

Hohenstein, Friederike; Kleickmann, Thilo; Zimmermann, Friederike; Köller, Olaf; Möller, Jens  
**Erfassung von pädagogischem und psychologischem Wissen in der  
Lehrmatsausbildung: Entwicklung eines Messinstruments**

*Zeitschrift für Pädagogik 63 (2017) 1, S. 91-113*



Quellenangabe/ Reference:

Hohenstein, Friederike; Kleickmann, Thilo; Zimmermann, Friederike; Köller, Olaf; Möller, Jens:  
Erfassung von pädagogischem und psychologischem Wissen in der Lehrmatsausbildung: Entwicklung  
eines Messinstruments - In: Zeitschrift für Pädagogik 63 (2017) 1, S. 91-113 - URN:  
urn:nbn:de:01111-pedocs-184826 - DOI: 10.25656/01:18482

<https://nbn-resolving.org/urn:nbn:de:01111-pedocs-184826>

<https://doi.org/10.25656/01:18482>

in Kooperation mit / in cooperation with:

**BELTZ JUVENTA**

<http://www.juventa.de>

#### Nutzungsbedingungen

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen. Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

#### Terms of use

We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

#### Kontakt / Contact:

peDOCS  
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation  
Informationszentrum (IZ) Bildung  
E-Mail: [pedocs@dipf.de](mailto:pedocs@dipf.de)  
Internet: [www.pedocs.de](http://www.pedocs.de)

# ZEITSCHRIFT FÜR PÄDAGOGIK

Heft 1

Januar/Februar 2017

■ *Thementeil*

**500 Jahre Reformation – Schlaglichter  
auf pädagogische Langzeitwirkungen**

■ *Allgemeiner Teil*

Die soziale Herkunft von Lehramtsstudierenden  
in Begabtenförderungswerken

Passungsverhältnisse: Methodologische und theoretische  
Reflexionen zur Interaktionsorganisation des Unterrichts

Erfassung von pädagogischem und psychologischem  
Wissen in der Lehramtsausbildung

## Inhaltsverzeichnis

### *Thementeil: 500 Jahre Reformation – Schlaglichter auf pädagogische Langzeitwirkungen*

*Marcelo Caruso*

500 Jahre Reformation – Schlaglichter auf pädagogische Langzeitwirkungen.  
Einführung in den Thementeil ..... 1

*Annette Scheunpflug*

Evangelische Bildung heute – 500 Jahre nach der Reformation ..... 5

*Friedrich Schweitzer/Hans-Georg Ziebertz*

Gestalt und Wirkung des Religionsunterrichts:  
Empirische Befunde und interkonfessionelle Perspektiven ..... 29

### *Allgemeiner Teil*

*Martin Drahm/Martin Rothland*

Die soziale Herkunft von Lehramtsstudierenden in Begabtenförderungswerken.  
Eine vergleichende Analyse ..... 52

*Matthias Martens/Barbara Asbrand*

Passungsverhältnisse: Methodologische und theoretische Reflexionen  
zur Interaktionsorganisation des Unterrichts ..... 72

*Friederike Hohenstein/Thilo Kleickmann/Friederike Zimmermann/  
Olaf Köller/Jens Möller*

Erfassung von pädagogischem und psychologischem Wissen  
in der Lehramtsausbildung: Entwicklung eines Messinstruments ..... 91

*Besprechung*

*Rudolf Tippelt*

Volker Lenhart (Hrsg.): Die erste Schrift zur Vergleichenden  
Erziehungswissenschaft/The First Treatise in Comparative Education.

Fredericus Augustus Hechtius: De re scholastica Anglica cum  
Germanica Comparata (1795–1798)

David Phillips: Investigating Education in Germany.

Historical studies from a British perspective ..... 114

*Dokumentation*

Pädagogische Neuerscheinungen ..... 118

Impressum ..... U3

## Table of Contents

*Topic: 500 Years of Reformation –  
Highlights on pedagogical long-term effects*

*Marcelo Caruso*

500 Years of Reformation – Highlights on pedagogical long-term effects.

An introduction ..... 1

*Annette Scheunpflug*

Protestant Education Today – 500 years after the Reformation ..... 5

*Friedrich Schweitzer/Hans-Georg Ziebertz*

Forms and Effects of Religious Education: Empirical findings  
and interdenominational perspectives ..... 29

### *Articles*

*Martin Drahm/Martin Rothland*

The Social Background of Teacher Students in Funding Organisations.

A comparative analysis ..... 52

*Matthias Martens/Barbara Asbrand*

Complementary Relations: Methodological and theoretical reflections

on the formal organisation of classroom-interaction ..... 72

*Friederike Hohenstein/Thilo Kleickmann/Friederike Zimmermann/*

*Olaf Köller/Jens Möller*

Measuring Pedagogical and Psychological Knowledge

in Teacher Education: Development of a test instrument ..... 91

Book Reviews ..... 114

New Books ..... 118

Impressum ..... U3

Friederike Hohenstein/Thilo Kleickmann/Friederike Zimmermann/  
Olaf Köller/Jens Möller

# Erfassung von pädagogischem und psychologischem Wissen in der Lehramtsausbildung: Entwicklung eines Messinstruments

**Zusammenfassung:** Professionelles Wissen von Lehrkräften stellt einen entscheidenden Einflussfaktor auf die Qualität des Unterrichts sowie die Leistung der Schülerinnen und Schüler dar. Zum Professionswissen wird neben den fachlichen und fachdidaktischen Inhalten auch pädagogisch-psychologisches Wissen gezählt. Bisher wurde in der Forschung vermehrt auf das Fachwissen sowie das fachdidaktische Wissen fokussiert. Wenige Studien konzentrieren sich auf das pädagogisch-psychologische Wissen der Lehrkräfte. In der vorliegenden Arbeit, die im Rahmen des KiL-Projekts (*Messung professioneller Kompetenzen in mathematischen und naturwissenschaftlichen Lehramtsstudiengängen*) entstanden ist, wurden unter Berücksichtigung vorangegangener Studien und unter Einbezug der Standards für die Lehrerbildung in dem Bereich der Bildungswissenschaften zehn Facetten des pädagogisch-psychologischen Wissens für Lehramtsstudierende postuliert und ein Leistungstest entwickelt. Es wurde ein Itempool zu den zehn Facetten generiert und in einer Pilotierungsstudie ( $N = 1276$ ) an zwölf Hochschulen in Deutschland getestet. Sieben Skalen zeigten eine hinreichende Reliabilität und Validität.

**Schlagerworte:** Pädagogisch-Psychologisches Wissen, Testentwicklung, Professionswissen, Lehramtsausbildung, Lehramtsstudierende

## 1. Einleitung

Professionelle Kompetenz von Lehrkräften umfasst professionelles Wissen, Überzeugungen, motivationale Orientierungen und selbstregulative Fähigkeiten (Baumert & Kunter, 2006). Das Professionswissen als zentraler Aspekt der professionellen Kompetenz von Lehrkräften hat einen großen Einfluss auf die Unterrichtsqualität und die Schülerleistungen (Baumert & Kunter, 2006; Baumert et al., 2010; Bransford, Darling-Hammond & LePage, 2005; Hill, Rowan & Ball, 2005). Daher wird der Erforschung des Professionswissens eine besondere Bedeutung zugemessen. Professionswissen wird üblicherweise in die Kategorien Fachwissen (Content Knowledge, CK), fachdidaktisches Wissen (Pedagogical Content Knowledge, PCK) sowie pädagogisches bzw. pädagogisch-psychologisches Wissen (Pedagogical Knowledge, PK bzw. Pedagogical/Psychological Knowledge, PPK; Voss, Kunter & Baumert, 2011) der Lehrkraft eingeteilt (Baumert & Kunter, 2013; Shulman, 1987). Als PPK definieren wir hier in Übereinstimmung mit Voss et al. (2011) schulpädagogisches und pädagogisch-psychologi-

sches Wissen (zur genaueren Erläuterung der einzelnen Facetten siehe Abschnitt 2). Die bisherige Forschung fokussierte primär auf die Untersuchung von PCK und CK (Loewenberg Ball, Thames & Phelps, 2008; Baumert et al., 2010; De Jong, Van Driel & Verloop, 2005; Hill, Schilling & Ball, 2004; Krauss et al., 2008), weniger Beachtung fand bisher PPK. Im Allgemeinen wird unter PPK Wissen gefasst, „das für eine erfolgreiche Gestaltung und Optimierung der Lehr-Lern-Situation in verschiedenen Unterrichtsfächern nötig ist [...]“ (Kunter et al., 2011, S. 194). In den Projekten COACTIV-R (*Studie zum Erwerb professioneller Kompetenz von Lehramtsanwärtern während des Vorbereitungsdienstes*), PROWIN (*Professionswissen in den Naturwissenschaften*), Bil-Wiss (*Bildungswissenschaftliches Wissen und der Erwerb professioneller Kompetenz in der Lehramtsausbildung*) und TEDS-M (*Teacher Education and Development Study in Mathematics*) wurden bereits Testinstrumente zur Erfassung von PPK entwickelt, die an Lehrkräften und angehenden Lehrkräften aus unterschiedlichen Schulformen erprobt wurden (König, Blömeke, Paine, Schwidt & Hsieh, 2011; Lenske, Thilmann, Wirth, Dicke & Leutner, 2015; Tepner et al., 2012; Voss et al., 2011). Voss, Kunter, Seiz, Hoehne und Baumert (2014) konnten nachweisen, dass PPK einen Prädiktor für Basisdimensionen guten Unterrichts (Monitoring, Störungsprävention, kognitive Aktivierung und konstruktive Unterstützung) darstellt. Dies untermauert die Bedeutung dieser Wissensfacette für den Unterricht. So scheint ein Aufbau dieses Wissens schon im Lehramtsstudium erstrebenswert. Zusätzlich wurden von der Kultusministerkonferenz (KMK) bundesweite Standards vorgegeben, die eine Vereinheitlichung der Hochschullehre im Lehramtsstudium fächerübergreifend und fachspezifisch ermöglichen sollen (KMK, 2004, 2014). Erste Tests, mit deren Hilfe die Wissenszuwächse im Laufe des Lehramtsstudiums erfasst werden können, liegen bereits vor (z. B. König & Blömeke, 2009, König & Seifert, 2012). Die diesen Tests zugrunde liegenden Modelle konzentrieren sich der o. g. Definition folgend auf Aspekte des PPK, die eng mit dem Unterrichten als Kerntätigkeit von Lehrkräften in Zusammenhang stehen. Das Tätigkeitsspektrum von Lehrkräften geht jedoch über das Unterrichten hinaus. Daher schließen aktuelle Modelle des PPK auch Wissen über Beratung, Kooperation, Innovation, Schulentwicklung und berufliches Wohlbefinden ein (Kunina-Habenicht et al., 2012; Kunter et al., 2011; Voss, Kunina-Habenicht, Hoehne & Kunter, 2015). Um das PPK entsprechend untersuchen zu können, müssten zusätzlich solche Wissensaspekte berücksichtigt werden, die für diese Tätigkeitsfelder von Lehrkräften relevant sind. So wird PPK auch in den Standards für die Lehrerbildung für den Bereich der Bildungswissenschaften breiter konzipiert (KMK, 2004). Die Befunde einer aktuellen Delphi-Studie legen ebenfalls nahe, dass PPK im Lehramtsstudium umfassender definiert werden müsste (Kunina-Habenicht et al., 2012).

Im vorliegenden Artikel stellen wir ein Testinstrument vor, das PPK erfassen soll und setzen es in einer Stichprobe ein, um erste Hinweise auf die Skalenqualitäten zu bekommen.

## 2. Messung von pädagogischem und psychologischem Wissen

Zur Messung von PPK bei ausgebildeten Lehrkräften wurden von verschiedenen Arbeitsgruppen in den Projekten COACTIV-R, PROWIN, TEDS-M und BilWiss bereits Testinstrumente entwickelt, deren Dimensionen hier kurz vorgestellt werden (vgl. Tab. 1). Im Rahmen der COACTIV-R-Studie entwickelten Voss et al. (2011) ein aus 39 Items bestehendes Testinstrument zum PPK. Es erfasst vier Subskalen: *Klassenführung*, *Lehrmethoden*, *Leistungsbeurteilung* und *Heterogenität von Schülerinnen und Schülern*. Im Rahmen von PROWIN wurden 23 Items zu den Subskalen *Klassenführung*, *Unterrichtsmethoden*, *Leistungsbeurteilung* und *individuelle Lernprozesse* bei Gymnasial- und Hauptschullehrkräften (Lenske et al., 2015; Tepner et al., 2012) eingesetzt. König et al. (2011) entwickelten im Rahmen des Projekts TEDS-M einen Test, der in Deutschland, den USA und Taiwan erprobt wurde. Erfasst wird Wissen zu den Subskalen *Adaptivität des Unterrichts*, *Struktur des Unterrichts*, *Klassenführung* und *Leistungsbeurteilung*.

Im BilWiss-Projekt wird konkret der bildungswissenschaftliche, fachunspezifische Teil der Lehramtsausbildung untersucht. Die bildungswissenschaftlichen Inhalte im Lehramtsstudium sollen theoretisch systematisiert werden. Vor diesem Hintergrund wurde eine Expertenbefragung (Delphi-Studie, Kunina-Habenicht et al., 2012) durchgeführt, die die wesentlichen Inhalte des bildungswissenschaftlichen Begleitstudiums der Lehramtsausbildung exzerpieren sollte. Insgesamt 49 Expertinnen und Experten wurden gebeten, die wichtigsten Kernthemen des bildungswissenschaftlichen Anteils des Lehramtsstudiums zu identifizieren: Von 213 vorgegebenen Themen wurden 104 Themen als relevant ausgewählt, dazu zählen die Oberthemen *Bildungssystem und Schulorganisation*, *Bildungstheorie*, *Lehrerberuf*, *Heterogenität und soziale Konflikte*, *Unterricht*, *Sozialisationsprozesse*, *Diagnostik und Evaluation*, *Lernprozesse* und *Entwicklungsprozesse*. Im Rahmen des BilWiss-Projekts wurde ein Testinstrument zur Erfassung des bildungswissenschaftlichen Wissens mit 280 Items in den Dimensionen *Unterrichtsdidaktik*, *Schulpädagogik*, *Bildungstheorie*, *Lernen/Entwicklung* und *Diagnostik/Evaluation* entwickelt (Kunina-Habenicht et al., 2013) und ein spezifisches Testverfahren für das psychologische Wissen der Lehramtsstudierenden konzipiert (Lohse-Bossenz, Kunina-Habenicht, Dicke, Leutner & Kunter, 2015). Das Testinstrument (101 Items) soll Wissen in den Dimensionen *Lernen*, *Entwicklung* und *Diagnostik* bei Lehrkräften im Vorbereitungsdienst messbar machen.

Im Rahmen der vorliegenden Testentwicklung soll ebenfalls die Erfassung des fachunabhängigen Wissens im Vordergrund stehen. Die Entwicklung unseres PPK-Tests für Lehramtsstudierende bezieht die vorherigen Testentwicklungen und insbesondere auch die Standards der Kultusministerkonferenz (KMK, 2004) für die Lehrerbildung im Bereich der Bildungswissenschaften (im Folgenden KMK-Standards genannt) ein. Die KMK-Standards gehen deutlich über unterrichtsbezogenes pädagogisches Wissen hinaus, wie es von Shulman (1987) als Wissen über Klassenführung und Instruktion beschrieben wurde (vgl. auch Voss et al., 2011). Hinzu kommen beispielsweise Inhalte zur Schulorganisation, Schulentwicklung, Bildungs- sowie Stress- und Belas-

tungsforschung. Neben der Berücksichtigung der KMK-Standards wird bei der Testkonstruktion auf die Ergebnisse der Delphi-Studie (Kunina-Habenicht et al., 2012) zurückgegriffen.

Im Verständnis der vorliegenden Studie besteht das PPK vor allem aus schulpädagogischen *und* pädagogisch-psychologischen Inhalten und enthält mehr Facetten als die anderen bereits vorgestellten Testverfahren, wobei die Auswahl der Facetten einem gewissen Grad an Subjektivität unterliegt. Bei der Auswahl der Facetten haben wir uns primär auf die KMK-Standards gestützt und dann im weiteren Verlauf andere Testverfahren (König et al., 2011; Tepner et al., 2012; Voss et al., 2011) berücksichtigt sowie die Ergebnisse der Delphi-Studie (Kunina-Habenicht et al., 2012) einbezogen. Die empirische Prüfung ergab, dass einige Facetten nicht adäquat erfasst werden konnten. Somit entstand im Prozess der Testentwicklung ein Instrument, das eher pädagogisches und psychologisches Wissen anstelle von (noch breiter verstandenem) bildungswissenschaftlichem Wissen erfasst. Somit integriert das vorliegende Verfahren die Dimensionen der bestehenden Testinstrumente (König et al., 2011; Lenske et al., 2015; Tepner et al., 2012; Voss et al., 2011) und wird erweitert um Inhalte, die in den KMK-Standards sowie der Delphi-Studie (Kunina-Habenicht et al., 2013) als bedeutsam für die Lehramtsausbildung erachtet werden. Tabelle 1 zeigt die zehn Facetten des PPK und die entsprechenden, inhaltlich verwandten Skalen der anderen Verfahren bzw. der KMK-Standards und den Ergebnissen der Delphi-Studie. In der Tabelle ist bspw. abzulesen, dass die Facette *Lehren* des vorliegenden Instruments in Analogie zu den zentralen PPK-Dimensionen *Lehrmethoden* aus der COACTIV-R-Studie, *Strukturierung von Unterricht* aus TEDS-M und *Unterrichtsmethoden* aus dem Projekt PROWIN entwickelt wurde sowie Bezüge zu der in den KMK-Standards formulierten Kompetenz 1 (Planung und Durchführung von Unterricht) aufweist.

Als Kernbereiche des PPK werden die fünf Facetten *Lehren, Lernen und Entwicklung, Klassenführung, Leistungsbeurteilung* und *Individuelle Besonderheiten* berücksichtigt. Fünf weitere Facetten werden zusätzlich aus den KMK-Standards, der Delphi-Studie des Projekts BilWiss und eigenen Vorarbeiten abgeleitet. In diesen Vorarbeiten sind die Curricula zum bildungswissenschaftlichen Begleitstudium an 16 Hochschulen untersucht worden (Hohenstein, Zimmermann, Kleickmann, Köller & Möller, 2014). Die fünf Facetten sind *Bildungssystem und Schulorganisation, Methoden und Ergebnisse bildungswissenschaftlicher Forschung, Reflexion und Weiterentwicklung der eigenen Professionalität, Schulentwicklung* und *Werte, Normen und soziale Konflikte*. Im Folgenden werden die Facetten detaillierter erläutert.

Die erste Facette *Lehren* umfasst Wissen über die Kerntätigkeit des Unterrichtens (Lipowsky, 2015). Der Fokus liegt auf der Gestaltung von Lerngelegenheiten für Schülerinnen und Schüler durch die Lehrkräfte. Diese Facette beinhaltet beispielsweise Items zu Unterrichtsmethoden oder zum Einsatz von Lernmedien im Unterricht (Lipowsky, 2015).

Die zweite Facette *Lernen und Entwicklung* umfasst psychologisches Wissen über Bedingungen, Prozesse und Formen des Lernens und der Entwicklung. Dazu zählen u. a. Items zum selbstregulierten Lernen (Bannach, 2002).

Facette	Exemplarische Bezüge zu Inhalten bestehender Ansätze
Lehren	<ul style="list-style-type: none"> <li>– Lehrmethoden (COACTIV-R)</li> <li>– Unterrichtsmethoden (PROWIN)</li> <li>– Planung und Durchführung von Unterricht (KMK)</li> <li>– Strukturierung von Unterricht (TEDS-M)</li> </ul>
Lernen und Entwicklung	<ul style="list-style-type: none"> <li>– Lernprozesse (COACTIV-R)</li> <li>– Metakognition und selbstreguliertes Lernen (Oberbegriff <i>Lernprozesse</i>, Delphi-Studie)</li> <li>– Wissenserwerb (Oberbegriff <i>Lernprozesse</i>, Delphi-Studie)</li> <li>– kognitive Entwicklung (Oberbegriff <i>Entwicklungsprozesse</i>, Delphi-Studie)</li> <li>– Unterstützung von Lernen, Förderung selbstbestimmten Lernens, soziale und kulturelle Lernvoraussetzungen, Diagnostizieren von Lernvoraussetzungen (KMK)</li> </ul>
Klassenführung	<ul style="list-style-type: none"> <li>– Klassenführung (COACTIV-R)</li> <li>– Klassenführung (PROWIN)</li> <li>– Klassenführung und Motivation (TEDS-M)</li> <li>– Forschung zur Klassenführung (Oberbegriff <i>Unterricht</i>, Delphi-Studie)</li> <li>– Unterricht als soziale Situation (Oberbegriff <i>Unterricht</i>, Delphi-Studie)</li> <li>– Konflikte und Kooperation im Unterricht (Oberbegriff <i>Unterricht</i>, Delphi-Studie)</li> <li>– Positives Lernklima/anregende Lernumgebung (Oberbegriff <i>Unterricht</i>, Delphi-Studie)</li> <li>– Vorbereitung variabler/differenzierter Unterrichtsverläufe (Oberbegriff <i>Unterricht</i>, Delphi-Studie)</li> </ul>
Leistungsbeurteilung	<ul style="list-style-type: none"> <li>– Leistungsbeurteilung (COACTIV-R)</li> <li>– Leistungsbeurteilung (TEDS-M)</li> <li>– Leistungsbeurteilung PROWIN</li> <li>– Formen der Schülerbeurteilung (Oberbegriff <i>Unterricht</i>, Delphi-Studie)</li> <li>– Gestaltung von Rückmeldung an die Schüler (Oberbegriff <i>Unterricht</i>, Delphi-Studie)</li> <li>– Diagnostizieren von Lernvoraussetzungen, Leistungsbeurteilung- und Rückmeldung (KMK)</li> </ul>
Individuelle Besonderheiten	<ul style="list-style-type: none"> <li>– Individuelle Besonderheiten (COACTIV-R)</li> <li>– Umgang mit Heterogenität (TEDS-M)</li> <li>– Sozialisationsprozesse (Delphi-Studie)</li> <li>– kulturelle und soziale Heterogenität (Oberbegriff <i>Heterogenität und soziale Konflikte</i>, Delphi-Studie)</li> <li>– Kenntnis über soziale und kulturelle Lebensbedingungen und Einfluss auf individuelle Entwicklung (KMK)</li> </ul>
Bildungssystem und Schulorganisation	<ul style="list-style-type: none"> <li>– Besondere Anforderungen des Lehrerberufs (KMK)</li> <li>– Bildungssystem und Schulorganisation (Delphi-Studie)</li> </ul>
Methoden und Ergebnisse bildungswissenschaftlicher Forschung	<ul style="list-style-type: none"> <li>– Beruf als Lernaufgabe (KMK)</li> <li>– Forschung zur Klassenführung (Oberbegriff <i>Lehrerberuf</i> Delphi-Studie)</li> <li>– Forschung zu Unterrichtsmethoden (Oberbegriff <i>Unterricht</i>, Delphi-Studie)</li> </ul>
Reflexion und Weiterentwicklung der eigenen Professionalität	<ul style="list-style-type: none"> <li>– Techniken des Stressmanagements (Oberbegriff <i>Lehrerberuf</i>, Delphi-Studie)</li> <li>– Theorien und Forschungsansätze zum Lehrerberuf (Oberbegriff <i>Lehrerberuf</i>, Delphi-Studie)</li> <li>– Beruf als Lernaufgabe (KMK)</li> </ul>
Schulentwicklung	<ul style="list-style-type: none"> <li>– Beteiligung an Planung und Umsetzung schulischer Projekte (KMK)</li> </ul>
Werte, Normen und soziale Konflikte	<ul style="list-style-type: none"> <li>– Konfliktarten und deren Bedingungen (Oberbegriff <i>Heterogenität und soziale Konflikte</i>, Delphi-Studie)</li> <li>– moralische Entwicklung (Oberbegriff <i>Entwicklungsprozesse</i>, Delphi-Studie)</li> <li>– Vermittlung von Werten und Normen und Unterstützung selbstbestimmten Urteils und Handelns (KMK), Lösungsansätze für Schwierigkeiten und Konflikte in Schule und Unterricht (KMK)</li> </ul>

Tab. 1: Inhalte des PPK und Bezüge zu bestehenden Ansätzen

Die dritte Facette *Klassenführung* berücksichtigt Wissen, das eine effiziente Steuerung des Unterrichts erlaubt. Besondere Bedeutung haben dabei störungspräventive Maßnahmen der Unterrichtsgestaltung, aber auch die Etablierung von Regelsystemen und Grundsätze der Kommunikation und des Umgangs miteinander (Seidel, 2015). Zu dieser Facette gehören Themen wie die Prävention von Unterrichtsstörungen (Kounin, 2006) oder auch der Umgang der Lehrkraft mit den Schülerinnen und Schülern in kritischen Unterrichtssituationen (Becker, 2006).

Die vierte Facette *Leistungsbeurteilung* thematisiert Wissen über Funktionen und Formen der Leistungsbeurteilung sowie Wissen über Gütekriterien von Lehrerurteilen und deren Problematik (Südkamp, Kaiser & Möller, 2012). Zu dieser Facette zählen Kenntnisse über die summative und formative Leistungsbeurteilung (Weinert, 2002) sowie über unterschiedliche Bezugsnormen und deren Auswirkungen (Dickhäuser & Rheinberg, 2003).

Die fünfte Facette *Individuelle Besonderheiten* erfasst Wissen über die Bedeutung und den Umgang mit individuellen Besonderheiten von Schülerinnen und Schülern. Erfasst wird Wissen über Lernvoraussetzungen wie Vorwissen, Intelligenz, Interesse oder sozioökonomischer Status und über Lern- und Entwicklungsstörungen. Abgefragt wird zum Beispiel Wissen über Geschlechterdifferenzen (Hannover & Kessels, 2011), über ADHS (Gawrilow, 2012) und über Begabtenförderung (Ziegler, 2002).

Aus den KMK-Standards und der Delphi-Studie lässt sich die sechste Facette *Bildungssystem und Schulorganisation* ableiten. Die Lehramtsstudierenden sollten demnach die Strukturen des Bildungssystems und rechtliche Rahmenbedingungen kennen. Bei der Entwicklung des vorliegenden Testinstruments werden zum Beispiel schulrechtliche Kenntnisse und Wissen über die Bildungsexpansion (Füssel & Leschinsky, 2008) berücksichtigt.

Als siebte Facette werden Kenntnisse der *Methoden und Ergebnisse bildungswissenschaftlicher Forschung* erfragt. Die KMK-Standards sehen vor, dass Lehramtsstudierende die Ergebnisse von Schulleistungsstudien verstehen und Forschungsergebnisse kritisch bewerten können. Wissen über zentrale Ergebnisse bildungswissenschaftlicher Forschung umfasst beispielsweise Ergebnisse nationaler und internationaler Schulleistungsstudien, wie der PISA-Studie (Drechsel, Prenzel & Seidel, 2015). Im vorliegenden Test werden zudem Grundlagen der Statistik und Grundlagenkenntnisse zu Forschungsmethoden (Imhof & Langfeldt, 2010) erfasst. Diese Inhalte wurden auch in der Delphi-Studie von Kunina-Habenicht et al. (2012) genannt.

Die achte Facette des vorliegenden Testinstruments wird *Reflexion und Weiterentwicklung der eigenen Professionalität* genannt. Konkrete Inhalte sind beispielsweise die Belastungs- und Stressforschung (Barth, 1992; Schaarschmidt, 2008) oder auch Ansätze der Fort- und Weiterbildung (Besser & Krauss, 2009).

Die neunte Facette des PPK bezieht sich auf das Thema *Schulentwicklung*. Wissen über Schulentwicklung umfasst das grundlegende Wissen über Ziele und Methoden der Schulentwicklung, über das Schulleben und die Umsetzung schulischer Projekte. Inhaltlich abgefragt wird zum Beispiel Wissen über Evaluation (Ditton, 2000), zur Schulinspektion oder auch zu Schulprogrammen (Köller, 2009).

Die zehnte Facette *Werte, Normen und soziale Konflikte* beinhaltet Wissen über zentrale, insbesondere demokratische Werte und Normen und deren Vermittlung. Dazu gezählt werden auch Aspekte der moralischen Erziehung. Ferner umfasst diese Facette Wissen über außerunterrichtliche soziale Konflikte und den adäquaten Umgang mit diesen. So werden Items zur Werte- und Moralerziehung (Standop, 2005) und zum Beispiel zur Schülermediation für die Konfliktlösung (Schmitt, 2005) eingesetzt. Im Folgenden wird die Testentwicklung näher beschrieben.

### 3. Entwicklung des Testinstruments und Fragestellungen

Zu den zehn Facetten des PPK wurde ein Itempool von insgesamt 266 Items generiert. Zur Aufgabenentwicklung wurden neun Expertinnen und Experten der Bildungsforschung (mindestens promovierte Wissenschaftlerinnen und Wissenschaftler aus den Bereichen Pädagogische Psychologie und Erziehungswissenschaften) gebeten, Aufgaben zu den einzelnen Facetten zu entwickeln. Als Vorlage zur Itementwicklung diente ein detaillierter Leitfaden, der neben der inhaltlichen Beschreibung der Facetten auch Hinweise zur Aufgabengenerierung enthielt. Dazu zählten beispielsweise die konkrete Beschreibung der gewünschten Itemformate (offen, halboffen, geschlossen) sowie mögliche Schwierigkeiten bei der Itementwicklung (z. B. Verwendung von passenden Distraktoren). Eine genaue Vorgabe, in welchen Anteilen Items in einzelnen Formaten entwickelt werden sollten, gab es nicht, lediglich die Aufforderung, dass aufgrund der Ökonomie des Tests die Mehrzahl der Items im Multiple-Choice-Format vorliegen sollte. Die Expertinnen und Experten wurden zusätzlich gebeten, bereits entwickelte Items zu bewerten. Hierzu wurde ein spezielles Bewertungsschema pro Aufgabe ausgefüllt. Die Aufgaben sollten hinsichtlich verschiedener Punkte bewertet werden (z. B. Passung zum Konstrukt, Schwierigkeit, Kodierung oder sprachliche Formulierung der Aufgabe). Die Bewertung erfolgte entweder anhand der Angabe von Ja (1) bzw. Nein (0) oder bei Rating-Fragen anhand der Zustimmung auf einer Skala von 1 (gar nicht) bis 4 (sehr). Die Schwierigkeit der Items sollte mit einer Abstufung von leicht (1), mittel (2) und schwer (3) bewertet werden.

Diese Items wurden an Studierenden des Lehramts ( $N > 500$  Studierende) aller Fächer der Christian-Albrechts-Universität zu Kiel präpilottiert. In den ersten Analysen wurden Items aufgrund nicht akzeptabler Kennwerte (z. B. Trennschärfe) ausgeschlossen, andere Items wurden modifiziert und optimiert. Die Analysen fanden auf Ebene der einzelnen Skalen statt und wurden nach der klassischen Testtheorie berechnet.

Ziel der vorliegenden Studie war es, Skalen zur Erfassung der Subfacetten des PPK zu entwickeln und erste Evidenz zu deren psychometrischer Qualität zu gewinnen. Beispielitems für die einzelnen Subfacetten des PPK sind in Tabelle 2 dargestellt. Vor diesem Hintergrund wird angenommen, dass die zehn vorgeschlagenen Skalen moderat positiv interkorrelieren. Um Hinweise zur diskriminanten Validität der Skalen zu erhalten, werden zudem negative Korrelationen zwischen den zehn entwickelten Skalen und der Abiturnote (je besser die Leistungen, desto niedriger (besser) die Abiturnote) erwartet.

Facette	Itemstamm	Antwortmöglichkeiten
Lehren	Welche Rolle spielt Methoden- vielfalt im Unterricht für die fach- lichen Leistungen der Schülerin- nen und Schüler? Bitte markieren Sie die richtige Aussage.	a) Die fachlichen Schulleistungen sind am höchsten, wenn möglichst viele unterschiedliche Unterrichtsmethoden zum Einsatz kommen. b) * Eine mittlere Zahl (drei bis vier) unterschiedlicher Me- thoden ist am leistungsförderlichsten. c) Es sollten möglichst nur zwei unterschiedliche Methoden, diese aber dafür konsequent, zum Einsatz kommen. d) Es gibt keinen Zusammenhang zwischen Methodenviel- falt und Leistung.
Lernen und Entwicklung	Welche der folgenden Aussagen zum Lernen und Erinnern sind korrekt? Bitte kreuzen Sie jeweils „richtig“ oder „falsch“ an.	a) * Wiederholung erhöht die Aktivationsstärke von Informa- tionen im kognitiven System. b) * Vorwissen beeinflusst die Verstehens- und Behaltens- leistung. c) Motivation ist ein bedeutsamerer Prädiktor für Lernerfolg als Vorwissen.
Klassenführung	Jacob Kounin hat bereits 1972 Prinzipien einer effektiven Klas- senführung beschrieben. Welche der folgenden Prinzipien wurden von ihm als Prinzipien einer effek- tiven Klassenführung benannt? Bitte kreuzen Sie jeweils „richtig“ oder „falsch“ an.	a) * Überlappung b) * Allgegenwärtigkeit c) Effizienz d) * Aktivierung der ganzen Schülergruppe e) Freundliche Klassenraumgestaltung
Leistungs- beurteilung	Welche Aussagen zur summa- tiven Leistungsbeurteilung tref- fen zu? Bitte kreuzen Sie jeweils „richtig“ oder „falsch“ an. Bei der summativen Leistungs- beurteilung...	a) ... erhalten die Lernenden während des laufenden Unter- richts Rückmeldung über ihre Leistungen und Kenntnisse und die Differenz zum zu erzielenden Kenntnisstand. b) * ... geben Lehrkräfte oder schulexterne Personen den Schülerinnen und Schülern nach Abschluss einer Lern- sequenz Rückmeldung über ihre Kenntnisse und Fähig- keiten. c) ... wird die Leistung der Selbsteinschätzung der Schüle- rinnen und Schüler insgesamt gemessen.
Bildungssystem und Schulorgani- sation	Die Rechte und Pflichten von Lehrkräften im Schulbereich wer- den festgelegt durch ... Bitte markieren Sie die richtige Aussage.	a) ... das Schulrecht in Deutschland. b) * ... die Schulgesetze der Bundesländer. c) ... die Standards der Kultusministerkonferenz. d) ... die länderspezifischen Richtlinien der Bildung.
Methoden und Ergebnisse bil- dungswiss. For- schung	In empirischen Untersuchungen wird zwischen verschiedenen Ar- ten von Variablen unterschieden; was kennzeichnet eine <u>unabhän- gige Variable</u> ? Bitte markieren Sie die richtige Aussage.	a) Sie hat beliebig viele Abstufungen im Gegensatz zu Kon- trollvariablen. b) * Ihre Auswirkungen auf die abhängige Variable werden untersucht. c) Sie variiert nicht systematisch und ist daher auch als Konstante zu bezeichnen. d) Anhand dieser werden Personen in Versuchsbedingun- gen eingeteilt.
Reflexion und Weiterentwick- lung der eigenen Professionalität	Die pädagogisch-psychologi- sche Forschung beschäftigt sich seit langem mit der Suche nach der „guten Lehrkraft“. Was kenn- zeichnet Professionalität und er- folgreiches Handeln von Lehr- kräften? Bitte kreuzen Sie jeweils „richtig“ oder „falsch“ an.	a) Das Potenzial, eine erfolgreiche Lehrkraft zu werden, bringt eine Person bereits mit: mancher hat die Fähigkeit dazu, mancher nicht. b) Gute Lehrkräfte zeichnen sich vor allem durch eine Reihe positiver Persönlichkeitseigenschaften (Freundlichkeit, Humor, etc.) aus. c) * Ob eine Lehrkraft im Unterricht erfolgreich handelt, hängt maßgeblich vom Wissen der Lehrkraft ab. d) * Kompetenzen für erfolgreiches Lehrerhandeln werden erlernt und entwickelt.

Anmerkung: \* = richtige Antwort.

Tab. 2: Itembeispiele für die zehn Facetten zur Messung des PPK, Fortsetzung siehe nächste Seite

Facette	Itemstamm	Antwortmöglichkeiten
Schulentwicklung	Welches Ziel wird mit Schulinspektion in Deutschland verfolgt? Bitte markieren Sie die richtige Aussage.	a) Vorbereitung eines innerdeutschen Vergleichs im Sinne eines „Länderrankings“ b) Überprüfung der pädagogischen Arbeit der einzelnen Lehrpersonen c) Unterstützung der Schulleitung bei Beförderungsentscheidungen d) * Impulsgebung für die Schul- und Unterrichtsentwicklung
Individuelle Besonderheiten	Maria hat einen Intelligenzquotienten (IQ) von 100. Welche der folgenden Aussagen trifft zu? Bitte markieren Sie die richtige Aussage.	a) Maria hat eine leicht überdurchschnittliche Intelligenz. b) Maria hat eine leicht unterdurchschnittliche Intelligenz. c) * Maria hat eine durchschnittliche Intelligenz. d) Maria ist sehr wahrscheinlich hochbegabt.
Werte, Normen und soziale Konflikte	In welchem der folgenden Fälle handelt es sich um legitime Fremdbestimmung? Bitte markieren Sie die richtige Antwort.	a) Eltern überlassen ihrem Kind, ob es seine Hausaufgaben macht oder nicht. b) * Die Lehrkraft entscheidet, dass ein Schüler aufgrund seiner Leistungen im Fach Mathematik nicht versetzt wird. c) Die Lehrkraft entscheidet, dass eine Schülerin nicht an der Klassenreise teilnehmen kann, weil die Eltern nicht für die Kosten aufkommen können. d) Da sich niemand zur Wahl gestellt hat, bestimmt die Lehrkraft, wer das Amt des Klassensprechers übernimmt.

Anmerkung: \* = richtige Antwort.

Tab. 2: Itembeispiele für die zehn Facetten zur Messung des PPK

Ebenfalls wird aus diesem Grund der Zusammenhang mit den kognitiven Grundfähigkeiten betrachtet. Es wird angenommen, dass ein moderat positiver Zusammenhang vorliegt, da eine gewisse kognitive Grundfähigkeit für die Bearbeitung des Tests erforderlich ist, aber mit den Skalen nicht die allgemeine Intelligenz abgebildet werden soll. Des Weiteren werden positive Korrelationen der PPK-Ergebnisse mit der Anzahl absolvierter Semester im bildungswissenschaftlichen Begleitstudium und der Anzahl an besuchten bildungswissenschaftlichen Lehrveranstaltungen erwartet, da davon auszugehen ist, dass sich das pädagogische und psychologische Wissen im Rahmen der aktiven Nutzung formaler Lerngelegenheiten im Lehramtsstudium entwickelt. Des Weiteren wird untersucht, ob Studierende des Gymnasiallehramts in dem Test besser abschneiden als Studierende des Grund-, Haupt-, Real- und Sonderschullehramts. Zudem werden positive Zusammenhänge zwischen den Skalen des entwickelten Testinstruments und der Studienwahlmotivation der Studierenden erwartet. Es kann angenommen werden, dass Studierende mit einer hohen Studienwahlmotivation eher gute Leistungen in pädagogischen und psychologischen Lerninhalten aufzeigen.

## 4. Methode

### 4.1 Durchführung

Die vorliegende Testentwicklung wurde im Rahmen des Projekts KiL durchgeführt, welches ein Kooperationsprojekt zwischen dem Leibniz-Institut für die Pädagogik der Naturwissenschaften und der Mathematik sowie der Abteilung Psychologie für Pädagogen der Christian-Albrechts-Universität zu Kiel ist. Im Rahmen von KiL sollen Tests zur Erfassung des fachlichen (CK) und fachdidaktischen (PCK) Professionswissens von Lehramtsstudierenden sowie für das fachunabhängige (PPK) Professionswissen in den Fächern Mathematik, Chemie, Biologie und Physik entwickelt werden. Zwischen Juni und August 2012 wurde an zwölf Hochschulen in Deutschland (Hamburg, Kassel, Potsdam, Oldenburg, Bremen, Regensburg, Berlin, Kiel, Freiburg, Rostock, Bielefeld und München) die Testung der Verfahren für PCK und CK sowie für PPK durchgeführt. Zur Rekrutierung der Stichprobe wurden an den jeweiligen Hochschulen Werbeposter ausgehängt und Flyer ausgelegt. In Vorlesungen wurde vom Lehrpersonal mit Power-Point-Folien unterstützte Werbung gemacht.

Die Testzeit betrug für jede Testperson 240 Minuten incl. Pausenzeiten. Diese Testzeit wurde in acht Blöcke von je 20 Minuten gegliedert. Den Fächern wurden unterschiedlich viele Blöcke – jeweils für PCK und CK – zugeteilt: In Mathematik und Biologie wurden 30 Blöcke, in Chemie und Physik 16 Blöcke getestet. Der PPK-Test wurde mit 20 Blöcken getestet. Abhängig von den fachspezifischen Tests und den für diese Fächer teilnehmenden Studierenden konnten Subsets von PPK-Skalen an unterschiedlich vielen Probanden getestet werden. Zusätzlich zu den Wissenstests wurden ein Fragebogen zur Abfrage soziodemographischer Daten und ein Intelligenztest eingesetzt (siehe dazu Abschnitt 4.3). Zur Testung des PPK konnte mittels dieses Rotationsdesigns kein gesamtes Testheft mit allen Facetten eingesetzt werden. So ergibt sich eine variierende Probandenzahl für die einzelnen Facetten (von  $N = 84$  bis  $N = 685$  Probanden pro Facette), bei denen es keinen Zusammenhang zwischen Standort und Booklet gibt. Die Testung wurde mittels einer standardisierten Instruktion durchgeführt: Die Testleitungen wurden angewiesen einen vorgegebenen Text vorzulesen, um die Studierenden nicht bei der Bearbeitung der Testbögen zu beeinflussen. Die Studierenden erhielten nach der Teilnahme eine Aufwandsentschädigung von 40 Euro.

### 4.2 Stichprobe

An den zwölf Hochschulen wurden  $N = 1276$  Lehramtsstudierende (davon 66.2% weiblich) der mathematisch-naturwissenschaftlichen Fächer getestet, die zum Erhebungszeitpunkt zwischen einem und 19 Semestern studierten. Das Alter der getesteten Stichprobe reichte von 17 bis 41 Jahre ( $M = 22.87$  Jahre,  $SD = 2.75$ ). Die durchschnittliche Abiturnote lag bei  $M = 2.3$  ( $SD = 0.66$ ).

Facette	N	M Alter (SD)	M Abinote (SD)	M Sem. BW (SD)
Lehren	685	22.2 (2.86)	2.27 (0.58)	4.31 (3.75)
Lernen und Entwicklung	518	22.7 (2.87)	2.26 (0.57)	4.30 (3.97)
Klassenführung	352	23.0 (2.63)	2.29 (0.64)	4.16 (2.86)
Leistungsbeurteilung	239	22.5 (2.36)	2.32 (0.69)	4.02 (2.79)
Bildungssystem & Schulorganisation	373	23.8 (2.49)	2.27 (0.53)	4.00 (2.79)
Methoden und Ergebnisse bildungswiss. Forschung	84	22.8 (2.94)	2.31 (0.58)	4.56 (3.09)
Reflexion und Weiterentwicklung der eigenen Professionalität	213	23.4 (2.40)	2.32 (0.57)	3.96 (2.66)
Schulentwicklung	210	23.3 (2.39)	2.32 (0.57)	4.52 (3.06)
Individuelle Besonderheiten	239	22.5 (2.38)	2.32 (0.69)	3.93 (2.62)
Werte, Normen und soziale Konflikte	84	22.8 (2.9)	2.32 (0.59)	4.01 (2.62)

Anmerkungen: N = Stichprobengröße; M Alter = Mittelwert Alter; M Abinote = Mittelwert der Abiturnote; M Sem. BW = Mittelwert der absolvierten Semester des bildungswissenschaftlichen Begleitstudiums; SD = Standardabweichung.

Tab. 3: Überblick über die Stichprobe pro Facette

Den Teilnehmern und Teilnehmerinnen wurden in einer vierstündigen Testung Tests zum CK und zum PCK für die jeweils studierten Fächer Chemie, Biologie, Physik und Mathematik sowie die Aufgaben zum PPK vorgelegt. Die jeweiligen Testteile wurden mittels eines Rotationsdesigns administriert, sodass nicht jeder Proband jeden Testteil bearbeitet hat. Tabelle 3 gibt einen Überblick über die Stichprobenbeschreibung pro Facette.

Die Studierenden der PPK-Stichprobe gaben an, durchschnittlich 4.35 ( $SD = 3.33$ ) Semester in dem bildungswissenschaftlichen bzw. erziehungswissenschaftlichen Begleitstudium der Lehramtsausbildung und im Mittel 7.44 ( $SD = 6.03$ ) Lehrveranstaltungen in den Bildungswissenschaften bzw. Erziehungswissenschaften absolviert zu haben.

### 4.3 Instrumente

Zur Validierung des Testinstruments zur Erfassung des PPK wurden weitere Variablen erhoben.

#### Studienwahlmotivation

Die Studienwahlmotivation wurde mit zwei Skalen des *Fragebogens zur Erfassung der Motivation für die Wahl des Lehramtsstudiums* (FEMOLA; Pohlmann & Möller, 2010) erfasst. Für die vorliegende Studie wurden die Skalen *Pädagogisches Interesse* (7 Items;  $\alpha = .74$ ; Itembeispiel: „Ich habe das Lehramtsstudium gewählt, weil es mir

Spaß macht, die Erziehung von Kindern und Jugendlichen mitzugestalten“) und *Fachliches Interesse* (6 Items;  $\alpha = .71$ ; Itembeispiel: „Ich habe das Lehramtsstudium gewählt, weil ich mich gern mit den Inhalten meiner Fächer beschäftige“) verwendet. Die Stichprobe gab ihre Selbsteinschätzung auf vierstufigen Likert-Skalen an (von *trifft überhaupt nicht zu* bis *trifft völlig zu*).

### **Kognitive Fähigkeiten und Abiturnote**

Zur Messung der kognitiven Fähigkeiten wurde die Subskala *Sprachliches Denken* des *Kognitiven Fähigkeitstest* (KFT 4-12+R; Heller & Perleth, 2000) verwendet. Der Proband/die Probandin muss in diesem Subtest (20 Items,  $\alpha = .72$ ) Gemeinsamkeiten von Begriffen suchen. Zum anderen wurde die Abiturnote der Studierenden erfragt.

### **Studienverlauf**

Die Stichprobe wurde gefragt, wie viele Semester sie im bildungswissenschaftlichen Begleitstudium zum Zeitpunkt der Befragung absolviert hat. Zudem wurde nach der Anzahl besuchter Lehrveranstaltungen im bildungswissenschaftlichen Begleitstudium gefragt.

## **4.4 Analysen**

Die Skalierung erfolgte im Sinne der Item Response Theory (IRT). Die Analyse der Reliabilität erfolgte mittels einer Powerskalierung. Dabei werden alle Items, die von Studierenden am Testende nicht erreicht wurden sowie alle fehlenden Werte als ‚falsch‘ gewertet, sodass ein Datensatz ohne fehlende Werte vorliegt. Dies ist ein übliches Verfahren in der Testkonstruktion (vgl. dazu König, Blömeke & Schwippert, 2013). Die Skalen wurden zunächst mit dem Statistikprogramm Conquest (Wu, Adams & Wilson, 1998) auf der Basis des Raschmodells skaliert und Items anhand der *discrimination* selektiert. Die *discrimination* gibt als Korrelation an, in welchem Maß die richtige Antwortkategorie mit den anvisierten Fähigkeiten zusammenhängt. Die Beurteilung der Reliabilität erfolgte anhand des EAP/PV-Werts (*Plausible Values based Expected a Posteriori Estimation*), welcher die messfehler- und schätzfehlerbereinigte Reliabilität angibt (vgl. Wu et al., 1998). Zur Konstruktvalidierung wurden im Programm Mplus (Muthén & Muthén, 2007) anhand von 2-parametrischen Item-Response-Modellen schließlich messfehlerbereinigte Zusammenhänge zwischen den Skalen des Testinstruments und o. g. Außenkriterien berechnet. Als Schätzer wurde der robuste Maximum-Likelihood-Schätzer (MLR) verwendet.

## 5. Ergebnisse

### 5.1 Skalierung

In einem ersten Schritt wurden nach der klassischen Testtheorie die Skalen- und Itemkennwerte (Mittelwerte, Standardabweichungen, Trennschärfen und Cronbachs  $\alpha$ ) betrachtet. Items mit geringer oder negativer Trennschärfe ( $r_{it} < .15$ ; vgl. zur Trennschärfeanalyse Lienert & Raatz, 1998, S. 202) wurden nicht zur Skalenbildung herangezogen. Zu schwierige und zu leichte Items wurden aus dem Test ausgeschlossen. Des Weiteren wurden Items nach inhaltlichen Gesichtspunkten betrachtet und durch die Expertinnen und Experten auf einem Bewertungsbogen dahingehend eingeschätzt, ob die Items den Inhalt der jeweiligen Facette abdecken. Items mit geringer eingeschätzter Prototypizität für die anvisierte Facette und geringer Repräsentation des für die Aufgabenlösung notwendigen Wissens im Lehramtsstudium wurden ausgeschlossen bzw. umformuliert. Die Anzahl der Items reduzierte sich in diesem ersten Schritt von 266 Items auf 92 Items (die Aufteilung zu den Dimensionen ist in Tab. 3 ersichtlich). Für die Skala *Individuelle Besonderheiten* wurde eine Reliabilität von lediglich .30 und für die Skala *Werte, Normen und soziale Konflikte* eine Reliabilität von nur .55 ermittelt. Diese Werte liegen nicht im akzeptablen Bereich (Rost, 2004), weshalb die Skalen für die weiteren Analysen ausgeschlossen wurden. Die beiden Skalen *Reflexion und Weiterentwicklung der eigenen Professionalität* und *Schulentwicklung* wurden zu einer reliablen Skala *Weiterentwicklung/Innovieren* zusammengefasst, da die jeweilige Reliabilität unter .55 lag und nicht akzeptabel war. Nach Betrachtung der Iteminhalte der beiden Skalen wurde eine Zusammenführung für vertretbar gehalten, da beide das Wissen rund um das Thema Weiterentwicklung erfassen sollen (dieser Aspekt wird weiter unten diskutiert).

Im Folgenden werden die Ergebnisse der Skalenanalyse der sieben Skalen *Lehren, Lernen und Entwicklung, Klassenführung, Leistungsbeurteilung, Bildungssystem und Schulorganisation, Methoden und Ergebnisse bildungswissenschaftlicher Forschung* und *Weiterentwicklung/Innovieren* präsentiert.

Für die sieben Skalen wurden Reliabilitäten von .65 bis .81 festgestellt (vgl. Tab. 4). Da die Skalen breite Wissensaspekte erfassen sollen, werden diese Kennwerte als akzeptabel erachtet (Valencia, Stallmann, Commeyras, Pearson & Hartman, 1991).

Analysen der messfehlerbereinigten Interkorrelationen der einzelnen Skalen zeigen größtenteils erwartungskonforme Zusammenhänge. Die Skalen *Lehren* und *Lernen und Entwicklung* weisen eine hohe Korrelation auf ( $r = .72$ , vgl. Tab. 5). Unerwarteter weise liegt die Korrelation der Skala *Klassenführung* mit der Skala *Leistungsbeurteilung* mit  $r = -.27$  im negativen Bereich; mit *Methoden bildungswissenschaftlicher Forschung* korreliert sie zudem nicht signifikant.

Skala	N	Aufgabenanzahl	EAP/PV Reliabilität	Trennschärfe
Lehren	685	14	.72	.18–.65
Lernen und Entwicklung	518	13	.72	.31–.70
Klassenführung	325	16	.78	.27–.55
Leistungsbeurteilung	239	11	.65	.40–.54
Bildungssystem & Schulorganisation	371	13	.66	.39–.55
Methoden bildungswissenschaftlicher Forschung	84	10	.81	.33–.66
Weiterentwicklung/Innovieren	210	15	.74	.27–.60

Anmerkungen: N = Stichprobengröße; Aufgabenanzahl = Items pro Dimension.

Tab. 4: Übersicht über die sieben entwickelten Skalen, Stichprobengröße, Aufgabenanzahl, Reliabilität und Itemtrennschärfe

Nr.	Skala	1	2	3	4	5	6
1	Lehren						
2	Lernen und Entwicklung	.72*					
3	Klassenführung	.30*	.40*				
4	Leistungsbeurteilung	.38*	.46*	–.27*			
5	Bildungssystem & Schulorganisation	.25*	.30*	.23*	.30*		
6	Methoden bildungswiss. Forschung	.20	.31*	.06	.20*	–	
7	Weiterentwicklung/Innovieren	.27*	.46*	.37*	.34*	.30*	–

Anmerkungen: \*  $p < .05$ . Da Dimension 6 nicht in allen Stichproben getestet wurde, sind Zusammenhangsanalysen mit den Dimensionen 5 und 7 nicht möglich.

Tab. 5: Interkorrelationen der Skalen des Instruments zur Erfassung des PPK

## 5.2 Zusammenhänge mit Außenkriterien

Die Korrelationen der Skalen mit den Außenkriterien stützen unsere Annahmen weitgehend (vgl. Tab. 6). Zu den Skalen *Lehren* sowie *Lernen und Entwicklung* finden sich positive Zusammenhänge des fachlichen und des pädagogischen Interesses. Entgegen den Erwartungen zeigt sich ein starker negativer Zusammenhang zwischen der Skala *Pädagogisches Interesse* und dem Abschneiden im Untertest *Klassenführung* (Tab. 6).

Die Abiturnote korreliert negativ mit allen sieben Skalen des Testinstruments: Studierende mit einer besseren Abiturnote schneiden in allen Skalen besser ab als Studierende mit einer schlechteren Abiturnote (Tab. 6). Darüber hinaus zeigten sich positive

Skala	N	Abitur- note	Sem. BW	LV BW	KFT-V	Gym vs. GRHS	Päd. In- teresse	Fachl. In- teresse
Lehren	820	-.31*	.22*	.44*	.40*	.16*	.38*	.31*
Lernen und Entwicklung	655	-.27*	.23*	.34*	.43*	.14*	.33*	.23*
Klassen- führung	361	-.18*	-.08	-.27*	.24*	.30*	-.55*	.06
Leistungsbe- urteilung	239	-.21*	.19	.11	.35*	.04	.04	-.07
Bildungssys- tem/Schul- organisation	136	-.25*	.41*	-.02	.19	.00	-.06	-.04
Methoden	84	-.25	.27*	.32*	.28*	-.03	-.03	.08
Weiterent- wicklung/ Innovieren	209	-.14*	.40*	.18	.33*	.37*	-.08	-.02

Anmerkungen: \*  $p < .05$ ; Sem. BW = besuchte Semesteranzahl des bildungswissenschaftlichen Studiums; LV BW = besuchte Lehrveranstaltungen der bildungswissenschaftlichen Studienanteile; KFT-V = Untertest Sprachliches Denken; Gym vs. GRHS = Vergleich zwischen Gymnasiallehramt und Grund-, Haupt- und Real- und Sonderschullehramt; Päd. Interesse = Untertest zum Pädagogischen Interesse des FEMOLAs (Pohlmann & Möller, 2010); Fachl. Interesse = Untertest zum Fachlichen Interesse des FEMOLAs (Pohlmann & Möller, 2010).

Tab. 6: Korrelationen der Skalen des Instruments zur Erfassung des PPK mit den Außenkriterien

Korrelationen der meisten Skalen des PPK-Tests mit dem Subtest *Sprachliches Denken* des KFT. Lediglich für die Skala *Bildungssystem und Schulorganisation* zeigte sich kein signifikanter Zusammenhang.

Die Anzahl absolvierter Semester im bildungswissenschaftlichen Begleitstudium hängt positiv mit dem Abschneiden im Test (mit Ausnahme der Skalen *Klassenführung* und *Leistungsbeurteilung*) zusammen (Tab. 6). Die Anzahl der absolvierten bildungswissenschaftlichen Lehrveranstaltungen korreliert ebenfalls positiv mit den Testergebnissen (eine Ausnahme bildet auch hier die Skala *Klassenführung*).

Studierende des Gymnasiallehramts erzielen in den vier Skalen *Lehren* ( $d = 0.76$ ), *Lernen und Entwicklung* ( $d = 0.23$ ), *Klassenführung* ( $d = 0.18$ ) und *Weiterentwicklung/Innovieren* ( $d = 0.33$ ) signifikant bessere Ergebnisse als Studierende der anderen Lehrämter. Auch bei Kontrolle der Abiturnote, der Anzahl absolvierter Semester und der Lehrveranstaltungen des bildungswissenschaftlichen Begleitstudiums, der Ergebnisse im KFT sowie der fachlichen und pädagogischen Studienwahlmotivation bleiben diese Unterschiede bestehen.

## 6. Diskussion

In der vorliegenden Untersuchung wurde ein Testinstrument zur Erfassung von pädagogischem und psychologischem Wissen von Lehramtsstudierenden vorgestellt, das sich an vorherige Testentwicklungen (König et al., 2011; Kunina-Habenicht et al., 2013; Tepner et al., 2012; Voss et al., 2011) anlehnt sowie zusätzliche Inhalte aus den KMK-Standards und Ergebnisse aus der Delphi-Studie (Kunina-Habenicht et al., 2012) aufgreift. Mittels Trennschärfeanalysen der Items und Reliabilitätsanalysen der Skalen konnten sieben reliable Subskalen identifiziert werden, die Facetten des pädagogischen und psychologischen Wissens abbilden

### 6.1 Validität der Skalen

Zunächst werden die Zusammenhänge zwischen den Skalen diskutiert, anschließend die Zusammenhänge mit den Außenkriterien. Die Zusammenhänge zwischen den einzelnen Skalen decken sich mit den Erwartungen an die Entwicklung eines Wissenstests weitgehend. Eine Ausnahme stellt der nicht signifikante Zusammenhang der Skala *Klassenführung* mit der Skala *Methoden und Ergebnisse bildungswissenschaftlicher Forschung*, und insbesondere der negative Zusammenhang mit der Skala *Leistungsbeurteilung* dar. Möglich ist, dass gerade die richtigen Antworten zu den Aufgaben aus der Skala *Klassenführung* vielen Lehramtsstudierenden kontraintuitiv erscheinen mögen.

Die Zusammenhangsanalysen mit den Außenkriterien geben erste Hinweise darauf, dass die entwickelten Skalen eine zufriedenstellende Validität aufweisen. Fast alle Skalen zeigen die erwarteten Zusammenhänge: Eine höhere Studienwahlmotivation hängt mit höherem pädagogischem und psychologischem Wissen zusammen. Das *pädagogische* und *fachliche Interesse* hängt vor allem mit dem Wissen in den Bereichen *Lehren und Lernen und Entwicklung* zusammen. Dagegen korreliert die Skala *Klassenführung* negativ mit dem *pädagogischen Interesse*, was darauf hinweist, dass mit der Skala eher pädagogische Überzeugungen und weniger Wissen über Klassenführung erfasst wird. Das Wissen zur Klassenführung scheint teilweise den erzieherischen Intentionen von Lehramtsstudierenden zu widersprechen.

Wie erwartet zeigten sich moderate positive Zusammenhänge der Skalen mit der Abiturnote. Darüber hinaus zeigten sich positive Korrelationen der Skalen des pädagogischen und psychologischen Tests mit dem Subtest *Sprachliches Denken* des KFT. Lediglich für die Skala *Bildungssystem und Schulorganisation* zeigte sich kein signifikanter Zusammenhang.

Auffallend ist der negative Zusammenhang zwischen der Skala *Klassenführung* und der Anzahl der besuchten Lehrveranstaltungen des bildungswissenschaftlichen Begleitstudiums. Eine mögliche Erklärung dafür ist, dass die Inhalte, die durch die Skala *Klassenführung* abgefragt werden, nicht in das Curriculum der Hochschulen implementiert wurden, und den Studierenden in Bezug auf Klassenführung eher Techniken der Klas-

senführung vermittelt werden, die nicht mit den Annahmen von Kounin (2006) übereinstimmen.

Es muss hervorgehoben werden, dass die Ergebnisse der Zusammenhangsanalysen mit Vorsicht zu interpretieren sind. Die Zusammenhänge der Skalen mit den kognitiven Fähigkeiten sind vielleicht weniger aussagekräftig als Validierungshinweis, da die Intelligenz ein wichtiger Prädiktor für den Studienerfolg ist (Schmidt & Hunter, 1998) und somit intelligentere Studierende generell im Studium erfolgreicher sind und mehr lernen. Wir interpretieren die Befunde im Sinne einer diskriminanten Validierung: moderat positive Zusammenhänge sind anzunehmen, da für die Testbearbeitung kognitive Fähigkeiten verlangt werden, allerdings ist es nicht das Ziel die allgemeine Intelligenz zu messen.

Die Studierenden des gymnasialen Lehramts schneiden in einigen Skalen besser ab als Studierende der nicht-gymnasialen Studiengänge. Dies widerspricht der Intention der KMK, die für alle Studiengänge dieselben Standards vorgegeben hat. Dieser Befund erscheint zunächst kontraintuitiv, da oft angenommen wird, dass in gymnasialen Lehramtsstudiengängen eher weniger pädagogisches und psychologisches Wissen vermittelt wird als in anderen Lehramtsstudiengängen. Unsere Ergebnisse deuten aber in dieselbe Richtung wie die von Krauss et al. (2008), nach denen Gymnasiallehrkräfte in Tests zu PCK und CK besser abschneiden als Lehrkräfte anderer Schularten. In längsschnittlichen Folgestudien mit paralleler Analyse der Curricula sollte geprüft werden, wie es zu diesen Leistungsvorteilen im gymnasialen Lehramtsstudium kommt.

## 6.2 Limitationen

Einschränkend ist zu berücksichtigen, dass aufgrund des Studiendesigns keine Dimensionsanalysen über alle Facetten durchgeführt werden konnten, da jede Person aufgrund des Testheftdesigns immer nur Items zu einzelnen oder wenigen Skalen bearbeitete. Dieses Manko wird in Folgeuntersuchungen behoben, in denen das vorliegende Instrument an einer großen Stichprobe bundesweit getestet wird, in dem jeder Proband/jede Probandin alle Items der verschiedenen Skalen bearbeiten soll. In der Folgeuntersuchung wird es auch möglich sein, Zusammenhänge zu PCK und CK zu analysieren. Die Zielsetzung dieser Studie war es, zunächst Skalen zu entwickeln, die dann in weitergehenden Untersuchungen eingesetzt werden können. Auch werden die Skalen *Individuelle Besonderheiten* und *Werte, Normen und soziale Konflikte*, die in der vorliegenden Studie keine akzeptablen Kennwerte aufgewiesen haben, in Folgestudien mit neuen Items erneut untersucht.

Eine weitere Einschränkung besteht in der Tatsache, dass die Stichprobe der vorliegenden Studie lediglich aus Lehramtsstudierenden der mathematisch-naturwissenschaftlichen Fächer besteht und Aussagen zu Studierenden anderer Studienfächer nicht möglich sind. Das pädagogisch-psychologische Wissen in der Lehramtsausbildung wird in dem Bereich der Bildungswissenschaften allerdings fächerübergreifend vermittelt, so dass nicht davon auszugehen ist, dass Lehramtsstudierende weiterer Fächer anderes

Wissen vermittelt bekommen. Das Wissen, welche die Lehramtsstudierenden in den Lerngelegenheiten im bildungswissenschaftlichen Begleitstudium erwerben sollen, wird durch die Standards für die Lehrerbildung für den Bereich der Bildungswissenschaften (KMK, 2004) festgelegt. Diese geben fächerübergreifend Kompetenzen und Inhalte vor, die die Lehramtsstudierenden nach Beendigung ihrer Ausbildung aufgebaut haben sollen.

Zudem wurde nicht explizit erfasst, inwieweit sich die Studierenden in ihrem Lehramtsstudium bereits mit den abgefragten Themen befasst haben, was ebenfalls eine Einschränkung der Interpretation der Ergebnisse darstellt. In einer Folgestudie soll eine differenzierte Abfrage der Lehrinhalte im Studium integriert werden.

In der vorliegenden Studie stand zunächst die Konzeption der Subskalen im Vordergrund. Entgegen der inhaltlichen Konzeption des Testinstruments konnten zwei Facetten empirisch nicht gestützt werden. Die Facetten *Individuelle Besonderheiten* und *Werte, Normen und soziale Konflikte* werden überarbeitet und in Folgestudien erneut überprüft. Auch in früheren Studien konnte die Facette *Individuelle Besonderheiten* nicht empirisch abgesichert werden (Voss et al., 2011). Die Facetten *Reflexion und Weiterentwicklung der eigenen Professionalität* und *Schulentwicklung* sind zu einer Skala zusammengefallen, was sich dadurch begründen lässt, dass beide Facetten inhaltliche Überschneidungspunkte aufweisen.

Des Weiteren ist zu beachten, dass durch das komplexe Rotationsdesign der Studie nicht alle Facetten mit derselben Stichprobengröße getestet werden konnten. Daher ergab sich bspw. für die Facette *Werte, Normen und soziale Konflikte* eine deutlich geringere Stichprobengröße als für die Facette *Lehren*. Dies ist dem Projektkontext der Studie geschuldet, da die einzelnen Studierenden sowohl einen fachunabhängigen als auch fachspezifischen Wissenstest bearbeiten sollten und demnach die zur Verfügung stehende Testzeit auf die einzelnen Testinstrumente aufgeteilt werden musste. Die geringe Stichprobengröße könnte einen Einfluss auf die geringen Skalen- und Itemkennwerte der Facette *Werte, Normen und soziale Konflikte* ausüben.

Die Lehramtsausbildung im Bereich der Bildungswissenschaften variiert von Hochschule zu Hochschule sehr stark (Hohenstein et al., 2014). Entsprechend unterscheiden sich für die Studierenden die Lerngelegenheiten deutlich. Dies deutet daraufhin, dass Dozentinnen und Dozenten kein einheitliches Verständnis von PPK, das in der Lehramtsausbildung vermittelt werden soll, teilen. Somit ist zu hinterfragen, ob der Einsatz eines Wissenstests, der sich zum Teil an den KMK-Standards orientiert, sinnvoll ist, wenn sich die Lehre an den Hochschulen so differenziert darstellt und schon die Umsetzung der KMK-Standards nicht durchgängig stattgefunden hat.

Zudem muss kritisch hinterfragt werden, inwieweit ein paper-pencil-Test tatsächlich prozedurales Wissen erfassen kann. In der vorliegenden Studie wurde dies bspw. in der Facette *Klassenführung* mit der umfassenden Beschreibung einer Unterrichtssituation versucht. Zukünftig wäre es empfehlenswert auch andere Formen der Wissenserfassung einzusetzen, wie zum Beispiel Videovignetten (Gaupp, 2008) oder computerbasierte Testverfahren wie den simulierten Klassenraum (Kaiser, Retelsdorf, Südkamp & Möller, 2013) und das Schülerinventar (Kaiser, Möller, Helm & Kunter, 2015).

Zusammenfassend lässt sich aber festhalten, dass mit der vorliegenden Studie sieben Skalen des PPK mit akzeptablen Reliabilitäten entwickelt wurden. In Folgeuntersuchungen sollen die Reliabilität und die Validität des Instruments eingehender untersucht werden. Dabei werden auch Skalen erneut einbezogen, die in der vorliegenden Studie empirisch nicht hinreichend gestützt werden konnten. Langfristiges Ziel ist es, mit Hilfe des Instruments die Entwicklungsbedingungen universitären pädagogischen und psychologischen Wissens zu beschreiben.

## Literatur

- Bannach, M. (2002). *Selbstbestimmtes Lernen. Freie Arbeit an selbst gewählten Themen*. Baltmannsweiler: Schneider Hohengehren.
- Barth, A.-R. (1992). *Burnout bei Lehrern. Theoretische Aspekte und Ergebnisse einer Untersuchung*. Göttingen: Hogrefe.
- Baumert, J., & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 9(4), 469–520.
- Baumert, J., & Kunter, M. (2013). Professionelle Kompetenz von Lehrkräften. In I. Gogolin, H. Kuper, H.-H. Krüger & J. Baumert (Hrsg.), *Stichwort: Zeitschrift für Erziehungswissenschaft* (S. 277–337). Wiesbaden: Springer VS.
- 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.
- Becker, G. E. (2006). *Lehrer lösen Konflikte. Handlungshilfen für den Schulalltag*. Weinheim: Beltz.
- Besser, M., & Krauss, S. (2009). Zur Professionalität als Expertise. In O. Zlatkin-Troitschanskaia, K. Beck, D. Sembill, R. Nickolaus, & R. Mulder (Hrsg.), *Lehrprofessionalität – Bedingungen, Genese, Wirkungen und ihre Messung* (S. 71–82). Weinheim/Basel: Beltz.
- Bransford, J., Darling-Hammond, L., & LePage, P. (2005). Introduction. In L. Darling-Hammond & J. Bransford (Hrsg.), *Preparing Teacher for a Changing World* (S. 1–39). San Francisco: Jossey-Bass.
- de Jong, O., van Driel, J. H., & Verloop, N. (2005). Preservice Teachers' Pedagogical Content Knowledge of Using Particle Models in Teaching Chemistry. *Journal of Research in Science Teaching*, 42(8), 947–964.
- Dickhäuser, O., & Rheinberg, F. (2003). Bezugsnormorientierung: Erfassung, Probleme, Perspektiven. In F. Rheinberg & J. Stiensmeier-Pelster (Hrsg.), *Diagnostik von Motivation und Selbstkonzept* (S. 41–55). Göttingen: Hogrefe.
- Ditton, H. (2000). Qualitätskontrolle und Qualitätssicherung in Schule und Unterricht. Ein Überblick zum Stand der empirischen Forschung. In A. Helmke, W. Hornstein & E. Terhart (Hrsg.), *Qualität und Qualitätssicherung im Bildungsbereich; Schule, Sozialpädagogik, Hochschule* (41. Beiheft, Zeitschrift für Pädagogik, S. 73–92). Weinheim: Beltz.
- Drechsel, B., Prenzel, M., & Seidel, T. (2015). Nationale und Internationale Schulleistungsstudien. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (S. 343–366). Berlin: Springer.
- Füssel, H.-P., & Leschinsky, A. (2008). Der institutionelle Rahmen des Bildungswesens. In K. S. Cortina, J. Baumert, A. Leschinsky, K. U. Mayer & L. Trommer (Hrsg.), *Das Bildungswesen in der Bundesrepublik Deutschland. Strukturen und Entwicklungen im Überblick* (S. 131–204). Reinbek: Rowohlt.

- Gaupp, N. (2008). Computerbasierte Erfassung sozialer Kompetenzen mit Video-Vignetten. In N. Jude, J. Hartig, & E. Klieme (Hrsg.), *Kompetenzerfassung in pädagogischen Handlungsfeldern. Theorien, Konzepte und Methoden* (S. 73–80). Bonn/Berlin: Bundesministerium für Bildung und Forschung.
- Gawrilow, C. (2012). *Lehrbuch ADHS*. München: Ernst Reinhardt.
- Hannover, B., & Kessels, U. (2011). Sind Jungen die neuen Bildungsverlierer? Empirische Evidenz für Geschlechterdisparitäten zuungunsten von Jungen und Erklärungsansätze. *Zeitschrift für Pädagogische Psychologie*, 25(2), 89–103.
- Heller, K. A., & Perleth, C. (2000). *Kognitiver Fähigkeitstest für 4.–12. Klassen: Revision (KFT 4-12+R)*. Göttingen: Hogrefe.
- Hill, H. C., Schilling, S. G., & Ball, D. L. (2004). Developing Measures of Teachers' Mathematics Knowledge for Teaching. *The Elementary School Journal*, 105(1), 11–30.
- Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of Teachers' Mathematical Knowledge for Teaching on Student Achievement. *American Educational Research Journal*, 42(2), 371–406.
- Hohenstein, 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.
- Imhof, M., & Langfeldt, H.-P. (2010). *Psychologie für Lehramtsstudierende*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Kaiser, J., Möller, J., Helm, F., & Kunter, M. (2015). Das Schülerinventar: Welche Schülermerkmale die Leistungsurteile von Lehrkräften beeinflussen. *Zeitschrift für Erziehungswissenschaft*, 18(2), 279–302.
- Kaiser, J., Retelsdorf, J., Südkamp, A., & Möller, J. (2013). Achievement and Engagement: How student characteristics influence teacher judgments. *Learning and Instruction*, 28, 73–84.
- KMK = Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (2004). *Standards für die Lehrerbildung: Bildungswissenschaften*. Beschluss der Kultusministerkonferenz vom 16.12.2004. [http://www.kmk.org/fileadmin/veroeffentlichungen\\_beschluesse/2004/2004\\_12\\_16-Standards-Lehrerbildung.pdf](http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_12_16-Standards-Lehrerbildung.pdf) [14.09.2016].
- KMK = Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (2014). *Ländergemeinsame inhaltliche Anforderungen für die Fachwissenschaften und Fachdidaktiken in der Lehrerbildung vom 16.10.2008 i. d. F. vom 10.09.2015*. [http://www.kmk.org/fileadmin/Dateien/veroeffentlichungen\\_beschluesse/2008/2008\\_10\\_16-Fachprofile-Lehrerbildung.pdf](http://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2008/2008_10_16-Fachprofile-Lehrerbildung.pdf) [30.09.2016].
- König, J., & Blömeke, S. (2009). Pädagogisches Wissen von angehenden Lehrkräften: Erfassung und Struktur von Ergebnissen der fachübergreifenden Lehrerausbildung. *Zeitschrift für Erziehungswissenschaft*, 12(3), 499–527.
- 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.
- König, J., Blömeke, S., & Schwippert, K. (2013). Pädagogisches Wissen im Studienverlauf. In S. Blömeke, A. Bremerich-Vos, G. Kaiser, G. Nold, H. Haudeck, J.-U. Keßler & K. Schwippert (Hrsg.), *Professionelle Kompetenzen im Studienverlauf. Weitere Ergebnisse zur Deutsch-, Englisch- und Mathematiklehrausbildung aus TEDS-LT* (S. 145–166). Münster: Waxmann.
- König, J., & Seifert, A. (Hrsg.) (2012). *Lehramtsstudierende erwerben pädagogisches Professionswissen. Ergebnisse der Längsschnittstudie LEK zur Wirksamkeit der erziehungswissenschaftlichen Lehrerausbildung*. Münster: Waxmann.
- Köller, M. (2009). *Konstruktion und Implementierung von Schulprogrammen: ein triangulativer Forschungsansatz*. Frankfurt a. M.: Peter Lang.
- Kounin, J. S. (2006). *Techniken der Klassenführung*. Münster: Waxmann.

- Krauss, S., Brunner, M., Baumert, J., Blum, W., Neubrand, M., & Jordan, A. (2008). Pedagogical Content Knowledge and Content Knowledge of Secondary Mathematics Teachers. *Journal of Educational Psychology, 100*(3), 716–725.
- Kunina-Habenicht, O., Lohse-Bossenz, H., Kunter, M., Dicke, T., Förster, D., Gößling, J., Schulze-Stocker, F., Schmeck, A., Baumert, J., Leutner, D., & Terhart, E. (2012). Welche bildungswissenschaftlichen Inhalte sind wichtig in der Lehrerbildung? *Zeitschrift für Erziehungswissenschaft, 15*(4), 649–682.
- 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., Baumert, J., Blum, W., Klusmann, U., Krauss, S., & Neubrand, M. (Hrsg.) (2011). *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms CO-ACTIV*. Münster: Waxmann.
- Lenke, G., Thilmann, H., Wirth, J., Dicke, T., & Leutner, D. (2015). Pädagogisch-psychologisches Professionswissen von Lehrkräften: Evaluation des ProWin-Tests. *Zeitschrift für Erziehungswissenschaft, 18*(2), 225–245.
- Lienert, G., & Ratz, U. (1998). *Testaufbau und Testanalyse* (6. Auflage). Weinheim: Psychologie Verlags Union.
- Lipowsky, F. (2015). Unterricht. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (S. 69–98). Berlin: Springer.
- Loewenberg Ball, D., Thames, M.H., & Phelps, G. (2008). Content Knowledge for Teaching: What makes it special? *Journal of Teacher Education, 59*(5), 389–407.
- Lohse-Bossenz, H., Kunina-Habenicht, O., Dicke, T., Leutner, D., & Kunter, M. (2015). Teachers' Knowledge About Psychology: Development and validation of a test measuring theoretical foundations for teaching and its relation to instructional behavior. *Studies in Educational Evaluation, 44*, 36–49.
- Muthén, L.K., & Muthén, B.O. (2007). *Mplus User's Guide* (5. Aufl.). Los Angeles, CA: Muthén & Muthén.
- Pohlmann, B., & Möller, J. (2010). Fragebogen zur Erfassung der Motivation für die Wahl des Lehramtsstudiums (FEMOLA). *Zeitschrift für Pädagogische Psychologie, 24*(1), 73–84.
- Rost, J. (2004). *Lehrbuch Testtheorie-Testkonstruktion*. Göttingen: Hans Huber.
- Schaarschmidt, U. (2008). Burnout im Lehrerberuf. In W. Schneider & M. Hasselhorn (Hrsg.), *Handbuch der Pädagogischen Psychologie* (S. 197–209). Göttingen: Hogrefe.
- Schmidt, F., & Hunter, J. (1998). The Validity and Utility of Selection Methods in Personnel Psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124*(2), 262–274.
- Schmitt, A. (2005). *Konfliktmediation in der Schule. Ergebnisse einer Evaluationsstudie*. Hamburg: Dr. Kovač.
- Seidel, T. (2015). Klassenführung. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie*. (S. 136–148). Berlin: Springer.
- Shulman, L. (1987). Knowledge and Teaching: Foundations of the new reform. *Harvard Educational Review, 57*(1), 1–23.
- Standop, J. (2005). *Werterziehung: Einführung in die wichtigsten Konzepte der Werteerziehung*. Weinheim: Beltz.
- Südkamp, A., Kaiser, J., & Möller, J. (2012). Accuracy of Teachers' Judgments of Students' Academic Achievement: A meta-analysis. *Journal of Educational Psychology, 104*, 743–762.
- Tepner, O., Borowski, A., Dollny, S., Fischer, H.E., Jüttner, M., Kirschner, S., Leutner, D., Neuhaus, B.J., Sandmann, A., Sumfleth, E., Thillmann, H., & Wirth, J. (2012). Modell zur Entwicklung von Testitems zur Erfassung des Professionswissens von Lehrkräften in den Naturwissenschaften. *Zeitschrift für Didaktik der Naturwissenschaften, 18*, 7–28.

- Valencia, S. W., Stallmann, A. C., Commeyras, M., Pearson, P., & Hartman, D. K. (1991). Four Measures of Topical Knowledge: A study of construct validity. *Reading Research Quarterly: RRQ*, 26, 204–233.
- Voss, T., Kunter, M., & Baumert, J. (2011). Assessing Teacher Candidates' General Pedagogical/Psychological Knowledge. Test construction and validation. *Journal of Educational Psychology*, 103(4), 952–969.
- Voss, T., Kunter, M., Seiz, J., Hoehne, V., & Baumert, J. (2014). Die Bedeutung des pädagogisch-psychologischen Wissens von angehenden Lehrkräften für die Unterrichtsqualität. *Zeitschrift für Pädagogik*, 60(2), 184–201.
- Voss, T., Kunina-Habenicht, O., Hoehne, V., & Kunter, M. (2015). Stichwort Pädagogisches Wissen von Lehrkräften: Empirische Zugänge und Befunde. *Zeitschrift für Erziehungswissenschaften*, 18(2), 187–223.
- Weinert, F. E. (Hrsg.) (2002). *Leistungsmessungen in Schulen*. Weinheim: Beltz.
- Wu, M. L., Adams, R. J., & Wilson, M. R. (1998). *ACER ConQuest: Generalized item response modeling software manual*. Melbourne: ACER.
- Ziegler, A. (2002). Hochbegabte und Begabtenförderung. In R. Tippelt (Hrsg.), *Handbuch Bildungsforschung* (S. 745–760). Wiesbaden: VS Verlag für Sozialwissenschaften.

**Abstract:** Teachers' professional knowledge represents a crucial factor concerning the quality of lessons and students' performance. Besides curricular and didactical contents, professional knowledge also consists of pedagogical-psychological knowledge. Research has so far focused on curricular and didactical knowledge; with few studies dedicated to the pedagogical-psychological knowledge of teachers. Considering existing studies and standards for teacher education, the present work – stemming from the KiL project (Measuring professional competencies in mathematical and natural-scientific teacher education) – postulates ten facets of pedagogical-psychological knowledge for teacher students and constructs a performance test. An item pool for the ten facets was generated and a pilot study ( $N = 1276$ ) was implemented in twelve universities in Germany. Seven scales indicate sufficient reliability and validity.

**Keywords:** Pedagogical-Psychological Knowledge, Test Construction, Professional Knowledge, Teacher Education, Teacher Students

**Anschrift der Autor\_innen**

Dipl.-Psych. Friederike Hohenstein, Christian-Albrechts-Universität zu Kiel,  
Arbeitseinheit Psychologie für Pädagogen,  
Olshausenstr. 75, 24118 Kiel, Deutschland  
E-Mail: f.hohenstein@psychologie.uni-kiel.de

Prof. Dr. Thilo Kleickmann, Christian-Albrechts-Universität zu Kiel,  
Lehrstuhl für Schulpädagogik,  
Olshausenstr. 75, 24118 Kiel, Deutschland  
E-Mail: kleickmann@paedagogik.uni-kiel.de

Prof. Dr. Friederike Zimmermann, Christian-Albrechts-Universität zu Kiel,  
Arbeitseinheit Psychologie für Pädagogen,  
Olshausenstr. 75, 24118 Kiel, Deutschland  
E-Mail: zimmermann@psychologie.uni-kiel.de

Prof. Dr. Olaf Köller,  
Leibniz Institut für die Pädagogik der Naturwissenschaften und Mathematik (IPN),  
Olshausenstr. 62, 24118 Kiel, Deutschland  
E-Mail: koeller@ipn.uni-kiel.de

Prof. Dr. Jens Möller, Christian-Albrechts-Universität zu Kiel,  
Arbeitseinheit Psychologie für Pädagogen,  
Olshausenstr. 75, 24118 Kiel, Deutschland  
E-Mail: jmoeller@psychologie.uni-kiel.de