

Der Erwerb von pädagogischem Wissen bei angehenden Lehrkräften in unterschiedlichen Ausbildungskontexten: Zur Wirksamkeit der Lehrerausbildung in Deutschland und Österreich

Johannes König · Stefan Klemenz

Eingegangen: 1. Oktober 2014 / Angenommen: 31. März 2015 / Online publiziert: 5. Mai 2015
© Springer Fachmedien Wiesbaden 2015

Zusammenfassung Im Rahmen der internationalen Vergleichsstudie TEDS-M wurde ein Testinstrument zur Erfassung von fächerübergreifendem, pädagogischem und auf die Kernaufgabe des Unterrichtens bezogenem Lehrerwissen entwickelt und eingesetzt, um Ergebnisse der Lehrerausbildung im internationalen Vergleich zu evaluieren. Als Fortführung wurde im Rahmen der Studie EMW begonnen, in Deutschland, Österreich und der Schweiz das pädagogische Wissen angehender Lehrkräfte von ihrem Eintritt in die Ausbildung an längsschnittlich zu erfassen, um differenziert Einblick in Lern- und Ausbildungsprozesse sowie Wirkungsweisen von Ausbildungsprogrammen und ihren Charakteristika zu erhalten. Der Beitrag verwendet einen Teildatensatz der EMW-Stichprobe von 1451 Lehramtsstudierenden aus Deutschland und Österreich, die in ihrem ersten und fünften Ausbildungssemester getestet wurden. Gefragt wird (1) nach einem Zuwachs an pädagogischem Wissen über die zwei Messzeitpunkte, (2) nach differenziellen Entwicklungen im pädagogischen Wissen in Abhängigkeit der landesspezifischen Lehrerausbildungssysteme sowie (3) nach der Bedeutung schulpraktischer Lerngelegenheiten als proximale Prädiktoren des Wissenszuwachses. Die Ergebnisse belegen für beide Länderstichproben einen statistisch signifikanten Wissenszuwachs, der große praktische Bedeutsamkeit besitzt. Eine Differenzierung des Tests in Dimensionen kognitiver Anforderungen verdeutlicht darüber hinaus einen stärkeren Wissenszuwachs im

Prof. Dr. J. König (✉) · S. Klemenz
Empirische Schulforschung, quantitative Methoden, Universität zu Köln,
Gronewaldstr. 2,
50931 Köln, Deutschland
E-Mail: johannes.koenig@uni-koeln.de

S. Klemenz
E-Mail: stefan.klemenz@uni-koeln.de

handlungsnahen Wissen in der österreichischen Stichprobe. In Mehrebenenanalysen lässt sich dieser durch den Einbezug von Prädiktoren, die Merkmale der schulpraktischen Lerngelegenheiten proximal erfassen, erklären. Implikationen für den Diskurs um den Theorie-Praxis-Bezug in der Lehrerbildung werden diskutiert.

Schlüsselwörter Lehrerbildung · Lehrerkompetenzen · Pädagogisches Wissen · Längsschnittstudie · Lerngelegenheiten

The acquisition of pedagogical knowledge by future teachers from a comparative perspective: on the effectiveness of teacher education in Germany and Austria

Abstract In TEDS-M, a comparative study, a test measuring pedagogical teacher knowledge related to teaching as the core task of teachers was developed and applied in order to evaluate the outcome of initial teacher education from a comparative perspective. To continue this, in Germany, Austria, and Switzerland the study EMW has started to longitudinally assess pedagogical knowledge of future teachers right from the start when they enter initial teacher education in order to get detailed insight into processes of learning and teacher training as well as the efficacy of teacher education programs and their characteristics. This paper uses a subset of the EMW-data focusing on a sample of 1451 student teachers from Germany and Austria that were tested in their first and their fifth semester. We ask (1) whether there is a learning gain in their pedagogical knowledge over the two measurement points, (2) whether there is a differential development in the pedagogical knowledge depending on the country-specific initial teacher education system, as well as (3) whether practical in-school opportunities to learn can function as proximal predictors of the learning gain. Findings show that there is a statistically significant learning gain in both country samples (large effect). Moreover, the differentiation of the test into subscales of cognitive demands shows a stronger learning gain in the procedural, more action-related knowledge in the Austrian student teacher sample. In multi-level analysis, this difference can be explained when predictors are included which proximally measure characteristics of practical in-school opportunities to learn. Implications for the discourse on the relationship of theory and practice in initial teacher education are discussed.

Keywords Longitudinal study · Opportunities to learn · Pedagogical knowledge · Teacher competence · Teacher education

1 Einleitung

Pädagogisches und auf die Kernanforderung des Unterrichtens bezogenes Wissen gilt heute übereinstimmend als zentraler Bereich des professionellen Wissens von Lehrerinnen und Lehrern (Shulman 1987; Bromme 1992; Terhart 1993) sowie als eine zentrale kognitive Komponente ihrer professionellen Kompetenz (Baumert und

Kunter 2006; Kunter et al. 2011). Neben rein fachlichen Anforderungen im Lehrerberuf (z. B. Strukturierung der Inhalte eines Faches) stehen Lehrkräfte vor zentralen Anforderungen, die, auch wenn sie teilweise fachspezifisch ausgestaltet werden müssen, für sämtliche Fächer relativ gleichförmig sind (z. B. Klassenführung). Für die erfolgreiche Bewältigung solcher Anforderungen wird ein pädagogisches Wissen vorausgesetzt (z. B. Shulman 1987; Bromme 1992; Voss et al. 2014). Im Rahmen der Lehrerbildung sollen angehende Lehrkräfte pädagogisches Wissen erwerben und somit auf den Lehrerberuf vorbereitet werden (z. B. Terhart 2012).

Während pädagogisches Wissen von angehenden wie berufstätigen Lehrkräften als Gegenstand empirischer Untersuchungen lange Zeit wenig Beachtung fand (Blömeke 2004; Baumert und Kunter 2006), ist in der jüngsten Zeit eine deutliche Zunahme an Forschungsansätzen zur standardisierten Erfassung bzw. Kompetenzmessung zu beobachten. So wurde – u. a. – international-vergleichend das pädagogische Wissen von angehenden Lehrkräften im Rahmen der IEA-Studie *Teacher Education and Development Study – Learning to Teach Mathematics* (TEDS-M) untersucht. Dafür wurden im Jahr 2008 in 17 Ländern weltweit das fachliche und das fachdidaktische Wissen sowie – ergänzend in den drei Teilnehmerländern Deutschland, USA und Taiwan – das pädagogische Wissen angehender Mathematiklehrkräfte für die Primarstufe und die Sekundarstufe I am Ende ihrer Ausbildung getestet (Blömeke et al. 2010a, b). TEDS-M stellt bislang die größte international angelegte und mit repräsentativen Länderstichproben arbeitende Studie zur Lehrerausbildung dar. Sie liefert, ähnlich wie die großen Schulleistungsstudien (z. B. PISA), präzise Informationen zu Ergebnissen der Lehrerausbildung eines Landes, versehen mit Referenzwerten aus dem internationalen Vergleich. Allerdings stellt sie, wie die großen Schulleistungsstudien meist auch, nur eine Momentaufnahme dar, da mit Querschnittstichproben angehender Lehrkräfte am Ende ihrer Ausbildung gearbeitet wird. So können Aussagen über die Wirksamkeit der unterschiedlichen institutionellen Lehr-Lernprozesse kaum getroffen werden, ein Defizit der empirischen Lehrerbildungsforschung, das schon lange moniert wird (z. B. Schaefers 2002; Terhart 2012).

An dieser Problemstellung setzt die Studie *Entwicklung von berufsspezifischer Motivation und pädagogischem Wissen in der Lehrerausbildung* (EMW) an, eine in den drei deutschsprachigen Ländern Deutschland, Österreich und der Schweiz laufende Längsschnittuntersuchung.¹ Mit ihr werden Studierende von bis zu 31 Universitäten bzw. Pädagogischen Hochschulen, die im Wintersemester 2011/2012 ein Lehramtsstudium aufnahmen, im Turnus von zwei Jahren zu ihrem pädagogischen Wissen mithilfe des in TEDS-M entwickelten Erhebungsinstruments getestet. In diesem Beitrag werden auf der Basis eines Teildatensatzes der EMW-Studie drei Fragestellungen untersucht. Erstens wird gefragt, ob es in den beiden Teilnehmerländern Deutschland und Österreich zu einem Zuwachs an pädagogischem Wissen über die im Rahmen der EMW-Studie bisher durchgeführten zwei Messzeitpunkte (d. h. innerhalb der ersten beiden Ausbildungsjahre) kommt. Zweitens wird nach differenziellen Entwicklungen im pädagogischen Wissen in Abhängigkeit der beiden Länder gefragt. Dafür wird, TEDS-M und den Studien im Kontext von TEDS-M folgend, der Test zur Erfassung von pädagogischem Wissen in drei Dimensionen bzw. Subskalen kognitiver Anforderungen unterteilt, die bei der Bearbeitung der Testaufgaben an die angehenden Lehrkräfte gestellt werden (vgl. Abschn. 2.2). Angesichts der

Variation in der organisatorisch-curricularen Gestaltung des Theorie-Praxis-Bezugs (z. B. Gewichtung, Verzahnung) der beiden Ausbildungssysteme in Deutschland und Österreich sowie der daraus resultierenden Unterschiede in den dargebotenen Lerngelegenheiten (vgl. Abschn. 3) werden Länderunterschiede im Zuwachs konzeptuell-analytischen Wissens sowie im Zuwachs von eher handlungsnahem pädagogischem Wissen erwartet. Da jedoch die Unterscheidung der beiden Ausbildungskontexte nur einen distalen Prädiktor für die Variation in der organisatorisch-curricularen Gestaltung des Theorie-Praxis-Bezugs darstellt, wird drittens mithilfe von Mehrebenenanalysen nach dem Einfluss von lernprozessbezogenen Tätigkeiten in schulpraktischen Lerngelegenheiten als proximale Prädiktoren gefragt. Dies geschieht mit dem Ziel, die Abhängigkeit des pädagogischen Wissens von bereichsspezifischen Lerngelegenheiten zu belegen und somit ein Argument für die Validität des getesteten pädagogischen Wissens zu liefern. Darüber hinaus erscheinen solche Analysen mit Daten aus zwei Ländern wichtig, da die jeweilige Lehrerbildung in Deutschland und Österreich mit sehr ähnlichen Problemlagen und Fragestellungen konfrontiert ist (vgl. Stadelmann und Rhyn 2011) und ein Vergleich der beiden Ausbildungskontexte neue Erkenntnisse für aktuelle Reformdiskurse erbringen kann.

2 Forschung zum Erwerb von pädagogischem Wissen

2.1 Konzeptualisierung von pädagogischem Wissen

Mit den einflussreichen theoretischen Arbeiten von Shulman (1987, S. 8) zur heuristischen Unterteilung von berufsspezifischem Lehrerwissen, d. h. dem „professionellen Wissen“ von Lehrkräften (Bromme 1992), kann pädagogisches Wissen (*general pedagogical knowledge*) zunächst verstanden werden als „knowledge, with special reference to those broad principles and strategies of classroom management and organization that appear to transcend subject matter“, wobei entscheidend sein dürfte, dass bei Shulman (1987) wie auch heute (vgl. Baumert und Kunter 2006; Baumert et al. 2010; Tatto et al. 2012) dieses Wissen von verschiedenen anderen Komponenten des professionellen Wissens abgegrenzt wird, insbesondere vom Fachwissen (*content knowledge*) und vom fachdidaktischen Wissen (*pedagogical content knowledge*). Während sich fachliches und fachdidaktisches Wissen eindeutig auf das zu unterrichtende Fach beziehen, ist pädagogisches Wissen losgelöst vom jeweiligen Fach. Fachdidaktisches Wissen kann als Ergebnis einer Transformation von fachlichem und pädagogischem Wissen zu einer neuen Kategorie des Lehrerwissens verstanden werden, wenn man Shulmans (1987, S. 8) Vorstellung von fachdidaktischem Wissen als „that special amalgam of content and pedagogy“ folgt. Demnach sind von diesen drei Wissensbereichen das fachliche und das pädagogische Wissen besonders weit voneinander entfernt.

Aktuelle kompetenztheoretische Ansätze schließen hier an, wenn diese Trias als drei wesentliche kognitive Komponenten professioneller Kompetenz von Lehrkräften konzeptualisiert wird (Baumert und Kunter 2006). Mit der Kompetenzorientierung (Weinert 2001) geraten Anforderungen in den Blick, die an Lehrkräfte in ihrem Beruf gestellt werden. In diesem Sinne wird auch in der vorliegenden Studie „professio-

nelle Kompetenz“ von Lehrkräften als die erfolgreiche Bewältigung zentraler Anforderungen definiert, die typisch für den Lehrerberuf sind. Im Speziellen fokussieren wir dabei auf das pädagogische Wissen, und zwar als kognitive Leistungsdisposition, die die erfolgreiche Bewältigung spezifischer Anforderungen (z. B. Klassenführung) wahrscheinlich macht (zuletzt Blömeke et al. 2015).

Ansätze zur definitorischen Klärung von pädagogischem Wissen und seiner Binnenstruktur weisen auf unterschiedliche Differenzierungen hin. Inhaltlich betrachtet werden etwa in einer von Baumert und Kunter (2006, S. 485) vorgelegten Übersicht vier Themen (Konzeptuelles bildungswissenschaftliches Grundlagenwissen, Allgemeindidaktisches Konzeptions- und Planungswissen, Unterrichtsführung und Orchestrierung von Lerngelegenheiten, Fachübergreifende Prinzipien des Diagnostizierens, Prüfens und Bewertens) mit Unteraspekten gelistet, welche verdeutlichen, dass pädagogisches Wissen auf Anforderungen des Unterrichtens (z. B. effektive Klassenführung) bezogen werden kann, aber auch über unterrichtliche Fragen hinausgeht (z. B. historische Grundlagen von Schule; vgl. auch bereits Terhart 1993). Hingewiesen wird ferner auf das Spannungsverhältnis von pädagogischem Wissen zwischen einer „disziplinbezogenen Betrachtung, die vom fächerübergreifenden bzw. erziehungs- und bildungswissenschaftlichen Curriculum der Lehrerbildung ausgeht, und einer anforderungsbezogenen Betrachtung, welche die typischen Herausforderungen des Lehrerberufs im Blick hat“ (König 2014, S. 620). Schließlich wird diskutiert, dass pädagogisches Wissen von Lehrkräften in unterschiedlichen Qualitäten und Formen auftreten und von Bedeutung sein könne (vgl. zuletzt Neuweg 2014). Zentral ist die Unterscheidung in deklaratives Wissen („Wissen, dass...“) und prozedurales Wissen („Wissen, wie...“), wobei letzteres jene Form darstellt, die besonders handlungsrelevant ist, während das ausschließliche Vorliegen deklarativen Wissens Schwierigkeiten bei der Umsetzung von Wissen in die Praxis mit sich führen kann (Gruber und Renkl 2000; Anderson et al. 2001). Prozedurales Wissen ist dadurch gekennzeichnet, dass es situations- und ablauforientiert organisiert ist („Handlungsschemata“). Das heißt in Bezug auf Lehrpersonen, dass sie Unterricht schrittweise anhand typischer Abfolgen wahrnehmen, planen und durchführen (Aebli 1983; Putnam 1987). Ein weiteres Begriffspaar ist das „praktische Wissen“, also jenes Wissen, das „the proximal guide for a good deal of a teacher’s classroom behavior“ darstellt (Berliner 2004, S. 206), und das theoretisch-formale Wissen, welches vor allem im Rahmen einer universitären Lehramtsausbildung schwerpunktmäßig erworben werden dürfte (Fenstermacher 1994; Baumert und Kunter 2006). Im nachfolgenden Abschn. 2.2 werden diese Begriffe beispielhaft auf die Messung von pädagogischem Wissen in der vorliegenden Studie bezogen.

In den vergangenen Jahren ist, gerade im deutschsprachigen Raum, ein deutlicher Zuwachs an Forschung zur standardisierten Erfassung von pädagogischem Wissen zu verzeichnen (u. a. Baer et al. 2007; Blömeke et al. 2008; Schulte et al. 2008; König und Blömeke 2009a; Voss et al. 2011; Kunina-Habenicht et al. 2013). Der vorliegende Beitrag bezieht sich auf die Arbeiten, die im Kontext von TEDS-M in den vergangenen Jahren entstanden sind und baut auf diese auf, u. a. da das im Rahmen von TEDS-M entwickelte Testinstrument zur Erfassung von pädagogischem Wissen auch in der hier in Rede stehenden EMW-Studie zum Einsatz gelangt.

2.2 Messung von pädagogischem Wissen in der vorliegenden Studie

Zur empirischen Untersuchung dieses Wissensbereichs liegt mit dem TEDS-M Instrument ein standardisierter Papier-Bleistift-Test vor, dessen Konzeption und Operationalisierung sich mittlerweile in verschiedenen empirischen Studien der nationalen und internationalen Lehrerbildungsforschung bewährt hat und selbst unter Kritikern deutliche Anerkennung findet (Lüders 2012).² Unterrichtliche Herausforderungen, wie sie mit der Forschung zur Unterrichtsqualität beschrieben werden können, stellen dabei den Ausgangspunkt für die inhaltliche Differenzierung des Testinstruments dar. Entsprechend kompetenztheoretischer Vorstellungen wird dann nach dem Wissen gefragt, welches nötig ist, um diese Anforderungen erfolgreich zu bewältigen. Der im Kontext von TEDS-M entwickelte Ansatz von König und Blömeke (2009a, S. 504) greift auf empirisch nachgewiesene Basisdimensionen von Unterrichtsqualität zurück und verknüpft diese mit Modellen der Allgemeinen Didaktik zu folgenden Inhaltsdimensionen pädagogischen Wissens: *Umgang mit Heterogenität, Strukturierung von Unterricht, Klassenführung, Motivierung, Leistungsbeurteilung*. Diese Dimensionen werden jeweils über mehrere komplexe Testaufgaben operationalisiert, die entsprechende unterrichtliche Herausforderungen zum Gegenstand haben (vgl. Abb. 1).

Quer zu diesen fünf berufsbezogenen Inhaltsbereichen weist das Testinstrument eine kognitionsbezogene Struktur auf (Abb. 2): Zur Unterscheidung verschiedener Qualitäten kognitiver Anforderungen, die bei der Bearbeitung der Testaufgaben an

Fünf berufliche Anforderungen von Lehrpersonen beim Unterrichten	Themen für die Operationalisierung
Strukturierung von Unterricht	<ul style="list-style-type: none"> • Komponenten- und prozessbezogene Planung • Analyse von Unterricht • Curriculare Strukturierung von Unterricht
Motivierung	<ul style="list-style-type: none"> • Leistungsmotivation • Motivierungsstrategien im Unterricht
Umgang mit Heterogenität	<ul style="list-style-type: none"> • Differenzierungsmaßnahmen • Methodenvielfalt
Klassenführung	<ul style="list-style-type: none"> • Störungspräventive Unterrichtsführung • effektive Nutzung der Unterrichtszeit
Leistungsbeurteilung	<ul style="list-style-type: none"> • Funktionen und Formen • zentrale Kriterien • Urteilsfehler

Abb. 1 Inhalte des Tests zum pädagogischen Wissen (König und Blömeke 2009a, S. 504)

Abb. 2 Inhaltsdimensionen und kognitive Prozesse als Testdesign-Matrix

	Erinnern	Verstehen/ Analysieren	Kreieren/ Generieren von Handlungsoptionen
Umgang mit Heterogenität			
Strukturierung von Unterricht			
Klassenführung			
Motivierung			
Leistungsbeurteilung			

die angehenden Lehrkräfte gestellt werden, wurde die von Anderson und Krathwohl (2001) revidierte und erweiterte Bloomsche Taxonomie kognitiver Prozesse aufgegriffen und auf die Testaufgaben bezogen. Im Fokus stehen dabei drei Dimensionen kognitiver Prozesse:

1. *Erinnern*. Zur Bearbeitung von Testaufgaben, die diesem kognitiven Prozess zugeordnet werden (vgl. als Beispiel Abb. 3), wird Wissen benötigt, das in Erinnerung gerufen werden muss (z. B. Lehrbuchwissen und bzw. oder Erfahrungswissen). Zur Bearbeitung werden kognitive Prozesse benötigt wie das Nennen bzw. Erkennen. Testaufgaben dieses Typs fordern bei der Befragung angehende Lehrer auf, ein Beispiel für eine Definition zu geben, eine Aufzählung von Elementen eines Phänomens, Begriffs oder Konzepts vorzunehmen oder einen Begriff oder ein Konzept zu erkennen oder zu identifizieren.
2. *Verstehen/Analysieren*. Zur Bearbeitung dieser Testaufgaben (vgl. als Beispiel Abb. 4) wird Wissen benötigt, das in Erinnerung gerufen werden muss (z. B. Lehrbuchwissen und bzw. oder Erfahrungswissen), aber darüber hinaus auch mit einer Problemstellung in Verbindung gebracht werden muss. Zur Bearbeitung werden kognitive Prozesse benötigt wie das Erklären bzw. Vergleichen. Bei der Bearbeitung der Testaufgaben dieser Dimension wird den angehenden Lehrpersonen abverlangt, einen Sachverhalt, ein Konzept, einen Aspekt zu charakterisieren oder zu erklären, einen Begriff oder ein Konzept auszuwählen und damit eine Erklärung zu liefern oder einen Sachverhalt, eine Situation, eine Falldarstellung, einen oder mehrere Begriffe zu vergleichen, zu kategorisieren, zu ordnen oder zu interpretieren.
3. *Kreieren/Generieren von Handlungsoptionen*. Zur Bearbeitung der Aufgaben, die dieser Dimension zugeordnet werden (als Beispiele vgl. Abb. 5 und 6), wird wiederum Wissen benötigt, das in Erinnerung gerufen werden muss (z. B. Lehrbuchwissen und bzw. oder Erfahrungswissen) und darüber hinaus zumindest ansatzweise auch mit einer Problemstellung und einer typischen Situation in Verbindung gebracht werden muss. Im Gegensatz zur vorherigen Dimension des Verstehens und Analysierens wird den Befragten für die Bearbeitung jedoch abverlangt, Handlungsoptionen zu entwickeln und zu formulieren, die ihnen für die Bewältigung einer typischen Situation zur Verfügung stehen, d. h. praktisches Wissen und Können, das als propositionale mentale Repräsentation beschrieben werden kann, darzulegen.

Diese Unterscheidung zielt darauf ab, soweit es im Rahmen einer Papier-Bleistift-Testung überhaupt möglich ist, spezifischen Anforderungen bei der Kompetenzmessung gerecht zu werden: Lehrkräfte müssen einen Großteil ihrer Arbeitszeit ausgehend von unterrichtlichen Situationen zukunftsorientiert und zielgerichtet planen und Handlungsoptionen entwerfen. Mit dem Testkonzept wird angenommen, dass Lehrpersonen Testaufgaben der dritten Anforderung des Kreierens bzw. Generierens von Handlungsoptionen erfolgreicher bearbeiten können, wenn sie ihr deklaratives Wissen über Unterricht vor dem Hintergrund eigener Unterrichtserfahrungen bereits hinreichend reflektieren konnten.

Die Konstruktvalidität des Tests konnte zum einen über die Prüfung der theoretisch angenommenen dimensional Struktur im Rahmen der Analysen von TEDS-M

nachgewiesen werden (Blömeke et al. 2010a, b). Zum anderen gelang der Nachweis der Konstruktrepräsentation über die Vorhersage von Item-Schwierigkeiten (König 2009). Konvergente und diskriminante Validität (Campbell und Fiske 1959) konnte für den Test vor allem anhand folgender Befunde belegt werden: Diskriminante Validität wurde u. a. erwartet gegenüber dem fachlichen Wissen und der Abiturnote, konvergente Validität hingegen gegenüber dem fachdidaktischen Wissen. Bestätigen ließen sich diese Annahmen dahingehend, dass das so gemessene pädagogische Wissen schwächer korreliert ist mit Fachwissen als mit fachdidaktischem Wissen bei angehenden Lehrkräften der Fächer Mathematik, Deutsch und Englisch (Blömeke et al. 2010a, b, 2011). Auch ist pädagogisches Wissen nur schwach mit der Abiturnote korreliert (König und Seifert 2012). Prognostische Validität zeigt sich bei berufstätigen Lehrkräften in Form positiver Korrelationen mit den Einschätzungen zur Unterrichtsqualität der von diesen Lehrkräften unterrichteten Schülerinnen und Schülern (Pflanzl et al. 2013). Angesichts dieser verschiedenen Befunde zur Validität des TEDS-M Instruments zur Erfassung von pädagogischem Wissen setzt dieser Beitrag einen Fokus auf eine Untersuchung der Veränderung des pädagogischen Wissens im Rahmen der Lehrerausbildung unterschiedlicher Ausbildungskontexte.

International wurde der Test im Rahmen der TEDS-M-Studie eingesetzt, um in einem Dreiländer-Vergleich den jeweiligen Wissensstand angehender Lehrkräfte am Ende ihrer Ausbildung in Deutschland, den USA und Taiwan abzuschätzen (Blömeke et al. 2010a, b). Mit den Ergebnissen werden spezifische Stärken und Schwächen der beteiligten Systeme abgebildet, etwa eine relative Stärke der US-amerikanischen angehenden Lehrkräfte im Bereich des Kreierens/Generierens von Handlungsoptionen, die bei angehenden Lehrkräften in Deutschland nicht gegeben war (bei Betrachtung der ipsativen Werte). In weiterführenden Studien mit österreichischen Stichproben (König und Blömeke 2009b; Hopf 2012) konnte gezeigt werden, dass der Test auch für den dortigen Kontext der Lehrerausbildung curricular valide ist, ohne dass sprachliche Anpassungen nötig gewesen wären.

2.3 Erwerb von pädagogischem Wissen

Empirische Erkenntnisse zur Entwicklung von Lehrerwissen stammen aus der Forschung zur Lehrerexpertise (u. a. Bromme 1992), mit der Entwicklungsstadien vom Novizen zum Experten beschrieben werden (Neuweg 1999; Berliner 2004). Zwar beziehen sich die Stadien vorwiegend auf den beruflichen Werdegang, ein Transfer auf die Lehrerausbildung ist jedoch möglich (Sabers et al. 1991): Lehramtsstudierende am Anfang ihrer Ausbildung sind „Novizen“ (*novice*), die erst im Verlauf ihrer Ausbildung professionelles Wissen und Handlungserfahrungen erwerben, sodass sie das Stadium des „Fortgeschrittenen Anfängers“ (*advanced beginner*) als das zweite von fünf Stadien mit dem Abschluss der zweiten Phase erreichen bzw. durchlaufen. Nach Berliner (2004) beginnt der (angehende) Lehrer auf dieser zweiten Stufe mit der Aneignung von „praktischem Wissen“ (vgl. Abschn. 2.1). Wesentlicher Fokus liegt auf der zunehmenden Kontext- und Situationseinbettung des professionellen Wissens. (Quasi-)längsschnittliche Untersuchungen zum Erwerb und zur Veränderung des pädagogischen Wissens von angehenden Lehrkräften während der Zeit ihrer Ausbildung (u. a. Baer et al. 2007; Blömeke et al. 2008; Schulte et al. 2008; Voss und

Kunter 2011), belegen, dass im Verlauf der Lehrerbildung trotz aller Kritik durchaus eine substantielle Zunahme im pädagogischen Wissen bei angehenden Lehrkräften erfolgt. Im Sinne der Expertiseforschung kann man annehmen, dass sich damit angehende Lehrkräfte vom Stadium des „Novizen“ hin zum Stadium des „Fortgeschrittenen Anfängers“ bewegen.

Diese Annahme wird auch durch empirische Befunde aus TEDS-M und der Studie *Längsschnittliche Erhebung pädagogischer Kompetenzen von Lehramtsstudierenden* (LEK) gestützt und differenziert. König und Seifert (2012) konnten im Rahmen der LEK-Studie anhand einer Längsschnittstichprobe von 261 Lehramtsstudierenden von vier unterschiedlichen deutschen Universitäten zeigen, dass zwischen dem ersten und dem vierten Ausbildungssemester der Erwerb von pädagogischem Wissen schwerpunktartig im Bereich des deklarativen Wissens liegt – und zwar bei allen vier in die LEK-Studie einbezogenen Universitäten auf ähnliche Weise. Lehramtsstudierende erwerben in dieser Zeit pädagogisches Wissen, das sie also in der Testsituation „Erinnern“ bzw. „Verstehen/Analysieren“ können. Handlungsnahe Wissen, operationalisiert über die kognitive Anforderung des „Kreierens/Generierens von Handlungsoptionen“, wird dagegen im Zeitraum der ersten zwei Jahre der universitären Lehrerbildung nur geringfügig erworben. Vergleichende Analysen mithilfe der TEDS-M Daten zeigen jedoch, dass Referendare kurz vor dem zweiten Staatsexamen im Vergleich mit den Studierenden des vierten Ausbildungssemesters nicht nur insgesamt über mehr pädagogisches Wissen verfügen, sondern ihnen insbesondere im Bereich des handlungsnahen Wissens überlegen sind (König 2013). Dies lässt sich mit den spezifischen Lerngelegenheiten der beiden Ausbildungsphasen, ihren theoretisch-formalen bzw. praktischen Akzentuierungen, gut erklären. Das Ergebnis liefert aber auch Hinweise, dass sich bei den angehenden Lehrkräften die in der Expertiseforschung beschriebene zunehmende Verknüpfung ihres Wissens mit Unterrichtserfahrungen (Berliner 2004) im Verlauf der Lehrerbildung vollzieht.

Mit nur vier Standorten eines Landes (Deutschland) ist allerdings auch die Aussagekraft der LEK-Studie begrenzt. Ihr Stichprobenumfang ist zu gering, um institutionelle Variation und Wirkungen in mehrerebenenanalytischen Regressionsmodellen stabil schätzen zu können. Dies sowie der über die beiden Messzeitpunkte der LEK-Studie an allen Standorten anzutreffende statistisch signifikante sowie praktisch bedeutsame Wissenszuwachs gibt Anlass, ein stark vergrößertes Design mit mehr Standorten und Teilnehmerländern anzuwenden, wie es mit der hier in Rede stehenden EMW-Studie umgesetzt wird.

3 Vergleichende Untersuchung des Erwerbs von pädagogischem Wissen

3.1 Forschung zur Lehrerbildung im internationalen Vergleich

Die international-vergleichende, empirische Forschung zur Wirksamkeit der Lehrerbildung ist eine neue Erscheinung der empirischen Bildungsforschung (Klieme 2012). Ihre Funktion ist es, besondere Leistungen eines Systems und seine Wirksamkeit unter Berücksichtigung der personellen, strukturellen, inhaltlichen sowie weiterer Bedingungen im systematischen Vergleich zu identifizieren (Blömeke 2011). Mit

der Erforschung von Wirksamkeit der Lehrerbildung ist insbesondere das Ziel verbunden, mit „stichhaltigen Daten (...) Inhalte, Methoden und Prozesse“ der Lehrerbildung darauf hin zu prüfen, ob „die gesetzten Ziele, nämlich die Ausbildung guter Lehrerinnen und Lehrer“, auch erreicht werden (Hascher 2011, S. 419). Die Wirksamkeitsforschung „soll systematisch prüfen, was gemeinhin als selbstverständlich bezüglich der Effekte oder Nicht-Effekte angenommen wird. Insofern unterscheidet sie sich klar von der Kritik der Lehrerbildung, die vor allem auf Meinungen und subjektiven Mängelzuschreibungen beruht“ (ebd.). Allerdings wird ein Mangel an international-vergleichenden Studien zum Lehrerberuf und zur Lehrerbildung beklagt (Blömeke 2011), eingeschlossen Vergleiche im deutschsprachigen Raum (Bosse et al. 2012). Narrativen Vergleichen von Strukturmerkmalen auf der Systemebene (u. a. Eurydice 2004; OECD 2005, 2009; Schleicher 2011) steht eine deutlich geringere Zahl von empirischen Studien gegenüber, in denen auf der Individualebene Daten mittels Fragebögen und Tests erhoben werden. Insbesondere die längsschnittliche Messung pädagogischen Wissens, welche Aussagen zur Wirksamkeit erlaubt, stellt in der international-vergleichenden Lehrerbildungsforschung „einen defizitären Forschungsbereich dar“ (Blömeke 2011, S. 346). Vor diesem Hintergrund begegnet also die hier in Rede stehende EMW-Studie – als eine in unterschiedlichen Ausbildungskontexten stattfindende Längsschnittuntersuchung mit Fokus auf das pädagogische Wissen als Element der sich während der Ausbildung entwickelnden professionellen Kompetenz zukünftiger Lehrpersonen – dieser ergänzungsbedürftigen Forschungslage (zu Limitierungen vgl. Abschn. 7.2).

3.2 Vergleich der Ausbildungskontexte Deutschland und Österreich

Die zwei im vorliegenden Beitrag einbezogenen deutschsprachigen Länder Deutschland und Österreich bieten als unterschiedliche Kontexte der Lehrerausbildung ein besonders großes Potenzial: Zum einen lassen sich die Probleme, die sich bei einer empirischen Vergleichsstudie etwa mit Blick auf die Konstruktvalidität und die empirische Validität der Erhebungsinstrumente ergeben (vgl. Blömeke und Paine 2008; Blömeke 2011; Klieme 2012), angesichts der gemeinsamen Sprache und der ähnlichen kulturellen Konnotationen mit einem vertretbaren Forschungsaufwand lösen. Zum anderen ist die Lehrerbildung in Deutschland und Österreich strukturell von sehr unterschiedlicher Prägung (vgl. Mayr und Neuweg 2009; Stadelmann und Rhyn 2011; Bosse et al. 2012), sodass man über Vergleichsanalysen zu ihrer Wirksamkeit im pädagogischen Bereich neue und wichtige Erkenntnisse für aktuelle, auch drängende Reformfragen der Lehrerbildung gewinnen kann.

Insbesondere die Variation in der organisatorisch-curricularen Gestaltung des Theorie-Praxis-Bezugs (z. B. Gewichtung, Verzahnung) bietet ein erhebliches Forschungspotenzial, ist es doch ein Thema, das die aktuellen Reformdebatten zur deutschen Lehrerbildung bestimmt (u. a. Expertenkommission 2007; Czerwenka 2011) und schon seit Jahrzehnten strittig diskutiert wird, ohne dass – von zahlreichen Studierenden-Umfragen einmal abgesehen – eine empirische Datenbasis vorliegen würde, die die tatsächliche Wirksamkeit des vielfach geforderten Praxisbezugs auf relevante Ergebnisvariablen wie das zu erwerbende professionelle Wissen der angehenden Lehrkräfte belegen könnte (vgl. Hascher 2011; Clift und Brady 2005). Die

Varianz in der Gestaltung von Lerngelegenheiten, die mit einem Vergleich der deutschen und der österreichischen Lehrerausbildung auf ihre Wirksamkeit hin untersucht werden kann, erlaubt hier wichtige, bislang nicht geleistete Analysen in Bezug auf die Anbahnung von praktischem Wissen bei Lehramtsstudierenden (u. a. Berliner 2004; Gruber und Rehl 2005; Clift und Brady 2005).

Kernmerkmal der bisherigen österreichischen Lehrerbildung (im Vergleich zur deutschen – hier bis auf Baden-Württemberg) ist die landesweite „Spaltung“ (Messner und Posch 2010, S. 177) in Universitäten und Pädagogische Hochschulen (PHs). Während die universitäre Ausbildung für das „Lehramt für Höhere Schulen“ in Österreich der Gymnasiallehrausbildung in Deutschland strukturähnlich ist (u. a. universitäre Verankerung, zweiphasige Ausbildung, fachliche Schwerpunktsetzung, Schulform des anspruchsvollsten Bildungsgangs im mehrgliedrigen Sekundarschulsystem), unterscheiden sich die Ausbildungsgänge der PHs in Österreich für die Lehrkräfte der sog. „Pflichtschulen“ (u. a. Volksschule, Hauptschule, Sonderschule) deutlich zu den vergleichbaren Ausbildungsgängen in Deutschland (d. h. Grundschule, Haupt-/Realschullehramt, Sonderpädagogik usw.). Österreich gehört bislang zu den letzten europäischen Ländern, in denen die Lehrerbildung für die aufgeführten Pflichtschullehrämter akademisiert wurde (vgl. Hopmann 2008). Jedoch erreichen die PHs auch bisher vermutlich nicht das akademische Niveau der Lehrerbildung anderer Länder oder jenes der PHs in Baden-Württemberg: Das österreichische PH-Studium beschränkt sich bislang auf ein sechssemestriges Bachelorstudium, enthält von Beginn an kontinuierlich Praktika (wöchentliche Praxistage, Blockpraktika) und mündet direkt in den Berufseinstieg (vgl. Mayr und Neuweg 2009; Bosse et al. 2012).³

3.3 Einfluss schulpraktischer Lerngelegenheiten auf den Erwerb pädagogischen Wissens

Modellhafte Vorstellungen zur empirischen Untersuchung der Bedingungen von professioneller Kompetenz im Sinne einer Wirksamkeit der Lehrerausbildung betonen die Bedeutung von *Prozessmerkmalen* zur Beschreibung der Lehr-Lern-Arrangements, die angehende Lehrkräfte unter der Zielperspektive durchlaufen, berufsrelevantes Wissen und Können zu erwerben und aufzubauen (Blömeke 2011). Der dafür in der empirischen Lehrerbildungsforschung verwendete Begriff der „Lerngelegenheiten“ – international *opportunities to learn* – bezieht sich auf institutionelle Ausbildungsangebote, von denen angenommen wird, dass sie zum Erwerb professioneller Kompetenzen einen Beitrag leisten (Tatto et al. 2008). Die Erhebung von Lerngelegenheiten und die differenzierte Analyse ihres Zusammenhangs zu Lernergebnissen bieten demnach wesentliche Einblicke in die Effektivität von Curricula und Möglichkeiten ihrer Qualitätssicherung.

Verschiedene Studien der empirischen Lehrerausbildungsforschung haben sich der Herausforderung gestellt, Lerngelegenheiten in standardisierter Form zu erfassen (vgl. z. B. Blömeke et al. 2010a, b), doch nur wenige dieser Untersuchungen haben dabei Lerngelegenheiten im erziehungswissenschaftlichen bzw. pädagogischen Studium über die einzelne Institution hinweg vergleichend erfasst (König et al. 2014). Fast keine Analyse existiert zur Erklärung des Erwerbs von pädagogischem Wis-

sen durch schulpraktische Lerngelegenheiten. Für diese würden jedoch zumindest zwei Gründe sprechen (vgl. Blömeke 2011): Zum einen ließe sich belegen, wie stark das pädagogische Wissen in Abhängigkeit spezifischer Lerngelegenheiten anwächst. Zum anderen wäre die Abhängigkeit des Wissenserwerbs von Lerngelegenheiten der Lehrerbildung auch ein Beleg für die curriculare Validität des getesteten Wissens.

4 Fragestellungen und Hypothesen

Im Folgenden werden drei Fragen bearbeitet, um den Erwerb von pädagogischem Wissen bei angehenden Lehrkräften in Deutschland und Österreich zu untersuchen:

1. Kann mit dem Instrument bei deutschen und österreichischen Lehramtsstudierenden ein Wissenszuwachs über zwei Messzeitpunkte (1. und 5. Ausbildungssemester) abgebildet werden?

Vermutet wird ein statistisch signifikanter Wissenszuwachs mit großer praktischer Bedeutsamkeit vom ersten zum zweiten Messzeitpunkt für beide Länderstichproben (H1), da zum einen der Test als curricular valide in Bezug auf die beiden Ausbildungskontexte erachtet wird und zum anderen von der Ausbildung in beiden Ländern Lernwirksamkeit angenommen wird. Diese Vermutung baut auf bisherigen mit dem Instrument erzielten und in den Abschn. 2.2 und 2.3 dargestellten Ergebnissen auf.

2. Ist der Erwerb abhängig von strukturell divergierenden, inhaltlich unterschiedlich entwickelten und ausgerichteten Ausbildungssystemen?

Angesichts der Variation in der organisatorisch-curricularen Gestaltung des Theorie-Praxis-Bezugs der beiden Ausbildungskontexte (Abschn. 3.2) werden Unterschiede im konzeptuell-analytischen sowie im handlungsnahen pädagogischen Wissenszuwachs erwartet. Wie in Abschn. 2.3 erläutert, zeigt die LEK-Studie für eine Längsschnittstichprobe von Lehramtsstudierenden in Deutschland, dass diese zwischen ihrem 1. und 4. Semester pädagogisches Wissen erwerben, dass dabei jedoch primär konzeptuell-analytisches und nur geringfügig handlungsnahes Wissen erworben wird, wobei der Erwerb handlungspraktischen Wissens jedoch durch schulpraktische Lerngelegenheiten unterstützt werden kann. Eine Stärke der universitär geprägten Ausbildungskontexte, wie sie in Deutschland (einschließlich der Pädagogischen Hochschulen, die sich ähnlich wie die deutschen Universitäten von der zweiten Ausbildungsphase abgrenzen) vorliegen, dürfte somit im Vergleich mit den Pädagogischen Hochschulen in Österreich in einer curricularen Schwerpunktlegung auf den Erwerb konzeptuell-analytischen Wissens im Bereich der Pädagogik sein. Wir vermuten daher differenzielle Wissenszuwächse in den Subskalen des Tests „Erinnern“ und „Verstehen/Analysieren“ (Abb. 2) zugunsten aller Studiengänge in Deutschland (H2a). Im Gegensatz dazu dürfte handlungsnahes pädagogisches Wissen in solchen Kontexten schwerpunktartig gelehrt und erworben werden, in denen von Anfang an die Nähe zum professionell-praktischen Diskurs betont wird und seminaristische Lehre mit schulpraktischem Training eng verflochten ist. Dies ist insbesondere in den an österreichischen Pädagogischen Hochschulen angebotenen Lehramtsstudiengängen der Fall. Wir vermuten daher differenzielle Wissenszuwächse in der Subskala des

Tests „Kreieren/Generieren von Handlungsoptionen“ zugunsten der Studiengänge an Pädagogischen Hochschulen in Österreich (H2b).

3. Korrespondiert die mit Frage (2) angenommene differenzielle Höhe des Zuwachses pädagogischen Wissens mit entsprechenden Unterschieden in den von den Studierenden wahrgenommenen Lerngelegenheiten?

Da Ausbildungssysteme und Studiengänge nur distale Indikatoren darstellen, werden ergänzend auch Skalen zur Erfassung von schulpraktischen Lerngelegenheiten (Selbstberichte der Studierenden) als proximale Indikatoren für die Variation in der organisatorisch-curricularen Gestaltung des Theorie-Praxis-Bezugs einbezogen. Wir vermuten, dass der Umfang gemachter Erfahrungen, die im Rahmen von schulpraktischen Lerngelegenheiten gesammelt werden konnten (wie die Planung, Durchführung und Analyse von pädagogischen Handlungssituationen, das Beziehen von Theorien auf praktische Situationen), die Veränderung der differenziellen Wissenszuwächse in der Subskala des Tests „Kreieren/Generieren von Handlungsoptionen“ zusätzlich erklären können (H3).

Längsschnittstudien wie die EMW-Studie haben verschiedene Einbußen hinsichtlich der Qualität ihrer Stichprobe in Kauf zu nehmen (detailliert Abschn. 5.1). Um die dargelegten Hypothesen dennoch untersuchen zu können, werden daher individuelle Eingangsvoraussetzungen (z. B. Schulabschlussnoten) kontrolliert. Die Auswahl dieser Kontrollvariablen geschieht auch auf der Basis von Analysen der LEK-Studie zu individuellen Eingangsvoraussetzungen von Lehramtsstudierenden, die für den Erwerb ihres pädagogischen Wissens ausschlaggebend sein können.

5 Untersuchungsmethode

5.1 Stichprobe

Für die Bearbeitung unserer Fragestellungen verwenden wir im Folgenden einen *Teildatensatz* der EMW-Studie, und zwar angehende Lehrkräfte aus Deutschland und Österreich, die zu beiden Messzeitpunkten im Wintersemester 2011/12 (1. Semester) und im Wintersemester 2013/14 (5. Semester) teilnahmen und einem der vier in Tab. 1 gelisteten Lehramtstypen der KMK (2009) zugeordnet werden konnten. Die vier Typen wurden ausgewählt, weil sie die für die nachfolgenden Analysen nötige Vergleichbarkeit aufweisen. Der Teildatensatz enthält insgesamt 33 Studiengänge in den 18 Standorten, wobei ein Studiengang eines Standortes im Durchschnitt von rund 44 Studierenden vertreten ist.

Zum ersten Messzeitpunkt der EMW-Studie wurden in Deutschland 18 Hochschulen aus neun Bundesländern einbezogen. Diese Hochschulen repräsentieren die weiteren, jeweils nicht einbezogenen Hochschulen dieser Bundesländer gut bzw. stellen im Einzelfall eine Vollerhebung auf Hochschulebene dar. In Nordrhein-Westfalen als einem dieser neun Länder wurde eine Zufallsstichprobe von Universitäten gezogen, in den übrigen Ländern wurde aus forschungsökonomischen Gründen die Auswahl der Universitäten letztlich verknüpft mit der Zusage eines Kooperationspartners vor Ort, um die Datenerhebung zu unterstützen.⁴ Die – zum ersten Messzeitpunkt – ein-

Tab. 1 Verteilung des hier verwendeten Teildatensatzes der EMW-Längsschnittstichproben von Studierenden und Hochschulen auf die Lehramtstypen der KMK (2009) sowie österreichische Bezeichnung in Klammern

KMK-Typ	Bezeichnung	Deutschland		Österreich		Gesamt	
		Studierende	Hochschulen	Studierende	Hochschulen	Studierende	Hochschulen
1	Lehrämter der Grundschule bzw. Primarstufe (Lehramt für Volksschulen)	208	5	386	6	594	11
3	Lehrämter für alle oder einzelne Schularten des Sekundarbereichs I (Lehramt für Hauptschulen)	239	5	194	5	433	10
4	Lehrämter für den Sekundarbereich II (allgemeinbildende Fächer) oder für das Gymnasium (Lehramt für höhere Schulen)	136	5	46	1	182	6
6	Sonderpädagogische Lehrämter (Sonderpädagogisches Lehramt)	207	4	35	2	242	6
	Gesamt	790	10	661	8	1451	18

bezogenen Bundesländer decken die Vielfalt der in Deutschland institutionalisierten Lehrerausbildung hinreichend ab (vertreten sind neue und alte Länder, Länder des südlichen, nord-westlichen und östlichen Bundesgebietes), allerdings ist kein Stadtstaat Teil der Länderstichprobe. In Österreich wurden zum ersten Messzeitpunkt neun Pädagogische Hochschulen und zwei Universitäten einbezogen, die ein repräsentatives Abbild der dortigen 22 Standorte mit Lehramtsausbildung (14 Pädagogische Hochschulen, acht Universitäten) darstellen. Auch hier war die Auswahl der elf Standorte aus forschungsökonomischen Gründen bedingt durch die Zusage einer Kooperation vor Ort, welche die Durchführung der Datenerhebung durch Projektpartner einschloss. Auf Individualebene konnten aus den 18 Universitäten in Deutschland 4402 Lehramtsstudierende an der EMW-Studie teilnehmen (Ausschöpfungsquote: 31,7%), während in Österreich 1585 von den elf Standorten teilnehmen konnten (Ausschöpfungsquote: 49,8%).

Vor Durchführung des *zweiten Messzeitpunktes* wurden die Kooperationspartner der einzelnen Standorte um ihre Einschätzung zur Erreichbarkeit der Studierenden ihres Standortes gebeten. An einigen Standorten wurde die Erreichbarkeit bereits im Vorfeld als praktisch nicht realisierbar eingeschätzt, sodass angesichts einer zu erwartenden sehr geringen Rücklaufquote diese Standorte ausgeschlossen werden mussten. Letztlich konnten zehn Standorte in Deutschland sowie acht in Österreich (sieben Pädagogische Hochschulen und eine Universität) wieder einbezogen werden. Auf institutioneller Ebene beträgt der Rücklauf des zweiten Messzeitpunktes demnach 56% in Deutschland und 73% in Österreich. Innerhalb der Standorte wurden Lehramtsstudierende entweder über Kontaktdaten (E-Mail-Adressen), welche sie beim ersten Messzeitpunkt zur Kontaktierung für Folgestudien hinterlassen hatten, zu einer erneuten Teilnahme eingeladen oder über Lehrveranstaltungen, die insbesondere für Lehramtsstudierende im 5. Ausbildungssemester vorgesehen waren, erreicht. Der tatsächliche Rücklauf konnte jedoch nur auf der Basis jener Studierenden berechnet werden, bei denen die Verknüpfung der Fragebögen des ersten und zweiten Messzeitpunktes über einen von ihnen – zu beiden Zeitpunkten – zu gene-

rierenden anonymisierten Code möglich war. Bezogen auf die Teilnehmer des ersten Messzeitpunkts *innerhalb* der Standorte beträgt die so ermittelte Rücklaufquote 27% in Deutschland und 59% in Österreich.

Angesichts des Stichprobenausfalls wurde eine Drop-out Analyse durchgeführt, in der auf der Basis der Stichprobe des ersten Messzeitpunktes die Teilnahme (kodiert als 1, Nichtteilnahme kodiert als 0) am zweiten Messzeitpunkt vorhergesagt wurde. Einbezogen wurden die Hintergrundvariablen Alter und Geschlecht, die Leistungsvariablen Schulabschlussnote und Score im Pädagogischen Wissen sowie Berufsmotivationskalen. Für die deutsche Stichprobe tendieren die Prädiktoren gegen Null (alle $|\beta| < 0,05$) und sind nicht statistisch signifikant. Für die österreichische Stichprobe zeigen sich kleine Zusammenhänge hinsichtlich des Alters ($-0,13$), des Geschlechts ($-0,09$), der Schulabschlussnote ($-0,08$), des pädagogischen Wissens ($0,10$) und der beruflichen Motivation, Lehramt als Verlegenheitslösung anzustreben ($0,10$). Die Ergebnisse der Drop-out Analyse weisen somit nicht auf eine Verzerrung der Stichproben hin, wie es anhand des Stichprobenausfalls insbesondere für die deutsche Stichprobe hätte erwartet werden können. Angesichts der Einschränkungen bei der Datenerhebung handelt es sich bei unserer Längsschnittstichprobe jedoch um eine Gelegenheitsstichprobe, die – ähnlich wie beim ersten Messzeitpunkt – zwar die Heterogenität der Lehrerbildung in beiden Ländern hinreichend abdeckt, jedoch nicht als repräsentativ für die beiden Länder bezeichnet werden kann (zu weiteren Einschränkungen vgl. Abschn. 7).

5.2 Instrumente

5.2.1 Testung des pädagogischen Wissens

Das pädagogische Wissen der angehenden Lehrkräfte wurde zu beiden Zeitpunkten mit dem Testinstrument erfasst, das im Rahmen der internationalen Vergleichsstudie TEDS-M entwickelt worden war (Abschn. 2.2). Zum zweiten Messzeitpunkt unserer Studie wurde das Instrument in voller Länge wie in TEDS-M eingesetzt. Abbildung 3 enthält eine Beispiel-Aufgabe der Inhaltsdimension „Leistungsbeurteilung“ und der kognitiven Anforderung „Erinnern“. Abbildung 4 ist eine Beispiel-Aufgabe zu entnehmen, die der Inhaltsdimension „Motivierung“ und der kognitiven Anforderung

Wenn diagnostische Urteile fair und genau sein sollen, dann müssen sie drei Gütekriterien erfüllen. Welche sind das?	
<i>Bitte nur <u>ein</u> Kästchen ankreuzen.</i>	
A.	Neutralität, Reliabilität, Veridikalität <input type="checkbox"/> 1
B.	Objektivität, Reliabilität, Validität <input checked="" type="checkbox"/> 2
C.	Objektivität, Reliabilität, Veridikalität <input type="checkbox"/> 3
D.	Neutralität, Reliabilität, Validität <input type="checkbox"/> 4

Abb. 3 Testaufgabe zur Erfassung von Wissen zur Leistungsbeurteilung mit der kognitiven Anforderung „Erinnern“ sowie korrekte Lösung (1 Punkt) (König und Blömeke 2009a)

Bei welchen der folgenden Fälle handelt es sich um eine intrinsische Motivation, bei welchen um eine extrinsische Motivation?

Kreuzen Sie ein Kästchen pro Zeile an.

	intrinsische Motivation	extrinsische Motivation
Ein Schüler lernt vor einer Mathematikarbeit, weil er...		
A. für eine gute Note eine Belohnung erwartet.	<input type="checkbox"/> ₁	<input checked="" type="checkbox"/> ₂
B. einen Tadel für eine schlechte Note vermeiden möchte.	<input type="checkbox"/> ₁	<input checked="" type="checkbox"/> ₂
C. an mathematischen Problemen interessiert ist.	<input checked="" type="checkbox"/> ₁	<input type="checkbox"/> ₂
D. seine Eltern nicht enttäuschen möchte.	<input type="checkbox"/> ₁	<input checked="" type="checkbox"/> ₂
E. seine gute Leistungsposition in der Klasse auch in Zukunft behalten möchte.	<input type="checkbox"/> ₁	<input checked="" type="checkbox"/> ₂

Abb. 4 Testaufgabe zur Erfassung von Wissen über Motivation mit der kognitiven Anforderung „Verstehen/Analysieren“ (jeweils 1 Punkt pro richtiger Antwort) (König und Blömeke 2009a)

Nennen Sie drei Unterrichtskonzepte bzw. Unterrichtsmethoden, die es den Schüler(inne)n grundsätzlich ermöglichen, über die Differenzierung im Unterricht mit zu entscheiden.

1)

2)

3)

Beispiel	Originalantworten	Bewertung
1.	1) Lernzirkel 2) Freiarbeit 3) Lehrervortrag	2 Punkte für Lernzirkel und Freiarbeit
2.	1) Wochenplan 2) Stationenlernen 3) Lernbuffet	3 Punkte

Abb. 5 Testaufgabe zur Erfassung von Wissen zur beruflichen Anforderung *Umgang mit Heterogenität* mit der kognitiven Anforderung „Kreieren/Generieren von Handlungsoptionen“ (Antwortkästen in verkleinerter Darstellung) sowie Originalantworten mit Punktvergabe (König und Blömeke 2009a)

derung des „Verstehens/Analysierens“ zugeordnet wird. Abbildung 5 und 6 enthalten Beispiel-Aufgaben der kognitiven Anforderung des „Kreierens/Generierens von Handlungsoptionen“. Angesichts des komplexen Aufbaus können aus einzelnen Testaufgaben mehrere dichotome Testitems resultieren, die anschließend Eingang in die Skalierungsanalysen finden. Die Kodierung der Antworten auf die offenen Testfragen wurde mithilfe der Kodieranleitungen aus TEDS-M von trainierten Kodierern

Abb. 6 Testaufgabe zur Erfassung von Wissen zur Strukturierung von Unterricht mit der kognitiven Anforderung „kreieren/generieren von Handlungsoptionen“ (König und Blömeke 2009a)

Stellen Sie sich vor, Sie helfen einer angehenden Lehrperson bei der Auswertung ihres Unterrichts, weil sie dies noch nie gemacht hat.
Welche Fragen würden Sie ihr stellen, damit sie ihren Unterricht angemessen analysiert?
Nennen Sie zehn zentrale Fragen und formulieren Sie diese bitte aus.

1)
2)
3)
...
10)

durchgeführt (Ergebnisse der Übereinstimmungsanalyse bei Doppelkodierungen in der EMW-Studie: durchschnittliches Kappa=0,8).

Anhand von vertiefenden Skalierungsanalysen auf der Basis der LEK- und TEDS-M-Daten konnte eine Kurzfassung des Tests entwickelt werden, die etwa die Hälfte aller Testaufgaben enthält, bei der jedoch sämtliche Zellen der Testdesign-Matrix abgedeckt sind (König und Blömeke 2010). Diese Testversion wurde aus forschungsökonomischen Gründen zum ersten Messzeitpunkt der EMW-Studie eingesetzt und sie erwies sich zum ersten Messzeitpunkt als reliabel. Mit dem vollständigen Test ist eine multidimensionale Skalierung der inhaltlichen Dimensionen und kognitiven Anforderungen möglich, die sich bei sehr großen Stichproben außerdem mithilfe der Skalierung im Ansatz virtueller Fälle (Rost 1996) und hinreichend vieler Anker-Items auch auf den ersten Messzeitpunkt übertragen lässt, selbst wenn zu diesem Zeitpunkt lediglich die Kurzfassung des Tests eingesetzt wurde (vgl. Wu et al. 1997).

5.2.2 Kontrollvariablen

Neben Alter und Geschlecht wurden die Abiturnote (Deutschland) bzw. die Abschlussnoten der höheren Schule in den Fächern Deutsch, Mathematik und Englisch (Österreich) der Studierenden erfragt. Abiturnote bzw. ein Durchschnittswert der drei Abschlussnoten werden als genereller Indikator für allgemeine kognitive Grundfähigkeiten einbezogen. Die soziale Herkunft der Studierenden wurde in der EMW-Studie erfasst, indem die Studierenden den elterlichen Berufsstatus (*International Socio-Economic Index of Occupational Status* (ISEI), Ganzeboom et al. 1992) angeben sollten, wobei nachfolgend der in der Familie höchste Status (HISEI) in die Analysen einbezogen wird.

5.2.3 Skalen zur Erfassung der lernprozessbezogenen Tätigkeiten in schulpraktischen Lerngelegenheiten

Verwendet wird ein im Rahmen der EMW-Studie entwickeltes und zum zweiten Messzeitpunkt eingesetztes Instrument zur Erfassung von lernprozessbezogenen Tätigkeiten in schulpraktischen Lerngelegenheiten. Es besteht aus insgesamt 74 Items, die sich zu fünf Skalen zusammenfassen lassen: Komplexität über forschungsmethodische Zugänge erkunden, pädagogische Handlungssituationen planen, pädagogische Handlungssituationen durchführen, Theorien auf Situationen beziehen, mit Situationen analytisch-reflexiv umgehen. Die Items besitzen ein dichotomes Antwortformat (ja/nein) und wurden im Fragebogen mit der Frage eingeleitet: „Haben

Tab. 2 Beispiel-Items zur Messung von Dimensionen schulpraktischer Tätigkeiten bei angehenden Lehrkräften (König et al. 2014, S. 14) und Reliabilität der Skalen (Cronbachs Alpha)

Dimension	Beispiel-Items	Items	Cronbachs Alpha
Komplexität über forschungs-methodische Zugänge erkunden	Ich habe Schülermeldungen protokolliert (z. B. Antworten eines Schülers mitgeschrieben)	9	0,70
	Ich habe den Lernstand eines Schülers diagnostiziert und schriftlich festgehalten		
	Ich habe eine Schülerbefragung durchgeführt mit einem Fragebogen		
Pädagogische Handlungssituationen planen	Ich habe Lernziele in Anlehnung an das Curriculum formuliert	12	0,83
	Ich habe Schülerinteressen bei der Unterrichtsplanung berücksichtigt		
	Ich habe die Planung einer weiteren Unterrichtsstunde auf den Unterrichtsverlauf aufgebaut		
Pädagogische Handlungssituationen durchführen	Ich habe die Anwesenheit der Schüler kontrolliert	31	0,90
	Ich habe leistungsdifferenzierte Arbeitsaufträge an Schüler vergeben		
	Ich habe Schülern gezeigt, wie sie ihren Lernweg selbst kontrollieren können		
Theorien auf Situationen beziehen	Ich habe Unterrichtsmethoden beobachtet, die ich in der Universität/der Pädagogischen Hochschule gelernt habe	11	0,83
	Ich habe für die Unterrichtsvorbereitung auf Material zurückgegriffen, das ich aus dem Studium kenne		
	Ich habe Konzepte des Unterrichts umgesetzt, die ich in der Universität/der Pädagogischen Hochschule gelernt habe		
Mit Situationen analytisch-reflexiv umgehen	Ich habe Hospitationsprotokolle von Unterricht alleine kriteriengeleitet ausgewertet	11	0,78
	Ich habe meinen Unterricht anhand didaktischer Theorien alleine analysiert		
	Ich habe mit einer Lehrperson Schlussfolgerungen für die künftige Unterrichtsgestaltung abgeleitet		

Sie in Ihrer bisherigen Schulpraxis/in Ihren bisherigen Schulpraktika die folgenden Tätigkeiten durchgeführt?“ Die einzelnen Dimensionen werden jeweils mit rund zehn Items gemessen, allein die Durchführung von pädagogischen Handlungssituationen wird über deutlich mehr Items erfasst, um ein besonders breites Spektrum pädagogischen Handelns abzubilden (vgl. Tab. 2). Die Skalen sind geeignet, um Varianz zwischen den Ländern abzubilden (vgl. die Mittelwerte und die Effektstärke η^2 der Ländervarianz in Tab. 3).⁵

5.3 Skalierung des Wissenstests

Die Skalierung des Wissenstest erfolgte zunächst getrennt für jeden Messzeitpunkt. Die Skalierung des Tests zum ersten Messzeitpunkt ist in König und Rothland (2013) dokumentiert. Die Skalierung des zweiten Messzeitpunktes folgte dem Vorgehen der LEK-Studie. Dabei mussten von den im TEDS-M Testinstrument enthaltenen Aufgaben einzelne Aufgaben bzw. die daraus resultierenden Items ausgeschlossen werden. Die in der LEK-Studie vorgenommene Skalierung des Tests bewährte sich letztlich auch für die Skalierung des zweiten Messzeitpunktes in der EMW-Studie. Insgesamt kommen 42 Items beim ersten und 79 Items beim zweiten Messzeitpunkt zum Ein-

Tab. 3 Deskriptive Kennwerte (Mittelwert, Standardfehler des Mittelwerts und Standardabweichung) für die Skalen zur Erfassung von lernprozessbezogenen Tätigkeiten in schulpraktischen Lerngelegenheiten (gegliedert nach Land) sowie Ländervarianz (η^2)

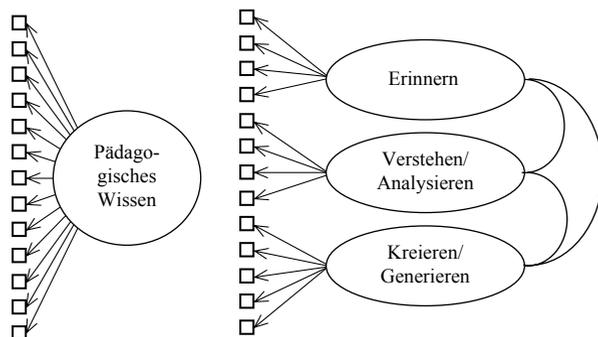
	Deutschland			Österreich			η^2
	M	SE	SD	M	SE	SD	
Komplexität über forschungsmethodische Zugänge erkunden	3,2	0,1	2,2	3,8	0,1	2,3	0,02
Pädagogische Handlungssituationen planen	5,2	0,1	3,0	8,5	0,1	2,4	0,26
Pädagogische Handlungssituationen durchführen	17,7	0,2	6,5	22,0	0,3	6,9	0,09
Theorien auf Situationen beziehen	5,2	0,1	3,0	8,0	0,1	2,3	0,22
Mit Situationen analytisch-reflexiv umgehen	4,5	0,1	2,7	6,9	0,1	2,5	0,18

satz. Die 42 Items des ersten Messzeitpunktes können gleichzeitig als Ankeritems zur Verknüpfung von beiden Messungen verwendet werden.

Für die hier vorzustellenden Analysen wurden in einem weiteren Schritt die Testdaten des ersten und zweiten Messzeitpunktes im Ansatz virtueller Fälle (Rost 1996) zunächst in einem eindimensionalen, dann in einem dreidimensionalen Modell skaliert. Entsprechend der grafischen Veranschaulichung in Abb. 7 wurde dabei die Passung eines Modells, das eine eindimensionale Struktur, und die Passung eines Modells, das zwischen drei Dimensionen kognitiver Anforderungen unterscheidet, mithilfe der Software *Conquest* IRT-skaliert und geschätzt (Wu et al. 1997). *Conquest* weist jedem Item aufgrund seiner empirischen Häufigkeit („Lösungsquote“) einen Itemparameter („Schwierigkeitsparameter“) und jeder befragten Person entsprechend des gezeigten Antwortverhaltens einen Parameter auf der latenten Variable Theta zu („Fähigkeitsparameter“); mit dem mehrdimensionalen *Random Coefficient Multinomial Logit*-Modell, das ebenfalls in *Conquest* implementiert ist, können mehrere latente Variablen gleichzeitig berücksichtigt werden, deren Zusammenhänge messfehlerbereinigt ausgegeben werden (Rost 1996; Wu und Adams 2006).

Für die Überprüfung der internen Konsistenz der drei Dimensionen ziehen wir die *Expected A Posteriori* Schätzung (EAP-Schätzung; de Ayala 1995) heran, die eine unverzerrte Beschreibung der Population liefert und die mehrdimensionale Modellstruktur berücksichtigt (Wu et al. 1997). Das eindimensionale Modell zeigt mit einem Wert von 0,84 eine gute Reliabilität (Theta-Varianz 0,71). Die Item-Fit-Statistiken verdeutlichen, dass die Item-Schwierigkeitswerte über einen Range von mehr als vier

Abb. 7 Schematische Darstellung des eindimensionalen Modells pädagogischen Wissens (*links*) und des dreidimensionalen Modells kognitiver Anforderungen (*rechts*)



Logits streuen (von $-2,36$ bis $2,28$). Item-Diskriminationswerte liegen durchschnittlich bei $0,34$, der *Weighted Mean Square* jedes Items fällt in einen akzeptablen Bereich zwischen $0,81$ und $1,16$, lediglich ein Item weist einen Wert von größer $1,20$ auf.⁶

Die drei Dimensionen des differenzierten Modells besitzen aufgrund der geringeren Item-Anzahl unterschiedlich hohe Reliabilitäten: „Erinnern“: $0,67$, Dimension „Verstehen/Analysieren“: $0,82$, Dimension „Kreieren/Generieren von Handlungsoptionen“: $0,59$ (jeweils vergleichbar mit Cronbachs Alpha), sie liegen somit in einem akzeptablen bis guten Bereich. Einen wichtigen Einblick erlauben die korrelativen Befunde der drei Dimensionen. Der höchste Zusammenhang ($0,59$) besteht zwischen der Dimension „Erinnern“ und der Dimension „Verstehen/Analysieren“, während die Dimension „Kreieren/Generieren von Handlungsoptionen“ weniger hoch mit den beiden anderen Dimensionen zusammenhängt ($0,46$ mit dem „Erinnern“, $0,43$ mit dem „Verstehen/Analysieren“).

Im Anschluss an diese Modellprüfung werden die durch die Skalierung ermittelten Leistungswerte (ein Gesamtscore, je ein Score für die drei Dimensionen kognitiver Anforderungen) in den nachfolgenden Analysen genutzt. Hierfür werden die Personenparameter aus der EAP-Schätzung verwendet. Diese werden für den Gesamtscore und für die drei Scores kognitiver Anforderungen getrennt im Ansatz virtueller Fälle auf einen Mittelwert von 50 und eine Standardabweichung von 10 transformiert. Anschließend erfolgt die Datenanalyse auf manifester Ebene mithilfe des Softwarepakets *Mplus* (Muthén und Muthén 1996–2006), welches u. a. Mehrebenenanalysen sowie einen adäquaten Umgang mit fehlenden Werten ermöglicht. In den nachfolgenden Analysen werden fehlende Werte mithilfe der Methode der *full-information-maximum-likelihood* modellbasiert geschätzt, d. h. der Behandlung fehlender Werte wird dadurch Rechnung getragen, dass Analyse- und Imputationsmodell identisch sind (vgl. Lüdtke et al. 2007).

6 Ergebnisse

6.1 Veränderung des pädagogischen Wissens

Bereits die in Tab. 4 dargestellten Mittelwerte verdeutlichen, dass es in beiden Ausbildungskontexten zu einem Zuwachs an Wissen kommt. Die Differenz der Gesamtscores in beiden Ländern liegt jeweils bei rund zehn Punkten und erreicht damit die Größe einer Standardabweichung. Tabelle 5 (obere Hälfte) sind ferner Kennwerte von mithilfe der Software *Mplus* durchgeführten Regressionsanalysen zu entnehmen, in denen das pädagogische Wissen der Lehramtsstudierenden im Ansatz virtueller Fälle und auf manifester Ebene als Kriteriumsvariable, der Messzeitpunkt als unabhängige Variable spezifiziert wurden. Im Falle der Differenzierung des pädagogischen Wissens in Subskalen wurden sämtliche Subskalen eines Modells als abhängige Variablen in einer Regressionsanalyse spezifiziert und der Effekt des Messzeitpunktes simultan auf diese Subskalen geschätzt.⁷ Ergänzend wurden die Regressionsmodelle auch unter Einschluss der Kontrollvariablen berechnet (Tab. 5, untere Hälfte). Da unter Einbezug der Kontrollvariablen die Effekte des Messzeitpunktes nicht geschmälert werden, beziehen wir uns nachfolgend auf die Ergebnisse der oberen Hälfte in Tab. 5.

Tab. 4 Deskriptive Kennwerte (Mittelwert, Standardfehler des Mittelwerts und Standardabweichung) für das pädagogische Wissen (Gesamtscore und Subdimensionen kognitiver Anforderungen), gegliedert nach Messzeitpunkt und Land

		t1			t2		
		M	SE	SD	M	SE	SD
Gesamtscore	Deutschland	45,1	0,3	8,1	54,8	0,3	9,2
	Österreich	44,8	0,4	8,9	55,2	0,3	8,0
Erinnern	Deutschland	44,7	0,3	7,6	54,9	0,4	9,6
	Österreich	45,2	0,3	8,6	55,1	0,3	8,1
Verstehen/Analysieren	Deutschland	45,3	0,3	8,6	55,7	0,3	8,3
	Österreich	44,4	0,4	9,1	54,8	0,3	7,3
Kreieren	Deutschland	47,0	0,3	9,2	51,1	0,4	10,2
	Österreich	47,0	0,4	9,7	54,1	0,4	9,0

Tab. 5 Regression des pädagogischen Wissens auf den Messzeitpunkt (Multi-Gruppen-Analyse für die Länder, Schätzung in *Mplus* unter Berücksichtigung des Ortes/Studiengangs als Stratifikationskriterien)

Modell	Eindimensional		Dreidimensional					
	Gesamtscore		Erinnern		Verstehen/ Analysieren		Kreieren/Generieren von Handlungsoptionen	
	β (SE)	R ² (SE)	β (SE)	R ² (SE)	β (SE)	R ² (SE)	β (SE)	R ² (SE)
Deutschland ^a	0,49 (0,02)	0,24 (0,02)	0,51 (0,02)	0,26 (0,02)	0,52 (0,02)	0,27 (0,02)	0,21 (0,02)	0,04 (0,01)
Österreich ^a	0,52 (0,02)	0,27 (0,02)	0,51 (0,02)	0,26 (0,02)	0,52 (0,02)	0,29 (0,02)	0,35 (0,02)	0,12 (0,02)
Deutschland ^b	0,49 (0,02)	0,24 (0,02)	0,51 (0,02)	0,27 (0,02)	0,52 (0,02)	0,28 (0,02)	0,21 (0,02)	0,05 (0,01)
Österreich ^b	0,52 (0,02)	0,29 (0,02)	0,51 (0,02)	0,28 (0,02)	0,53 (0,02)	0,29 (0,02)	0,35 (0,02)	0,15 (0,02)

^a Ohne Kontrollvariablen (Geschlecht, Alter, Schulabschlussnote, HISEI)

^b Mit Kontrollvariablen (Geschlecht, Alter, Schulabschlussnote, HISEI)

Der Messzeitpunkt als Prädiktor zur Erklärung des Gesamtscores in der Regressionsanalyse ist statistisch signifikant und von großer praktischer Bedeutsamkeit: 24% der Varianz in der deutschen und 27% der Varianz in der österreichischen Stichprobe können durch den Zeitpunkt erklärt werden. Auch die Subskalen der kognitiven Anforderungen können durch den Messzeitpunkt statistisch signifikant vorhergesagt werden – doch hier zeigen sich Unterschiede nach Subskala, die auf unterschiedliche Schwerpunkte des Wissenserwerbs schließen lassen. Im Bereich des „Erinnerns“ sowie im Bereich des „Verstehens/Analysierens“ lassen sich sowohl für die deutsche Stichprobe (Varianzaufklärung 26 bzw. 27%) als auch für die österreichische Stichprobe (Varianzaufklärung 26 bzw. 29%) Unterschiede durch den Zeitpunkt jeweils mit großer praktischer Bedeutsamkeit erklären. Es zeigen sich keine Unterschiede zugunsten der deutschen Studierenden. Dagegen fällt der Mittelwertunterschied im Bereich des „Kreierens/Generierens von Handlungsoptionen“ für die deutsche Stichprobe deutlich kleiner aus (4% der Varianz im pädagogischen Wissen können durch den Messzeitpunkt erklärt werden) als für die österreichische Stichprobe (12% erklärte Varianz). Die deutschen Studierenden weisen innerhalb ihrer ersten beiden Studienjahre einen statistisch signifikant geringeren Wissenszuwachs

auf als die österreichischen, wie ein Vergleich der unterschiedlich hohen Anteile erklärter Varianz verdeutlicht ($t=3,58, p<0,001$). Die erreichten Mittelwerte zum zweiten Messzeitpunkt sind bei den österreichischen Studierenden statistisch signifikant höher als bei den deutschen ($t=-5,66, p\leq 0,001$), während sie sich zum ersten Messzeitpunkt nicht statistisch signifikant unterscheiden ($t=-0,09, n. s.$).

6.2 Befunde aus Mehrebenenanalysen

Zur weiteren Bearbeitung unserer Fragestellungen beziehen wir Eingangsvoraussetzungen als Kontrollvariablen und die Skalen zur Erfassung von lernprozessbezogenen Tätigkeiten in schulpraktischen Lerngelegenheiten als proximale Prädiktoren zur Erklärung der Veränderung im pädagogischen Wissen ein. Hierfür spezifizieren wir Zwei-Ebenen-Modelle mit den Lehramtsstudierenden auf der ersten Ebene (Individualdatenebene) und ihrem Studiengang innerhalb ihrer Hochschule als zweite Ebene (Institutionenebene). Auf der Individualebene kontrollieren wir die jeweilige Kriteriumsvariable zum ersten Messzeitpunkt, Geschlecht, Alter, Schulabschlussnote und den höchsten in der Familie vorhandenen Erwerbsstatus (HISEI). Auf der Institutionenebene kontrollieren wir die Studiengänge als Dummy-Variablen (Referenz: Gymnasiallehramt bzw. Lehramt für höhere Schulen) und fügen das Land als Prädiktor ein.

Wie in Tab. 6 anhand der Befunde aus den vier Modellen für den Gesamtscore sowie die drei Scores kognitiver Anforderungen ersichtlich ist, wird bei Kontrolle von individuellen Eingangsvoraussetzungen sowie der Studiengänge der Länderprädiktor nicht signifikant – mit Ausnahme des vierten Modells mit dem „Kreieren/Generieren von Handlungsoptionen“ als abhängige Variable. Mit $\beta=0,37$ ($p\leq 0,05$) hat der österreichische Ausbildungskontext einen Einfluss auf den Wissenszuwachs in diesem Bereich pädagogischen Wissens, während bei den beiden anderen Scores kognitiver Anforderungen kein überzufälliger Länderunterschied auf die Wissensveränderung festgestellt werden kann.

Um zu prüfen, ob die Länderunterschiede im Wissenszuwachs im Bereich „Kreieren/Generieren von Handlungsoptionen“ auf Unterschiede in den schulpraktischen Lerngelegenheiten zurückgeführt werden können, wurden – ausgehend vom vierten Modell (M4) in Tab. 6 – zunächst auf der Individualebene schrittweise die fünf Skalen lernprozessbezogener Tätigkeiten in schulpraktischen Lerngelegenheiten eingeführt (Tab. 7).⁸ Da sie am Gesamtmittelwert zentriert sind, enthalten sie nicht nur Varianzanteile der Individual- sondern auch der Institutionenebene (Bryk und Raudenbush 1992). Mit ihrer Einführung ist der Länderprädiktor nicht mehr statistisch signifikant.⁹ Demnach lassen sich bereits mit dem am Gesamtmittelwert zentrierten Prädiktoren lernprozessbezogener Tätigkeiten in schulpraktischen Lerngelegenheiten Veränderung im Bereich des „Kreierens/Generierens von Handlungsoptionen“ erklären. Die Veränderungseffekte sind statistisch signifikant, allerdings nur für die drei Skalen „Pädagogische Handlungssituationen durchführen“, „Theorien auf Situationen beziehen“ und „Mit Situationen analytisch-reflexiv umgehen“ auch von praktischer Bedeutsamkeit ($\beta\geq 0,1$). Studierende, die über umfangreichere schulpraktische Erfahrungen in diesen Bereichen berichten, zeigen einen stärkeren Wissenszuwachs im Bereich des „Kreierens/Generierens von Handlungsoptionen“ bei der Testung pädagogischen Wissens.

Tab. 6 Befunde aus Mehrebenenanalysen zum pädagogischen Wissen (standardisierte Regressionskoeffizienten und Standardfehler)

	Gesamtscore		Erinnern		Verstehen/ Analysieren		Kreieren/Ge- nerieren von Handlungsoptionen	
	M1		M2		M3		M4	
	β	SE	β	SE	β	SE	β	SE
<i>Individualebene^a</i>								
Kriterium t1	0,22***	0,03	0,11**	0,03	0,21***	0,03	0,22***	0,03
Geschlecht	-0,04	0,03	-0,09**	0,03	0,03	0,03	-0,10**	0,03
Alter	0,01	0,04	0,04	0,03	0,01	0,04	-0,01	0,05
Schulabschlussnote	-0,18***	0,03	-0,18***	0,04	-0,19***	0,03	-0,10***	0,03
HISEI	-0,02	0,03	0,00	0,03	-0,02	0,03	-0,02	0,05
<i>Institutionenebene</i>								
Lehramtstyp (Referenz: Gymnasiallehramt bzw. Lehramt für höhere Schulen)								
Lehrämter der Grund- schule bzw. Primar- stufe (Lehramt für Volksschulen)	0,78***	0,19	0,59*	0,24	0,93***	0,16	0,49**	0,19
Lehrämter für alle oder einzelne Schul- arten des Sekundar- bereichs I (Lehramt für Hauptschulen)	0,20	0,26	0,20	0,26	0,27	0,26	-0,02	0,25
Sonderpädagogische Lehrämter (Sonderpäd- agogisches Lehramt)	0,47*	0,24	0,35	0,28	0,63**	0,20	0,22	0,24
Land (Kodierung: Deutschland = 0, Öster- reich = 1)	-0,05	0,17	-0,17	0,20	-0,20	0,16	0,37*	0,18

* $p \leq 0,05$; ** $p \leq 0,01$; *** $p \leq 0,001$

^a Prädiktoren auf Individualebene am Gesamtmittelwert zentriert

7 Zusammenfassung und Diskussion

7.1 Zusammenfassung und Diskussion

Der vorliegende Beitrag hatte zum Ziel, am Beispiel der beiden Ausbildungskontexte Deutschland und Österreich Annahmen zum Erwerb von pädagogischem Wissen im Rahmen der Lehrerausbildung zu formulieren. Solche Vergleichsanalysen sind wichtig, da sich die beiden deutschsprachigen Länder mit analogen oder identischen Problemlagen und Fragestellungen konfrontiert sehen (vgl. Stadelmann und Rhyn 2011). Für das verwendete Instrument zur Erfassung von pädagogischem Wissen, welches ursprünglich für internationale Vergleichsanalysen im Rahmen von TEDS-M entwickelt worden war, wurde curriculare Validität in Bezug auf die beiden hier einbezogenen Ausbildungskontexte angenommen. Die Ergebnisse bestätigen diese sowie die vermutete Lernwirksamkeit der deutschen und österreichischen Lehrerausbildung mit Blick auf das getestete pädagogische Wissen: In beiden Ausbildungskontexten ist ein statistisch signifikanter Wissenszuwachs bei Lehramtsstudierenden zwischen ihrem ersten und fünften Ausbildungssemester feststellbar, was die Hypothese zur

Tab. 7 Befunde aus Mehrebenenanalysen zum pädagogischen Wissen im Bereich des „Kreierens/Generierens von Handlungssituationen“ (standardisierte Regressionskoeffizienten und Standardfehler)

	M1		M2		M3		M4		M5	
	β	SE								
<i>Individualebene^a</i>										
Kriterium t1	0,21***	0,04	0,21***	0,04	0,21***	0,03	0,21***	0,04	0,20***	0,04
Geschlecht	-0,10**	0,04	-0,09**	0,04	-0,10**	0,04	-0,10*	0,04	-0,09*	0,04
Alter	0,01	0,05	0,01	0,05	0,05	0,05	0,02	0,05	0,02	0,05
Schulabschlussnote	-0,12***	0,03	-0,11***	0,03	-0,12***	0,03	-0,12***	0,03	-0,11***	0,03
HISEI	-0,02	0,03	-0,01	0,03	-0,01	0,04	0,00	0,04	-0,01	0,03
<i>Schulpraktische Lerngelegenheiten</i>										
Komplexität über forschungsmethodische Zugänge erkunden	0,06*	0,03								
Pädagogische Handlungssituationen planen			0,07*	0,03						
Pädagogische Handlungssituationen durchführen					0,14***	0,03			0,18***	0,04
Theorien auf Situationen beziehen										
Mit Situationen analytisch-reflexiv umgehen										
<i>Institutionenebene</i>										
Lehramtstyp (Referenz: Gymnasiallehramt (Lehramt für höhere Schulen)										
Lehrämter der Grundschule bzw. Primarstufe (Lehramt für Volksschulen)	0,37*	0,18	0,37*	0,17	0,17	0,21	0,30	0,17	0,39*	0,16
Lehrämter für alle oder einzelne Schularten des Sekundarbereichs I (Lehramt für Hauptschulen)	-0,14	0,23	-0,20	0,23	-0,34	0,26	-0,28	0,25	-0,18	0,25
Sonderpädagogische Lehrämter (Sonderpädagogisches Lehramt)	0,02	0,23	0,02	0,25	-0,16	0,23	0,04	0,24	0,07	0,23
Land (Kodierung: Deutschland = 0, Österreich = 1)	0,20	0,23	0,14	0,23	0,13	0,25	-0,01	0,25	0,04	0,24

* $p \leq 0,05$; ** $p \leq 0,01$; *** $p \leq 0,001$

^a Prädiktoren auf Individualebene am Gesamtmittelwert zentriert

ersten Fragestellung (H1) bestätigt und bisherige Ergebnisse aus Längsschnittstudien zum Erwerb von pädagogischem Wissen repliziert, insbesondere aus der LEK-Studie. Ferner zeigen sich höhere Zuwächse im Bereich des handlungsnahen Wissens (Dimension „Kreieren/Generieren von Handlungsoptionen“) bei österreichischen Studierenden im Vergleich zu den deutschen. Die unterschiedlich hohen Zuwächse können auf die unterschiedlichen schulpraktischen Lerngelegenheiten zurückgeführt werden, welche ihrerseits laut Angabe der Studierenden umfangreicher in Österreich sind als in Deutschland (vgl. Tab. 3). Dieses Ergebnis bestätigt damit unsere Hypothesen, dass mit dem österreichischen Lehrerbildungssystem (H2b) sowie ferner mithilfe proximaler Prädiktoren zur Erfassung der schulpraktischen Lerngelegenheiten (H3) der unterschiedlich hohe Wissenszuwachs im Bereich des „Kreierens/Generierens von Handlungsoptionen“ erklärt werden kann. Nicht belegen hingegen ließ sich die Vermutung (H2a), dass Lehramtsstudierende in Deutschland aufgrund des stärker akademisch geprägten Studiums im Bereich des konzeptuell-analytischen Wissens (Dimensionen „Erinnern“ und „Verstehen/Analysieren“) zu einem höheren Wissenszuwachs gelangten als österreichische Studierende. In unseren Stichproben gelangen sowohl die deutschen als auch die österreichischen Studierenden zu ähnlich hohen Wissenszuwächsen.

Die hier durchgeführten Analyseergebnisse können als wichtiger Beitrag zur Frage nach (fehlender) Wirksamkeit der Lehrerbildung gesehen werden, die sich gerade in den deutschsprachigen Ländern seit Jahren wiederkehrend stellt (u. a. Oser und Oelkers 2001; Blömeke 2004; Terhart 2012). Insbesondere die Kohärenz zwischen den beiden Ausbildungsorten Hochschule und Schulpraxis wird – bereits für die ersten Semester der Lehramtsausbildung an Universitäten bzw. Pädagogischen Hochschulen – als wichtige Voraussetzung für eine gelingende Verknüpfung zwischen Theorie und Praxis gesehen (KMK 2005). Unsere Ergebnisse deuten darauf hin, dass eine solche Kohärenz in Österreich eher gegeben ist als in Deutschland, jedenfalls mit Blick auf die Einschätzung der Studierenden, Theorien auf praktische Situationen beziehen zu können (Tab. 3) bzw. mit Blick auf die Veränderungswirkung dieser lernprozessbezogenen Dimension auf das handlungsnahes pädagogische Wissen (Tab. 7, M4). Dies lässt sich weiterführend im Zusammenhang betrachten mit Ergebnissen, die bereits in TEDS-M für die Situation angehender Mathematiklehrkräfte in Deutschland erbracht wurden, welche nämlich ihre Ausbildung im internationalen Vergleich besonders wenig kohärent erlebten (Hsieh et al. 2011).

Neben unterschiedlichen Ausbildungskontexten dürfte es außerdem auf der Individualenebene angehender Lehrkräfte zu erheblichen Unterschieden in der Wahrnehmung schulpraktischer Lerngelegenheiten bzw. in der konkreten Nutzung durch die Studierenden kommen. Speziell im bildungswissenschaftlichen Begleitstudium (Lohmann et al. 2011) und in Schulpraktika (König 2012) besteht innerhalb von Ausbildungsgängen und Standorten Wahlfreiheit für Studierende. Professionstheoretische, sich in der curricularen Gestaltung niederschlagende Vorstellungen zur Lehrerbildung, z. B. als ein berufsbiographischer (Terhart 2011), „auf die Selbstverantwortung der Studierenden gerichteter Prozess“ (Messner und Posch 2010, S. 185), können die individualisierte Nutzung von Lerngelegenheiten durch die Studierenden zusätzlich verstärken. Auch dies kann als Erklärung für die aufgetretenen Zusammenhänge zwischen Lerngelegenheiten und Wissenszuwachs (Tab. 7) herangezogen werden.

Schließlich verdeutlichen die Ergebnisse, dass eine Differenzierung des pädagogischen Wissens in handlungsnahes und konzeptuell-analytisches von Bedeutung ist, um Lern- und Ausbildungsprozesse angehender Lehrkräfte sowie Wirkungsweisen von Ausbildungsprogrammen und ihren Charakteristika genauer beschreiben und analysieren zu können. Die alleinige Verwendung eines Gesamtscores pädagogischen Wissens in der vorliegenden Studie hätte von einer solchen Differenzierung abstrahiert. Ebenfalls stellt sich die Frage, ob Tests, die zum Beispiel aus forschungsökonomischen Gründen hauptsächlich geschlossene Testaufgaben einsetzen und auf die Verwendung von aufwändigen offenen Testaufgabenformaten (s. Abb. 5 und 6) verzichten, sensitiv für die besonderen Wirkungen von schulpraktischen Lerngelegenheiten in der Lehrerbildung sind. Wiederholt wurde in den vergangenen Jahren verlangt, Testungen von Lehrerwissen stärker an ihrem Handeln in der Praxis zu orientieren (Shavelson 2010; Blömeke et al. 2015). Auch wenn weitere Möglichkeiten der Kompetenzmessung bestehen und derzeit genutzt werden (z. B. video-basierte Tests) scheint doch die Arbeit mit offenen und komplexen Testfragen im Rahmen eines Papier-Bleistift-Tests eine möglicher Weg zu sein, dieser Forderung zu begegnen.

7.2 Limitierungen

Als zentrale Limitierung unserer Studie ist die eingeschränkte Stichprobenqualität zu nennen, vor allem fehlende Repräsentativität für das jeweilige Land, aber auch eine niedrige Rücklaufquote in Deutschland zum zweiten Messzeitpunkt. Sollte entgegen den Ergebnissen unserer Drop-out Analyse doch eine positive Auslese an Studierenden vorliegen, könnte dies Varianz einschränken, d. h. möglicherweise wird das Niveau pädagogischen Wissens zum zweiten Messzeitpunkt in der deutschen Stichprobe überschätzt. Der Wissenszuwachs wäre demnach geringer als die hier erbrachten Ergebnisse zeigen. Da jedoch die Zuwächse in Österreich nicht geringer, sondern eher höher ausfallen, kann angenommen werden, dass mit einer höheren Stichprobenqualität die hier angetroffenen Unterschiede nicht geringer, sondern vermutlich sogar größer (zugunsten Österreichs) sein dürften. Dies ist als offene Frage für zukünftige Studien zu verstehen, denen es gelingt, eine Qualität von Länderstichproben nach den in TEDS-M vorliegenden Standards der IEA zu erreichen. Allerdings können die hier erzielten Ergebnisse dadurch etwas abgesichert werden, da sie die Ergebnisse der LEK-Studie relativ präzise replizieren, wenn auch für deutlich mehr Standorte. Im Falle der österreichischen Stichprobe konnte ebenfalls das Ergebnis aus einem früheren, quasi-längsschnittlichen Vergleich von König und Blömeke (2009b) nun anhand einer echten Längsschnittstichprobe mit mehreren Standorten repliziert werden, da der Leistungszuwachs in der vorliegenden Studie ähnlich hoch ausfällt wie der in der früheren Studie berichtete Mittelwertunterschied zwischen Studierenden des ersten und fünften Semesters.

Ferner ist die Erfassung von Lerngelegenheiten kritisch zu diskutieren, auch wenn mit dem verwendeten Instrument eine Verbesserung gegenüber früheren Instrumenten erzielt worden sein dürfte. Dies betrifft einerseits die Schwierigkeit, Inhalte einzelner Items unterschiedlicher Dimensionen besonders eindeutig voneinander abzugrenzen, andererseits neben einer Quantifizierung von Lerngelegenheiten auch

ihrem multidimensionalen Charakter gerecht zu werden. Zumindest mit Blick auf die erste der fünf Skalen (Komplexität über forschungsmethodische Zugänge erkunden) wäre zukünftig eine extensivere Erfassung sinnvoll, um zu differenzierten Aussagen etwa zu den verschiedenen Formen forschenden Lernens (vgl. Schneider und Wildt 2009) zu gelangen.

7.3 Ausblick auf zukünftige Forschung

Die vorliegende Fragestellung zielte auf die Untersuchung des Wissenszuwachses in unterschiedlichen Ausbildungskontexten. Die in unserer Untersuchung für schulpraktische Lerngelegenheiten berichtete Variation sollte mithilfe weiterer Merkmale der Lehrerausbildungssysteme analysiert werden. Im Rahmen der EMW-Studie werden zahlreiche Maße zum intendierten und implementierten Curriculum erfasst, sodass weiterführende Analysen zum Einfluss konkreter Merkmale auf die Veränderung des pädagogischen Wissens durchgeführt werden. Weitere individuelle Eingangsvoraussetzungen (z. B. motivationale Merkmale) werden ebenfalls geprüft. Während in der vorliegenden Studie nur zwei Messzeitpunkte berücksichtigt werden konnten, stellt sich die Frage nach Verlaufsanalysen mit weiteren Messzeitpunkten. Hierfür wird im Herbst 2015 ein dritter Messzeitpunkt anvisiert. Während sich Lehramtsstudierende in Deutschland zu diesem Zeitpunkt im Master-Studium befinden, verfügt die Vergleichsstichprobe in Österreich über bis zu einem Jahr Berufserfahrung als voll berufstätige Lehrkraft. Inwieweit sich die hier beobachtbaren differenziellen Entwicklungen im handlungsnahen pädagogischen Wissen fortsetzen, ist eine Frage, die für die Gestaltung der Lehrerbildung, insbesondere mit Blick auf neue schulpraktische Bestandteile wie etwa das Praxissemester an deutschen Hochschulen, wichtige Implikationen birgt und somit zukünftig unter vergleichender Perspektive nicht nur von der EMW-Studie bearbeitet werden dürfte.

Anmerkungen

- 1 Die EMW-Studie wird von der Rhein-Energie-Stiftung Köln (Projektnummer: W-13-2-003) gefördert und von Prof. Dr. Johannes König (Universität zu Köln) und Prof. Dr. Martin Rothland (Universität Siegen) geleitet.
- 2 So thematisiert Lüders (2012) beispielsweise, dass nicht sämtliche Wissensbestände der Allgemeinen Didaktik von dem Test abgedeckt seien, erkennt aber wertschätzend an, dass Inhalte abgedeckt seien, die eine breite Einsetzbarkeit des Tests ermöglichen.
- 3 Während diese Situation noch für die im Rahmen der EMW-Studie untersuchten angehenden Lehrkräfte zutrifft, sei allerdings ergänzend auf den aktuell laufenden Reformprozess „Pädagoginnenbildung NEU – die Zukunft der Pädagogischen Berufe“ in Österreich verwiesen, mit dem in naher Zukunft wesentliche Veränderungen der österreichischen Lehramtsausbildung einhergehen dürften.
- 4 Die EMW-Studie wird getragen von einem Netzwerk von Kooperationspartnern der teilnehmenden Hochschulen, die vor Ort die Datenerhebung unterstützen. Für die Mithilfe bei der Durchführung dieser Studie danken wir herzlich diesen Kooperationspartnern sowie allen weiteren Personen, die dabei mitgewirkt haben. Außerdem gilt unser Dank allen Lehramtsstudierenden, die sich die Mühe gemacht haben, an der Befragung bzw. Testung teilzunehmen.
- 5 Die Skalen wurden – nach Prüfung der faktoriellen Struktur – lediglich durch Aufsummierung der dichotomen Items gebildet und nicht zusätzlich durch die Anzahl der Items geteilt.

- 6 Aus theoretischen Gründen wurde dieses Item jedoch nicht ausgeschlossen. In der anschließenden dreidimensionalen Skalierung weist es zudem einen weitaus besseren Fit auf (*Weighted Mean Square* 1,07).
- 7 Alternativ wurden die Effekte auf latenter Ebene mit dem Programm *Conquest* geschätzt. Insgesamt führen die auf diese Weise erzielten Ergebnisse jedoch nicht zu einer anderen Interpretation, sodass im Folgenden allein die Ergebnisse aus *Mplus* berichtet werden, welche unter Berücksichtigung der Stratifikationskriterien der Stichprobe (Universität bzw. Pädagogische Hochschule, Studiengang) erzielt werden konnten und damit zu korrekten Standardfehlern und Teststatistiken gelangen.
- 8 Aufgrund relativ hoher Korrelationen der fünf Skalen lernprozessbezogener Tätigkeiten in der deutschen ($0,34 \leq r \leq 0,62$) und österreichischen Stichprobe ($0,38 \leq r \leq 0,72$) verzichten wir mit Blick auf mögliche Probleme der Multikollinearität darauf, die fünf Skalen gemeinsam in eine Regressionsanalyse einzubeziehen.
- 9 Auf eine zusätzliche Einführung des aggregierten Merkmals wird daher verzichtet, zumal angesichts der großen Mittelwertunterschiede Probleme der Multikollinearität auftreten können und die Einführung weiterer Prädiktoren auf Institutionenebene angesichts der kleinen Einheitenstichprobe ($n=33$) nicht ratsam erscheint.

Literatur

- Aebli, H. (1983). *Zwölf Grundformen des Lehrens*. Stuttgart: Klett-Cotta.
- Anderson, L. W., & Krathwohl, D. R. (Hrsg.) (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. New York: Longman.
- Baer, M., Dörr, G., Fraefel, U., Kocher, M., Küster, O., Larcher, S., Müller, P., Sempert, W., & Wyss, C. (2007). Werden angehende Lehrpersonen durch das Studium kompetenter? *Unterrichtswissenschaft*, 35, 15–47.
- Baumert, J., & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 9(4), 469–520.
- Berliner, D. C. (2004). Describing the behavior and documenting the accomplishments of expert teachers. *Bulletin of Science, Technology & Society*, 24, 200–212.
- Blömeke, S. (2004). Empirische Befunde zur Wirksamkeit der Lehrerbildung. In S. Blömeke, P. Reinhold, G. Tulodziecki, & J. Wildt (Hrsg.), *Handbuch Lehrerbildung* (S. 59–91). Bad Heilbrunn: Klinkhardt.
- Blömeke, S. (2011). Forschung zur Lehrerbildung im internationalen Vergleich. In E. Terhart, H. Bennewitz, & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (S. 345–361). Münster: Waxmann.
- Blömeke, S., & Paine, L. (2008). Getting the fish out of the water: Considering benefits and problems of doing research on teacher education at an international level. *Teaching and Teacher Education*, 24(4), 2027–2037.
- Blömeke, S., Felbrich, F., & Müller, A. (2008). Messung des erziehungswissenschaftlichen Wissens angehender Lehrkräfte. In S. Blömeke, G. Kaiser, & R. Lehmann (Hrsg.), *Professionelle Kompetenz angehender Lehrerinnen und Lehrer* (S. 171–193). Münster: Waxmann.
- Blömeke, S., Kaiser, G., & Lehmann, R. (Hrsg.) (2010a). *TEDS-M 2008– Professionelle Kompetenz und Lerngelegenheiten angehender Primarstufenlehrkräfte im internationalen Vergleich*. Münster: Waxmann.
- Blömeke, S., Kaiser, G., & Lehmann, R. (Hrsg.) (2010b). *TEDS-M 2008– Professionelle Kompetenz und Lerngelegenheiten angehender Mathematiklehrkräfte im internationalen Vergleich*. Münster: Waxmann.
- Blömeke, S., Gustafsson, J.-E., & Shavelson, R. (2015). Beyond dichotomies: Competence viewed as a continuum. *Zeitschrift für Psychologie*, 223, 3–13.
- Blömeke, S., Bremerich-Vos, A., Haudeck, H., Kaiser, G., Lehmann, R., Nold, G., Schwippert, K., & Willenberg, H. (Hrsg.) (2011). *Kompetenzen von Lehramtsstudierenden in gering strukturierten Domänen. Erste Ergebnisse aus TEDS-LT*. Münster: Waxmann.
- 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, 133–180.

- Bosse, D., Criblez, L., & Hascher, T. (Hrsg.) (2012). *Reform der Lehrerbildung in Deutschland, Österreich und der Schweiz. Teil 1: Analysen, Perspektiven und Forschung*. Immenhausen: Prolog.
- Bromme, R. (1992). *Der Lehrer als Experte: zur Psychologie des professionellen Wissens*. Bern: Huber.
- Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models*. Newbury Park: Sage.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81–105.
- Clift, R. T., & Brady, P. (2005). Research on methods courses and field experiences. In M. Cochran-Smith & K. M. Zeichner (Hrsg.), *Studying teacher education. the report of the AERA panel on research and teacher education* (S. 309–424). Mahwah: Erlbaum.
- Czerwenka, K. (2011). Schulpädagogik und schulpraktische Studien: Der Theorie-Praxis-Bezug in der Lehrerbildung. In S. Hellekamps, W. Plöger, & W. Wittenbruch (Hrsg.), *Handbuch Erziehungswissenschaft. Bd. 3: Schule* (S. 665–674). Paderborn: Schöningh.
- De Ayala, R. J. (1995). *An investigation of the standard errors of expected a posteriori ability estimates*. Paper presented at the Annual Meeting of the AERA. San Francisco.
- Eurydice (2004). *Der Lehrerberuf in Europa. Profil, Tendenzen und Anliegen. Bericht 1: Lehrerausbildung und Maßnahmen für den Übergang in das Berufsleben*. Brüssel: Eurydice.
- Expertenkommission (2007). *Ausbildung von Lehrern in NRW*. Düsseldorf: MSW.
- Fenstermacher, G. D. (1994). The Knower and the Known: The Nature of Knowledge in Research on Teaching. In L. Darling-Hammond (Hrsg.), *Review of Research in Education* (Vol. 20, S. 3–56). Washington.
- Ganzeboom, H. B., De Graaf, P. M., & Treiman, D. J. (1992). A standard international socio-economic index of occupational status. *Social science research*, 21(1), 1–56.
- Gruber, H. & Renkl, A. (2000). Die Kluft zwischen Wissen und Handeln: Das Problem des trägen Wissens. In G. H. Neuweg (Hrsg.), *Wissen – Können – Profession. Ausgewählte Verhältnisbestimmungen* (S. 155–174). Innsbruck.
- Gruber, H., & Rehr, M. (2005) Praktikum statt Theorie? Eine Analyse relevanten Wissens zum Aufbau pädagogischer Handlungskompetenz. *Forschungsbericht Nr. 15*. Universität Regensburg.
- Hascher, T. (2011). Forschung zur Wirksamkeit der Lehrerbildung. In E. Terhart, H. Bennewitz, & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrberuf* (S. 384–406). Münster: Waxmann.
- Hopf, B. (2012). *Pädagogische Handlungskompetenz ohne pädagogische Ausbildung?* Hamburg: Dr. Kovac Verlag.
- Hopmann, St. T. (2008). Lehrer/innenbildung in internationaler Perspektive. *Journal für lehrerinnen- und lehrerbildung*, 8(3), 6–13.
- Hsieh, F. J., Law, C. K., Shy, H. Y., Wang, T. Y., Hsieh, C. J., & Tang, S. J. (2011). Mathematics teacher education quality in TEDS-M: Globalizing the views of future teachers and teacher educators. *Journal of Teacher Education*, 62(2), 172–187.
- Klieme, E. (2012). Internationales large scale assessment in der Lehrerbildung: Anmerkungen zu einem neuen Paradigma der vergleichenden Bildungsforschung. *Zeitschrift für Pädagogik*, 58, 492–499.
- KMK (2005) = Kultusministerkonferenz. (2005). *Eckpunkte für die gegenseitige Anerkennung von Bachelor- und Masterabschlüssen in Studiengängen, mit denen die Bildungsvoraussetzungen für ein Lehramt vermittelt werden*. Bonn: KMK.
- KMK (2009) = Kultusministerkonferenz. (2009). *Information des Sekretariats über die Regelungen des KMKBeschlusses vom 22.10.1999 'Gegenseitige Anerkennung von Lehramtsprüfungen und Lehramtsbefähigungen'* (Stand : 05.02.2009). Bonn: Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland.
- König, J. (2009). Zur Bildung von Kompetenzniveaus im Pädagogischen Wissen von Lehramtsstudierenden: Terminologie und Komplexität kognitiver Bearbeitungsprozesse als Anforderungsmerkmale von Testaufgaben? *Lehrerbildung auf dem Prüfstand*, 2(2), 244–262.
- König, J. (2012). Zum Einfluss der Schulpraxis im Lehramtsstudium auf den Erwerb von pädagogischem Wissen: Spielen erste Unterrichtsversuche eine Rolle? In T. Hascher & G. H. Neuweg (Hrsg.), *Forschung zur (Wirksamkeit der) LehrerInnenbildung* (S. 143–159). Wien: LIT-Verlag.
- König, J. (2013). First comes the theory, then the practice? On the acquisition of general pedagogical knowledge during initial teacher education. *International Journal of Science and Mathematics Education*, 11(4), 999–1028.
- König, J. (2014). Forschung zum Erwerb von pädagogischem Wissen angehegender Lehrkräfte in der Lehrerausbildung. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrberuf* (2. überarb. und erw. Auflage, S. 615–641). Münster: Waxmann.

- König, J., & Blömeke, S. (2009a). Pädagogisches Wissen von angehenden Lehrkräften: Erfassung und Struktur von Ergebnissen der fachübergreifenden Lehrerbildung. *Zeitschrift für Erziehungswissenschaft*, 12(3), 499–527.
- König, J., & Blömeke, S. (2009b). Pädagogisches Wissen von österreichischen Lehramtsstudierenden. *Erziehung & Unterricht*, 159(1/2), 175–186.
- König, J., & Blömeke, S. (2010). *Pädagogisches Unterrichtswissen (PUW). Dokumentation der Kurzfassung des TEDS-M-Testinstruments zur Kompetenzmessung in der ersten Phase der Lehrerbildung*. Berlin: Humboldt-Universität.
- König, J., & Rothland, M. (2013). Pädagogisches Wissen und berufsspezifische Motivation am Anfang der Lehrerbildung. Zum Verhältnis von kognitiven und nicht-kognitiven Eingangsmerkmalen von Lehramtsstudierenden. *Zeitschrift für Pädagogik*, 59(1), 43–65.
- König, J., & Seifert, A. (Hrsg.) (2012). *Lehramtsstudierende erwerben pädagogisches Professionswissen. Ergebnisse der Längsschnittstudie LEK zur Wirksamkeit der erziehungswissenschaftlichen Lehrerbildung*. Münster: Waxmann.
- König, J., Tachtsoglou, S., Darge, K., & Lünemann, M. (2014). Zur Nutzung von Praxis: Modellierung und Validierung lernprozessbezogener Tätigkeiten von angehenden Lehrkräften im Rahmen ihrer schulpraktischen Ausbildung. *Zeitschrift für Bildungsforschung*, 4(1), 3–22.
- Kunina-Habenicht, O., Schulze-Stocker, F., Kunter, M., Baumert, J., Leutner, 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 COACTIV*. Münster: Waxmann.
- Lohmann, V., Seidel, V., & Terhart, E. (2011). Bildungswissenschaften in der universitären Lehrerbildung: Curriculare Strukturen und Verbindlichkeiten. *Lehrerbildung auf dem Prüfstand*, 4(2), 271–302.
- Lüders, M. (2012). „Pädagogisches Wissen“ – Eine Testkritik. *Zeitschrift für Erziehungswissenschaft*, 15(4), 775–791.
- Lütke, O., Robitzsch, A., Trautwein, U., & Köller, O. (2007). Umgang mit fehlenden Werten in der psychologischen Forschung. Probleme und Lösungen. *Psychologische Rundschau*, 58, 103–117.
- Mayr, J., & Neuweg, G. H. (2009). Lehrer/innen als zentrale Ressource im Bildungssystem: Rekrutierung und Qualifizierung. In W. Specht (Hrsg.), *Nationaler Bildungsbericht Österreich 2009* (Bd. 2, S. 99–119). Graz: Leykam.
- Messner, R., & Posch, P. (2010). Lehrerbildung in Österreich: Die Reform sollte zu Ende geführt werden! *Erziehung & Unterricht*, 160(1–2), 177–197.
- Muthén, B. O., & Muthén, L. K. (1996–2006). *Mplus (Version 4.2)* [Computer software]. Los Angeles, CA: Muthén & Muthén.
- Neuweg, G. H. (1999). *Könnerschaft und implizites Wissen. Zur lehr-lerntheoretischen Bedeutung der Erkenntnis- und Wissenstheorie Michael Polanyis*. Münster: Waxmann.
- Neuweg, H. G. (2014). Das Wissen der Wissensvermittler. Problemstellungen, Befunde und Perspektiven der Forschung zum Lehrerwissen. In E. Terhart, H. Bennewitz, & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (S. 451–477). Münster: Waxmann Verlag.
- OECD (2005) = Organisation für wirtschaftliche Zusammenarbeit und Entwicklung. (2005). *Teachers matter. Attracting, developing and retaining effective teachers*. Paris: OECD.
- OECD (2009) = Organisation für wirtschaftliche Zusammenarbeit und Entwicklung. (2009). *Equally prepared for life? How 15-year-old boys and girls perform in school*. Paris: OECD.
- OECD (2009) = Organisation für wirtschaftliche Zusammenarbeit und Entwicklung. (2009). *Creating effective teaching and learning environments. First results from TALIS*. Paris: OECD.
- Oser, F., & Oelkers, J. (Hrsg.) (2001). *Die Wirksamkeit der Lehrerbildungssysteme*. Chur: Rüegger.
- Pflanzl, B., Thomas, A., & Matischek-Jauk, M. (2013). Pädagogisches Wissen und pädagogische Handlungskompetenz. *Erziehung und Unterricht*, 1/2, 1–8.
- Putnam, R. T. (1987). Structuring and adjusting content for students: A study of live and simulated tutoring of addition. *American Educational Research Journal*, 24(1), 13–48.
- Rost, J. (1996). *Lehrbuch Testtheorie, Testkonstruktion*. Bern: Huber.
- Sabers, D. S., Cushing, K. S., & Berliner, D. C. (1991). Differences among teachers in a task characterized by simultaneity, multidimensionality, and immediacy. *American Educational Research Journal*, 28, 63–88.

- Schaefers, C. (2002). Forschung zur Lehrerbildung in Deutschland—eine bilanzierende Übersicht der neueren empirischen Studien. *Schweizerische Zeitschrift für Bildungswissenschaften*, 24(1), 65–90.
- Schleicher, A. (2011). *Building a high-quality teaching profession. Lessons from around the world*. Paris: OECD.
- Schneider, R., & Wildt, J. (2009). Forschendes Lernen in Praxisstudien – Wechsel eines Leitmotivs. In: B. Roters, R. Schneider, B. Koch-Priewe, J. Thiele, & J. Wildt (Hrsg.), *Forschendes Lernen im Lehramtsstudium. Hochschuldidaktik – Professionalisierung Kompetenzentwicklung* (S. 8–36). Bad Heilbrunn: Klinkhardt.
- Schulte, K., Bögeholz, S., & Watermann, R. (2008). Selbstwirksamkeitserwartungen und Pädagogisches Professionswissen im Verlauf des Lehramtsstudiums. *Zeitschrift für Erziehungswissenschaft*, 11, 268–287.
- Shavelson, R. J. (2010). On the measurement of competency. *Empirical Research in Vocational Education and Training*, 2(1), 82–103.
- Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Research*, 57, 1–22.
- Stadelmann, W., & Rhyn, H. (2011). Zusammenarbeit zwischen Deutschland, Österreich und der Schweiz in Bildungsfragen am Beispiel Lehrerbildung. *Beiträge zur Lehrerbildung*, 29, 109–114.
- Tatto, M. T., Schwille, J., Senk, S., Ingvarson, L., Peck, R., & Rowley, G. (2008). *Teacher education and development study in Mathematics (TEDS-M): Conceptual framework*. East Lansing: Michigan State University.
- Tatto, M. T., Schwille, J., Senk, S., Ingvarson, L., Rowley, G., Peck, R., Bankov, K., Rodriguez, M., & Reckase, M. (2012). Policy, practice, and readiness to teach primary and secondary mathematics in 17 countries. Findings from the IEA teacher education and development study in mathematics (TEDS-M).
- Terhart, E. (1993). Pädagogisches Wissen. Überlegungen zu seiner Vielfalt, Funktion und sprachlichen Form am Beispiel des Lehrwissens. In J. Oelkers & H.-E. Tenorth (Hrsg.), *Pädagogisches Wissen* (S. 129–141). Weinheim: Beltz.
- Terhart, E. (2011). *Lehrerberuf und Professionalität: Gewandeltes Begriffsverständnis – neue Herausforderungen*. (Zeitschrift für Pädagogik: 57. Beiheft, S. 202–224).
- Terhart, E. (2012). Wie wirkt Lehrerbildung? *Zeitschrift für Bildungsforschung*, 2, 2–23.
- 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.
- Weinert, F. E. (2001). Concept of Competence: A Conceptual Clarification. In D. S. Rychen & L. H. Salganik (Hrsg.), *Defining and Selecting Key Competencies* (S. 45–56). Göttingen.
- Wu, M., & Adams, R. (2006). Modelling mathematics problem solving item responses using a multidimensional IRT model. *Mathematics Education Research Journal*, 18(2), 93–113.
- Wu, M. L., Adams, R. J., & Wilson, M. R. (1997). *ConQuest: Multi-aspect test software* [computer program]. Camberwell: Australian Council for Educational Research.