforderungen für die Fachwissenschaften und Fachdidaktiken in der Lehrerbildung*. Beschluss der Kultusministerkonferenz vom 16.10.2008 in der Fassung vom 16.09.2010.
- König, J., Blömeke, S., Paine, L., Schwidt, W., & Hsieh, F. (2011). General Pedagogical Knowledge of Future Middle School Teachers: On the Complex Ecology of Teacher Education in the United States, Germany, and Taiwan. *Journal of Teacher Education*, 62(2), 188-201.
- Krauss, S., Kunter, M., Brunner, M., Baumert, J., Blum, W., Neubrand, M., Jordan, A. & Löwen, K. (2004). *COACTIV: Professionswissen von Lehrkräften, kognitiv aktivierender Mathematikunterricht und die Entwicklung von mathematischer Kompetenz*.
- Krauss, S., Baumert, J., Blum, W. (2008). Secondary mathematics teachers' pedagogical content knowledge and content knowledge: validation of the COACTIV constructs. *ZDM Mathematics Education*, (40), 873-892.
- Krauss, S., Neubrand, M., Blum, W., Baumert, J., Brunner, M., Kunter, M., & Jordan, A. (2008). Die Untersuchung des professionellen Wissens deutscher Mathematik-Lehrerinnen und-Lehrer im Rahmen der COACTIV-Studie. *Journal für Mathematik-Didaktik*, 29(3-4), 233-258.

- Krauss, S., Blum, W., Brunner, M., Neubrand, M., Baumert, J., Kunter, M., Besser, M. & Elsner, J. (2011). Konzeptualisierung und Testkonstruktion zum fachbezogenen Professionswissen von Mathematiklehrkräften. *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV*, 135-161.
- Kunina-Habenicht, O., Schulze-Stocker, F., Kunter, M., Baumert, J., Leutner, D., Förster, D., Lohse-Bossenz, H., & Terhart, E. (2013). Die Bedeutung der Lerngelegenheiten im Lehramtsstudium und deren individuelle Nutzung für den Aufbau des bildungswissenschaftlichen Wissens. *Zeitschrift für Pädagogik*, 59(1), 1-23.
- Kunter, M., Brunner, D. P. M., Baumert, J., Klusmann, D. P. U., Krauss, S., Blum, W., ... & Neubrand, M. (2005). Der Mathematikunterricht der PISA-Schülerinnen und -Schüler. *Zeitschrift für Erziehungswissenschaft*, 8(4), 502-520.
- Kunter, M., Baumert, J., & Blum, W. (Eds.). (2011a). *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV*. Waxmann Verlag.
- Kunter, M., Kleickmann, T., Klusmann, U. & Richter, D. (2011b). Die Entwicklung professioneller Kompetenz von Lehrkräften. *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV*, 1, 55-68.
- Kunter, M. (2011). Motivation als Teil der professionellen Kompetenz: Forschungsbefunde zum Enthusiasmus von Lehrkräften. In *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV*. Waxmann. 259-275.
- Kurdiat, C. (2013). *Untersuchung der kognitiven Validität eines Testinstruments zur Messung professionellen Wissens angehender Physiklehrkräfte*. Unveröffentlichte Masterarbeit, Christian-Albrechts-Universität zu Kiel.
- Lampert, M. (1984). Teaching about thinking and thinking about teaching. *Journal of Curriculum Studies*, 16 (1). 1-18.
- Lederman, N. G. (1992). Students' and teachers' conceptions of the nature of science: A review of the research. *Journal of research in science teaching*, 29(4), 331-359.
- Leisen, J. (2004). Einheitliche Prüfungsanforderungen Physik. *Mathematische und Naturwissenschaftliche Unterricht*, 57(3), 155-159.
- Lenhard, W. & Lenhard, A. (2015). *Calculation of Effect Sizes*. available: http://www.psychometrica.de/effect_size.html. Bibergau (Germany): Psychometrica.
- Linacre, J.M. (2002). What do Infit and Outfit, Mean-square and Standardized mean? *Rasch Measurement Transactions*, 16(2), p.878.
- Linacre, J. M. (2015). WINSTEPS Rasch measurement computer program. *Chicago: Winsteps.com*.
- Liu, X. (2010). Using and Developing Measurement Instruments in Science Teaching: A Rasch Modeling Approach.
- Loughran, J., Berry, A., Mulhall, P. (2006). *Understanding and developing science teachers' pedagogical content knowledge*. Rotterdam, The Netherlands: Sense Publishers.

- Magnusson, S., Krajcik J. & Borko H. (1999). Nature, sources, and development of pedagogical content knowledge. In J. Gess-Newsome & N.G. Lederman (Hrsg.), *Examining pedagogical content knowledge*, 95-132.
- Maloney, D. P., O’Kuma, T. L., Hieggelke, C. J., & Van Heuvelen, A. (2001). Surveying students’ conceptual knowledge of electricity and magnetism. *American Journal of Physics*, 69(S1), S12-S23.
- Matzdorf, R. & Düchs, G. (2013). Immer mehr Parkstudierende – Statistiken zum Physikstudium an den Universitäten in Deutschland 2013. *Physik Journal*, Ausgabe August/September, 12(2013), S.29-33.
- Meltzer, D. E. (2011). Research on the education of physics teachers. In: *Teacher Education in Physics – Research, Curriculum, and Practice*. American Physical Society, 3-14.
- Merzyn, G. (1994). *Physikschulbücher, Physiklehrer und Physikunterricht*. Kiel: IPN.
- Messick, S. (1989). Validity. In Linn, R.L. (Hrsg.), *Educational measurement*, 3. Auflage. New York: Macmillan. S. 13-103.
- Messick, S. (1995). Validity of Psychological Assessment – Validation of Inferences From Persons’ Responses and Performances as Scientific Inquiry Into Score Meaning. In: *American Psychologist*, Vol. 50, 741-749.
- MHB (2011). *Modulhandbuch für Lehramt an Gymnasien und Gesamtschulen (Bachelor)*. Universität Duisburg-Essen. Stand 2011.
- MHDB (2013). *Modulhandbuch Bachelorstudiengang Physik auf Basis der Prüfungs- und Studienordnung vom 31. Januar 2007*. Ludwig-Maximilians-Universität München, Stand 13.06.2013.
- Moosbrugger, H., & Kelava, A. (2008). Qualitätsanforderungen an einen psychologischen Test (Testgütekriterien). In: *Testtheorie und Fragebogenkonstruktion* (pp. 7-26). Springer Berlin Heidelberg.
- Neubrand, M. (2006). Professionalität von Mathematik-Lehrerinnen und-Lehrern. *Beiträge zum Mathematikunterricht 2006*.
- Neumann, I., Neumann, K. & Nehm, R. (2011). Evaluating instrument quality in science education: Rasch-based analyses of a nature of science test. *International Journal of Science Education*, 33(10), 1373-1405.
- Nienhaus, G. U. (2010). Fast 10 000 Neueinschreibungen. *Physik Journal*, 9(8/9), 27.
- Nolting, W. (2006a). *Grundkurs Theoretische Physik 1 – Klassische Mechanik*. Springer-Verlag, 8. Auflage.
- Nolting, W. (2006b). *Grundkurs Theoretische Physik 2 – Analytische Mechanik*. Springer-Verlag, 7. Auflage.
- Nolting, W. (2007). *Grundkurs Theoretische Physik 3 – Elektrodynamik*. Springer-Verlag, 5. Auflage.
- Nolting, W. (2005a). *Grundkurs Theoretische Physik 4 – Spezielle Relativitätstheorie, Thermodynamik*. Springer-Verlag, 6. Auflage.
- Nolting, W. (2009). *Grundkurs Theoretische Physik 5/1 – Quantenmechanik – Grundlagen*. Springer-Verlag, 7. Auflage.

- Nolting, W. (2006c). *Grundkurs Theoretische Physik 5/2 – Quantenmechanik – Methoden und Anwendungen*. Springer-Verlag, 7. Auflage.
- Nolting, W. (2007). *Grundkurs Theoretische Physik 6 – Statistische Physik*. Springer-Verlag, 6. Auflage.
- Nolting, W. (2005b). *Grundkurs Theoretische Physik 7 – Viel-Teilchen-Theorie*. Springer-Verlag, 5. Auflage.
- 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.
- Oelkers, J. (2003). Wie man Schule entwickelt. *Eine bildungspolitische Analyse nach PISA*, 2.
- Olszewski, J. (2010). *The Impact of Physics Teachers' Pedagogical Content Knowledge on Teacher Action and Student Outcomes*. Dissertationsschrift.
- Oser, F., & Oelkers, J. (2001). Die Wirksamkeit der Lehrerbildungssysteme in der Schweiz. *Chur Zürich*.
- Oser, F. (2002). Unter dem Durchschnitt: Einige Resultate und Stärken bzw. Schwächen unterschiedlicher Darstellungsweisen von Ländervergleichen am Beispiel IEA-CivicEducation. *Schweizerische Zeitschrift für Bildungswissenschaften*, 24(3), 487-505.
- Pant, H. A., Stanat, P., Schroeders, U., Roppelt, A., Siegle, T., & Pöhlmann, C. (2013). IQB-Ländervergleich 2012. *Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I*.
- Paris, S.G., Lipson, M. Y., Wixson, K.K. (1983). Becoming a Strategic Reader. *Contemporary Educational Psychology* 8, 293-316.
- Park, S., & Oliver, J. S. (2008). Revisiting the conceptualisation of pedagogical content knowledge (PCK): PCK as a conceptual tool to understand teachers as professionals. *Research in Science Education*, 38(3), 261-284.
- Park, S., & Chen, Y. C. (2012). Mapping out the integration of the components of pedagogical content knowledge (PCK): Examples from high school biology classrooms. *Journal of Research in Science Teaching*, 49(7), 922-941.
- POH (2013). *Gemeinsame Prüfungsordnung für den fächerübergreifenden Bachelorstudiengang an der Gottfried Wilhelm Leibniz Universität Hannover in der Fassung vom 09.09.2013*. Verkündungsblatt der Gottfried Wilhelm Leibniz Universität Hannover vom 01.09.2014.
- R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Wien, Österreich. URL: <http://www.R-project.org/>.
- Rasch, G (1960). *Probabilistic Models for Some Intelligence and Attainment Tests*. Copenhagen, Denmark: Danmarks Paedagogiske Institut.
- Riese, J (2009). Professionelles Wissen und professionelle Handlungskompetenz von (angehenden) Physiklehrkräften. *Studien zum Physik- und Chemielernen*, Band 97.
- Riese, J, Reinhold, P (2010). Empirische Erkenntnisse zur Struktur professioneller Handlungskompetenz von angehenden Physiklehrkräften. *Zeitschrift für Didaktik der Naturwissenschaften*, Jg. 16.

- Riese, J., Gramzow, Y., Reinhold, P., Oettinghaus, L., & Korneck, F. (2011). Unterschiede im Professionswissen bei angehenden Physiklehrkräften-Ein Vergleich verschiedener Lehramtszugänge. *PhyDid B-Didaktik der Physik-Beiträge zur DPG-Frühjahrstagung*.
- Riese, J., & Reinhold, P. (2012). Die professionelle Kompetenz angehender Physiklehrkräfte in verschiedenen Ausbildungsformen. *Zeitschrift für Erziehungswissenschaft*, 15(1), 111-143.
- Rossee, Y., Oberski, D., Byrnes, J. Vanbrabant, L. & Savalei, Y. (2013). Lavaan: Latent variable analysis [Software]. URL <http://cran.r-project.org/package=lavaan> (R package version 0.5-23.1097).
- Sadler, P.M., Sonnert, G., Coyle, H.P., Cook-Smith, N., Miller, J.L. (2013). The Influence of Teachers Knowledge on Student Learning in Middle School Physical Science Classrooms. *American Educational Research Journal*. Vol. 50, No. 5, 1020-1049.
- Schilcher, K. (2010). *Theoretische Physik kompakt für das Lehramt*. Oldenbourg Verlag.
- Schlesinger, C. 2013. *PISA-Aufsteiger Deutschland – Die Streber-Republik*. Wirtschafts Woche WiWo.de. 03.12.2013.
- Schmelzing, S., Wuesten, S., Sandmann, A., & Neuhaus, B. (2010). Measuring declarative and reflective components of biology teachers' pedagogical content knowledge. *Contemporary science education research: preservice and inservice teacher education*. Ankara, Turkey: Pegem Akademi, 71-77.
- Schmelzing, S., van Driel, J. H., Jüttner, M., Brandenbusch, S., Sandmann, A., & Neuhaus, B. J. (2013). DEVELOPMENT, EVALUATION, AND VALIDATION OF A PAPER-AND-PENCIL TEST FOR MEASURING TWO COMPONENTS OF BIOLOGY TEACHERS' PEDAGOGICAL CONTENT KNOWLEDGE CONCERNING THE "CARDIOVASCULAR SYSTEM". *International Journal of Science and Mathematics Education*, 11(6), 1369-1390.
- Schmüser, P. (2012a). *Theoretische Physik für Studierende des Lehramts 1: Quantenmechanik*. Springer-Verlag.
- Schmüser, P. (2012b). *Theoretische Physik für Studierende des Lehramts 2: Elektrodynamik und Spezielle Relativitätstheorie*. Springer-Verlag.
- SPO (2015). *Studien- und Prüfungsordnung des Fachbereichs Physik der Freien Universität Berlin für den Bachelorstudiengang Physik für das Lehramt und für das 60-Leistungspunkte-Modulangebot Physik im Rahmen anderer Studiengänge*. Freie Universität Berlin, Stand 11.08.2015.
- 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.
- Schön, D. A. (1983). *The reflective practitioner: How professionals think in action* (Vol. 5126). Basic books.
- Schön, D. (1987). Educating the reflective practitioner.
- Shavelson R. J., Ruiz-Primo, M. A. & Wiley E. W. (2005). Windows in the mind. In: *Higher Education* (2005), 49, 413-430.
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15 (2), 4-14.

- Shulman, L. S. (1987). Knowledge and teaching of the new reform. *Harvard Educational Review*, 57, 1-22.
- Smith, D.C., Neale, D.C. J (1989). The construction of subject matter knowledge in primary science teaching. *Teaching and Teacher Education*, 5, 1-20.
- Sternberg R.J., Wagner, R.K., Williams W.M., Horvath, J.A. (1995). Testing common sense. *American Psychologist*, Vol 50(11), 912-927
- Swackhamer, G., & Hestenes, D. (2005). An energy concept inventory. *Arizona State University*.
- Tamir, P. (1988). Subject matter and related pedagogical knowledge in teacher education. *Teaching and Teacher Education*, 4, 99-110.
- Tepner, O., & Dollny, S. (2011). Chemistry teachers' content knowledge and its correlation to pedagogical content knowledge. Retrieved September, 19, 2013.
- Tepner, O., Borowski, A., Dollny, S., Fischer, H. E., Jüttner, M., Kirschner, S., Leutner, S., Neuhaus, B. J., Sandmann, A., Sumfleth, E., Thillmann, H. & Wirth, J. (2012). Modell zur Entwicklung von Testitems zur Erfassung des Professionswissens von Lehrkräften in den Naturwissenschaften. In: *Zeitschrift für Didaktik der Naturwissenschaften*, 18, 7-28.
- Terhart, E. (2000). Qualität und Qualitätssicherung im Schulsystem. Hintergründe-Konzepte-Probleme. *Zeitschrift für Pädagogik*, 46(6), 809-829.
- Terhart, E. (2002). Standards für die Lehrerbildung. Eine Expertise für die Kultusministerkonferenz.- Münster: ZKL-Verlag.
- Terhart, E. (2005). Standards für die Lehrerbildung-ein Kommentar. *Zeitschrift für Pädagogik*, 51(2), 275-279.
- Thelk, A.D., Hoole, E.R., Lottridge, S.M. (2006). What Are You Thinking? Postsecondary Student Think-Alouds of Scientific and Quantitative Reasoning Items. *The Journal of General Education*. Vol. 55, No. 1.
- Tipler, P. A., & Mosca, G. (2014). *Physik: für Wissenschaftler und Ingenieure*. Springer-Verlag.
- Van Someren, M.W., Barnard, Y.F., Sandberg, J.A.C. (1994). *The Think Aloud Method. A practical guide to modeling cognitive processes*. Academic Press London.
- Vogelsang, C., Reinhold, P. (2013). Zur Handlungsvalidität von Tests zum professionellen Wissen von Lehrkräften, *Zeitschrift für Didaktik der Naturwissenschaften*, Jg. 19, 129-157.
- Vogelsang, C. (2014). *Validierung eines Instruments zur Erfassung der professionellen Handlungskompetenz von (angehenden) Physiklehrkräften - Zusammenhangsanalysen zwischen Lehrerkompetenz und Lehrerperformanz* (Vol. 174). Logos Verlag Berlin GmbH.
- Voss, T., & Kunter, M. (2011). Pädagogisch-psychologisches Wissen von Lehrkräften. *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV*, 193-214.
- Voss, T., Kleickmann, T., Kunter, M., & Hachfeld, A. (2011a). Überzeugungen von Mathematiklehrkräften. In *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV*, 235-257.

- Voss, T., Kunter, M., & Baumert, J. (2011b). Assessing teacher candidates' general pedagogical/psychological knowledge. Test construction and validation. *Journal of educational psychology*, 103(4), 952-969.
- Weinert, F. E., & Helmke, A. (1996). *Der gute Lehrer: Person, Funktion oder Fiktion?* (pp. 223-233).
- Weinert, F. E. (2001). Concept of competence: A conceptual clarification.
- Wiesner, H., Hopf, M., & Schecker, H. (Eds.). (2011). *Physikdidaktik kompakt*. Aulis-Verlag.
- Wirtz, M. A., & Caspar, F. (2002). *Beurteilerübereinstimmung und Beurteilerreliabilität: Methoden zur Bestimmung und Verbesserung der Zuverlässigkeit von Einschätzungen mittels Kategoriensystemen und Ratingskalen*. Hogrefe.
- Woitkowski, D., Riese, J., Reinhold, P. (2011). Modellierung fachwissenschaftlicher Kompetenz angehender Physiklehrkräfte, *Zeitschrift für Didaktik der Naturwissenschaften*, Jg. 17.
- Woitkowski, D. (2015). *Fachliches Wissen Physik in der Hochschulausbildung – Konzeptualisierung, Messung, Niveaubildung*. Logos Berlin.
- Wright, B.D., Masters, G.N. (1982). *Rating Scale Analysis. Rasch Measurement*. MESA Press, Chicago.
- Wright, B.D., Stone, M.H. (1979). *Best Test Design. Rasch Measurement*. MESA Press, Chicago.
- Wuttiptom S., Sharma M.D., Johnston I.D., Chitaree R. & Soankwan C. (2009). Development and Use of a Conceptual Survey in Introductory Quantum Physics. In: *International Journal of Science Education*. Vol. 31, No. 5, 631-654.

Eidesstattliche Erklärung

Hiermit erkläre ich, dass die vorliegende Dissertation – abgesehen von der Beratung durch meine Betreuer – nach Inhalt und Form meine eigene Arbeit ist. Die vorliegende Dissertation wurde weder ganz noch zum Teil einer anderen Stelle im Rahmen eines Prüfungsverfahrens vorgelegt, wurde nicht veröffentlicht oder zur Veröffentlichung eingereicht. Teile der Befunde der vorliegenden Dissertation wurden bereits in Form von Zeitschriftenartikeln oder Tagungsbandbeiträgen publiziert. Die vorliegende Arbeit ist unter Einhaltung der Regeln guter wissenschaftlicher Praxis der Deutschen Forschungsgemeinschaft entstanden. Ich erkläre ferner, dass mir kein akademischer Grad entzogen wurde.

Kiel, den _____

Jochen Kröger