

S. Krauss, A. Lindl, A. Schilcher, M. Fricke, A. Göhring,
B. Hofmann, P. Kirchhoff, R. H. Mulder (Hrsg.)

FALKO

FACHSPEZIFISCHE LEHRERKOMPETENZEN

Konzeption von Professionswissenstests
in den Fächern Deutsch, Englisch, Latein, Physik,
Musik, Evangelische Religion und Pädagogik



WAXMANN

Stefan Krauss, Alfred Lindl, Anita Schilcher,
Michael Fricke, Anja Göhring, Bernhard Hofmann,
Petra Kirchhoff, Regina H. Mulder (Hrsg.)

FALKO: Fachspezifische Lehrerkompetenzen

Konzeption von Professionswissenstests in den Fächern
Deutsch, Englisch, Latein, Physik, Musik,
Evangelische Religion und Pädagogik

Mit einem Vorwort von Jürgen Baumert
Mit neuen Daten aus der COACTIV-Studie



Waxmann 2017
Münster • New York

Bibliografische Informationen der Deutschen Nationalbibliothek

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

Print-ISBN 978-3-8309-3445-5

E-Book-ISBN 978-3-8309-8445-0

© Waxmann Verlag GmbH, 2017

www.waxmann.com

info@waxmann.com

Umschlaggestaltung: Inna Ponomareva, Jena

Umschlagbild: Johannes Wild, Regensburg

Satz: Stoddart Satz- und Layoutservice, Münster

Druck: Hubert & Co., Göttingen

Gedruckt auf alterungsbeständigem Papier,
säurefrei gemäß ISO 9706



Printed in Germany

Alle Rechte vorbehalten. Nachdruck, auch auszugsweise, verboten.

Kein Teil dieses Werkes darf ohne schriftliche Genehmigung des Verlages in irgendeiner Form reproduziert oder unter Verwendung elektronischer Systeme verarbeitet, vervielfältigt oder verbreitet werden.

Inhalt

Jürgen Baumert

Vorwort.....7

Stefan Krauss, Alfred Lindl, Anita Schilcher & Oliver Tepner

1. Das Forschungsprojekt FALKO – ein einleitender Überblick.....9

Markus Pissarek & Anita Schilcher

2. FALKO-D: Die Untersuchung des Professionswissens von
Deutschlehrenden. Entwicklung eines Messinstruments zur
fachspezifischen Lehrerkompetenz und Ergebnisse zu dessen Validierung.....67

Petra Kirchoff

3. FALKO-E: Fachspezifisches professionelles Wissen von
Engischlehrkräften
Entwicklung und Validierung eines domänenspezifischen Testinstruments... 113

Alfred Lindl & Harald Kloiber

4. *artifices docendi*
FALKO-L: Modellierung und Messung domänenspezifischer
Kompetenzen von Lateinlehrkräften 153

Anja Schödl & Anja Göhring

5. FALKO-P: Fachspezifische Lehrerkompetenzen im Fach Physik
Entwicklung und Validierung eines Testinstruments zur Erfassung
des fachspezifischen Professionswissens von Physiklehrkräften..... 201

Gabriele Puffer & Bernhard Hofmann

6. FALKO-M: Entwicklung und Validierung eines Testinstruments zum
domänenspezifischen Professionswissen von Musiklehrkräften 245

Michael Fricke

7. FALKO-R: Professionswissen von Religionslehrkräften
Entwicklung eines Messinstruments zur fachspezifischen
Lehrerkompetenz..... 291

Regina H. Mulder, Susanne Sauer & Franziska Kempka

8. FALKO-PA: Ein Instrument aus flexibel einsetzbaren Vignetten
zur Erfassung pädagogischer Kompetenzen 337

Alfred Lindl & Stefan Krauss

- 9.** Transdisziplinäre Perspektiven auf domänenspezifische
Lehrerkompetenzen
Eine Metaanalyse zentraler Resultate des Forschungsprojektes FALKO 381

Anita Schilcher, Stefan Krauss, Karsten Rincke & Sven Hilbert

- 10.** Ausblick – Aus FALKO wird FALKE
Fachspezifische Lehrerkompetenz im Erklären..... 439

Verzeichnis der Autorinnen und Autoren 453

Die Items der Testinstrumente können unter www.waxmann.com/buch3445
kostenlos abgerufen werden. Passwort: FALKO_wax#3445#

Vorwort

In der vorliegenden Publikation zum FALKO-Projekt haben sich verschiedene Didaktiken der Universität Regensburg in beispielhafter Weise zusammengefunden, um auf der Grundlage eines geteilten theoretischen Rahmenmodells das Professionswissen von Lehrkräften und angehenden Lehrkräften, die an unterschiedlichen Sekundarschulformen unterrichten und unterrichtet werden, in den Fächern Deutsch, Englisch, Latein, Physik, Musik und Evangelische Religionslehre zu untersuchen. Gemeinsames Ziel der beteiligten Wissenschaftlerinnen und Wissenschaftler war, zuverlässige und valide Instrumente zur Erfassung domänenspezifischen Professionswissens in einer theoretisch begründeten gemeinsamen transdisziplinären Perspektive zu entwickeln und empirisch zu erproben, und zwar auch für schwach strukturierte Domänen. Das vorliegende Buch ist der Erfolgsbericht eines bislang einzigartigen Unternehmens.

Ein großes Verdienst von FALKO ist es, dass auch Fachdidaktiken für dieses Projekt gewonnen werden konnten, die traditionell eher in einer gewissen Distanz zur empirischen Bildungsforschung stehen. Berücksichtigt man, dass es in den Unterrichtsfächern Musik und Latein oder für den konfessionellen Religionsunterricht bislang nicht nur keine Professionswissenstests, sondern überhaupt nur wenige empirische Studien gibt, dokumentieren Testkonstruktionen und ihre empirische Prüfung eine beachtliche zurückgelegte Wegstrecke. Die gleichzeitige Online-Veröffentlichung sämtlicher Items ermöglicht einen Diskurs über die Befunde auf breiter Ebene.

Das gemeinsame konzeptuelle Vorgehen bei der Operationalisierung des Fachwissens und des fachdidaktischen Wissens in Anlehnung an die mittlerweile einschlägigen Tests aus der COACTIV-Studie für Mathematiklehrkräfte erlaubt es, die Befunde zu den einzelnen Fächern aus metaanalytischer Perspektive zu vergleichen. In dem Buch wird dabei die Frage beantwortet, welche der mittlerweile vielzitierten Ergebnisse der COACTIV-Studie – beispielsweise die großen Schulformunterschiede im Professionswissen oder auch der fehlende Zusammenhang mit der Berufserfahrung – rein mathematikspezifisch sind und in welchen Fächern sich vergleichbare Befunde zeigen. Dies ist vor dem Hintergrund der großen Heterogenität der Fächer eine wichtige, die Ausbildungsstruktur betreffende Frage.

Das vorliegende Buch, dessen Beiträge über die Testkonstruktion hinausgehen, legt Zeugnis dafür ab, dass interdisziplinäre Zusammenarbeit auch zwischen heterogenen Fächern theoretisch und empirisch gelingen kann und widerspricht allen unfruchtbaren Polarisierungen. Der Regensburger Erfolg in der „Qualitätsoffensive Lehrerbildung“ des BMBF bestätigt die Forschungsinitiative und gibt ihr zusätzlichen Rückhalt. Die geplante theoretische und empirische Untersuchung von Gemeinsamkeiten und Unterschieden „guten Erklärens“ in elf Unterrichtsfächern hat weltweit nur wenige Konkurrenten.

Prof. Dr. Drs. h.c. Jürgen Baumert
(Direktor Emeritus des Max-Planck-Instituts für Bildungsforschung
und ehemaliger Vizepräsident der Max-Planck-Gesellschaft)

Das Forschungsprojekt FALKO – ein einleitender Überblick

1. Von PISA über COACTIV zu FALKO: Eine kurze Historie	13
1.1 PISA 2000: Ein „empirischer Weckruf“ für die Didaktiken	13
1.2 COACTIV 03/04	14
1.3 Effekte von Lehrermerkmalen auf Unterrichtsqualität und unterrichtliche Zielkriterien	17
1.4 Effekte des fachbezogenen Professionswissens (Fachwissen und fachdidaktisches Wissen)	20
1.5 Zwischenfazit	23
2. Fachwissen und fachdidaktisches Wissen: Testkonstruktionen und der zugrunde liegende Begriff der professionellen Kompetenz	25
2.1 Testkonstruktionen zum Fachwissen und zum fachdidaktischen Wissen in COACTIV und FALKO	26
2.2 Testkonstruktionen zum Fachwissen und zum fachdidaktischen Wissen anderer Arbeitsgruppen	31
2.3 Professionelles Wissen = professionelles Können = professionelles Handeln?	33
2.4 Wissensbegriff bei COACTIV und FALKO	38
3. Aufbau der Ergebnisteile im vorliegenden Band (exemplarisch illustriert anhand von Daten aus der COACTIV-Studie)	40
3.1 Stichprobe und Durchführung	41
3.2 Psychometrische Gütekriterien	41
3.3 Deskriptive Ergebnisse	43
3.3.1 Skalenmittelwerte und Interkorrelationen	43
3.3.2 Gruppenunterschiede zwischen Studierenden und Lehrkräften	44
3.3.3 Schulformunterschiede (Ausbildungssensitivität)	45
3.3.4 Zusammenhänge des fachbezogenen Professionswissens mit Indikatoren des Schul- und Studienerfolges sowie der Berufs- und Studiendauer	48
3.4 Dimensionalität des fachbezogenen Professionswissens (CFA)	49
3.5 Das Professionswissen von Fachdidaktikern	50
4. Das Lehrerbildungssystem in Bayern	51
4.1 Die bayerische Lehramtsprüfungsordnung und universitäre Ausgestaltungen	51
4.2 Modularisierte Studienstruktur	53
4.3 Gymnasium (GY, „vertieft“) versus Nicht-Gymnasium (NGY, „nicht vertieft“)	54
Register der Abkürzungen für die Tests aus Tabelle 4 (alphabetisch)	55
Literatur	56

Das professionelle Wissen ist in vielen Berufen und Expertisedomänen entscheidend für die erfolgreiche Bewältigung komplexer professionsspezifischer Aufgaben, weshalb eine empirische Erfassung dieses Wissens auf der Basis einer Anforderungsanalyse wertvolle Informationen über Stärken und Schwächen eines Berufsstandes, aber auch über interindividuelle Kompetenzunterschiede bei Mitgliedern der Profession liefern kann. Eine solche Bestandsaufnahme im größeren Rahmen (im Sinne eines Monitoringprozesses) ermöglicht darüber hinaus einen Abgleich von normativen Forderungen oder Erwartungen an eine Profession von Seiten der Gesellschaft, Politik oder Wissenschaft mit empirischen Realitäten und kann somit auch Auslöser für evidenzbasierte Innovationen zur Entwicklung von Professionen im Rahmen von Aus- oder Fortbildung sein.

Im vorliegenden Band werden die psychometrischen Testkonstruktionen zum Fachwissen und zum fachdidaktischen Wissen von Sekundarschullehrkräften in den Fächern Deutsch, Englisch, Latein, Physik, Musik und Ev. Religion¹ der interdisziplinären Forschungsgruppe FALKO (**F**achspezifische **L**ehrer**k**ompetenzen) vorgestellt, die zur vielfältigen Weiternutzung, zum Beispiel für Grundlagenforschung oder in der Lehreraus- und -fortbildung, einsetzbar sind. Trotz der augenscheinlichen Heterogenität der Unterrichtsfächer konnte bei FALKO ein gemeinsames und in weiten Teilen paralleles konzeptuelles, technisches und methodisches Vorgehen aller sechs beteiligten Fachdidaktiken verwirklicht werden (Kap. 2–7). Dies bezieht sich nicht nur auf die Testkonstruktionen selbst, die jeweils in dezidiertem Anlehnung an Shulman (1986; 1987) und an COACTIV² (Kunter, Baumert et al., 2011; vgl. auch Abschn. 1) erfolgten, sondern auch auf die Administration der Tests in ersten Validierungsstudien, die Kodierung der offenen Antworten der getesteten Studienteilnehmer sowie die statistischen Analysen.

Ein Kapitel zur Konstruktion von Vignetten zum pädagogischen Wissen (Kap. 8), das aufgrund der dort thematisierten fachunabhängigen Lehrerkompetenzen gewissermaßen „quer“ zu den anderen Kapiteln steht, rundet den mittlerweile oft zitierten „Dreiklang“ der Kernkategorien des Professionswissens von Lehrkräften ab (Fachwissen, im Folgenden kurz „FW“; fachdidaktisches Wissen, kurz „FDW“; pädagogisches Wissen, kurz „PW“). Darüber hinaus tragen zum vorliegenden Band noch weitere Didaktiken der Universität Regensburg als Autoren bei, wodurch insgesamt folgende Disziplinen beteiligt sind: Chemiedidaktik, Deutschdidaktik, Englischdidaktik, Ev. Religion, Lateindidaktik, Mathematikdidaktik, Musikpädagogik, Naturwissenschaft und Technik, Physikdidaktik, allgemeine Pädagogik und Psychologie.

Die domänenspezifischen Operationalisierungen beider fachbezogener Wissensbereiche (FW und FDW) auf der Grundlage einer gemeinsamen fachübergreifenden Konzeptualisierung in Anlehnung an COACTIV (Fachwissen als „vertieftes Hintergrundwissen zum Schulstoff“; Wissen über „Erklären & Repräsentieren“ sowie über „typische Schülerfehlvorstellungen und -schwierigkeiten“ als zwei von drei Facetten fachdidaktischen Wissens) erlauben in einer abschließenden Metaanalyse der Er-

- 1 Für dieses Schulfach gibt es je nach Bundesland die unterschiedlichen Bezeichnungen „Religion“, „Religionslehre“ oder „Religionsunterricht“ (vgl. Fricke, 2017).
- 2 „COACTIV“: Cognitive Activation in the Classroom. Ein Register für sämtliche im Fließtext verwendeten Projekt- und Testbezeichnungen findet sich unter Abschn. 5.1.

gebnisse (Kap. 9) auch die Inklusion der Daten aus der COACTIV-Studie zum Fach Mathematik. Interessant ist in diesem Zusammenhang, welche der COACTIV-Ergebnisse (z. B. stark ausgeprägte Schulformunterschiede vor allem im Fachwissen, keine Zunahme des Professionswissens mit der Berufserfahrung als Lehrkraft etc.) rein mathematikspezifisch sind und welche sich auch in anderen Fächern nachweisen lassen.

Eine weitere Besonderheit der vorliegenden Publikation besteht darin, dass sämtliche Items aller Tests den Lesern online (www.waxmann.com/buch3445) zur Verfügung gestellt werden. Neben der im Rahmen der Veröffentlichung von Testergebnissen in der empirischen Bildungsforschung oft eingeforderten Transparenz soll dies auch verdeutlichen, dass die mit den FALKO-Tests erzielten Ergebnisse keinerlei abschließenden Anspruch erheben, sondern vielmehr ein Diskurs über fachspezifische professionelle Kompetenzen von Lehrkräften in einer großen Bandbreite sehr heterogener Fächer angestoßen werden soll. Da wissenschaftliche, aber auch subjektive Theorien der Autoren aus abstrakten Konzeptualisierungen nicht immer zu erkennen sind, soll dadurch eine offene Diskussion auch auf der Ebene der Operationalisierung ermöglicht werden, weshalb neben den online veröffentlichten Items im Anhang der jeweiligen Kapitel auch jeweils dazugehörige zentrale Ergebnisse auf Itemebene präsentiert werden. Gerade im Hinblick auf Unterrichtsfächer, in denen es generell (also auch über Professionswissenstests von Lehrkräften hinaus) nur sehr wenig empirische Forschung gibt (wie in Latein, Musik oder Religion), erscheint dieser Aspekt relevant.

Das vorliegende Einleitungskapitel besteht im Wesentlichen aus vier Teilen. Da die COACTIV-Studie, in der das Professionswissen von Mathematiklehrkräften untersucht wurde, Motivation und Ansatzpunkt für die Gründung der interdisziplinären Forschungsgruppe FALKO an der Universität Regensburg war, soll zunächst ein Einblick in Genese, zentrale Untersuchungsgegenstände und wichtige Ergebnisse dieser Studie gegeben werden. Dabei werden insbesondere die Konzeptualisierung und die Operationalisierung von FW und FDW in der COACTIV-Studie beleuchtet, die die Vorlage für die Konstruktion der entsprechenden FALKO-Tests waren, sowie Untersuchungen zur prädiktiven Validität dieser beiden Wissensbereiche für Unterrichtsqualität und den Leistungszuwachs von Schülern vorgestellt.

Mittlerweile gibt es – vor allem im deutschsprachigen Raum – zahlreiche weitere Ansätze zur Erfassung des Professionswissens von Lehrkräften in verschiedenen Unterrichtsfächern (vgl. Tab. 4). Der zweite Teil des Einleitungskapitels motiviert noch einmal die Beschäftigung mit diesem Thema und gibt einen kurzen Überblick über derzeitige Strömungen mit besonderem Fokus auf aktuellen Diskussionen zum Wechselspiel zwischen Wissen, Können und tatsächlichem Unterrichtshandeln. Für spezifische Ergebnisse zu bisherigen Professionswissenstests in Unterrichtsfächern, die auch im vorliegenden Band vertreten sind, sei auf die jeweiligen Beiträge verwiesen (Deutsch, Kap. 2: Pissarek & Schilcher, 2017; Englisch, Kap. 3: Kirchhoff, 2017; Physik, Kap. 5: Schödl & Göhring, 2017). Viele der bislang konzipierten Tests wurden dabei vor allem bei Studierenden beziehungsweise Referendaren eingesetzt, sodass bei insgesamt fünf von sechs FALKO-Disziplinen (nämlich bei allen außer Physik) erstmalig praktizierende Lehrkräfte mit fachbezogenen Professionswissenstests untersucht wurden (vgl. Abschn. 2). Für drei in diesem Band thematisierte Unterrichtsfächer (Latein,

Kap 4: Lindl & Kloiber, 2017; Musik, Kap. 6: Puffer & Hofmann, 2017; Ev. Religion, Kap. 7: Fricke, 2017) liegen noch keine vergleichbaren Testkonstruktionen vor. Eine Diskussion bisheriger Tests zum (fachunabhängigen) pädagogischen Wissen findet sich in Beitrag 8 (Pädagogik: Mulder, Sauer & Kempka, 2017).

Die zentralen Hypothesen für die vorliegenden Testkonstruktionen betreffen vor allem die Ausbildungssensitivität. Entsprechend dem Begriff „Professionswissen“ sollten FW und FDW im Rahmen der Professionalisierung (universitäres Studium und Vorbereitungsdienst) erworben werden, sodass deutliche Unterschiede zwischen Lehramtsstudierenden und Lehrkräften im Schuldienst feststellbar sein sollten. Weiterhin ist aufgrund des üblicherweise sehr fachintensiven gymnasialen Studiums und des vergleichsweise hohen fachlichen Niveaus in dieser Schulform (GY) zu erwarten, dass gymnasiale Lehrkräfte über deutlich mehr fachbezogenes Wissen – zumindest aber über mehr Fachwissen – verfügen als ihre jeweiligen Kollegen aus nichtgymnasialen Sekundarschulformen (NGY). Vergleichbare Unterschiede werden auch bereits zwischen Studierenden des gymnasialen beziehungsweise nichtgymnasialen Lehramts erwartet.

Im vorliegenden Band sind die Ergebnisteile bezüglich des gemessenen FW und FDW in den sechs Unterrichtsfächern identisch aufgebaut (siehe Kap. 2–7). Nach einem Überblick über psychometrische Gütekriterien werden die mit den Tests untersuchten domänenspezifischen Stichproben in jedem Kapitel für die deskriptiven und die inferenzstatistischen Analysen entsprechend zunächst nach *beruflichem Status* (Studierende vs. Lehrkräfte) aufgeteilt und die beiden Teilgruppen in einem zweiten Schritt jeweils noch nach *Schulform* (GY vs. NGY) getrennt betrachtet. Auf die Analyse der paarweisen Unterschiede in beiden Wissensbereichen zwischen den vier Gruppen (gymnasiale Lehrkräfte und Studierende sowie nichtgymnasiale Lehrkräfte und Studierende, kurz GY-L, GY-S, NGY-L und NGY-S) folgen jeweils klassische lineare Regressionen der Leistungen im FDW und im FW auf die Prädiktoren Status und Schulform unter Kontrolle von Geschlecht und Abiturnote (letztere als Approximation für die kognitive Leistungsfähigkeit). Untersuchungen zum Zusammenhang des gemessenen Professionswissens mit Indikatoren des Schul- beziehungsweise Studienerfolgs und der Berufs- beziehungsweise Studiendauer (Produkt-Moment-Korrelationen nullter Ordnung; kriteriale Validität) sowie eine konfirmatorische Faktorenanalyse zur Überprüfung der Dimensionalität des Professionswissens (faktorielle Validität) schließen die Ergebnisteile jeweils ab. Im dritten Teil der Einleitung werden (beispielhaft anhand entsprechender Resultate der COACTIV-Studie) Einzelheiten zum Aufbau der Ergebnisteile in den Kapiteln 2–7 erläutert.

Dieser dritte Teil hat demnach noch eine weitere Funktion: Die an die Konzeptualisierungen der COACTIV-Studie angelehnten domänenspezifischen Operationalisierungen des Professionswissens in FALKO legen prinzipiell eine Berücksichtigung der Daten der COACTIV-Studie in einer gemeinsamen Metaanalyse nahe (Kap. 9). Hierzu musste jedoch erst noch eine für eine solche Metaanalyse fehlende Stichprobe nichtgymnasialer Lehramtsstudierender für das Fach Mathematik mit den COACTIV-Tests nacherhoben werden (alle anderen Stichproben waren bereits erhoben, Krauss et al., 2011; Kleickmann et al., 2013, verwenden bei einer nichtgymnasialen Studieren-

denstichprobe einen modifizierten Test). Somit dient Teil 3 nicht nur der Illustration des Aufbaus sämtlicher Ergebnisteile, sondern es werden im Rahmen des vorliegenden Bandes dadurch auch neue, bislang noch nicht veröffentlichte Befunde zum Professionswissen von (angehenden) Mathematiklehrkräften präsentiert.

Da der überwiegende Teil der Studienteilnehmer (Lehrkräfte und Lehramtsstudierende), der im Rahmen von FALKO die Tests zum FDW und zum FW, aber auch die Vignetten zum PW bearbeitete, aus Bayern stammt, werden im vierten und letzten Teil des Einleitungskapitels Spezifika des bayerischen Lehramtsausbildungssystems zusammengefasst. Dabei soll aufgrund der anvisierten Zielgruppe der Sekundarschullehrkräfte besonderes Augenmerk auf die Ausbildung zum Haupt-/Mittelschul-, Real-schul- beziehungsweise Gymnasiallehrer³ gelegt werden.

1. Von PISA über COACTIV zu FALKO: Eine kurze Historie

Im Folgenden werden – ausgehend von PISA 2000 – die Genese⁴ der COACTIV-Studie, deren theoretischer Rahmen, wichtige untersuchte Konstrukte sowie zentrale Ergebnisse dieser Studie vorgestellt und erläutert, inwiefern dies der Ansatzpunkt zur Gründung der FALKO-Forschungsgruppe war.

1.1 PISA 2000: Ein „empirischer Weckruf“ für die Didaktiken

In der ersten PISA-Studie („Programme for International Student Assessment“) im Jahr 2000 schnitten die deutschen Schüler der neunten Jahrgangsstufe in den Bereichen Mathematik, Naturwissenschaften und Lesen im internationalen Vergleich nur unterdurchschnittlich ab (Baumert et al., 2001). Dieses heute als „PISA-Schock“ bekannte Ereignis führte sowohl in den Medien als auch in der Wissenschaft schnell zu einer heftigen Diskussion über mögliche Ursachen, war man doch der Meinung, dass Deutschland, das Land der Dichter und Denker, gerade in Bezug auf die Qualität seines Bildungs- und Schulsystems vorbildlich sei.

Bei PISA 2000 wurden neben den viel zitierten Tests in Lesen, Mathematik und Naturwissenschaften von den Schülern per Fragebogen auch zahlreiche Indikatoren des Schulerlebens wie zum Beispiel die wahrgenommene Unterrichtsqualität erhoben, aber auch Eingangsvoraussetzungen wie Intelligenz, sozioökonomischer Status, elterliche Unterstützung oder Merkmale der Lebenswelt der Jugendlichen erfasst (Baumert et al., 2001). Trotz dieser Fülle an Daten von mehr als 50 000 deutschen Schülern aus knapp 1500 Schulen fehlten bei PISA 2000 aber offensichtlich Daten über die Hauptprotagonisten der Unterrichtsgestaltung: die Lehrkräfte. Am Max-Planck-Institut für Bildungsforschung in Berlin, das federführend bei der Durchführung der ersten PISA-

3 Um die Lesbarkeit zu verbessern, werden im vorliegenden Beitrag nur die männlichen Bezeichnungen verwendet, wobei jedoch jeweils beide Geschlechter gleichermaßen gemeint sind.

4 Teile dieser Schilderung – vor allem der Anfang von Abschnitt 1 – sind Krauss und Schilcher (2016) entnommen.

Studie in Deutschland war, entstand im Zuge der Diskussionen um das überraschend schlechte Abschneiden deutscher Schüler demzufolge die Idee, bei der nächsten PISA-Studie 2003 auch die Lehrkräfte mit in die Untersuchungen einzubeziehen.

1.2 COACTIV 03/04

Dies war die Geburtsstunde der COACTIV-Studie („Professionswissen von Lehrkräften, kognitiv aktivierender Mathematikunterricht und die Entwicklung mathematischer Kompetenz“), einer breit angelegten Lehrerstudie, die technisch und konzeptuell an PISA 2003 gekoppelt war und in der zusätzlich die Lehrkräfte der getesteten Klassen umfassend untersucht wurden (Abb. 1; Kunter, Baumert et al., 2011). Bei PISA werden zwar grundsätzlich Daten zu den drei Domänen Lesen, Mathematik und Naturwissenschaften erhoben, aber im Zuge des dreijährigen PISA-Zyklus wechselt der Hauptfokus regelmäßig. Da bei PISA 2003 der primäre Fokus auf die Mathematik gerichtet war (bei der aktuell veröffentlichten PISA-Studie 2015 waren dies die Naturwissenschaften; Reiss et al., 2016), lag es nahe, in der COACTIV-Studie vor allem die Mathematiklehrkräfte der untersuchten Klassen in den Blick zu nehmen.⁵

Während die Schüler im Rahmen von PISA 2003 an den Testtagen vormittags sowohl die Tests als auch die Fragebögen absolvierten, bearbeiteten deren Mathematiklehrkräfte an den entsprechenden Nachmittagen – ebenfalls unter der Aufsicht geschulter Testleiter – die COACTIV-Instrumente (Fragebögen und Tests). Genau wie Schüler bei PISA unter anderem biografische Daten angaben, über ihre Motivationen und (schulischen) Interessen berichteten sowie ihren Unterricht beurteilten, beantworteten auch die „COACTIV-Lehrkräfte“ biografische und motivationale Fragen und legten Rechenschaft über Aspekte ihres Unterrichts und ihres Berufserlebens ab. Bei der Konzeption der Fragebogeninstrumente für die Lehrkräfte konnte auf eine lange Tradition der pädagogisch-psychologischen Unterrichtsforschung zurückgegriffen werden: Seit etwa 100 Jahren werden – mit unterschiedlichen Akzentuierungen – Eigenschaften und Merkmale von Lehrkräften empirisch untersucht. Diese Akzentuierungen werden oft unter den Begriffen „Persönlichkeitsparadigma“, „Prozess-Produkt-Paradigma“ und „Experten-Paradigma“ zusammengefasst (Tab. 1).

5 Seit PISA 2003 wird auf nationaler Ebene in Deutschland regelmäßig ein – wenn auch im Vergleich zu COACTIV wieder stark verkürzter – Lehrerfragebogen eingesetzt. Die OECD führte den Lehrerfragebogen als internationale Option erst zur aktuellen Erhebungsrunde 2015 ein (an der 18 von 72 Staaten teilnahmen; vgl. Jude, Klieme, Kuger & Zehner, 2016). Themen des Lehrerfragebogens bei PISA 2015 waren z.B. Informationen zur Stellensituation, zur Berufswahl und zu aktuellen Arbeitsbedingungen an der Schule. Weiterhin wurden soziodemografische Merkmale, Aus-, Fort- und Weiterbildungsaspekte, Kooperation im Kollegium, Handeln im Unterricht sowie Zufriedenheit mit dem Beruf und speziell mit dem Unterrichten an der Schule erfasst (Jude et al., 2016). Da die COACTIV-Studie nicht genuiner Teil von PISA 2003 war, sondern eine von DFG und Max-Planck-Gesellschaft finanzierte Satellitenstudie (Kunter, Baumert et al., 2011), wurde das Professionswissen in den künftigen Lehrerfragebögen nicht mehr erfasst.

Tabelle 1: Paradigmen der Lehrerforschung (nach Krauss & Bruckmaier, 2014)

	Persönlichkeits-Paradigma	Prozess-Produkt-Paradigma	Experten-Paradigma
Zeit	ca. 1900–1960 (verstärkt empirisch etwa ab 1940)	ca. ab 1960 (bis heute)	ca. ab 1985 (heute zentral)
Beeinflusst durch	Eigenschaftsorientierte Persönlichkeitstheorien (etwa ab 1940 auch Persönlichkeitstests)	Behaviorismus (Verhalten des Lehrers)	Kognitivismus (Fokus auf Denken und Wissen des Lehrers)
Untersuchungs- methode	Tests und Fragebögen (Labor), Persönlichkeit des Lehrers im Vordergrund	Unterrichtsbeobachtung (später auch mit Videotech- nik), Handeln des Lehrers im Vordergrund	– Integration bisheriger Forschungsmethoden – Entwicklung von Professionswissenstests für Lehrer
Bemerkung	Nur wenige und oft schwache bzw. triviale Zusammenhänge	– Erste robuste und stabile Befunde – Unterrichtsaspekte objektiv, reliabel und valide messbar	– Systemische Sicht – Schwerpunkt wieder auf Person des Lehrers – Professionswissen ent- scheidend

Während sich insbesondere in der ersten Hälfte des letzten Jahrhunderts das Forschungsinteresse vor allem auf Persönlichkeitseigenschaften von Lehrkräften richtete (Persönlichkeitsparadigma), wechselte der Fokus mit dem Aufkommen des Behaviorismus in den 1950er-Jahren von der *Lehrerpersönlichkeit* auf das *Lehrerverhalten* (Prozess-Produkt-Paradigma), also auf Unterrichtsprozesse und entsprechende „Produkte“ auf Seiten der Schüler (z. B. Leistungszuwächse).⁶ In den letzten Jahrzehnten rückte die Person der Lehrkraft wieder mehr in den Mittelpunkt des Interesses (Expertenparadigma), doch diesmal stehen dabei nicht Charaktermerkmale, sondern vielmehr Wissen und Fertigkeiten im Vordergrund. Im derzeit aktuellen Expertenparadigma werden die vorherigen Strömungen im Wesentlichen wieder integriert und die Lehrkraft wird auf der Basis der Befunde der kognitionspsychologischen Expertiseforschung als Experte⁷ für die komplexen unterrichtlichen Anforderungen gesehen (Bromme, 1992; Krauss & Bruckmaier, 2014).

Im Rahmen dieser Paradigmen wurde eine Reihe von (mittlerweile in der Unterrichtsforschung bewährten) Untersuchungsinstrumenten entwickelt, die leistungsrelevante Unterrichtsaspekte wie beispielsweise Strukturiertheit und Klarheit der inhaltlichen Darstellung oder Facetten der Klassenführung betreffen. Viele der Fragen zum Unterricht, wie zum Beispiel zum Umgang mit Schülerfehlern oder zu Bewertungsmaßstäben, wurden sowohl den Schülern bei PISA als auch den Lehrkräften im Rahmen von COACTIV vorgelegt (Abb. 1, Mitte), was eine wechselseitige Vali-

6 Gelegentlich wird auch das „Prozess-Mediations-Produkt-Paradigma“ genannt, um hervorzuheben, dass Lehrkräfte nur Lerngelegenheiten bereitstellen, die von Schülern genutzt werden können (z. B. Helmke, 2010).

7 Auch wenn es in Deutschland alltagssprachlich unüblich ist, Lehrkräfte als „Experten“ zu bezeichnen (in einer Studie halten selbst in einer Stichprobe von Lehramtsstudenten nur 40% diese Bezeichnung für angemessen, Bromme, 1993), sind die Fachbegriffe „Experte“ und „Expertise“ in der wissenschaftlichen Diskussion über „gute Lehrer“ weit verbreitet.

dierung der Antworten ermöglichte. Die größte Herausforderung bei COACTIV bestand aber nicht in der Konzipierung beziehungsweise Adaption von Fragebögen für die Lehrkräfte – hierfür gab es genügend Vorlagen –, sondern in der Beantwortung des Problems, wie ein „Leistungstest“ für Mathematiklehrkräfte aussehen sollte, der den PISA-Leistungstests entsprach, die die Schüler zu bearbeiten hatten. Ein Befund, der in der kognitionspsychologischen Expertiseforschung schon für viele – gerade akademische – Domänen wiederholt verifiziert werden konnte, ist die entscheidende Rolle des professionellen Wissens (z. B. Gruber & Mandl, 1996). Experten zeichnen sich in vielen Bereichen wie zum Beispiel Medizin, Rechtsprechung oder auch kognitiv anspruchsvollen Spielen wie Schach gerade dadurch aus, dass sie über eine sehr breite und gut vernetzte Wissensbasis in ihrer Expertisedomäne verfügen und dieses Wissen schnell und flexibel einsetzen können (ausführlich zum Expertenparadigma siehe Abschn. 2.3).

Es lag also nahe, für die COACTIV-Studie das professionelle Wissen von Mathematiklehrkräften zu konzeptualisieren und zu operationalisieren, um es so einer Messung zugänglich zu machen. In der Literatur gab es als Grundlage zwar bereits theoretische Taxonomien des Lehrerwissens, beispielsweise die mittlerweile viel zitierte Unterscheidung von Fachwissen, fachdidaktischem Wissen und pädagogischem Wissen nach Lee Shulman aus den 1980er-Jahren (siehe hierzu ebenfalls Abschn. 2),⁸ eine Operationalisierung dieses Wissens für spezifische Unterrichtsfächer verbunden mit einer psychometrischen Testkonstruktion war zu diesem Zeitpunkt in der empirischen Lehr-Lernforschung aber noch Forschungsneuland.

Darüber hinaus wurden bei COACTIV möglichst umfassend auch relevante und bewährte Befragungsinstrumente für Lehrkräfte eingesetzt, die sich thematisch einem (oder mehreren) der drei Paradigmen zuordnen lassen (Tab. 1) und deren psychometrische Güte (Objektivität, Reliabilität und Validität) beziehungsweise prädiktive Validität (für Unterrichtsqualität bzw. Leistungszuwachs von Schülern) in früheren Studien bereits nachgewiesen worden waren. Insgesamt bearbeiteten die an der COACTIV-Studie beteiligten Mathematiklehrkräfte (über ein Jahr verteilt) Untersuchungsinstrumente mit einem zeitlichen Umfang von über zwölf Stunden, wobei die *Testinstrumente*, zu deren Items es normativ richtige Antworten gab, jeweils am PISA-Testtag nachmittags unter Aufsicht eines Testleiters bearbeitet werden mussten, während Fragebögen, zu deren Fragestellungen es nur subjektiv richtige Antworten gibt, von den Lehrkräften zuhause bearbeitet werden konnten (für eine Dokumentation der Erhebungsinstrumente siehe Baumert et al., 2009). Weiterhin reichten die bei COACTIV untersuchten Lehrkräfte Mathematikaufgaben ein, die sie in ihrer Klasse im Unterricht, als Hausaufgabe oder in Klausuren verwendet hatten (vgl. Abb. 1, Mitte unten). Insgesamt wurden bei COACTIV pro Mathematiklehrkraft über 1000 Variablen erhoben.

8 Shulman (1986, 1987) nennt noch curriculares Wissen, Wissen um die Psychologie des Lerner, Organisationswissen sowie erziehungphilosophisches, bildungstheoretisches und bildungshistorisches Wissen. Die drei genannten Bereiche haben sich jedoch mittlerweile als die Kernkategorien des Lehrerwissens etabliert (Lipowski, 2006; Blömeke, Kaiser & Lehmann, 2008; Helmke, 2010; Voss, Kunina-Habenicht, Hoehne & Kunter, 2015).

Der besondere Charme der COACTIV-Studie liegt darin, dass nicht nur Merkmale und Fähigkeiten von Lehrkräften gemessen wurden und diese beispielsweise zwischen verschiedenen Schulformen verglichen werden konnten, sondern dass es durch die enge Verzahnung mit PISA zudem möglich war, diese Lehreigenschaften mit dem Lernzuwachs der Schüler, aber auch mit weiteren unterrichtlichen Zielkriterien in Beziehung zu setzen. Abbildung 1 verdeutlicht dabei die den Analysen zugrunde gelegte Kausalstruktur: Lehrercharakteristika beeinflussen das Unterrichtsgeschehen, das wiederum einen Effekt auf die Entwicklung von Schülermerkmalen hat. In Deutschland wurde PISA 2003 zu einer Längsschnittstudie erweitert, das heißt, die getesteten Schüler der neunten Klassen hatten am Ende der zehnten Jahrgangsstufe abermals PISA-Fragebögen und Tests zu beantworten und auch die Lehrkräfte wurden 2004 wieder am Nachmittag des Testtages befragt und getestet (deshalb in Abb. 1 PISA 03/04 bzw. COACTIV 03/04). Somit war es möglich, tatsächlich die Bedeutung von Lehrer- beziehungsweise Unterrichtsvariablen für den Leistungszuwachs von Schülern (bzw. für den *Zuwachs* der erlebten Lernfreude oder die *Reduktion* von Leistungsangst) innerhalb eines Jahres zu modellieren und statistisch abzuschätzen (zum methodischen Vorgehen sowie zur Kontrolle von Variablen auf der individuellen Schülerebene siehe z. B. Kunter, Klusmann et al., 2013).

1.3 Effekte von Lehrermerkmalen auf Unterrichtsqualität und unterrichtliche Zielkriterien

Mittlerweile wurden mit den COACTIV-Daten verschiedene Strukturgleichungsmodelle spezifiziert, mit denen der Einfluss von Lehrervariablen auf Aspekte der Unterrichtsqualität oder auf unterrichtliche Zielkriterien analysiert wurde. Für die FALCO-Tests sind solche Modelle noch Zukunftsmusik, da diese Analysen aber die Motivation für die Beschäftigung mit dem professionellen Wissen von Lehrkräften auch in anderen Unterrichtsfächern waren, werden die entsprechenden COACTIV-Ergebnisse hier kurz zusammengefasst.

Alle bislang gerechneten Strukturgleichungsmodelle beziehen sich im Wesentlichen auf vier übergeordnete Kompetenzaspekte (Professionswissen, professionelle Überzeugungen, motivationale Orientierungen und selbstregulative Fähigkeiten, vgl. Abb. 1; für Details siehe Abb. 2), drei zentrale Merkmale der Unterrichtsqualität (kognitive Aktivierung, konstruktive Lernunterstützung, Klassenführung) beziehungsweise auf drei wichtige unterrichtliche Zielkriterien (Leistungszuwachs, Zunahme der Freude an Mathematik, Reduktion der Leistungsangst; vgl. Abb. 1).

Die untersuchten Strukturgleichungsmodelle unterscheiden sich im Wesentlichen zum einen darin, ob *direkte* Effekte von $L \rightarrow S$ (sog. „Black-box“-Modelle), über die Unterrichtsqualität mediiert von $L \rightarrow U \rightarrow S$ (sog. „Mediationsmodelle“) oder „lediglich“ von $L \rightarrow U$ oder von $U \rightarrow S$ geschätzt wurden (für L, U und S siehe Abb. 1 oben), und zum anderen dadurch, wie viele der Konstrukte (bzw. welche Teilbereiche wie z. B. Fachwissen oder fachdidaktisches Wissen) aus Abbildung 1 dabei integriert wurden (da sich die Komplexität der Modelle und somit die Zahl der zu schätzenden

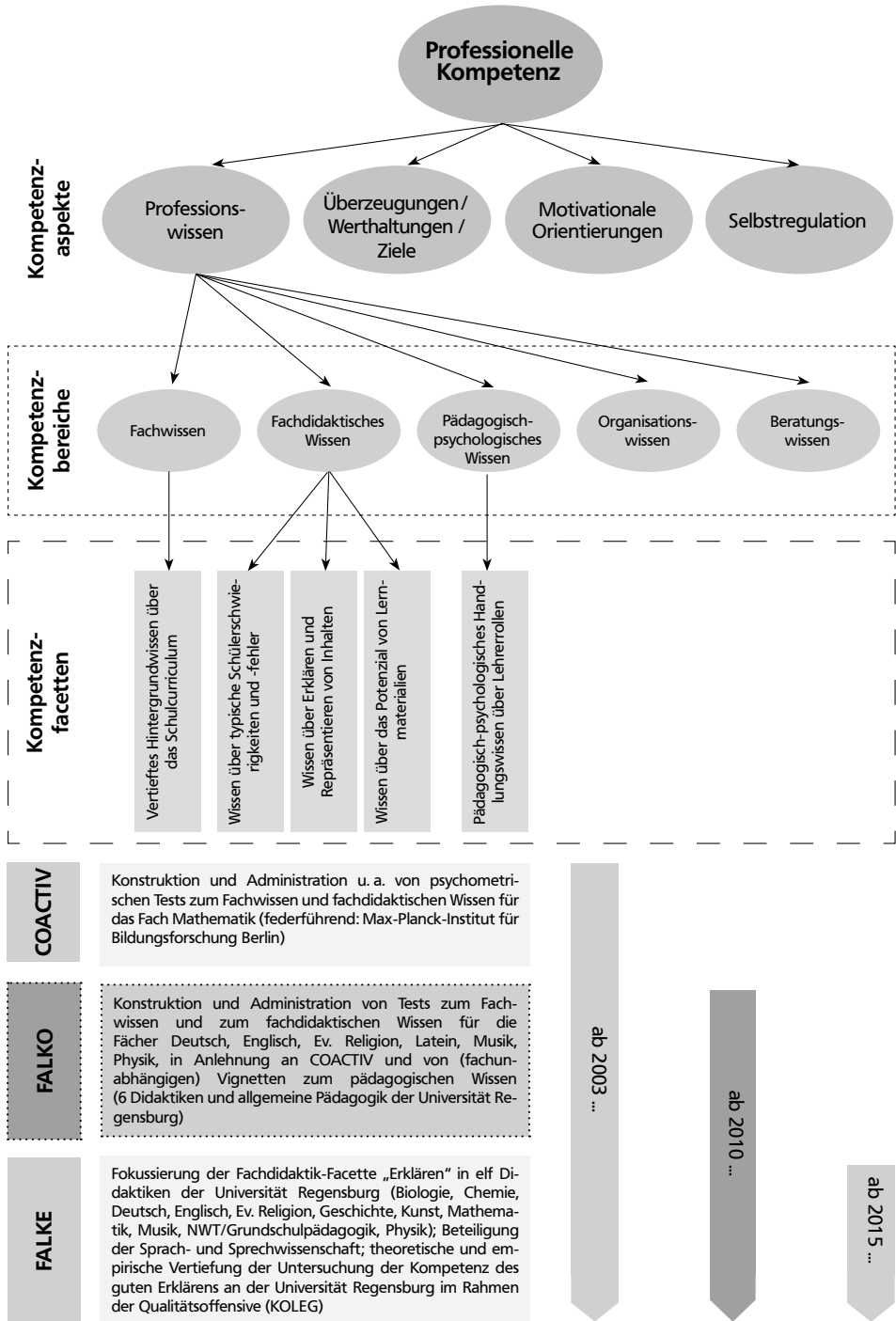
Abbildung 1: Die Verzahnung der COACTIV-Studie mit PISA 03/04 und entsprechendes Kausalitätsmodell (speziell unter Berücksichtigung von Konstrukten, die in Strukturgleichungsmodellen implementiert wurden)

<p>COACTIV 03/04 (Lehrerfragebögen und Tests)</p> <p>Lehrer (L) <i>Professionelle Kompetenz</i></p>	<p>COACTIV 03/04 (Lehrerfragebögen) PISA 03/04 (Schülerfragebögen)</p> <p>Unterricht (U) <i>Unterrichtsqualitätsmerkmale</i></p>	<p>PISA 03/04 (Schülerfragebögen und Tests)</p> <p>Schüler (S) <i>Unterrichtliche Zielkriterien</i></p>
<p><i>Professionswissen</i> (z. B. Fachwissen, fachdidaktisches Wissen, pädagogisches Wissen, diagnostische Kompetenz)</p>	<p><i>Kognitive Aktivierung</i> (z. B. über Anregungspotential der im Unterricht eingesetzten Aufgaben; vgl. unten)</p>	<p>Kernkriterium jeglichen Unterrichts: <i>Leistungszuwachs der Schüler</i></p>
<p><i>Professionelle Überzeugungen</i> (z. B. konstruktivistische oder transmissive Überzeugungen)</p>	<p><i>Konstruktive Lernunterstützung</i> (z. B. durch Schülerorientierung, Unterstützung bei Fehlern oder Geduld)</p>	<p><i>Steigerung der Lernfreude</i></p>
<p><i>Motivationale Orientierungen</i> (z. B. Enthusiasmus für das Fach oder für das Unterrichten)</p>	<p><i>Klassenführung</i> (z. B. durch Aufrechterhaltung von Disziplin oder adäquate Zeitnutzung)</p>	<p><i>Reduktion von Leistungsangst</i></p>
<p><i>Selbstregulative Fähigkeiten</i> (z. B. burn-out präventives Verhalten und Einstellungen)</p>	<p><i>Bem. 1:</i> Teilweise Items für Lehrer und Schüler parallel formuliert <i>Bem. 2:</i> Von Lehrkräften wurden darüber hinaus (insgesamt 40 000) Aufgaben eingesammelt und analysiert, die diese in den PISA-Klassen eingesetzt hatten (Jordan et al., 2008)</p>	<p><i>Bem.:</i> Es wurden zahlreiche weitere Konstrukte erhoben (Ramm et al., 2006; Prenzel et al., 2004)</p>
<p><i>Bem.:</i> Es wurden zahlreiche weitere Konstrukte erhoben (Kunter, Baumert et al., 2011; Baumert et al., 2009)</p>		

Parameter sehr schnell erhöht, gibt es hierbei gewisse Limitationen, z. B. Kline, 2011). Zunächst soll auf Modelle der Art $U \rightarrow S$ eingegangen werden.

Das theoriebasierte Modell der Unterrichtsqualität, das in Kunter und Voss (2011) ausführlich abgeleitet wird, wurde anhand dreier latenter Konstrukte (kognitive Aktivierung, konstruktive Lernunterstützung, Klassenführung) spezifiziert (vgl. Abb. 1, Mitte), wobei alle drei Dimensionen durch eine Vielzahl manifester Indikatoren identifiziert wurden. Die kognitive Aktivierung im Unterricht wurde beispielsweise über das Anregungspotential der von den Lehrkräften in ihren PISA-Klassen verwendeten Aufgaben abgeschätzt (z. B. ob mathematisches Argumentieren oder Modellieren erforderlich war), die konstruktive Lernunterstützung wurde über Schülerskalen zu Interaktionsaspekten wie Sozialorientierung, Umgang mit Fehlern oder Geduld der Lehrkraft operationalisiert und als Indikatoren für die Klassenführung dienten sowohl Lehrer- (COACTIV) als auch Schülerskalen (PISA) zu Disziplin, Unterrichtsunterbrechungen oder Zeitmanagement (für Details siehe Voss & Kunter, 2011).

Abbildung 2: Modell professioneller Lehrerkompetenzen von COACTIV (oben) sowie die Folgeprojekte FALKO (vorliegender Band) und FALKE (siehe Ausblick).



In drei getrennten Modellen ($U \rightarrow S$) wurde zunächst der gemeinsame Einfluss dieser drei übergeordneten Unterrichtsmerkmale auf den Leistungszuwachs (Modell 1), auf die Erhöhung der Lernfreude (Modell 2) und auf die Reduktion von Leistungsangst (Modell 3) berechnet (jeweils „Drei-Prädiktoren“-Modelle). Während kognitive Aktivierung und Klassenführung vor allem den Leistungszuwachs vorhersagen konnten, hatte die konstruktive Lernunterstützung einen signifikanten Einfluss auf die Reduzierung der Leistungsangst. Konstruktive Unterstützung und Klassenführung hatten weiterhin einen positiven Effekt auf die Lernfreude (Kunter & Voss, 2011).

Um noch einen Schritt in der Kausalkette zurückzugehen, wurde ein theoretisches und mittlerweile weit verbreitetes Modell professioneller Kompetenz von Lehrkräften postuliert (Abb. 2, für Details siehe Baumert & Kunter, 2011a). Die vier übergeordneten Kompetenzaspekte teilen sich jeweils in Kompetenzbereiche und schließlich weiter in Facetten auf (in Abb. 2 ist diese Unterteilung für das Professionswissen illustriert).

Zunächst wurden „Ein-Prädiktor“-Modelle spezifiziert, in denen oftmals zwei konkurrierende oder komplementäre Kompetenzbereiche eines übergeordneten Kompetenzaspekts (L) jeweils einzeln als potentielle Prädiktoren entweder für Unterrichtsmerkmale (U) oder für spezifische unterrichtliche Zielkriterien (S) angenommen wurden. So konnte zum Beispiel gezeigt werden, dass konstruktivistische, aber nicht transmissive Überzeugungen⁹ für bestimmte Zielkriterien prädiktiv valide sind (Voss & Kunter, 2011) und Enthusiasmus für das Unterrichten, jedoch nicht für das Fach alle drei Unterrichtsqualitätsmerkmale positiv beeinflusst (Kunter et al., 2008, 2011). Zudem hat diagnostische Kompetenz bezüglich der Bewertung von Aufgabenschwierigkeiten, allerdings nicht bezüglich der Einschätzung der Leistungsrangreihe der eigenen Schüler spezifische prädiktive Validität (Anders, Kunter, Brunner, Krauss & Baumert, 2010; Brunner et al., 2011). Weiterhin können sogar selbstregulative Fähigkeiten die Unterrichtsqualität beeinflussen (Klusmann, Kunter, Trautwein, Lüdtke & Baumert, 2008; Klusmann, 2011).

1.4 Effekte des fachbezogenen Professionswissens (Fachwissen und fachdidaktisches Wissen)

In Bezug auf FALKO sind vor allem die Modelle zum Fachwissen und zum fachdidaktischen Wissen interessant. FW wurde in COACTIV über 13 offene Items zum vertieften Hintergrundwissen über Inhalte des Schulcurriculums gemessen. Die Gesamtskala des FDW bestand aus neun Items zum Erklären und Repräsentieren mathematischer Inhalte (Subskala „E & R“), sieben Items zu Schülerfehlvorstellungen oder typischen Schwierigkeiten (Subskala „Schk“) und vier Items zum multiplen Lösungspotential von Aufgaben (Subskala „Pot“). Alle 20 Items des FDW-Tests hatten ebenfalls offenes Antwortformat. Die Testkonstruktion sowie wichtige diesbezügliche Ergebnisse wer-

⁹ Für die Konzeptualisierung der hier genannten Kompetenzbereiche sowie auf detaillierte Ergebnisse sei auf das jeweilige Buchkapitel des COACTIV-Sammelbandes verwiesen (Kunter, Baumert et al., 2011), die Operationalisierungen können in den meisten Fällen der Dokumentation der COACTIV-Erhebungsinstrumente entnommen werden (Baumert et al., 2009).

den ausführlich in Abschnitt 2 respektive Abschnitt 3 vorgestellt (Beispielitems sind Tab. 3a und b im nächsten Abschnitt zu entnehmen).

Baumert und Kunter (2011b) beziehungsweise Baumert et al. (2010) konnten mit Black-box-Modellen und mit Mediationsmodellen zeigen, dass das FDW die Unterrichtsqualität (genauer die kognitive Aktivierung und die Lernunterstützung) sowie den Lernzuwachs von Schülern in hohem Ausmaß erklären kann. Um die Größenordnung des Effekts abschätzen zu können, kann man sich zwei vollständig identische Schulklassen vorstellen, die von Lehrkräften unterrichtet werden, deren fachdidaktisches Wissen sich um zwei Standardabweichungen unterscheidet. Die Daten sagen für die eine Klasse einen Lernfortschritt von ungefähr $d = 0,2$ und für die andere Klasse von etwa $d = 0,5$ voraus. Dieser Unterschied von circa $d = 0,3$ entspricht in etwa dem durchschnittlichen Lernzuwachs eines ganzen Schuljahres (Baumert & Kunter, 2011b, Ehmke, Blum, Neubrand, Jordan & Ulfing, 2006).

Von besonderem Interesse ist laut Baumert und Kunter (2011b) ferner der Befund, dass auch konstruktive Lernunterstützung in Mathematik auf fachdidaktisches Wissen angewiesen ist und nicht (allein) als Ausdruck einer Haltung der sozialen Fürsorge verstanden werden darf. Als Hinweis auf die diskriminante Validität des fachdidaktischen Wissens lässt sich die Unabhängigkeit zur Klassenführung interpretieren. Eine geordnete Klassenführung – die wohl mehr vom generischen pädagogisch-psychologischen Wissen abhängt – ist grundsätzlich auch bei unzureichendem fachdidaktischen Wissen denkbar (Baumert & Kunter, 2011b).

Für das FW konnten keine vergleichbaren signifikanten Effekte auf Aspekte der Unterrichtsqualität oder auf die Schülerleistung nachgewiesen werden. Dies heißt jedoch keinesfalls, dass Fachwissen – gefasst als ein tiefes konzeptuelles Verständnis des mathematischen Unterrichtsstoffs – unwichtig sei. Die Daten legen vielmehr nahe, dass das Fachwissen den Entwicklungsraum des fachdidaktischen Wissens festlegt und damit auch indirekt die Unterrichtsqualität. Auch zahlreiche qualitative Studien haben gezeigt, dass Mängel im Fachwissen die Entwicklungsmöglichkeiten fachdidaktischer Ressourcen limitieren (vgl. hierzu Baumert & Kunter, 2011b). Die hohe Korrelation von FW und FDW in der COACTIV-Studie (vgl. Abschn. 3) legt (zumindest im Fach Mathematik) die Interpretation nahe, Fachwissen als *Prädiktor für das FDW* aufzufassen, da eine reichhaltige fachwissenschaftliche Basis eine wichtige Quelle für die Entwicklung von FDW ist (Krauss et al., 2008c; Lindl & Krauss, 2017).

Kunter und Kollegen (2013) haben die bis hierher gewonnenen Erkenntnisse genutzt, um ein komplexes Strukturgleichungsmodell mit den COACTIV-Daten aufzubauen (Abb. 3). Hierzu wurden vier Prädiktoren auf Lehrerseite ausgewählt, für die in bisherigen Modellen (einzeln) jeweils bedeutsame Effekte auf Unterrichtsqualitätsaspekte beziehungsweise auf Schülermerkmale nachgewiesen werden konnten: Fachdidaktisches Wissen, konstruktivistische Überzeugungen, Unterrichtsenthusiasmus und selbstregulative Fähigkeiten. Zudem wurden die drei zentralen Unterrichtsqualitätsmerkmale als Mediatoren modelliert und zwei Zielkriterien gewählt, nämlich der Leistungszuwachs und die emotional-motivationale Lernfreude. Zunächst haben Kunter et al. (2013) schrittweise fünf Black-box-Modelle spezifiziert ($L \rightarrow S$), in denen die vier Prädiktoren jeweils zuerst separat die beiden Zielkriteri-

en vorhersagen sollten (Modelle 1a–1d; Tab. 2) und schließlich ein gemeinsames Black-box-Modell zur Vorhersage beider Zielkriterien durch alle vier Prädiktoren (Modell 2). Tabelle 2 fasst die Ergebnisse dieser Black-box-Modelle zusammen.

Tabelle 2: Black-box-Modelle zur Vorhersage von Leistungszuwachs und Steigerung der Lernfreude (nach Kunter, Klusmann et al., 2013)

Modell	Prädiktor (L)		Leistungszuwachs (S)	Lernfreude (S)
			<i>Standardisierte Regressionskoeffizienten</i>	
1a	Fachdidaktisches Wissen	→	0,62*	- 0,03
1b	Konstruktivistische Überzeugungen	→	0,18*	0,04
1c	Unterrichtsenthusiasmus	→	0,20*	0,38*
1d	Selbstregulative Fähigkeiten	→	- 0,05	0,09
2	Fachdidaktisches Wissen	→	0,38*	- 0,05
	Konstruktivistische Überzeugungen		- 0,03	- 0,01
	Unterrichtsenthusiasmus		0,20*	0,37*
	Selbstregulative Fähigkeiten		- 0,07	0,04

Bem.: Signifikantes Regressionsgewicht: * $p \leq 0,05$.

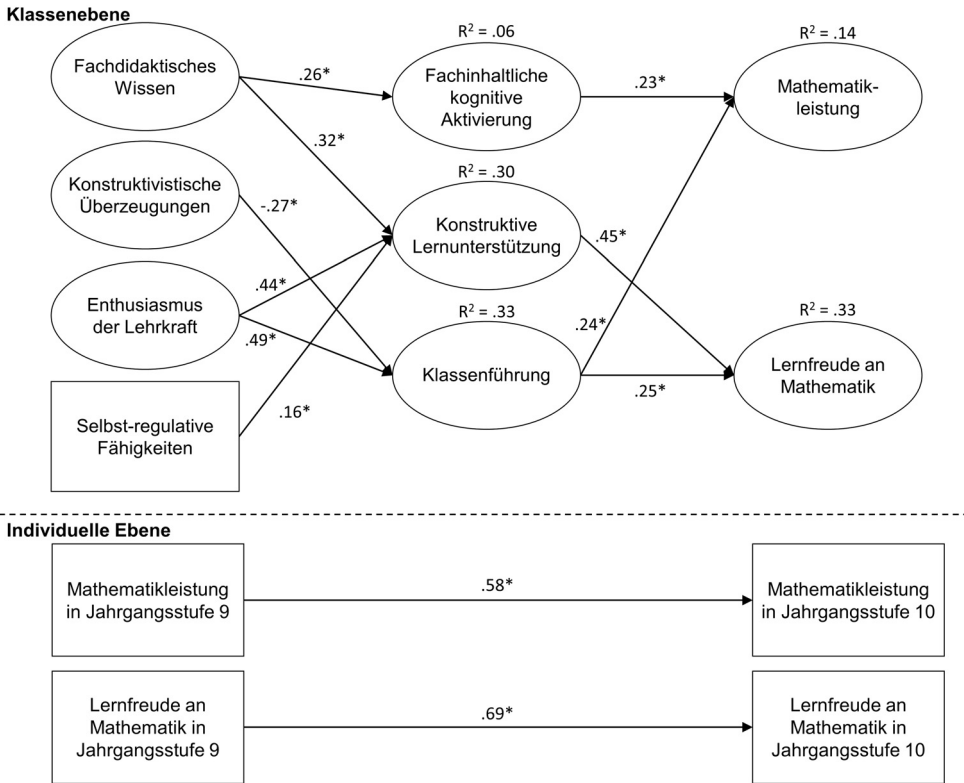
Die abgebildeten standardisierten Regressionskoeffizienten sind so zu interpretieren, dass beispielsweise eine Steigerung des FDW um eine Standardabweichung mit einem durchschnittlichen Zuwachs von 0,62 Standardabweichungen in der Schülerleistung einhergeht (Tab. 2). Augenscheinlich ist im Vergleich der vier „Ein-Prädiktor-Modelle“ das FDW der stärkste Prädiktor für den Leistungszuwachs (Modelle 1a–1d), der entsprechende standardisierte Regressionskoeffizient ist deskriptiv etwa dreimal so hoch wie diejenigen für konstruktivistische Überzeugungen oder Unterrichtsenthusiasmus.

Während der Unterrichtsenthusiasmus und die selbstregulativen Fähigkeiten (Modell 2) offensichtlich orthogonal sind und auch unbeeinflusst von den anderen beiden Prädiktoren variieren (die entsprechenden Regressionskoeffizienten ändern sich bei Hinzunahme der drei anderen Prädiktoren kaum), teilen FDW und konstruktivistische Überzeugungen gemeinsame Varianz, weshalb ihr Beitrag zum Leistungszuwachs bei wechselseitiger Kontrolle reduziert ist. Interessanterweise verschwindet der Einfluss der konstruktivistischen Überzeugungen in einem gemeinsamen Modell sogar gänzlich, der Einfluss des FDW bleibt jedoch nach wie vor deskriptiv am größten (Modell 2). Die Lernfreude wird in allen fünf Modellen nur vom Unterrichtsenthusiasmus der Lehrkraft beeinflusst.

Abbildung 3 illustriert das komplexeste Modell, das bislang mit den COACTIV-Daten gerechnet wurde (Kunter, Klusmann et al., 2013). In diesem Modell (Modell 3) wird Modell 2 noch um die Mediation durch die Unterrichtsmerkmale erweitert ($L \rightarrow U \rightarrow S$).

Ein erstaunliches Ergebnis, das hier jedoch nicht näher diskutiert werden soll, ist der negative Einfluss der konstruktivistischen Überzeugungen auf die Klassenführung (für eine Diskussion siehe Kunter, Klusmann et al., 2013). Im vorliegenden Zusammenhang ist vor allem der Befund von Bedeutung, dass der einzige Weg in Modell 3,

Abbildung 3: Integratives Mediationsmodell (Modell 3) mit vier Kompetenzbereichen (L), drei Unterrichtsqualitätsmerkmalen (U) und zwei Zielkriterien (S) (nach Kunter, Klusmann et al., 2013)



$N_{\text{Schüler}} = 4353$; $N_{\text{Lehrkräfte}} = 194$

Bem.: Als Ellipsen dargestellte Variablen basieren auf mehreren, rechteckige auf einzelnen Indikatoren; * $p \leq 0,05$.

über *fachinhaltliche* Aktivierung eine Leistungssteigerung zu erzielen, vom fachdidaktischen Wissen ausgeht. Unterrichtsenthusiasmus trägt zwar ebenfalls zur Leistungssteigerung bei, dieser Pfad wird aber durch die Klassenführung mediiert. Dies bedeutet, dass Schüler von motivierten und enthusiastischen Lehrkräften zwar ebenfalls mehr lernen, aber vermutlich nicht, weil Fachinhalte besser vermittelt werden, sondern weil der Enthusiasmus zur Sicherung der allgemeinen Lernbedingungen (Klassenführung) beiträgt. Ein weiteres bemerkenswertes Ergebnis ist, dass FDW nicht nur die Schülerleistung, sondern auch – vermittelt über konstruktive Lernunterstützung – die Freude der Schüler an Mathematik positiv beeinflusst.

1.5 Zwischenfazit

In der Gesamtschau aller vorliegenden Erkenntnisse zur prädiktiven Validität der bei COACTIV betrachteten Konstrukte scheint das Postulat der Expertiseforschung, dass das Wissen der erklärungsmächtigste Faktor für effektives professionelles Han-

deln ist – zumindest für das Fach Mathematik – auch auf die Lehrerdomäne zuzutreffen. Motivationale Orientierungen wie beispielsweise Enthusiasmus können zwar nominell größere Effekte haben (vgl. z.B. Abb. 3), aus dieser Quelle lässt sich aber kein *fachinhaltlich*, sondern nur über die Klassenführung mediierter Einfluss auf Schülerleistungen nachweisen. Darüber hinaus werden Motivation und Begeisterung oftmals gar nicht zur eigentlichen „Expertise“ gerechnet (z.B. Bromme, 2008), da sie im Gegensatz zum professionellen Wissen nur als bedingt veränderbar oder durch Ausbildung beeinflussbar gelten (z.B. Kunter, Kleickmann, Klusmann & Richter, 2011; vgl. auch Abschn. 2). In Black-box-Modellen, in denen nicht der zusätzliche Versuch unternommen wird, auch noch die spezifische Wirkungsweise von Lehrerkompetenzen aufzuschlüsseln, ist das FDW jedenfalls der deskriptiv stärkste Prädiktor für das Lernen der Schüler. Die Ergebnisse der COACTIV-Studie legen also nahe, sich auch in anderen Unterrichtsfächern mit dem professionellen Wissen von Lehrkräften zu beschäftigen.

Weiterhin ist auch die deutliche Verbesserung deutscher Schüler in den letzten PISA-Studien (Deutschland belegte 2012 und 2015 jeweils den 16. Platz von ca. 70 Ländern, also einen Platz im oberen Mittelfeld; Prenzel, Sälzer, Klieme & Köller, 2013; bzw. Reiss et al., 2016) ein eindrucksvoller Beleg für die Wirksamkeit von zahlreichen seit PISA 2000 implementierten Veränderungen. Beispielhaft seien hier Initiativen auf der Ebene von Schulstrukturen (z.B. Ganztagschulen), durch Leistungssetzungen (z.B. Bildungsstandards), durch Unterrichtsentwicklungsprogramme (z.B. SINUS für mathematisch-naturwissenschaftlichen Unterricht) oder durch direkte bildungspolitische Steuerung von Lerninhalten (z.B. über schulische Curricula oder Änderungen in der Lehramtsausbildung) genannt (Baumert, 2016). Mitverantwortlich für diese Verbesserung ist auch die empirische Bildungsforschung, nicht nur deshalb, weil sie auf Defizite hingewiesen hat, sondern auch, da in Folge des „PISA-Schocks“ eine evidenzbasierte Diskussion über professionelles Wissen von Lehrkräften – vor allem in mathematisch-naturwissenschaftlichen und in sprachlichen Fächern – angestoßen wurde.

Ein gelegentlich unter den Stichworten *Allgemeinbildung* oder *Grundbildung* geäußelter Kritikpunkt gegenüber PISA betrifft die Frage nach dem Verhältnis zwischen dem literacy-Konzept von PISA (Baumert et al., 2001) und dem Bildungsauftrag der Schule. In einem aktuellen Sonderheft der Zeitschrift für Erziehungswissenschaft zur empirischen Bildungsforschung (Baumert & Tillmann, 2016), in dem Befürworter und Gegner der empirischen Bildungsforschung zu Wort kommen, strukturiert Tillmann (2016) typische Einwände thematisch nach bildungstheoretischen Diskussionen über den Allgemeinbildungsbegriff, disziplinpolitischen Diskussionen über das Selbstverständnis der Erziehungswissenschaft, steuerungstheoretischen Diskussionen bezüglich Änderungen in Bildungssystemen, schulpädagogischen Diskussionen aus Sicht der Unterrichtspraxis, medienpolitischen Diskussionen in Bezug auf die journalistische und öffentliche Wahrnehmung, gerechtigkeits-theoretischen Diskussionen über Bildungschancen und der Humankapitaldiskussion bezüglich der Einflussnahme der OECD auf nationale Bildungspolitik.

Da augenscheinlich mit „empirische Bildungsforschung“ meist Forschung zu mathematischem oder naturwissenschaftlichem beziehungsweise neusprachlichem Unterricht gemeint ist, sieht sich diese gelegentlich dem bildungstheoretischen Reduk-

tionismusvorwurf (vgl. Tillmann, 2016) ausgesetzt: Durch die Fixierung bei PISA auf drei Disziplinen würde ein Großteil der Unterrichtsrealität (man denke an Fächer wie Erdkunde, Musik, Religion, Kunst, Sport, Latein oder Geschichte) gar nicht tangiert. Damit verbunden ist der Vorwurf, dass hiermit offenbar eine „Fixierung auf das Meßbare“ vorgenommen wurde (Tillmann, 2016, S. 11). Ein weiterer bildungstheoretischer Einwand betrifft die Nivellierung beziehungsweise Missachtung von (Teil-)Disziplinen, da beispielsweise durch die ausschließliche Konzentration auf Lesekompetenz literarisches Verstehen oder Schreibkompetenz außer Acht gelassen werden.

Vorwürfen dieser Art begegnet der vorliegende Band in zweierlei Hinsicht: Zum einen werden Fächer wie Latein, Musik oder Religion, die traditionell eher distal zur empirischen Bildungsforschung stehen, in ein gemeinsames didaktisches Forschungsprogramm integriert. So kann grundsätzlich auch überprüft werden, ob es sich bei mathematisch-naturwissenschaftlichem Unterricht tatsächlich um „gut strukturierte“ Disziplinen, bei sprachlichen Fächern um tendenziell geringer strukturierte Domänen (z. B. Blömeke et al., 2011) und bei Unterrichtsfächern, die auch eine ästhetische oder spirituelle Weltbegegnung vermitteln (z. B. Musik bzw. Religion), um „schlecht strukturierte“ Disziplinen handelt oder ob sich professionelle Kompetenzen von Lehrkräften in allen Fächern vergleichbar reliabel, valide und objektiv messen lassen (siehe hierzu Lindl & Krauss, 2017). Zum anderen wird bei den FALKO-Testkonstruktionen in den sprachlichen Fächern Deutsch, Englisch und Latein im Hinblick auf die Operationalisierung beider Professionswissensbereiche explizit auch Wert auf kulturelle beziehungsweise literarische Aspekte gelegt, da diese auch in schulischen Lehrplänen verankert sind.

In steuerungstheoretischen Diskussionen haben die Didaktiken bislang im Vergleich zu den „Big Playern“ aus der Pädagogik eine eher untergeordnete Rolle gespielt (bei PISA war die Fachdidaktik Deutsch trotz der Untersuchung von Lesekompetenz bislang nicht einmal beteiligt, vgl. Pissarek & Schilcher, 2017). Auch hier kann und sollte durch gemeinsame Anstrengungen in Forschung und Lehre (vgl. dazu auch den Ausblick des vorliegenden Bandes, Schilcher et al., 2017) Boden gut gemacht werden. Dass dies lohnend sein kann, belegt sicherlich auch die Tatsache, dass PISA 2015 nach bislang ausnahmslos pädagogisch-erziehungswissenschaftlicher Federführung (Baumert et al., 2001; Prenzel et al., 2004; 2007; 2013; Klieme et al., 2010) erstmals unter der Gesamtleitung einer Didaktik durchgeführt wurde (Reiss et al., 2016).

2. Fachwissen und fachdidaktisches Wissen: Testkonstruktionen und der zugrunde liegende Begriff der professionellen Kompetenz

Mittlerweile gibt es eine ganze Reihe von Forschungsansätzen, die domänenspezifische Messinstrumente für das Professionswissen (angehender) Lehrkräfte in unterschiedlichen Domänen vorgelegt haben. Psychometrische Testkonstruktionen für die drei Kernkategorien des Lehrerverwissens (FW, FDW, PW) sind interessanterweise überwiegend ein Phänomen des deutschsprachigen Raumes (vgl. Voss, Kunina-Habenicht, Hoehne & Kunter, 2015). Dies dürfte sicherlich nicht zuletzt auch auf den

„PISA-Schock“ zurückzuführen sein, wodurch der empirischen Bildungsforschung im Allgemeinen und den PISA-Studien im Speziellen gerade in der deutschen Öffentlichkeit und Wissenschaft große Aufmerksamkeit entgegengebracht wurde (vgl. Tillmann, 2016).

Im Folgenden sollen Historie und theoretischer Hintergrund der Testkonstruktionen bei COACTIV und FALKO erläutert und anschließend ein kurzer Überblick über weitere existierende fachbezogene Professionswissenstests bezüglich FW und FDW von (angehenden) Lehrkräften gegeben werden (siehe Tab. 4; für einen Überblick zum pädagogisch-psychologischen Wissen siehe Voss, Kunina-Habenicht, Hoehne & Kunter, 2015, sowie Mulder, Kempka & Sauer, 2017). Abschließend folgt eine Reflexion aktueller Diskussionen zur Unterscheidung von Wissen, Können und Handeln im Hinblick auf den Begriff der professionellen Kompetenz.

2.1 Testkonstruktionen zum Fachwissen und zum fachdidaktischen Wissen in COACTIV und FALKO

Bezüglich des Unterrichtsfachs Mathematik gab es neben COACTIV etwa zeitgleich noch die Testentwicklungen der Michigan-Arbeitsgruppe (Ball, Hill & Bass, 2005) sowie MT21 als Vorläuferstudie von TEDS-LT (Blömeke, Kaiser & Lehmann, 2008), weswegen dem Fach Mathematik hier oft eine Vorreiterrolle zugeschrieben wird (zu den Projektabkürzungen vgl. Abschn. 5.1). Obwohl die drei genannten Projekte unabhängig voneinander arbeiteten und darüber hinaus auch unterschiedliche Zielgruppen im Blick hatten (COACTIV: Sekundarschullehrkräfte; Michigan: Grundschullehrkräfte; MT21: Lehramtsstudierende und Referendare), gibt es zwei auffällige Parallelen, und zwar eine methodische und eine inhaltliche: Erstens finden sich in den Tests aller drei Gruppen Items, die als „hypothetische Unterrichtsszenarios“ formuliert sind, in denen Lehrkräfte in typische didaktische Anforderungssituationen versetzt werden (vgl. z. B. Tab. 3a). Zweitens werden in einem Teil der Items das „Erklären von Fachinhalten“ und „fachbezogene Schülervorstellungen“ explizit oder implizit thematisiert (ein Vergleich der genannten Ansätze findet sich in Krauss et al., 2008a). Diese beiden Aspekte werden auch in den meisten Testkonstruktionen zum fachdidaktischen Wissen aus jüngerer Zeit aufgegriffen (als Überblick z. B. Park & Oliver, 2008; Depaeppe, Verschaffel & Kelchtermans, 2013).

Der theoretische Vordenker, der bereits lange vor diesen ersten Operationalisierungen wichtige Konzeptualisierungsarbeit leistete und auf den sich heute nahezu alle Arbeitsgruppen berufen, die sich mit dem professionellen Wissen von Lehrkräften beschäftigen, ist Lee Shulman (1986, 1987). Shulman zeichnete nicht nur durch eine theoretische Anforderungsanalyse die taxonomische Strukturierung in die (heutigen) Kernkategorien FW, FDW sowie PW vor (für weitere Kategorien vgl. Anm. 8), sein Verdienst besteht auch in der Mahnung, die lange Zeit hauptsächlich generische Unterrichtsforschung *generell* stärker auf fachbezogene Aspekte auszurichten:

In their necessary simplification of the complexities of classroom teaching, investigators ignored one central aspect of classroom life: the subject matter. This omission also characterized most other research paradigms in the study of teach-

ing. Occasionally subject matter entered into the research as a context variable – a control characteristic for subdividing data sets by content categories (e. g., ‘When teaching 5th grade mathematics, the following teacher behaviors were correlated with outcomes. When teaching 5th grade reading, ...’). But no one focused on the subject matter content itself. [...] Why this sharp distinction between content and pedagogical process? (Shulman, 1986, S. 6)

In der Folge charakterisiert Shulman die beiden wesentlichen fachbezogenen Professionswissenskategorien, nämlich das *pedagogical content knowledge* (PCK, fachdidaktisches Wissen)¹⁰ und das *content knowledge* (CK, Fachwissen). Beide Charakterisierungen sollen ausführlich angeführt werden, da sie im Kern genau die Konzeptualisierungen von COACTIV und FALKO wiedergeben:

Fachdidaktisches Wissen beziehungsweise *pedagogical content knowledge*

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

Fachwissen beziehungsweise *content knowledge*

To think properly about content knowledge requires going beyond knowledge of the facts or concepts of a domain. It requires understanding the structures of the subject matter [...]. For Schwab (1978) the structures of a subject include both the substantive and syntactic structure. The substantive structures are the variety of ways in which the basic concepts and principles of the discipline are organized to incorporate its facts. The syntactic structure of a discipline is the set of ways in which truth or falsehood, validity or invalidity, are established. [...]. The teacher need not only to understand that something is so, the teacher must further understand why it is so, on what grounds its warrant can be asserted, and under what circumstances our belief in its justification can be weakened and even denied. (Shulman, 1986, S. 9)

Aus diesen drei Zitaten wird deutlich, dass hier die Fachdidaktiken gefordert sind und entsprechende Testkonstruktionen nicht Aufgabe der Erziehungswissenschaften oder der Psychologie sein können. Insbesondere die Charakterisierung des fachdidakti-

10 Zunächst wurde *pedagogical content knowledge* auch als „fachspezifisches pädagogisches Inhaltswissen“ übersetzt (z. B. Bromme, 1992), da die von Shulman beschriebenen Inhalte im deutschsprachigen Raum universitär aber im Rahmen der Fachdidaktik institutionalisiert sind, hat sich später die klarere Bezeichnung „fachdidaktisches Wissen“ durchgesetzt.

schen Wissens („... the ways of representing and formulating the subject that make it comprehensible to others.“, „the conceptions and preconceptions that students [...] bring with them ...“) wird in der einen oder anderen Form in den meisten gängigen Operationalisierungen zum FDW aufgegriffen. In COACTIV – und somit auch in allen FALKO-Testkonstruktionen – bilden das Wissen über „Erklären und Repräsentieren“ und das Wissen über „Schülerkognitionen und fachliche Fehlvorstellungen“ zwei explizite Subfacetten fachdidaktischen Wissens, zu denen jeweils Items entwickelt wurden. Während im Zuge des Konstruktivismus in den letzten Jahrzehnten die Thematisierung von Schülervorstellungen oder Fehlkzepten in der Lehr-Lernforschung ganz natürlich scheint, war die Untersuchung des „guten Erklärens“ in den Didaktiken in den letzten Jahren interessanterweise eher in den Hintergrund gedrängt (vgl. Ausblick, Schilcher et al., 2017).

Eine dritte didaktische Facette wurde bei COACTIV respektive FALKO je fachspezifisch, aber ebenfalls vor einem gemeinsamen theoretischen Hintergrund bestimmt: So dominieren beispielsweise Aufgaben das Unterrichtsgeschehen in Mathematik und empirische Studien belegen, dass gerade durch die Thematisierung verschiedener Lösungsmöglichkeiten kognitive Vernetzung und somit qualitativvolles Lernen ermöglicht werden können (Große, 2005; Silver et al., 2005). Die Aufgabe für die Lehrkräfte bezüglich der zu dieser Facette gehörigen Items im COACTIV-Test bestand darin, möglichst viele verschiedene und für Schüler gangbare Lösungswege von typischen Mathematikaufgaben (und somit das „Potential“ dieser Aufgaben für qualitativvolles Lernen) anzugeben. Beispielitems für die COACTIV-Tests finden sich in Tabelle 3a.

Diese „Potential“-Facette kann sich in unterschiedlichen FALKO-Fächern auf andere Lernmaterialien beziehen und auch auf andere Art und Weise realisiert werden. So wurde sie in Deutsch, Englisch und Latein beispielsweise als „Wissen über das Potential von Texten und (inter-)kulturellen Gegenständen“ operationalisiert (Pissarek & Schilcher, 2017; Kirchhoff, 2017; Lindl & Kloiber, 2017) oder in Physik als das „Wissen über Messen und Experimentieren“ (Schödl & Göhring, 2017). In allen sieben Disziplinen (COACTIV und sechs FALKO-Fächer) wird das FDW somit theoretisch in die drei Facetten „Wissen über Erklären und Repräsentieren“ (E & R), „Wissen über Schülerkognitionen, typische Fehler und Schwierigkeiten“ (Schk) und „Wissen über das Potential von Lernmaterialien“ (Pot) untergliedert.

Bezüglich des FW werden gelegentlich verschiedene Wissensniveaus vorgeschlagen, zum Beispiel (1) Alltagswissen, (2) Schulwissen, (3) vertieftes Hintergrundwissen zum Schulstoff sowie (4) rein universitäres Wissen (z. B. Krauss et al., 2011). Gerade in Mathematik, wo sich universitäres und Schulwissen sehr stark unterscheiden, hat sich auch die explizite Trennung von rein akademischem Fachwissen und Universitätswissen mit schulischem Bezug als fruchtbar erwiesen (Heinze, Dreher, Lindmeier & Niemand, 2016). Bei COACTIV und FALKO wird Fachwissen im Wesentlichen im Bereich zwischen Niveau 2 und 3 operationalisiert (für einen Überblick siehe Lindl & Krauss, 2017).

Tabelle 3a: Beispielitems aus den COACTIV-Tests zum Fachwissen und zum fachdidaktischen Wissen von Mathematiklehrkräften (aus Krauss et al., 2011; für Lösungsskizzen siehe Tab. 3b)

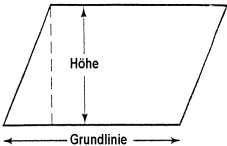
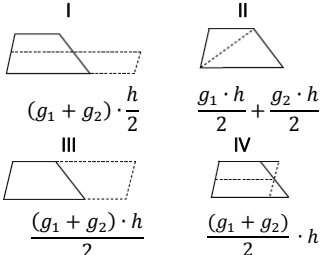
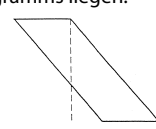
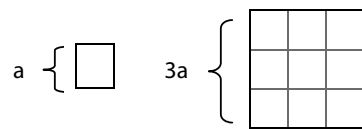
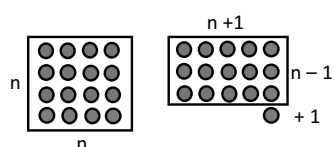
Kategorie	Beispielitem	
Fachdidaktisches Wissen (FDW) „Zugänglichmachen: Erklären und Repräsentieren“ (E & R)	„Minus 1 mal minus 1“ Eine Schülerin sagt: Ich verstehe nicht, warum $(-1) \cdot (-1) = 1$ ist. Bitte skizzieren Sie kurz möglichst viele verschiedene Wege, mit denen Sie der Schülerin diesen Sachverhalt klar machen könnten.	„Trapez“ Die folgenden Formeln liefern alle den Flächeninhalt eines Trapezes. $\text{I} \quad (g_1 + g_2) \cdot \frac{h}{2}$ $\text{II} \quad \frac{g_1 \cdot h}{2} + \frac{g_2 \cdot h}{2}$ $\text{III} \quad \frac{(g_1 + g_2) \cdot h}{2}$ $\text{IV} \quad \frac{(g_1 + g_2)}{2} \cdot h$ Welchen didaktischen Nutzen kann die Betrachtung dieser einzelnen Formeln haben? Bitte begründen Sie Ihre Antwort.
Fachdidaktisches Wissen (FDW) „Schülerkognitionen: Typische Fehler und Schwierigkeiten“ (Schk)	„Parallelogramm“ Die Fläche eines Parallelogramms lässt sich berechnen aus der Länge der Grundlinie mal Länge der Höhe.  Geben Sie bitte ein Beispiel eines Parallelogramms (anhand einer Skizze), an dem Schüler bei dem Versuch, diese Formel anzuwenden, eventuell scheitern könnten.	„Gleichung“ Bitte stellen Sie sich folgende Situation vor: Eine Schülerin berechnet für die Gleichung $(x-3) \cdot (x-4) = 2$ die Lösungen $x=5$ oder $x=6$ Was hat diese Schülerin vermutlich gerechnet?
Fachdidaktisches Wissen (FDW) „Multiples Lösungspotential von Aufgaben“ (Aufg)	„Quadrat“ <i>Wie ändert sich der Flächeninhalt eines Quadrats, wenn man die Seitenlänge verdreifacht?</i> <i>Begründe deine Antwort!</i> Bitte schreiben Sie möglichst viele verschiedene Lösungsmöglichkeiten (Begründungen) zu dieser Aufgabe kurz auf.	„Nachbarzahlen“ <i>Luca behauptet: „Das Quadrat einer natürlichen Zahl ist immer um 1 größer als das Produkt ihrer beiden Nachbarzahlen.“</i> <i>Stimmt Lucass Behauptung?</i> Bitte schreiben Sie möglichst viele verschiedene Lösungsmöglichkeiten zu dieser Aufgabe kurz auf.
Fachwissen (FW)	„Primzahl“ „Ist $2^{1024} - 1$ eine Primzahl?“ <i>(Dieses Item wurde der Testpilotierungsphase entnommen.)</i>	„Unendlicher Dezimalbruch“ Gilt $0,999999\dots = 1$? Bitte begründen Sie.

Tabelle 3b: Exemplarische Lösungsskizzen zu den Items aus Tab. 3a

Kategorie	Musterlösungen zu den Beispielitems aus Tabelle 3a	
Fachdidaktisches Wissen (FDW) „Zugänglichmachen: Erklären und Repräsentieren“ (E & R)	„Minus 1 mal minus 1“ Auch wenn die Permanenzreihe die Aussage nicht beweist, kann ihr Einsatz hier konzeptuelles Verständnis fördern und Verbindungen von Sachzusammenhängen herstellen: $ \begin{array}{r} 2 \cdot (-1) = -2 \\ 1 \cdot (-1) = -1 \\ 0 \cdot (-1) = 0 \\ (-1) \cdot (-1) = 1 \end{array} $	„Trapez“ Die vier Formeln entsprechen vier verschiedenen Varianten, die Trapezfläche herzuleiten (g_1 = Unterseite, g_2 = Oberseite, h = Höhe). 
Fachdidaktisches Wissen (FDW) „Schülerkognitionen: Typische Fehler und Schwierigkeiten“ (Schk)	„Parallelogramm“ Eine mögliche Lehrerantwort könnte hier zum Beispiel thematisieren, dass Schüler Schwierigkeiten bekommen könnten, wenn die Höhenfußpunkte außerhalb des Parallelogramms liegen: 	„Gleichung“ Die Schülerin übergeneralisiert vermutlich ein (nur für die Zahl 0 richtiges) Schema. Sie glaubt fälschlicherweise, dass allgemein für alle k gelten würde: Aus $(x - a)(x - b) = k$ folgt: $x - a = k$ oder $x - b = k$
Fachdidaktisches Wissen (FDW) „Multiples Lösungspotential von Aufgaben“ (Aufg)	„Quadrat“ Algebraisch: Flächeninhalt des Ursprungsquadrats: a^2 Flächeninhalt des „neuen“ Quadrats: $(3a)^2 = 9a^2$, also 9-mal so viel. Geometrisch: 9-mal das Ursprungsquadrat	„Nachbarzahlen“ Algebraisch: Sei n eine beliebige natürliche Zahl. $(n - 1) \cdot (n + 1) = n^2 - 1$, das ist um 1 kleiner als n^2 . Geometrisch:
		
Fachwissen (FW)	„Primzahl“ Nein, denn es gilt: $a^2 - b^2 = (a - b)(a + b)$. Demnach lässt sich $2^{1024} - 1$ zerlegen in $(2^{512} - 1) \cdot (2^{512} + 1)$.	„Unendlicher Dezimalbruch“ Sei $0,999\dots = a$ Dann ist $10a = 9,99\dots$, und deshalb gilt: $ \underbrace{10a - a}_{9a} = \underbrace{9,99\dots - 0,999\dots}_9 $ Also ist $0,999999\dots = 1$.