

Z Erziehungswiss (2015) 18:375–404
DOI 10.1007/s11618-015-0625-7

ALLGEMEINER TEIL

ZfE

Analyse von schriftlichen Unterrichtsplanungen: Empirische Befunde zur didaktischen Adaptivität als Aspekt der Planungskompetenz angehender Lehrkräfte

Johannes König · Christiane Buchholtz · Dieter Dohmen

Eingegangen: 8. September 2014 / Angenommen: 1. April 2015 / Online publiziert: 3. Mai 2015
© The Author(s) 2015. Dieser Artikel ist auf Springerlink.com mit Open Access verfügbar

Zusammenfassung Planungsentscheidungen von Expertenlehrpersonen sind untrennbar mit den spezifischen situativen Anforderungen, mit der Lerngruppe, verbunden. In dieser Untersuchung soll die Kontextsituierung als Ausgangspunkt aufgegriffen werden, indem ein ausgewählter Ausschnitt der in schriftlichen Unterrichtsplanungen von angehenden Lehrkräften realisierten Planungskompetenz theoriegeleitet als Konstrukt konzipiert und empirisch erfasst wird. Als zentrale generische Anforderung fokussieren wir die *didaktische Adaptivität*, d. h. die fachunspezifische Passung von kognitiven Voraussetzungen einer Lerngruppe und den Aufgaben für die geplante Unterrichtsstunde. Die Konzeption wird über Indikatoren (Analysekriterien) operationalisiert, um Planungsentscheidungen aus schriftlichen Unterrichtsplanungen zu rekonstruieren und zu quantifizieren. Datengrundlage bilden schriftliche Unterrichtsplanungen (Lehrproben) von 106 angehenden Lehr-

Zusatzmaterial online

Zusätzliche Information ist in der online Version dieses Artikels
(doi: [10.1007/s11618-015-0625-7](https://doi.org/10.1007/s11618-015-0625-7)) enthalten.

Prof. Dr. J. König (✉)
Humanwissenschaftliche Fakultät, Universität zu Köln,
Gronewaldstr. 2,
50931 Köln, Deutschland
E-Mail: johannes.koenig@uni-koeln.de

Dr. C. Buchholtz
Professional School of Education, Humboldt-Universität zu Berlin,
Hausvogteiplatz 5-7,
10099 Berlin, Deutschland
E-Mail: christiane.buchholtz@staff.hu-berlin.de

Dr. D. Dohmen
Forschungsinstitut für Bildungs- und Sozialökonomie,
Reinhardtstr. 31,
10117 Berlin, Deutschland
E-Mail: D.Dohmen@fibs.eu

kräften in Berlin zu zwei Zeitpunkten (Beginn und Ende der zweiten Ausbildungsphase bzw. Referendariat). Die Ergebnisse zeigen, dass unterschiedlich komplexe Analysekriterien – verwendet als Items bei einer Rasch-Skalierung – erwartungsgemäß unterschiedlich hohe Lösungshäufigkeiten aufweisen, dass die betrachtete Kompetenz reliabel gemessen werden kann und dass es wie erwartet im Verlauf der zweiten Ausbildungsphase zu einem Zuwachs an gemessener Planungskompetenz kommt. Interkorrelative Analysen mit Außenkriterien (Abiturnote, Note des Ersten Staatsexamens, professionelle Überzeugungen zum Lehren und Lernen) belegen die Konstruktvalidität zusätzlich. Eine kleine Analyse an einer zusätzlichen Stichprobe von 22 Referendaren aus der Ausbildungsregion Köln zum Zusammenhang zwischen Planungskompetenz und Schülerurteilen zur gehaltenen Unterrichtsstunde geben erste Hinweise auf prognostische Validität. Limitierungen des methodischen Vorgehens und Implikationen für die Gestaltung der Lehrerausbildung werden diskutiert.

Schlüsselwörter Adaptivität · Aufgabe · Didaktik · Lehrerbildung · Lehrerkompetenzen · Längsschnittstudie · Referendariat · Unterrichtsplanung

An analysis of written lesson plans to examine pedagogical adaptivity in the planning competences of trainee teachers

Abstract Planning decisions of expert teachers are inextricably linked to the specific teaching demands of the learning group. This study focuses on situational aspects through an examination of selected extracts of written lesson plans of trainee teachers, which are constructively conceptualized and then empirically measured. We highlight pedagogical adaptivity (i.e. the ways in which the cognitive requirements of the learning group fit with the assignments of the respective lesson) as a central generic task. This conceptualization is operationalized using several indicators (analysis criteria) to reconstruct and quantify planning decisions from written lesson plans. The data set comprises the written plans of demonstration lessons of 106 trainee teachers in Berlin at two points in time (the beginning and end of the second phase of initial teacher training). Findings confirm that differently complex analysis criteria used in a Rasch analysis show different item difficulty. Planning competence can be measured in a reliable way and, as expected, increases during the second phase of initial teacher training. In addition, inter-correlative analyses using external criteria (grades from second and third level education, professional beliefs about teaching and learning) prove the construct validity. An additional sample of 22 teacher candidates from the Cologne region, used to analyze the relation between planning competence and student ratings on the lesson taught, provide some initial insights into the measure's predictive validity. Limitations of the methodological approach and implications for teacher education are discussed.

Keywords Adaptivity · Didactics · Lesson planning · Longitudinal study · Task · Teacher competence · Teacher education · Teacher induction

1 Einleitung

Die Planung von Unterricht ist zentraler Bestandteil der beruflichen Aufgaben und Tätigkeiten von Lehrkräften (z. B. Baumert und Kunter 2006). Sie ist insofern auch zentraler Bestandteil der Lehrerausbildung, weil angehende Lehrkräfte in dieser Tätigkeit wissenschaftlich ausgebildet und in der Praxis geschult werden sollen (KMK 2004). Unterricht planen zu können ist grundsätzlich lehr- und lernbar. Indem (zukünftige) Lehrpersonen Unterricht planen, kann sich ihre Fähigkeit (weiter-)entwickeln, zunächst in Gedanken Handlungsschemata zu erstellen, bevor gehandelt wird. Die Fähigkeit zu planen wird als „Kernfähigkeit handlungsfähiger Menschen“ betrachtet (Peterßen 2003, S. 60).

Die empirische Untersuchung und Analyse von Planungskompetenz ist bislang nur vereinzelt angegangen worden (Bromme 1981; Haas 1998; Jacobs et al. 2008). Vorliegende Studien zeichnen zumeist die Entwicklung von Planungskompetenz nach, zugleich zeigt sich aber, dass kein geteiltes Konzept dieser Kompetenz in der Literatur erkennbar ist und unterschiedliche forschungsmethodische Zugänge nebeneinander stehen. Ansätze zur standardisierten Erfassung liegen unseres Wissens nicht vor. Dies überrascht, da Forschung zur Unterrichtsplanung einen direkten Blick auf die kognitiven Aktivitäten von Lehrkräften bietet und die tatsächliche Organisation und Strukturierung des Unterrichts durch schriftliche Unterrichtsplanungen vorhergesagt werden können (Clark und Peterson 1984; Seel 1997). Zwar werden aktuell in standardisierten Tests zur Erfassung von pädagogischem oder fachdidaktischem Wissen Testaufgaben eingesetzt, bei denen angehende bzw. berufstätige Lehrkräfte ihr Wissen zu Planungsaspekten unter Beweis stellen sollen (z. B. König und Blömeke 2009, 2010; Döhrmann et al. 2010; Baer et al. 2011). Diesen Ansätzen ist jedoch gemein, dass sie von den Befragten nicht verlangen, ihr Wissen auf eine konkrete, reale Lerngruppe zu beziehen. Planungsentscheidungen, sofern sie bei diesen Ansätzen überhaupt Gegenstand der Messung sind, werden also weitgehend dekontextualisiert sowie entkoppelt von den spezifischen situativen Anforderungen erfasst. Der Einbezug situativer Bedingungen in die unterrichtliche Planung ist jedoch eine grundlegende Anforderung des Lehrerberufs und damit ein substanzieller Bestandteil von Planungskompetenz.

Dies manifestiert sich beispielhaft in der *didaktischen Adaptivität*, wie sie als Aspekt der Planungskompetenz von angehenden Lehrkräften in dieser Untersuchung zum Gegenstand gemacht werden soll. Gemeint ist ein fachunspezifischer Aspekt der Planungskompetenz – die fachunspezifische (also didaktische und nicht fachdidaktische) Fähigkeit, einen adaptiven Unterricht zu gestalten, indem Vorwissensunterschiede und kognitive Heterogenität der Lerngruppe bei der Unterrichtsplanung explizit berücksichtigt werden. Die zu leistende Adaptivität kann sich auf verschiedene Momente der Planung beziehen, etwa auf die verschiedenen Entscheidungsfelder der Unterrichtsplanung wie der Intention, der Thematik, der Methodik und der Medien (Heimann 1965; Heimann et al. 1965). In dieser Untersuchung fokussieren wir jedoch auf die didaktische Adaptivität von kognitiven Voraussetzungen einer Lerngruppe und den Aufgaben, welche die Lernenden in der geplanten Unterrichtsstunde bearbeiten sollen. Dabei wird die Aufgabe als Integrationsbegriff der Unterrichtsplanung verstanden (ausführlich Abschn. 2.2), aber auch rein zeitlich betrachtet

ist die Bearbeitung von Aufgaben durch Schülerinnen und Schüler wesentlicher Bestandteil von Unterricht (vgl. z. B. Hiebert et al. 2003 für den Mathematikunterricht). Da die Frage nach der so definierten Adaptivität bei der Planung typischerweise und – auf einer allgemeinen Ebene – unabhängig vom Schulfach und der Schulform in den Vordergrund rückt, betrachten wir sie als zentrale generische Anforderung. Hierzu präsentieren wir eine theoriegeleitete Konzeption. Diese operationalisieren wir anschließend über Indikatoren (Analysekriterien), um Planungsentscheidungen aus schriftlichen Unterrichtsplanungen zu rekonstruieren und zu quantifizieren. Auf der Basis des Datenmaterials von schriftlichen Unterrichtsplanungen (Lehrproben) angehender Lehrkräfte zu zwei Zeitpunkten – Beginn und Ende der zweiten Ausbildungsphase (Vorbereitungsdienst bzw. Referendariat) – wird empirisch geprüft, ob die Indikatoren eine latente Variable „didaktische Adaptivität als Aspekt der Planungskompetenz“ messen. Wir vermuten, dass sich unterschiedlich komplexe Analysekriterien bei einer Rasch-Skalierung über unterschiedlich hohe Lösungshäufigkeiten der verwendeten Indikatoren niederschlagen und dass es im Verlauf der zweiten Ausbildungsphase zu einem Zuwachs an gemessener Planungskompetenz kommt. Ferner vermuten wir, dass sich die so gemessene Kompetenz mit relevanten Außenkriterien validieren lässt.

2 Theoretischer Rahmen der Untersuchung von Planungskompetenz

Angesichts der Komplexität einer Messung von möglichen Kompetenzaspekten in Hinblick auf die Planung von Unterricht und aus forschungsökonomischen Gründen bildet lediglich eine zentrale Herausforderung bzw. Problemstellung den Ausgangspunkt unserer Studie: die *didaktische Adaptivität*. Von dieser nehmen wir jedoch an, dass sie typischerweise im Planungsprozess auftritt, und zwar als ein Problem, das gelöst werden muss (vgl. Yinger 1977; Bromme 1981), so dass der Begriff der „Kompetenz“ im Sinne der aktuellen Verwendung in der empirischen Bildungsforschung (Klieme und Leutner 2006, S. 879) gerechtfertigt ist, der zu folge es sich um „kontextspezifische kognitive Leistungsdispositionen“ handelt, „die sich funktional auf Situationen und Anforderungen in bestimmten Domänen beziehen“. Die hier im Mittelpunkt stehende Messung von didaktischer Adaptivität als Aspekt der Planungskompetenz angehender Lehrkräfte ist demnach eine kognitive Leistungsdisposition, die im Kontext der Planung von Unterricht vonnöten ist. Sie bezieht sich auf den Umgang mit der spezifischen, zu unterrichtenden Lerngruppe, deren Leistungsheterogenität bei der Planung von Unterricht angemessen berücksichtigt werden muss, um der Anforderung eines qualitätsvollen Unterrichts nachzukommen. Diese kognitive Leistungsdisposition schließt situationales und konzeptuelles Wissen ein (vgl. De Jong und Ferguson-Hessler 1996): situationales Wissen um die Lerngruppe und ihre kognitiven Lernvoraussetzungen, die entscheidend für die zu planende Unterrichtssituation sind; konzeptuelles Wissen über didaktische Maßnahmen und Strategien, die einen adaptiven Unterricht ermöglichen. Dies wird im Folgenden genauer erläutert und für die vorliegende Studie spezifiziert.

Grundsätzlich wird Unterricht für eine ganz bestimmte Lerngruppe geplant, wobei die Lerngruppe die situativen Bedingungen des Unterrichtens maßgeblich

determiniert. Im Mittelpunkt des Unterrichts steht, konstruktivistischen Überlegungen folgend, die eigenständige Auseinandersetzung der Lernenden mit dem Gegenstand. Dieser wird von Lehrkräften in der Regel über die Material- und Sachfrage bedacht, manifestiert sich in „Aufgaben“ bzw. „Aufgabenstellungen“ (wie wir definitorisch nachfolgend noch ausführen werden, vgl. Abschn. 2.2) und wird mit Zielstellungen des Unterrichts funktional verknüpft. Die Aufgabe ist somit ein zentrales didaktisches Gestaltungsmittel für das von den Lehrkräften intendierte Lernhandeln (Bromme 1981; Shavelson 1987), das besonders die sogenannte „Erarbeitungsphase“ des Unterrichts bestimmt. Vor- und Nachbereitung dieser Phase (einleitendes Lehrer-Schüler-Gespräch oder einleitender Lehrervortrag, anschließender Transfer oder anschließende Plenumsdiskussion usw.) stehen in der Funktion, die Erarbeitungsphase vorzubereiten, weiterzuführen bzw. generell zu stützen und zu rahmen (Hill et al. 1983). Wir fokussieren daher im Folgenden auf zwei Bereiche der Unterrichtsplanung: die Lerngruppe als wesentliche Größe situativer Bedingungen für die einzelne Unterrichtsstunde und die Aufgabe als die zentrale Einheit der Unterrichtsplanung, in der ein Großteil der von der Lehrkraft für die einzelne Unterrichtsstunde getroffenen Planungsentscheidungen zusammenläuft und sich entfaltet.

2.1 Zur Bedeutung der Lernvoraussetzungen einer Lerngruppe bei der Unterrichtsplanung

Aus allgemeindidaktischer Perspektive sind bei der *Planung von Unterricht* Bedingungen zu bedenken und Entscheidungen zu treffen (Heimann 1965; Heimann et al. 1965). Modelle der Allgemeinen Didaktik zur Unterrichtsplanung enthalten unter anderem die Elemente Schülervoraussetzungen, Inhalte und Ziel, welche von der planenden Lehrperson im Zuge der Planung von Unterricht verknüpft und geordnet werden sollen (z. B. Blankertz 1969; Jank und Meyer 2002; Tulodziecki et al. 2004; Plöger 2008). Als leitend für diese Kategorien der Unterrichtsplanung werden überfachliche allgemeindidaktische Modelle angesehen, während fachdidaktische Spezifika sich erst auf konkreteren Analyseebenen entfalten (Arnold und Koch-Priewe 2010).

Aus der Expertiseforschung ist bekannt, dass Expertenlehrpersonen prozessorientiert planen, indem sie das Vorwissen ihrer Schülerinnen und Schüler und die kognitiven Anforderungen der gewählten Aufgaben für den Unterricht stringent aufeinander beziehen. So untersuchte Westermann (1991, S. 292) Entscheidungsprozesse bei der Unterrichtsplanung im Rahmen eines Experten-Novizen-Vergleichs und konnte feststellen: „The expert teachers thought about learning from the perspective of the student and performed a cognitive analysis of each learning task during planning, which adapted to the needs of students during teaching. In contrast, novices used specific lesson objectives to form structured lesson plans that they did not adapt to meet student needs during teaching.“ Das Bewusstsein der Lehrperson für die Lernenden und das Einbeziehen der Lernenden in den Planungsprozess stellt also ein wesentliches Unterscheidungsmerkmal zwischen Experten und Novizen dar (neben Westermann 1991, vgl. auch Housner und Griffey 1985; Borko und Livingston 1989; Kagan 1992; Berliner 2004). Expertenlehrkräfte fokussieren stärker auf ihre Schülerinnen und Schüler. Sie nehmen Schülervoraussetzungen als Schlüsselement für ihren Unter-

richt wahr (Smith und Strahan 2001). Ihr Planungshandeln zeichnet sich durch eine gelungene Wissensintegration von situationalem und konzeptuellem Wissen aus (vgl. De Jong und Ferguson-Hessler 1996), indem die bedeutsamen situativen Merkmale erfasst und zunehmend „intuitiv“ mit darauf abgestimmten unterrichtlichen Lehr- und Lernaktivitäten zusammengeführt werden (Bromme 1992; Cook 1992; Ericsson 1996).

2.2 Zur Bedeutung der Aufgabe bei der Unterrichtsplanung

In der Allgemeinen Didaktik ist der *Aufgabe* (*learning task*) lange Zeit nur wenig Beachtung geschenkt geworden. Erst im Zuge des handlungs- und aufgabenorientierten Unterrichts gewann sie wieder an Bedeutung sowie in den letzten Jahren als Diagnoseinstrument im Zusammenhang mit Schulleistungsstudien (Blömeke et al. 2006; Blumschein 2014). Aus allgemeindidaktischer Sicht wird sie vor allem als methodisches Gestaltungsmittel eines schülerorientierten Unterrichts begriffen, d. h. der Blick auf Aufgaben ist eher methodisch-organisatorischer Natur, während „eine vertiefte Auseinandersetzung mit der Inhalts- und Lernprozessbedeutung von Aufgaben“ nur vereinzelt stattfindet (Kleinknecht et al. 2011, S. 62). Der von uns verwendete Aufgabenbegriff grenzt sich von diesem in der Didaktik verbreiteten Begriff ab, indem wir der Definition von Aufgaben als „einer Aufforderung zum Lern-Handeln“ folgen (z. B. Bruder 2008, S. 19) und dabei besonders das Potenzial der Aufgabe in den Blick nehmen, ein in kognitiver, motivationaler und kommunikativer Hinsicht vielfältiges Lernhandeln anzuregen. Die Aufgabe verstehen wir als Integrationsbegriff (vgl. Zahorik 1975; Yinger 1977), der mindestens die allgemeindidaktische Inhalts- und Zielfrage bündelt und als „Lernaufgabe“ zur Weiterentwicklung von Wissen und Kompetenzen intendiert ist (Leisen 2006; Kleinknecht et al. 2011; Schmit et al. 2014). Aufgaben in diesem Sinne sind fast ausnahmslos komplex, insofern als dass ihre Lösung zumeist mehrere Wissensarten und verschiedene kognitive Fertigkeiten erfordert (z. B. Bremerich-Vos 2008). Klassifikationssysteme und Analysemodelle von Aufgaben, die sowohl aus überfachlicher Sicht als auch aus fachdidaktischer Perspektive vorliegen, entfalten differenziert die einzelnen Anforderungsdimensionen von Aufgaben im wissensbezogenen, kognitiven und motivationalen Bereich (Anderson und Krathwohl 2001; Neubrand 2002; Johnson 2003; Blömeke et al. 2006; Maier et al. 2010). Zugleich bilden sie das Spektrum zur Variation des Schwierigkeitsgrades der Aufgabe in mehreren Dimensionen ab, um sie auf die bestehenden Lernvoraussetzungen abzustimmen.

Empirische Studien zeigen, dass die Aufgabe bei Lehrkräften den Fokus der Unterrichtsplanung und -gestaltung bildet. So konnte z. B. Zahorik (1975) in einer Befragung von 194 Lehrkräften nach ihren Planungsentscheidungen feststellen, dass sich Lehrpersonen mit Abstand am häufigsten über die Aufgaben Gedanken machen, während die Entscheidung, welche am häufigsten als erstes getroffen wird, die Inhalte betrifft, gefolgt von der Zielfrage als zweite Entscheidung. Die Dominanz der Inhalte im Planungsprozess und die nachgeordneten Überlegungen zur den Unterrichtszielen konnte Seel (1997) gleichermaßen für angehende Lehrkräfte finden. Auch Peterson et al. (1978) kamen über die Transskripte vom lauten Planungsdenken von Lehrkräften zu Ergebnissen, die jene von Zahorik (1975) stützen: Den größten Anteil ihrer Pla-

nungszeit verbrachten die Lehrkräfte mit den zu unterrichtenden Inhalten, daraufhin mit den Schüleraktivitäten und erst zum Schluss mit der Zielfrage. Allerdings gilt es bei der Studie von Peterson et al. (1978) zu berücksichtigen, dass die Probanden die zu unterrichtenden Schüler nicht kannten, was erklären könnte, warum speziell in dieser Studie die Frage der Schülervoraussetzungen unberücksichtigt blieb.

Im Folgenden sei „Aufgabe“ über die Aufgabe bzw. Aufgabenstellung definiert, die die Schülerinnen und Schüler in der geplanten Phase der „Erarbeitung“ erhalten. Es handelt sich also um den in der Unterrichtsplanung formulierten Arbeitsauftrag, mit dem sich die Lernenden in der Phase der Erarbeitung (i. d. R. Stillarbeit, Partner- oder Gruppenarbeit) auseinandersetzen. In einer schriftlichen Unterrichtsplanung liegt die Aufgabe in der Regel in Form eines Aufgabenblattes bzw. in Form von mehreren Aufgabenblättern (ggf. mit Material) vor. Unter *Aufgabenstellung* sind damit auch die unmittelbaren Arbeitsaufträge für die Schülerinnen und Schüler gemeint, welche ihre Lernaktivitäten bzw. den Hauptteil ihrer Lernaktivitäten in der Unterrichtsstunde steuern sollen.

2.3 Didaktische Adaptivität als zentrale Anforderung für Lehrkräfte

Unbenommen weiterer Anforderungen, welche bei der Planung von Unterricht eine Rolle spielen können und von der Lehrkraft erfolgreich bewältigt werden müssen, nehmen wir an, dass Lehrpersonen der „Planungssituation“ nicht entkommen können, in der sie das Vorwissen ihrer Schülerinnen und Schüler und die kognitiven Anforderungen der gewählten Aufgaben für den Unterricht aufeinander beziehen. In der Gestaltung von kompetenzorientiertem Unterricht besitzt die Aufgabe dafür eine „Schlüsselfunktion“ (Müller et al. 2013, S. 130), da sie in besonderem Maße das Potential bietet, in unterschiedlicher Weise an verschiedene Motivlagen und kognitive Kompetenzstände der Schülerinnen und Schüler anzuschließen. Aufgaben können hinsichtlich ihres kognitiven Aktivierungsgrades, der instruktionalen Vorgabe und des Umfangs- und Schwierigkeitsgrades abgestuft sowie in inhaltlicher Hinsicht variiert werden. In der Verbindung von Lerngruppe und Aufgabe eröffnet sich damit ein besonderer Gestaltungsraum didaktischer Adaptivität, sodass sich an dieser Schnittstelle Planungskompetenz untersuchen lassen dürfte. Zur Operationalisierung der Kompetenz, wie sich qualitativ gute Planungsentscheidungen in der Bereitstellung von Aufgaben für die spezifische Lerngruppe niederschlagen, eignet sich aus unserer Sicht die Frage, inwieweit *heterogene Lernvoraussetzungen* wahrgenommen und bei der Unterrichtsplanung berücksichtigt werden. Derartige Adaptivität im Unterricht, also ein „intelligenter Umgang mit Heterogenität durch Differenzierung von Zielsetzungen, Individualisierung von Aufgabenstellungen und Variation von Methoden und Sozialformen“ (Baumert et al. 2004, S. 317), gilt unstrittig als zentrale Dimension der Unterrichtsqualität (z. B. Helmke 2010). Ebenso wirkt sich sowohl aus motivationstheoretischer als auch lerntheoretischer Sicht ein individuell angemessenes Schwierigkeitsniveau der Aufgabe förderlich auf die Leistungsmotivation und das Leistungsergebnis aus (Wygotski 1978; Heckhausen 1989). Folglich benötigen Lehrkräfte heutzutage professionelle Kompetenzen, um auf eine heterogene Schülerklientel angemessen einzugehen (Horstkemper 2004; Helmke 2010), und zwar in der Grundschule ebenso wie im gegliederten Sekundarstufenschulwesen.

Dabei geht es um eine – auf einzelne Schüler oder auf Gruppen von Schülern bezogene – Individualisierung der Lehr-Lernprozesse im Unterricht als zentrale Aufgabe von Lehrpersonen. Individualisierung wird durch Maßnahmen innerer Differenzierung (Binnendifferenzierung) ermöglicht, die sich auf unterschiedliche Methoden, Lerninhalte, Lernmaterialien, unterschiedliche Lernzielniveaus und Techniken der Motivierung von Schülern beziehen können (Schwarzer und Steinhagen 1975; Weinert 1997; Bönsch 2004; Tulodziecki et al. 2004). Im Allgemeinen werden hinsichtlich der „Heterogenität“ von Schülerinnen und Schülern unterschiedliche Kriterien diskutiert (u. a. Horstkemper 2004; Bönsch 2004; Fend 2008). Gleichwohl abstrahiert die Diskussion vielfach auf Lern- und Leistungsunterschieden unter Schülern im Klassenverband, da dieses Kriterium für Lehrkräfte bei der Planung, Durchführung und Auswertung von Unterricht wie auch der Beurteilung und Beratung von zentraler Bedeutung ist und außerdem das Kriterium der Lern-/Leistungsunterschieden teilweise mit anderen Kriterien der Heterogenität – wie z. B. der ethnischen – kovariieren kann (vgl. Stanat und Christensen 2006).

Auf Lern- und Leistungsunterschieden von Schülerinnen und Schülern können Lehrkräfte sehr unterschiedlich reagieren (vgl. Weinert 1997; Helmke 2010): Werden die Lern- und Leistungsunterschiede ignoriert (passive Reaktionsform) und sich an einem Durchschnittsniveau orientiert, kann dies die Leistungsunterschiede vergrößern. Durch schulorganisatorische Maßnahmen zur Homogenisierung von Schulklassen (Exklusion, äußere Differenzierung, Zurückstellung vom Schulbesuch, Klassenwiederholung) erfolgt eine Anpassung der Schüler an den Unterricht (substitutive Reaktionsform). Der Unterricht kann hingegen auch an lernrelevante Unterschiede zwischen den Schülern angepasst werden (aktive Reaktionsform) oder es können sogar einzelne Schüler durch eine adaptive Gestaltung des Unterrichts gefördert werden (proaktive Reaktionsform), indem etwa differenzielle Lernziele gesetzt werden. Neben empirischen Befunden (z. B. Hattie 2009), die zeigen, dass sich differenzierte Lernangebote direkt leistungsfördernd auswirken können, da mehr Lerner ihr individuelles Potential entfalten können, zeigen empirische Studien, dass differenzierte Lernangebote auch mittelbar auf Leistung wirken können, z. B. über ein verbessertes Leistungsfähigkeitskonzept (Lipowsky et al. 2011), oder indem sie das Interesse am Fach befördern (Fauth et al. 2014).

Wir werden nachfolgend diesen Bereich von Unterrichtsqualität bei der Analyse von schriftlichen Unterrichtsplanungen fokussieren und über seine Rekonstruktion auf Planungskompetenz bei angehenden Lehrkräften schließen.

3 Fragestellungen

Gegenstand unserer Untersuchung ist die standardisierte Messung eines relevanten Aspektes von Planungskompetenz angehender Lehrkräfte über die Analyse von schriftlichen Unterrichtsplanungen. Im Einzelnen wollen wir folgende Fragestellungen beantworten:

1. Lässt sich die didaktische Adaptivität als Aspekt von Planungskompetenz, operationalisiert über Indikatoren (Analysekriterien), skalieren und damit als reliable latente Fähigkeitsvariable modellieren?

Wir vermuten, dass die Analyse von schriftlichen Unterrichtsplanungen anhand von spezifischen Indikatoren eine Rekonstruktion von Planungskompetenz bei angehenden Lehrkräften erlaubt und damit Aussagen zu ihrem situationalen und konzeptuellen Wissen (vgl. De Jong und Ferguson-Hessler 1996), das bei der betrachteten Planungsanforderung relevant ist, getroffen werden können. Da die hier untersuchte Kompetenz über unterschiedlich schwierige Analysekriterien operationalisiert wird (z. B. einfache Beschreibung der Lerngruppe vs. komplexe Verknüpfungen von situationalem und konzeptuellem Planungswissen, s. Abschn. 4.2), nehmen wir an, dass sich Fähigkeitswerte und Item-Schwierigkeitsparameter im Rahmen einer IRT-Skalierung auf einem Kontinuum abbilden lassen (vgl. Rost 1996).

1. Lässt sich über die Messung zu zwei Zeitpunkten in der zweiten Ausbildungsphase (Vorbereitungsdienst bzw. Referendariat) bei angehenden Lehrkräften eine Zunahme ihrer Planungskompetenz abbilden?

Da die Planung, Durchführung und Analyse von Unterricht zentraler Lehr- und Lerngegenstand der zweiten Phase der hiesigen Lehrerausbildung ist und wir davon ausgehen, dass angehende Lehrkräfte einschlägige Lerngelegenheiten zum Erwerb von Unterrichtsexpertise erhalten, vermuten wir, dass die Analysekriterien in den Planungen am Ende des Referendariats in höherem Maße erfüllt sind als zu Beginn.

1. Kann das Konstrukt über Außenkriterien validiert werden?

Da die Veränderung der Personenmittelwerte über die Zeit (Fragestellung 2) Einblick in die *curriculare Validität* erlaubt, jedoch kein zwingendes inhaltliches Validitätsargument ist, sind weitere Belege zur Konstruktvalidität zu erbringen. Zur Beantwortung der Frage verwenden wir zum einen Indikatoren für kognitive Grundfertigkeiten (Abiturnote) und Studienerfolg (Gesamtnote des Ersten Staatsexamens). Zum anderen verwenden wir als professionelles Personenmerkmal pädagogische Überzeugungen zum Lehren und Lernen. Die Betrachtung von Aufgaben als Anregung zur eigenständigen Auseinandersetzung der Lernenden mit dem Unterrichtsgegenstand sowie zur Aufforderung zum selbsttätigen Lern-Handeln (s. Abschn. 2.2) dürften im Zusammenhang mit entsprechenden Überzeugungen der angehenden Lehrkräfte stehen. Unter den pädagogischen Überzeugungen (*beliefs*) von Lehrkräften ist die zweidimensionale Erfassung einer transmissionsorientierten und einer konstruktivistisch orientierten Überzeugung weit verbreitet (König 2012), sodass diese genutzt wird. Konstruktivistisch orientierte Lehrpersonen dürften eher geneigt sein, binnendifferenzierenden Unterricht zu planen, während transmissionsorientierte Lehrpersonen vermutlich auf ein einheitliches Aufgabenangebot setzen, das weniger auf Individualisierung zielt. Wir vermuten, dass im Sinne einer *diskriminanten Validität* Korrelationen zwischen Außenkriterien und Planungskompetenz niedrig ausfallen ($|r| < 0,3$). Im Sinne einer *konvergenten Validität* sollten jedoch statistisch signifikante Zusammenhänge nachweisbar sein, wobei Noten und die Transmissionsorientierung negativ sowie die konstruktivistische Orientierung positiv mit der Planungskompetenz korreliert sein sollten.

Ferner vermuten wir, dass im Sinne einer *prognostischen Validität* die hier untersuchte Planungskompetenz nicht unabhängig von dem tatsächlich umgesetzten Unterricht ist. Mithilfe einer Analyse an einer kleinen zusätzlichen Stichprobe gehen wir daher der Frage nach, ob sich die bei Referendaren gemessene didaktische Adaptivität als Aspekt der Planungskompetenz in Urteilen der von ihnen unterrichteten Schülerinnen und Schüler über binnendifferenzierende Maßnahmen als Qualitätsaspekt von Unterricht niederschlägt. Angesichts bisheriger Erkenntnisse zum Zusammenhang von Planung und Durchführung von Unterricht, etwa dass die tatsächliche Organisation und Strukturierung des Unterrichts durch schriftliche Unterrichtsplanungen vorhergesagt werden können (vgl. Abschn. 1), vermuten wir eine positive Korrelation mit praktischer Bedeutsamkeit zwischen Kompetenz und Unterrichtsqualität.

4 Untersuchungsmethode

4.1 Stichproben

Für die Studie wurden 211 Unterrichtsentwürfe von 106 Lehramtsanwärterinnen und Lehramtsanwärttern analysiert, je ein zu Beginn und ein zum Ende des Vorbereitungsdienstes erstellter Entwurf (eine Person reichte nur einen Entwurf ein). Die Entwürfe umfassen Unterrichtsplanungen in insgesamt 19 Fächern für die Klassen 1 bis 13 verschiedener Schulformen. Die Unterrichtsentwürfe wurden im Kontext der *Evaluation des Berliner Vorbereitungsdienstes* durch das *Forschungsinstitut für Bildungs- und Sozialökonomie* gesammelt (vgl. Buchholz et al. 2012) und uns für die vorliegende Analyse zur Verfügung gestellt. Die Lehramtsanwärterinnen und Lehramtsanwärter, die die Entwürfe verfassten, absolvierten ihren Vorbereitungsdienst von 2010 bis 2012 an acht verschiedenen Berliner Seminarstandorten. Von den 22 Schulpraktischen Seminaren, die für die Evaluationsstudie ausgewählt worden waren, wurden diese acht für die vorliegende Studie zur Planungskompetenz herangezogen. Die Auswahl erfolgte einerseits über die Form der Ausbildung: Ausgewählt wurden Seminare, die Lehramtsanwärterinnen und Lehramtsanwärter noch in einem zweijährigen Vorbereitungsdienst ausbildeten. Andererseits wurden aus Gründen des Feldzugangs nur solche Seminargruppen einbezogen, die mehr als zwei Personen unserer Stichprobe enthielten. Die Rücklaufquote der auf diese Weise einbezogenen angehenden Lehrkräfte beträgt 68 %. Dabei verteilen sich die Rückläufe gleichmäßig über die acht Seminargruppen, so dass wir nicht von einem Selektionseffekt durch Standorte ausgehen.

Laut gültiger Werte¹ sind 64 % Personen der Stichprobe weiblich und 36 % männlich. Zu Beginn der zweiten Ausbildungsphase waren 10 % der Probanden 25 Jahre oder jünger, 62 % zwischen 26 und 30 Jahre alt, 18 % zwischen 31 und 35 und 10 % 36 Jahre oder älter. Die Abiturnote variiert zwischen 1,0 und 3,7 ($M=2,3$, $SD=0,6$), die Note des Ersten Staatsexamens zwischen 1,0 und 3,6 ($M=2,0$, $SD=0,6$). 29 % der Probanden strebten das Grundschullehramt, 11 % das kombinierte Lehramt für Primarstufe/Sekundarstufe I und 60 % das Gymnasiallehramt an.

Da in der Studie *Evaluation des Berliner Vorbereitungsdienstes* keine Maße zur Qualität des von den Referendaren gehaltenen Unterrichts erhoben wurden, können

mit diesem Datensatz keine Hinweise auf prognostische Validität der gemessenen Kompetenz gegeben werden (Fragestellung 3). Aus diesem Grund greifen wir auf eine zusätzliche, jedoch kleine Stichprobe von 22 Referendaren der Ausbildungsregion Köln zurück, die im Schuljahr 2013/2014 an der von der DFG-geförderten Studie *Längsschnittliche Erhebung pädagogischer Kompetenzen von Lehramtsstudierenden und ReferendarInnen* (LEK-R, DFG-Gz. KO3947/3–2) teilgenommen haben. Aus LEK-R liegen für diese 22 angehenden Lehrkräfte schriftliche Unterrichtsplanungen und Beurteilungen der von ihnen unterrichteten Schülerinnen und Schüler in Bezug auf die gehaltene Unterrichtsstunde vor, welche korrelative Analysen zur Planungskompetenz und Unterrichtsqualität ermöglichen. Die Referendarinnen und Referendare wurden jeweils zu Beginn ihrer zweiten Phase untersucht. 16 sind weiblich (73 %), 6 männlich. Im Durchschnitt waren sie 27 Jahre alt ($SD=3,4$). Vier Referendare strebten das Grundschullehramt (18 %), acht das Haupt-/Realschullehramt (36 %), sieben ein Lehramt für Gymnasien (32 %) und drei das sonderpädagogische Lehramt (14 %) an. Der Schülerdatensatz der 22 Referendare besteht aus 449 Schülerinnen und Schülern (54,3 % weiblich), im Durchschnitt sind die befragten Schülerinnen und Schüler knapp 12 Jahre alt ($SD=2,41$). Pro Klasse wurden im Durchschnitt rund 20 Schülerinnen und Schüler befragt ($SD=6,83$).

4.2 Instrument zur Messung von Planungskompetenz

Um die in unserer Studie fokussierte Anforderung der angehenden Lehrkräfte, situative Bedingungen des Unterrichts mit den zu treffenden Entscheidungen über Aufgabe und Zieldimension in Verbindung zu bringen, zu operationalisieren, entwickelten wir verschiedene Kriterien, die erfassen, inwiefern das Aufgabenangebot in der Erarbeitungsphase ihrer Unterrichtsstunde auf das Vorwissen der Lerngruppe abgestimmt wird. Für die Analyse der schriftlichen Unterrichtsplanungen nutzten wir insgesamt elf Kriterien (vgl. Tab. 1). Sechs Kriterien dienen der Erfassung deskriptiver Aspekte und ihrer Qualität, weitere fünf Items erfassen komplexere Aspekte der Anwendung bzw. Verknüpfung. Wenn wir das jeweilige Kriterium erfüllt sahen, wurde ein Punkt vergeben.

Die Zusammenstellung dieser Items erfolgte aufgrund theoretischer Überlegungen zu relevanten Kriterien der Passung von Lerngruppe und Lernangebot. Sie bilden damit einen bestimmten Ausschnitt von Planungskompetenz ab, ohne dass der Anspruch erhoben wird, Planungskompetenz vollständig zu erfassen (vgl. auch unsere abschließende Diskussion). Im Folgenden soll es darum gehen, diese Kategorien mit typischen Beispielen aus den analysierten Entwürfen zu belegen und damit ihre Relevanz zu demonstrieren. Das zum vorliegenden Beitrag zusätzliche Online-Material enthält ferner die vollständige Dokumentation von zwei schriftlichen Unterrichtsplanungen, von denen die eine alle elf Kriterien (zusätzliches Online-Material, Abb. 2), die andere lediglich drei der elf Kriterien erfüllt (zusätzliches Online-Material, Abb. 3). Diese Planungen stehen somit stellvertretend für den oberen und den unteren Bereich der Leistungsverteilung unseres noch darzustellenden Gesamtindikators zur Planungskompetenz und ermöglichen im Rahmen dieses Artikels eine ganzheitliche Betrachtung der auf diese Planungen angewendeten Analysekatoren. Auf beide Planungen nehmen wir nachfolgend Bezug.

Tab. 1 Kriterien zur Analyse von schriftlichen Unterrichtsplanungen

Deskription		Anwendung
<i>Beschreibung situativer Bedingungen (Lerngruppe)</i>		<i>Anwendung erbrachter Beschreibungen auf die Situation</i>
10	Die Lerngruppe wird beschrieben	31 Bisheriger spezifischer Lernstand der Schüler (<i>gesamte Lerngruppe</i>) hinsichtlich der für die Unterrichtsstunde gewählten Aufgaben wird beschrieben – im Sinne des „nächsten Lernschrittes“
11	Unterschiede in der Lerngruppe hinsichtlich kognitiver Voraussetzungen werden beschrieben	32 Bisheriger spezifischer Lernstand der Schüler (<i>Differenzierung auf einzelne Schüler oder Schülergruppen</i>) hinsichtlich der für die Unterrichtsstunde gewählten Aufgaben wird beschrieben – im Sinne des jeweiligen „nächsten Lernschrittes“
12	Unterschiede in der Lerngruppe hinsichtlich motivationaler Voraussetzungen werden beschrieben	
<i>Beschreibung der Aufgabe</i>		<i>Verknüpfung von Elementen der Planung</i>
20	Die Aufgabenstellung wird beschrieben	33 Aufgabenstellung und Ziel der Unterrichtsstunde werden verknüpft
21	Die Aufgabenstellung (nicht: Nutzung oder Bearbeitung durch die Schüler) erfolgt <i>differenziert</i> hinsichtlich <i>kognitiver</i> Unterschiede der Schüler	34 Gruppierung der Schüler (<i>grouping students</i>) wird verknüpft mit der Differenzierung von Aufgaben
22	Die unter [21] vorgenommene Differenzierung der Aufgabenstellung wird <i>begründet</i>	35 Die durch die Aufgabenstellung erzielten Ergebnisse werden kontrolliert

4.2.1 Items zur Beschreibung situativer Bedingungen (Lerngruppe)

Die ersten drei Analyse Kriterien unseres Instruments (Items 10, 11 und 12 in Tab. 1) erfassen, ob die Lehrperson die situativen Bedingungen des Unterrichts – die Lerngruppe – gedanklich berücksichtigt und Lernvoraussetzungen einschätzt bzw. beurteilt.

Das erste Kriterium (Item 10) sehen wir erfüllt, sobald an irgendeiner Stelle der schriftlichen Planung die Lerngruppe beschrieben wird. Dies ist zum Beispiel in Form der folgenden Beschreibung gegeben: „Die Lerngruppe besteht aus 23 Schülern im Alter von 15 bis 18 Jahren, darunter zwei Schüler ohne Migrationshintergrund und acht Jungen. Zwei Jugendliche wiederholen die zehnte Klasse. Die Schüler haben in der siebten Klasse mit dem Erlernen des Französischen als zweite Fremdsprache begonnen.“ Die beiden im zusätzlichen Online-Material dokumentierten Planungen erfüllen ebenfalls dieses Kriterium (zusätzliches Online-Material, Abb. 2, Zeilen 174–177; zusätzliches Online-Material, Abb. 3, Zeilen 122–133).

Ein weiteres Kriterium (Item 11) sehen wir erfüllt, wenn speziell kognitive bzw. leistungsbezogene Unterschiede beschrieben werden, also auf Leistungsheterogenität der Lerngruppe Bezug genommen wird. Entscheidend ist hierbei der Fokus auf die *Lernenden als Gruppe*, während eine Erläuterung, die sich lediglich auf einen einzelnen Schüler bezieht, nicht ausreicht. Das Kriterium ist zum Beispiel durch folgende Beschreibung erfüllt: „Die Lerngruppe ist als heterogen zu bezeichnen. Hervorzuheben sind die Leistungsunterschiede zwischen den Jungen und Mädchen.“ Die

Planungen im zusätzlichen Online-Material unterscheiden sich in diesem Kriterium. Plan A (zusätzliches Online-Material, Abb. 2) enthält eine solche Beschreibung (Zeilen 178–179), Plan B nicht.

Um nicht allein kognitive, sondern auch motivationale Unterschiede der Lernenden herauszustellen, verwenden wir in Analogie zum vorhergehenden Item 11 ein weiteres Kriterium, das wir erfüllt sehen, wenn speziell motivationale Unterschiede der Lernenden herausgestellt werden (Item 12). Beispielsweise ist das Kriterium durch folgende Beschreibung erfüllt: „Seit (...) der Wahl der Grund- und Leistungskurse für die Oberstufe ist die Lernbereitschaft und -motivation zahlreicher Schüler spürbar gesunken. Lediglich eine Handvoll Schüler hat sich für eine Fortsetzung des Erlernens der französischen Sprache entschieden.“ Wie bereits bei Item 11 muss auch bei Item 12 mindestens eine soziale Differenzierung der Lernenden sprachlich erkennbar sein (z. B. „Die meisten Schüler..., drei von ihnen hingegen...“); eine Betrachtung der gesamten Lerngruppe reicht nicht aus. Plan A (zusätzliches Online-Material, Abb. 2) enthält eine solche Beschreibung (Zeilen 180–181), Plan B wiederum nicht.

4.2.2 *Items zur Beschreibung der Aufgabe*

Drei weitere Analyse Kriterien (Items 20, 21 und 22) erfassen Aspekte der Aufgabe (s. Definition unter Abschn. 2).

Mit dem ersten Kriterium (Item 20) wird erfasst, ob die Aufgabenstellung, die die Schülerinnen und Schüler erhalten werden, in der schriftlichen Unterrichtsplanung beschrieben wird. Nicht ausreichend ist z. B. ein Vermerk in der Verlaufsplanung, dass die Lernenden ein Arbeitsblatt erhalten und dieses dann im Anhang des Entwurfs aufgeführt wird, sondern es muss – zumindest ansatzweise – eine Beschreibung der Aufgabe vorliegen. Beide Pläne (zusätzliches Online-Material, Abb. 2, Zeilen 214–215 und Abb. 3, Zeilen 109–114) enthalten die Beschreibung der Aufgabenstellung.

Ein weiteres Kriterium (Item 21) erfasst, ob die Aufgabenstellung, die die Schüler und Schülerinnen zur Phase der „Erarbeitung“ erhalten werden, in Hinblick auf kognitive Unterschiede bei den Schülerinnen und Schülern differenziert. Gefragt ist also eine in Variationen konstruierte und auf unterschiedliche kognitive Voraussetzungen abgestimmte Aufgabenstellung. Nicht gemeint ist hier eine einheitliche Aufgabenstellung, die von den Schülern unterschiedlich bearbeitet werden kann („Differenzierung von unten“), da dies prinzipiell bei jeder Aufgabenstellung im Sinne der hier betrachteten „Lernaufgabe“ möglich ist. Hinsichtlich der Darstellung ist es ausreichend, wenn die in Variationen konstruierte Aufgabenstellung stellvertretend mit z. B. drei unterschiedlichen Schülern beschrieben wird. Wird eine Aufgabenstellung vorgenommen, die inhaltlich differenziert (z. B. ein Thema wird unterteilt, jede Gruppe bearbeitet ein anderes Teil-Thema), aber nicht auf unterschiedliche Leistungsniveaus der Schüler zielt, so sehen wir dieses Kriterium ebenfalls nicht erfüllt. Während Plan B (s. zusätzliches Online-Material) diesem Kriterium nicht gerecht wird, enthält Plan A die folgende Passage, welche konkret auf Erfüllung des Kriteriums schließen lässt (zusätzliches Online-Material, Abb. 2, Zeilen 262–264): „Ich habe mich für drei unterschiedliche Detektivgeschichten entschieden, welche ich in differenzierter Form aufbereitete.“ Auf die differenzierte Aufgabenstellung wird in

Plan A darüber hinaus an weiteren Stellen Bezug genommen, sodass insgesamt das Kriterium erfüllt ist. Beispiele aus zwei anderen Planungen, die ebenfalls auf Erfüllung des Kriteriums schließen lassen, lauteten wie folgt: „Die dreifach differenzierten Aufgabenblätter (...). Als zusätzliches Arbeitsmaterial steht allen Schülerinnen und Schülern das Aufgabenblatt ‚Erfinde selbst schöne Päckchen‘ zur Verfügung.“ sowie in einer weiteren Planung „(...) Für das Erstellen der Wandzeitung ist er [ein bestimmter Schüler; Anm. der Verf.] zusammen mit zwei anderen Kindern mit der Aufgabe betraut, Baderegeln aus einem Faltblatt der DLRG auf ein Plakat zu übertragen. Beim Erstellen der Wandzeitung hat er [ein anderer Schüler; Anm. der Verf.] zusammen mit einem anderen Jungen die Aufgabe, Inhalte und Ergebnisse eines Experiments zum Thema ‚Schwimmen und Sinken‘ auf einem Plakat darzustellen.“⁴²

Schließlich verwenden wir ein drittes Kriterium (Item 22). Dieses erfasst aufbauend auf Item 21 (bzw. Item 221, vgl. Anmerkung ii), ob die vorgenommene Differenzierung begründet wird. Dieses Kriterium sehen wir z. B. erfüllt, wenn ein Hinweis wie der folgende gegeben wird: „Die dreifach differenzierten Aufgabenblätter ermöglichen, dass schwächere Schülerinnen und Schüler in der Aufgabenbewältigung gefördert und stärkere gefordert werden, wodurch das Prinzip des individuellen Lernens berücksichtigt wird.“ Plan A erfüllt dieses Kriterium (zusätzliches Online-Material, Abb. 2, Zeilen 216–220), Plan B nicht.

4.2.3 Items zur Passung zwischen Lerngruppe und Aufgabe

Zwei Analysekriterien (Items 31 und 32) erfassen, inwieweit die in den Unterrichtsplanungen formulierte Aufgabe bzw. Aufgabenstellung in Anpassung an kognitive Voraussetzungen der Lerngruppe erfolgt. Dies wird über die Frage operationalisiert, ob der spezifische Lernstand hinsichtlich der Aufgabe beschrieben wird und die Aufgabe einen möglichen und sinnvollen „nächsten Lernschritt“ der Schüler darstellt. Es muss also ersichtlich sein, dass die Aufgaben am zuvor Erreichten bzw. Erlernten der Lerngruppe angeschlossen werden und dass die Aufgaben auf die situativen Bedingungen (Voraussetzungen der Lerngruppe) bezogen werden. Nicht ausreichend ist hierbei, wenn allein Bezug auf die inhaltliche Gliederung der Unterrichtseinheit genommen wird und auf die inhaltliche Verortung der Stunde in der Struktur der Unterrichtseinheit verwiesen wird.

Item 31 erfasst dabei, ob dies für die *gesamte Lerngruppe* erfolgt, Item 32 erfasst, ob dies *differenziert* für bestehende kognitive Unterschiede in der Lerngruppe vorgenommen wird. Im Fall von Item 32 muss die Aufgabe bzw. Aufgabenstellung Differenzierungsmaßnahmen für den Unterricht in Abstimmung auf die Lernvoraussetzungen der Schülerinnen und Schüler enthalten. Die planende Lehrperson muss also unter Beweis stellen, dass sie geeignete Differenzierungsmaßnahmen kennt und deren Anwendung für eine spezifische Unterrichtsstunde beschreibt.

In Plan B (zusätzliches Online-Material, Abb. 3) ist keines der beiden Kriterien erfüllt. In Plan A (zusätzliches Online-Material, Abb. 2) ist das erste der beiden Kriterien (Item 31) erfüllt (Zeilen 271–301), indem bezogen auf die Darstellung des „Stands der Kompetenzentwicklung“ der Lerngruppe die Aufgabentätigkeit der Schülerinnen und Schüler angeschlossen wird. Das zweite der beiden Kriterien (Item 32) ist erfüllt, weil der Lernstand für drei Schülergruppen tabellarisch dargelegt (Zei-

len 116–147) und die Aufgabenstellung für diese drei Gruppen differenziert wird (Zeilen 95–113).

4.2.4 Items zur Verknüpfung von Aufgaben und weiteren Planungselementen

Mit einem weiteren Kriterium (Item 33) wird erfasst, inwieweit eine Verknüpfung der Aufgabe bzw. Aufgabenstellung mit der Zielsetzung der Unterrichtsstunde oder mit einer Zielstellung des Unterrichts beschrieben bzw. angesprochen wird. Mit „Standard“ wird im sprachlichen Duktus vieler schriftlicher Planungen Äquivalentes zu „Ziel“ oder „Zielstellung“ ausgedrückt, sodass im Zuge der hier vorzunehmenden Kodierungen „Standard“ und „Ziel“ als Synonyme betrachtet werden. Die mit Item 33 geforderte Verknüpfung wird hierbei als Indikator für ein analytisches Planungsverständnis aufgefasst, das herausstellt, an welcher Stelle die Aufgabe in der Funktion steht, bestimmte Ziele, die i. d. R. kongruent zu Rahmenvorgaben sind, zu verfolgen. Fehlt diese Verknüpfungsleistung, so könnte vermutet werden, dass die Aufgabenwahl in Hinblick auf eine übergeordnete Zielebene bzw. Kompetenzentwicklung nicht hinreichend analytisch und reflektierend erfolgt ist.

Zur technischen Umsetzung wurden zwei Aspekte definiert, die in der Planung enthalten sein mussten, um das Kriterium als erfüllt zu betrachten: Die schriftliche Darstellung der Verknüpfung sollte von zwei Seiten aus erfolgen: Einerseits deduktiv *vom Ziel zur Aufgabe*. Dies ist häufig durch eine tabellarische Darstellung der „Standards“ und der „Standardkonkretisierung“ sichtbar (zusätzliches Online-Material, Abb. 2, Zeilen 66–92). Andererseits induktiv *von der Aufgabe zum Ziel*. Eine solche Verknüpfung findet sich i. d. R. an anderer Stelle im Fließtext (zusätzliches Online-Material, Abb. 2, Zeilen 232–233: „Das Ziel der Arbeitsphase ist es, dass sich die Lernenden selbständig mit ihrem Text auseinandersetzen und ihre Lesefähigkeit unter Beweis stellen.“). Plan B erfüllt dieses Kriterium nicht, da nur der zweite Aspekt (zusätzliches Online-Material, Abb. 3, Zeilen 31–34: „Die SchülerInnen fertigen eine Zeichnung nach ihrem Tonmodell eines Raumes aus dem vergangenen Halbjahr an. Hierbei soll besonders die Raumdarstellung durch Hell-Dunkel erlernt und geübt werden.“), nicht jedoch der erste enthalten ist.

Ein weiteres Kriterium (Item 34) zielt darauf ab, zu erfassen, inwieweit eine Verknüpfung der Aufgabe bzw. Aufgabenstellung mit der sozialen Ordnung des Unterrichts beschrieben wird. Konkret geht es dabei um die Frage, ob die Gruppierung der Schülerinnen und Schüler (*grouping students*) funktional verknüpft wird mit der Differenzierung von Aufgaben. Beispiele hierfür sind die folgenden Sätze: „Die Partnergruppen sind bewusst homogen gebildet worden, damit sich beide Gruppenmitglieder gleichermaßen angesprochen fühlen und gleichermaßen tätig werden.“ sowie „In der Anwendungsphase setzen die Schülerinnen und Schüler das erarbeitete Aufgabenformat eigenständig um. Die Anwendung erfolgt in Einzelarbeit, damit jedes Kind gemäß seinen individuellen Kenntnissen ungestört Lernerfahrungen machen kann und nicht durch leistungsstärkere oder -schwächere Mitschüler in seinem Arbeitsprozess gehemmt wird.“ Auch in Plan A findet sich eine solche Verknüpfung (zusätzliches Online-Material, Abb. 2, Zeilen 223–225: „Die Erarbeitungsphase wird durch die Veränderung der Sitzordnung eingeleitet, indem sich die Kinder je nach

aufgeklebter Nummer auf ihrem Sitzplatz in nummernhomogene Tischreihen umsetzen (...“), in Plan B hingegen nicht.

Schließlich wird mit einem Kriterium (Item 35) erfasst, inwieweit eine Kontrolle der Ergebnisse, welche durch die Aufgabenstellung erzielt wurden, in der Planung vorgesehen ist.

Wird eine solche Kontrolle allein auf eine nachfolgende Stunde verlagert, so muss das Item mit „Nein“ beantwortet werden, d. h. dieses Item stellt einen Indikator für die zeitlich unmittelbar an die Bearbeitung oder ggf. vorläufige Bearbeitung der Aufgabe angeschlossene Ergebniskontrolle dar. Insbesondere wenn differenzierte Aufgabenstellungen erfolgen, ist eine gemeinsame Ergebnissicherung aus didaktischer Sicht von entscheidender Bedeutung (Tulodziecki et al. 2004). Ein Beispiel für das Vorliegen einer solchen Verknüpfung ist: „Bereitgestellte Lösungsblätter ermöglichen die Selbstkontrolle der Schülerinnen und Schüler.“ Sowohl Plan A (zusätzliches Online-Material, Abb. 2, Zeilen 239–241) als auch Plan B (zusätzliches Online-Material, Abb. 3, Zeilen 93–97) erfüllen dieses Kriterium.

4.3 Instrumente zur Erfassung der Außenkriterien

Zur Bearbeitung der Fragestellung 3 werden als Außenkriterien neben der Abiturnote sowie der Note des Ersten Staatsexamens, welche direkt erfragt wurden, pädagogische Überzeugungen zum Lehren und Lernen verwendet, um professionelle Personenmerkmale einzubeziehen. Um Anschlussfähigkeit zu sichern, kommen zwei aus der Literatur bekannte und vielseitig eingesetzte Skalen zum Einsatz, die zwei Dimensionen erfassen: die Transmissionsorientierung und die konstruktivistische Orientierung. Entsprechende Skalen wurden dem technischen Bericht von Seidel et al. (2003) entnommen und leicht umformuliert, um eine Anpassung an alle Unterrichtsfächer zu gewährleisten. Die Transmissionsorientierung wurde mit neun Items gemessen (z. B. „Schülerinnen und Schüler sollten Anwendungsaufgaben, denen unterschiedliche inhaltliche Zusammenhänge zugrunde liegen, nicht vermischen, sondern nacheinander behandeln.“ $\alpha=0,70$), die konstruktivistische Orientierung mit acht Items (z. B. „Schülerinnen und Schüler lernen Inhalte am besten, indem sie selbst Wege zur Lösung von Problemen entdecken.“ $\alpha=0,68$). Eingeleitet wurde die Item-Batterie mit der Frage „Wie stark stimmen Sie mit folgenden Aussagen zum Lehren und Lernen überein?“ und für die Beantwortung stand ein vierstufiges Antwortformat zur Verfügung („stimmt nicht“ bis „stimmt genau“).

Im Rahmen der zusätzlichen Stichprobe der Studie LEK-R wurden die von den Referendaren unterrichteten Schülerinnen und Schüler zu ihrer Einschätzung von verschiedenen Aspekten der Unterrichtsqualität befragt, darunter über eine Skala zur Binnendifferenzierung im Unterricht, die für die hier zu leistende, kleine Ergänzungsanalyse zur prognostischen Validität verwendet wird. Sie stammt aus PISA 2003 (Ramm et al. 2006, S. 200) und umfasst vier Items: „Die einzelnen Schüler/Schülerinnen haben oft verschiedene Aufgaben.“, „Schnellere Schüler/Schülerinnen können schon zum Nächsten übergehen.“, „Die Lehrerin/der Lehrer stellt unterschiedlich schwere Fragen, je nachdem, wie gut ein Schüler/eine Schülerin ist.“, „Die Lehrerin/der Lehrer verlangt von den guten Schülern/Schülerinnen deutlich mehr.“ Die einleitende Frage lautete „Wie geht es bei euch im Unterricht zu?“, wobei die Schülerinnen

und Schüler sich explizit auf die eine gehaltene Unterrichtsstunde beziehen sollten, deren Unterrichtsplanungen ebenfalls analysiert wurden. Für die Beantwortung der Items stand ein vierstufiges Antwortformat zur Verfügung („Ja, trifft zu“ bis „Nein, trifft nicht zu“). Eine konfirmatorische Faktorenanalyse auf Schülerebene weist einen guten Fit für die Skala aus ($\chi^2=1,08$, $df=2$, n.s.; $RMSEA<0,01$), und die Skala weist eine akzeptable Reliabilität auf ($\alpha=0,60$). Die Variation der Skala über die 22 Schulklassen ist substanziell ($ICC=0,17$), sodass eine mehrebenenanalytische Modellierung angebracht ist.

4.4 Datenanalyse

Die Kodierung der schriftlichen Unterrichtsplanungen erfolgte anhand der dargelegten Kriterien. Zur Verfeinerung der Kriterien und Sammlung von geeigneten Beispielen wurden in einem ersten Schritt rund ein Drittel der Unterrichtsplanungen von den Verf. unabhängig voneinander kodiert. Diese Planungen waren zuvor zufällig aus dem Datensatz gezogen worden. Die Kategoriensysteme wurden im mehrfachen Wechsel zwischen deduktiver und induktiver Herangehensweise weiterentwickelt. Dies gewährleistete einerseits theoretisch begründete Kategorien und andererseits die erschöpfende Nutzung der in den Planungen vorzufindenden Informationen. Bei fehlender Übereinstimmung der von den Verf. unabhängig voneinander kodierten Planungen wurde eine Einigung erzielt, die letztlich in die hier dargestellten Datenanalysen einfließen.

Die Kodierungen der Antworten der restlichen Planungen erfolgten in einem zweiten Schritt jeweils durch zwei geschulte Rater,³ die ebenfalls unabhängig voneinander die offenen Antworten kodierten. Als Übereinstimmungsmaß wurde Cohen's Kappa berechnet (Wirtz und Caspar 2002). Werte größer 0,75 gelten als sehr gute Übereinstimmung. Die für die verwendeten Indikatoren berechneten Kappa-Werte variieren zwischen 0,67 und 0,93 mit einem Mittelwert von $M=0,83$. Unterhalb der Grenze von 0,75 liegt lediglich ein Indikator, sodass die für die Kodierung entwickelten Kategoriensysteme als bewährt gelten können. Bei fehlender Übereinstimmung wurden in gemeinsamer Diskussion Einigungen erzielt. Schwierig zu kodierende Antworten wurden als „Grenzfälle“ dokumentiert, um im Anschluss ein konsistentes Vorgehen mit ähnlichen Antworten zu gewährleisten.

5 Ergebnisse

5.1 Ergebnisse aus Skalierungsanalysen zur Planungskompetenz

Die Skalierung der Daten erfolgte im vorliegenden Fall auf der Grundlage des dichotomen Raschmodells (Rost 1996). Die für die Skalierung genutzte Software *Conquest* (Wu et al. 1997) weist jedem Item aufgrund seiner Lösungsquote einen Schwierigkeitsparameter und jeder befragten Person entsprechend der gezeigten Leistung einen Fähigkeitsparameter zu. Von den für die Kodierung verwendeten Variablen wurden 11 als Items in die Skalierung einbezogen. Item 221 wurde ausgeschlossen, da es eine Kodierungshäufigkeit von nur 1 % aufwies (vgl. Anmerkung 2).

Für die Skalierung greifen wir auf ein Verfahren zurück, das in der Literatur als Ansatz virtueller Personen beschrieben (Rost 1996; Hartig und Kühnbach 2005) bzw. als „konkurrenente“ Skalierung bezeichnet wird (z. B. van Davier et al. 2006). Hierbei stellen echte Längsschnittprobanden, d. h. Personen, für die eine zweimalige Messung vorliegt, zwei Fälle in der Skalierungsdatei dar (ein „echter“ und ein „virtueller“). Mit 106 Probanden des ersten und 105 Probanden des zweiten Messzeitpunktes (von einer Person lag uns nur eine Planung vor) ergibt sich hiermit eine Skalierung an 211 Fällen. Der entscheidende Vorteil dieses Vorgehens liegt darin begründet, dass im Falle kleinerer Stichproben wie in der hier vorliegenden Studie ($n < 150$) dennoch eine Skalierung nach der Item-Response-Theorie (IRT) möglich ist (vgl. Bond und Fox 2007). In Verbindung mit der genutzten Software *Conquest* erhält man differenzierte Informationen über die einbezogenen Items (s. die Ergebnisse der Skalierung in Abb. 1 und Tab. 2), welche über die üblichen Informationen aus der Anwendung von Skalierungsverfahren der klassischen Testtheorie deutlich hinausgehen.⁴

Die Skala erweist sich als reliabel (EAP-Reliabilität 0,703, Theta-Varianz 1,631), die Items streuen hinreichend über das Fähigkeitsspektrum. Die Diskriminationswerte liegen, von einem Item mit einer besonders hohen Lösungshäufigkeit (Item 35) abgesehen, insgesamt hoch genug ($> 0,3$). Die gewichteten Abweichungsquadrate liegen alle im angemessenen Bereich (zwischen 0,80 und 1,20).

5.2 Deskriptive Kennwerte zur Veränderung der Planungskompetenz

Von jeder angehenden Lehrperson wurden zwei Planungen analysiert: eine Planung einer Lehrprobe zu Beginn (t1) und eine Planung einer Lehrprobe vom Ende (t2) des Vorbereitungsdienstes bzw. Referendariats. Wie Tab. 3 entnommen werden kann, steigt die Lösungshäufigkeit für jedes der elf Items im Verlauf der zweiten Ausbildungsphase an, d. h. die mit den Indikatoren definierten Qualitätsmerkmale schriftlicher Unterrichtsplanungen werden in zunehmendem Maße erfüllt. Besonders deutlich ($> 25\%$; $p \leq 0,001$) ist dies für Item 21 zur differenzierten Aufgabenstellung und für Item 33 zur Verknüpfung von Aufgaben- und Zielstellung des Unterrichts. Deutlich ($\geq 15\%$; $p \leq 0,01$) ist aber auch eine Zunahme bei der Wahrnehmung kognitiver Unterschiede der Lerngruppe (Item 11) und bei der Berücksichtigung des spezifischen Lernstands der Lerngruppe (31) sowie Teilgruppen der Lerngruppe (32) für die gewählten Aufgaben. Nur geringfügige ($< 6\%$; n. s.) Unterschiede lassen sich für deskriptive Kategorien (Beschreibung der Lerngruppe und Aufgabenstellung, Items 10 und 20) wie auch für die Sicherung (Item 35) feststellen. Insgesamt bezieht sich also der Kompetenzzuwachs auf die anspruchsvolleren Analyse Kriterien, und hier wiederum auf typische Herausforderungen der Unterrichtsplanung, auf die bereits die Expertiseforschung verweist (Berücksichtigung der spezifischen Lerngruppe, Verknüpfung von Elementen der Unterrichtsplanung).

Die durchschnittlich erreichte Punktzahl steigt von Beginn (t1: $M = 5,32$, $SE = 0,20$, $SD = 2,06$) zum Ende (t2: $M = 6,86$, $SE = 0,19$, $SD = 1,93$) deutlich an. Der Mittelwertunterschied ist statistisch signifikant ($t(104) = -6,66$, $p < 0,001$) und von hoher praktischer Bedeutsamkeit ($d = 0,8$). Für den gewählten Ausschnitt von Planungskompetenz können wir also belegen, dass diese von den angehenden Lehrkräften im

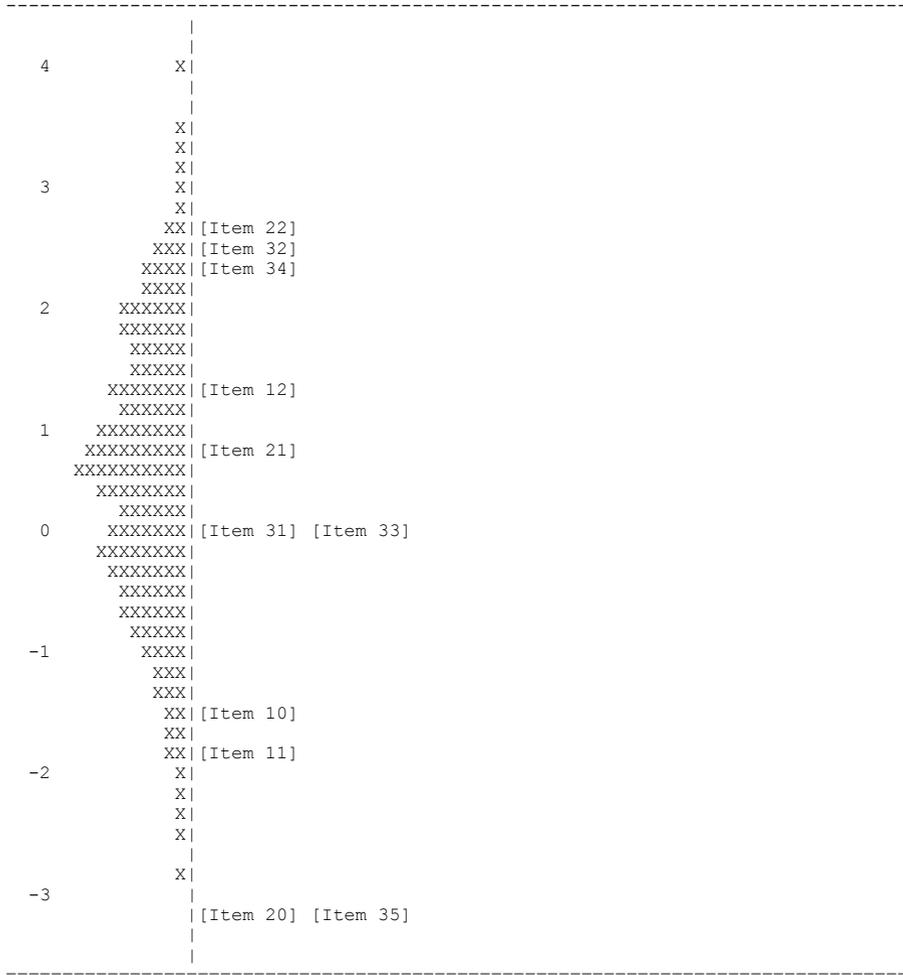


Abb. 1 Skalierung der Items zur Planungskompetenz (Ein ‚X‘ steht für 1,3 Fälle, die Beschriftungen rechts zeigen die Lage der Items auf dem Kontinuum an)

betrachteten Zeitabschnitt des Vorbereitungsdiensts erworben bzw. weiterentwickelt wird.

5.3 Belege zur Konstruktvalidität

Die Prüfung der Konstruktvalidität erfolgt einerseits über die Abiturnote als Indikator für kognitive Grundfertigkeiten sowie der Note des Ersten Staatsexamens als Indikator für Studienerfolg. Andererseits beziehen wir relevante pädagogische Überzeugungen zum Lehren und Lernen ein (vgl. Abschn. 4.3). Um einen adäquaten Umgang mit fehlenden Werten bei der Berechnung von Interkorrelationen zu gewährleisten, wurde ein Modell mit diesen Variablen sowie der Variablen der Planungskompetenz

Tab. 2 Item-Schwierigkeitsparameter der Skalierung für die Items zur Messung von Planungskompetenz sowie Angaben zum *Weighted Item Fit* (Mean Square, Konfidenzintervall, t-Wert) und zum Discrimination-Index

Item	Kriterium	Item-Schwierigkeits-Parameter (SE)	Mean Square (Konfidenzintervall)	t-Wert	Discrimination Index	Lösungshäufigkeit
10	Die Lerngruppe wird beschrieben	-1,483 (0,174)	0,92 (0,77, 1,23)	-0,7	0,54	82,5
11	Unterschiede in der Lerngruppe hinsichtlich kognitiver Voraussetzungen werden beschrieben	-1,792 (0,184)	0,88 (0,74, 1,26)	-0,9	0,58	85,8
12	Unterschiede in der Lerngruppe hinsichtlich motivationaler Voraussetzungen werden beschrieben	1,217 (0,147)	1,13 (0,86, 1,14)	1,7	0,47	37,0
20	Die Aufgabenstellung wird beschrieben	-3,205 (0,243)	0,94 (0,48, 1,52)	-0,1	0,36	95,3
21	Die Aufgabenstellung (nicht: Nutzung oder Bearbeitung durch die Schüler) erfolgt differenziert hinsichtlich kognitiver Unterschiede der Schüler	0,757 (0,144)	1,01 (0,87, 1,13)	0,1	0,56	45,5
22	Die unter [21] vorgenommene Differenzierung der Aufgabenstellung wird begründet	2,656 (0,178)	0,91 (0,76, 1,24)	-0,7	0,53	15,6
31	Bisheriger spezifischer Lernstand der Schüler (<i>gesamte Lerngruppe</i>) hinsichtlich der für die Unterrichtsstunde gewählten Aufgaben wird beschrieben – im Sinne des „nächsten Lernschrittes“	0,132 (0,145)	1,07 (0,86, 1,14)	1,0	0,53	57,4
32	Bisheriger spezifischer Lernstand der Schüler (<i>Differenzierung auf einzelne Schüler oder Schülergruppen</i>) hinsichtlich der für die Unterrichtsstunde gewählten Aufgaben wird beschrieben – im Sinne des jeweiligen „nächsten Lernschrittes“	2,531 (0,174)	1,03 (0,77, 1,23)	0,2	0,45	17,1
33	Aufgabenstellung und Ziel der Unterrichtsstunde werden verknüpft	0,055 (0,145)	0,94 (0,86, 1,14)	-0,8	0,64	58,8
34	Gruppierung der Schüler (<i>grouping students</i>) wird verknüpft mit der Differenzierung von Aufgaben	2,337 (0,168)	1,10 (0,79, 1,21)	1,0	0,37	19,4
35	Die durch die Aufgabenstellung erzielten Ergebnisse werden kontrolliert	-3,205 (0,545)	1,05 (0,48, 1,52)	0,3	0,24	95,3

Tab. 3 Lösungshäufigkeiten der Items zu Beginn (t1) und zum Ende (t2) des Referendariats sowie Ergebnisse (p -Werte) aus non-parametrischen Tests zur Überprüfung der gleichen Verteilung der Werte zu t1 und t2

Item	Kriterium	t1	t2	p
10	Die Lerngruppe wird beschrieben	80,2	84,8	0,405
11	Unterschiede in der Lerngruppe hinsichtlich kognitiver Voraussetzungen werden beschrieben	78,3	93,3	0,005
12	Unterschiede in der Lerngruppe hinsichtlich motivationaler Voraussetzungen werden beschrieben	33,0	41,0	0,243
20	Die Aufgabenstellung wird beschrieben	92,5	98,1	0,109
21	Die Aufgabenstellung (nicht: Nutzung oder Bearbeitung durch die Schüler) erfolgt <i>differenziert</i> hinsichtlich <i>kognitiver</i> Unterschiede der Schüler	32,1	59,0	<0,001
22	Die unter [201] vorgenommene Differenzierung der Aufgabenstellung wird <i>begründet</i>	10,4	21,0	0,054
31	Bisheriger spezifischer Lernstand der Schüler (<i>gesamte Lerngruppe</i>) hinsichtlich der für die Unterrichtsstunde gewählten Aufgaben wird beschrieben – im Sinne des „nächsten Lernschrittes“	48,1	66,7	0,004
32	Bisheriger spezifischer Lernstand der Schüler (<i>Differenzierung auf einzelne Schüler oder Schülergruppen</i>) hinsichtlich der für die Unterrichtsstunde gewählten Aufgaben wird beschrieben – im Sinne des jeweiligen „nächsten Lernschrittes“	9,4	24,8	0,001
33	Aufgabenstellung und Ziel der Unterrichtsstunde werden verknüpft	45,3	72,4	<0,001
34	Gruppierung der Schüler (<i>grouping students</i>) wird verknüpft mit der Differenzierung von Aufgaben	12,3	26,7	0,012
35	Die durch die Aufgabenstellung erzielten Ergebnisse werden kontrolliert	92,5	98,1	0,109

zu beiden Zeitpunkten mithilfe der Software *Mplus* (Muthén und Muthén 1998–2006) geschätzt, die eine modellbasierte Imputation fehlender Werte über das FIML-Verfahren (*Full Information Maximum Likelihood*) ermöglicht. Ausbildungsgänge innerhalb von Studienstandorten wurden als Stratifikationsvariable modelliert (unter Zuhilfenahme der Analyseoption *Type = Complex*), um der Clusterstruktur der Stichprobe gerecht zu werden und korrekte Standardfehler und Teststatistiken zu erhalten. Die Planungskompetenz wurde zu beiden Messzeitpunkten als latente Variable durch die 11 Indikatoren spezifiziert, welche als kategoriale Daten definiert wurden. Insgesamt weist das so geschätzte Modell einen akzeptablen Fit auf ($n=106$; $\chi^2/df=1,19$, $p=0,014$; RMSEA=0,042).

Wie Tab. 4 entnommen werden kann, zeigen sich negative Korrelationen mit den Noten zu Beginn des Vorbereitungsdiensts. Im Falle der Abiturnote ($-0,31$, $p \leq 0,01$) ist dieser Zusammenhang statistisch signifikant und substantiell. Im Falle der Note des Ersten Staatsexamens ist dieser Zusammenhang deutlich niedriger ($-0,13$, $p \leq 0,10$). Am Ende des Vorbereitungsdiensts lassen sich keine linearen Zusammenhänge mit praktischer Bedeutsamkeit zwischen den Noten und der Planungskompetenz nachweisen ($|r| < 0,1$). Ebenfalls kann Tab. 4 entnommen werden, dass wie erwartet die Transmissionsorientierung negativ, die konstruktivistische Orientierung positiv mit der Planungskompetenz korreliert. Zu Beginn des Vorbereitungsdiensts ist der Zusammenhang mit der konstruktivistischen Orientierung, zum Ende des Vor-

Tab. 4 Ergebnisse aus interkorrelativen Analysen zur Konstruktvalidität

	Abiturnote	Note des Ersten Staatsexamens	Transmissionsorientierung	Konstruktivistische Orientierung	Planungskompetenz t1
Planungskompetenz t1	-0,31**	-0,13 [#]	-0,07	0,20*	–
Planungskompetenz t2	0,06	-0,05	-0,22*	0,10	0,40***

*** $p \leq 0,001$; ** $p \leq 0,01$; * $p \leq 0,05$; [#] $p \leq 0,10$

bereitungsdiensts ist der Zusammenhang mit der Transmissionsorientierung statistisch signifikant (jeweils $p \leq 0,05$).

5.4 Hinweise auf prognostische Validität

Um erste Hinweise auf prognostische Validität zu erhalten, wurde an einer zusätzlichen Stichprobe von 22 Referendarinnen und Referendaren das Analyseinstrument eingesetzt. Ihre schriftlichen Unterrichtsplanungen wurden analog zu der hier dargelegten Vorgehensweise kodiert. Um Vergleichbarkeit zu gewährleisten, wurden diese 22 Fälle zusammen mit der zuvor genutzten Skalierungsdatei von 211 Fällen (Abschn. 5.1) kombiniert. Aus dieser Skalierung bezogen wir WLE-Personenfähigkeitswerte zur Planungskompetenz dieser 22 Referendare. In einer Mehrebenenanalyse mit Schülerinnen und Schülern auf Ebene 1 und den 22 Referendaren bzw. Schulklassen auf Ebene 2 wurde die Skala „Binnendifferenzierung“ (Abschn. 4.3) als abhängige Variable, die gemessene Planungskompetenz als unabhängige Variable spezifiziert. Aufgrund der geringen Anzahl von Einheiten auf Ebene 2 wurde die Regressionsanalyse auf Basis manifester Variablen berechnet.⁵ Auf Ebene 2 der Referendare bzw. Schulklassen ist der Prädiktor „Planungskompetenz“ statistisch signifikant ($\beta = 0,369$, $p \leq 0,05$) und kann 13,6% der Varianz auf Ebene 2 erklären. Höhere Fähigkeitkeitswerte der Referendarinnen und Referendare in der Planungskompetenz gehen demnach einher mit einer höheren Zustimmung der von ihnen unterrichteten Schülerinnen und Schüler hinsichtlich des Qualitätsaspekts „Binnendifferenzierung“ im Unterricht.

6 Diskussion

Wie lässt sich Planungskompetenz messen? Wir haben uns mit dem Aspekt der „didaktischen Adaptivität“ auf nur eine einzelne, aber hoch relevante Problemstellung konzentriert, die bei der Planung von Unterricht typischerweise und – auf einer allgemeinen Ebene – unabhängig vom Schulfach und der Schulform auftritt und gelöst werden muss, also als zentrale generische Anforderung für Lehrkräfte gelten kann. Jede Schulstunde enthält die Auseinandersetzung der Schülerinnen und Schüler mit dem jeweiligen Unterrichtsgegenstand, der sich als Aufgabe eingebunden in Inhalts-, Material- und Zielentscheidungen manifestiert. Lehrkräfte müssen ihr Wissen über Inhalte, Materialien und Ziele zur Gestaltung von Unterricht auf eine praktisch-konkrete, individuelle Situation anwenden, um ihren Unterricht auf die eigene Lerngruppe angemessen abzustimmen. Für die Planung von „gutem“ Unter-

richt ist also entscheidend, ob die Unterrichtsgestaltung und die situativen Bedingungen sinnvoll zusammengeführt werden. In dieser Dimension des Planungsprozesses, der Anwendung von theoretischem Wissen auf die praktische Situation, zeigt sich unseres Erachtens ein zentraler Aspekt von Planungskompetenz als eine wichtige Voraussetzung unterrichtlicher Handlungskompetenz von Lehrerinnen und Lehrern.

Somit lautete unsere erste Fragestellung, ob sich der gewählte Aspekt von Planungskompetenz, operationalisiert über Indikatoren, skalieren und damit als reliable latente Fähigkeitsvariable modellieren lässt. Insgesamt kann diese Frage positiv beantwortet werden. Die verwendeten Indikatoren erweisen sich als trennscharf, die Skala ist reliabel, und die Varianz der latenten Variable ist hinreichend groß. Auch wenn der mittlere Schwierigkeitsbereich sinnvollerweise noch durch weitere Items abgedeckt werden sollte, so zeigt sich dennoch, dass die Items angemessen über das Fähigkeitsspektrum streuen. Die gute Passung des Skalierungsmodells auf die Daten bestätigt unsere Annahme zur Machbarkeit der standardisierten Messung von Planungskompetenz: zum einen hinsichtlich des gewählten Ausschnittes, zum anderen hinsichtlich der für diesen gewählten Aspekt relevanten Indikatoren. Angesichts der Komplexität von Unterrichtsplanung als Gegenstand der Messung und der Heterogenität der analysierten Unterrichtsplanungen über Schulstufen, -formen und -fächer hinweg waren diese guten psychometrischen Kennwerte der Skalierung nicht notwendigerweise zu erwarten. Einschränkend gilt allerdings zu berücksichtigen, dass wir angesichts der für IRT-Skalierungsanalysen relativ kleinen Stichprobe auf die Modellierung von möglicherweise abhängigen Items als *partial-credit* verzichtet haben, während wir die entsprechenden dichotomen Items aufgrund der mit ihnen verbundenen relevanten Informationen nicht von der Analyse ausschließen wollten. Dieser technischen Limitierung unseres Vorgehens sollte in zukünftigen Untersuchungen begegnet werden, etwa indem eine deutlich größere Stichprobe von schriftlichen Unterrichtsplanungen in die Analyse einbezogen wird. Anwendung finden könnten bei Längsschnittanalysen auch *state-of-the-art*-Modellierungen wie *Latent-Change*-Modelle, da sie die Modellierung der (Residual-)Korrelationen und damit die Abhängigkeiten zwischen einzelnen Items ermöglichen (vgl. Steyer et al. 1997).

Im Sinne einer curricularen Validierung der so gemessenen Planungskompetenz lautete unsere zweite Frage, ob sich über die Messung zu zwei Zeitpunkten in der zweiten Ausbildungsphase (Referendariat) bei angehenden Lehrkräften eine Zunahme ihrer Planungskompetenz abbilden lässt. Dies vermuteten wir, da die Planung von Unterricht zentraler Lerngegenstand der zweiten Ausbildungsphase ist. Die Ergebnisse der Messwiederholung verweisen darauf, dass der beobachtete Kompetenzzuwachs eine professionelle Weiterentwicklung abbildet. Die Entwürfe erfüllen mit fortschreitender Ausbildung zunehmend die Kriterien der Berücksichtigung des spezifischen Lernstands der Lerngruppe in Bezug auf die gestellte Aufgabe, des Lernstands von Teilgruppen der Lerngruppe in Bezug auf die gestellte Aufgabe, die daran anschließende differenzierende Gestaltung der Lernaufgabe und deren Anbindung an das Unterrichtsziel. Die Kategorien, in denen die größten Zuwächse zu verzeichnen sind, betreffen also Verknüpfungsleistungen von verschiedenen Elementen der Planung bzw. von verschiedenen Wissensinhalten und Wissensarten, was als besonderes Charakteristikum von Expertenwissen gilt.

Zusätzlich zur Analyse einer Veränderung der Personenmittelwerte über die Zeit (Fragestellung 2) galt es, weitere Belege zur Konstruktvalidität zu erbringen (Fragestellung 3). Über interkorrelative Analysen konnte dargelegt werden, dass zu Beginn der zweiten Ausbildungsphase die Ausprägung der Planungskompetenz nicht unabhängig von kognitiven Grundfertigkeiten (gemessen über die Abiturnote) und dem bisherigen Ausbildungserfolg (gemessen über die Note des Ersten Staatsexamens) ist. Im weiteren Verlauf nahmen diese Zusammenhänge ab, vermutlich weil die Planungskompetenz einer starken Veränderung und eher niedrigen Stabilität unterliegt, möglicherweise aber auch, weil es sich bei der Planungskompetenz um eine spezifische professionelle Kompetenz handelt, die mit allgemeinen kognitiven Fähigkeiten (Abiturnote) bzw. dem universitären Wissen (Note des Ersten Staatsexamens) umso schwächer korreliert, je stärker sie ausgeprägt ist. Pädagogische Überzeugungen, die zu Beginn der zweiten Ausbildungsphase erfragt worden waren, korrelierten zu beiden Zeitpunkten in erwarteter Richtung (Transmissionsorientierung negativ, konstruktivistische Orientierung positiv). Diese Ergebnisse zu Zusammenhängen mit Außenkriterien werten wir als Belege der konvergenten Validität (signifikante Korrelation der Abiturnote mit Planungskompetenz zu Beginn des Referendariats, signifikante Korrelationen zwischen pädagogischen Überzeugungen und Planungskompetenz), während ihre jeweilige Höhe bzw. der im Verlauf der zweiten Ausbildungsphase schwächer werdende Zusammenhang zwischen Planungskompetenz und vorherigen Noten als Belege für die diskriminante Validität betrachtet werden können.

Erste Hinweise auf die prognostische Validität der hier gemessenen Planungskompetenz erhielten wir über die Verwendung einer zusätzlichen Stichprobe von Referendarinnen und Referendaren aus der Studie LEK-R. In einer Mehrebenenanalyse konnte ihre Planungskompetenz die Urteile der von ihnen unterrichteten Schülerinnen und Schülern hinsichtlich des Qualitätsmerkmals „Binnendifferenzierung“ im Unterricht statistisch signifikant vorhersagen. Fällt die didaktische Adaptivität in der Planung höher aus, so nehmen die Schülerinnen und Schüler in der durchgeführten Stunde auch einen stärker binnendifferenzierenden Unterricht wahr. Obgleich es sich hierbei um eine ergänzende Analyse an einer kleinen Stichprobe von angehenden Lehrkräften handelt, dürfte dieses Ergebnis den bislang mageren Forschungsstand zum Zusammenhang von Planung und Durchführung von Unterricht dennoch bereichern. Während bislang hauptsächlich Hinweise zur Vorhersage der Organisation und Strukturierung des Unterrichts vorliegen (Clark und Peterson 1984; Seel 1997), kann nun vermutet werden, dass binnendifferenzierende Maßnahmen im Unterricht durch die didaktische Adaptivität der Lehrkräfte, welche sie bei der Planung unter Beweis stellen, bedingt sind.

6.1 Limitierungen des methodischen Zugangs

Die Analyse von schriftlichen Unterrichtsplanungen ermöglicht Aussagen über die Qualität von Unterrichtskonzeptionen als Voraussetzung einer Handlungssituation im Sinne von professioneller Planungskompetenz, sie stößt jedoch an Grenzen in Hinblick auf die Beschreibung von Handlungskompetenz. Nicht umsonst sprechen bereits Peterson et al. (1978, S. 418) von dem Begriffspaar *planning decisions* und *interactive decisions*. Die Divergenz zwischen Planung und ihrer tatsächlichen

Umsetzung fällt größer aus, je komplexer die Situation (Wahl 1991) und je unerfahrener die Lehrperson ist (Roth et al. 2005). Bei Novizen steht im Gegensatz zu Experten eine höhere Bemühung um planungsgerechte Umsetzung einem höheren Ausmaß an tatsächlichen Planungsmodifikationen gegenüber (z. B. Westermann 1991). Diese zeigen sich vor allem in Weglassungen oder Änderungen aufgrund falscher Zeitplanung und aufgrund von nicht antizipierten Schülerreaktionen (Seel 1996). Aber auch erfahrene Lehrpersonen sehen sich gezwungen, ihre Pläne durch ihr Handeln zu ergänzen und gegebenenfalls improvisierend zu füllen. Gleichwohl kann insofern von einem engen Zusammenhang zwischen (guter) Planung und (gutem) Unterrichtshandeln ausgegangen werden, als dass Planungen für die Unterrichtsdurchführung handlungsleitend sind und die Handlungssituationen strukturieren und entlasten. Vor diesem Hintergrund besteht also ein sehr großes Forschungsdesiderat darin, den Zusammenhang zwischen Planung und Durchführung von Unterricht genauer zu untersuchen. Die ergänzende Analyse auf Basis der Daten der Studie LEK-R gibt lediglich Anlass zu Vermutungen, die in zukünftigen Studien jedoch vertiefend analysiert werden müssen. Aufschlussreich dürften weitere Datenquellen (z. B. Videoaufnahmen des Unterrichts) sein. Auch Informationen über das tatsächlich vorliegende Ausmaß der Heterogenität der jeweiligen Schulklasse (z. B. aus Schulleistungstests) wären wichtig einzubeziehen – unsere Analyse beruht allein auf die im Rahmen einer schriftlichen Unterrichtsplanung vorgenommenen Darstellung der jeweiligen angehenden Lehrperson.

Eine weitere Einschränkung besteht darin, dass oftmals weder alles, was Lehrpersonen planen, von ihnen auch als schriftliche Planung verzeichnet ist, noch alle Planungsentscheidungen darin expliziert sind (z. B. McCutcheon 1980; Borko und Livingston 1989). Mit Blick auf angehende Lehrkräfte gilt zudem zu beachten, dass schriftliche Unterrichtsplanungen spezifischen Vorgaben der Lehrerbildung unterliegen, was zur Folge haben dürfte, dass auch ganz andere Ziele verfolgt werden (z. B. Einhalten von Prüfungsanforderungen, unreflektierte Übernahme von persönlichen Präferenzen der prüfenden Seminarleiter). Allerdings geben hierzu die Ergebnisse zu interkorrelativen Analysen mit Außenkriterien sowie die Ergebnisse der ergänzenden Analyse zur prognostischen Validität (welche zudem eine Stichprobe eines völlig anderen Ausbildungskontextes nutzte) doch wichtige Hinweise, die darauf schließen lassen, dass die hier gemessene Planungskompetenz nicht einfach nur ein Ergebnis normativer Erwartungen oder Vorgaben von Auszubildenden der zweiten Phase sein kann.

Neben diesen Limitierungen, die sich auf das Konstrukt und das Datenmaterial beziehen, ist anzumerken, dass wir im Gegensatz zu einer Testung von planungsbezogenem Wissen angehender Lehrkräfte (z. B. König und Blömeke 2009, 2010; Döhrmann et al. 2010; Baer et al. 2011) in unserer Studie das Wissen nicht direkt erfassen, sondern über die Verwendung von Indikatoren zur Analyse von schriftlichen Planungen auf das Vorhandensein von Wissen bzw. erfolgten Wissensverknüpfungen schließen. Der damit einhergehenden Ungenauigkeit der Messung (hervorgerufen z. B. durch Unvollständigkeit oder Mängel in der Darstellung von Planungen) steht jedoch der große Vorteil gegenüber, Kompetenzen von angehenden Lehrkräften situativ und im spezifischen Kontext ihrer Lerngruppen zu erfassen. Insbesondere das situationale Wissen über die spezifische Lerngruppe sowie das konzeptuelle Wissen über didaktische

Maßnahmen zur Adaptivität dürften hierbei eine entscheidende Rolle spielen. Aus diesem Grund nimmt der von uns gewählte Zugang die genannten Limitierungen in Kauf. Eine weitere Einschränkung der Untersuchung an sich besteht darin, dass lediglich ein spezifischer Aspekt der Planungskompetenz angehender Lehrkräfte analysiert wurde. Die Teilkompetenz der didaktischen Adaptivität besitzt zwar zentrale Bedeutung. Mit ihr wird jedoch Planungskompetenz keineswegs hinreichend beschrieben.

6.2 Ausblick auf zukünftige Forschung

Weiterführend wäre somit zu überlegen, analog zu dem hier dargelegten Vorgehen weitere Teilkompetenzen mithilfe spezifischer Items zu konstruieren, z. B. zum (zeitlichen) Stundenverlauf, zur kognitiven Aktivierung durch die gewählten Inhalte und Themen oder zur geplanten Kommunikationsstruktur zwischen Lehrperson und Schülerinnen und Schülern. Wichtig erscheint dabei, die hier vorgenommene generische Analyse von Aufgaben um eine fachdidaktische Analyse der Aufgaben zu ergänzen, um beispielsweise Annahmen zur Binnenstrukturierung von Planungskompetenz in fachbezogene und fächerübergreifende Anteile zu prüfen und somit die Anschlussfähigkeit des methodischen Zugangs an aktuelle Vorstellungen über die analytische Differenzierung des kognitiven Bereichs professioneller Kompetenz von Lehrkräften in fachdidaktisches und pädagogisches Wissen (Baumert und Kunter 2006) zu sichern. Mit der weiterführenden Konstruktion und Prüfung von ergänzenden Teilkompetenzen sollte auch die Entwicklung einer theoretischen Rahmenkonzeption zur Planungskompetenz von Lehrkräften verbunden sein.

Forschung zur Planungskompetenz, wie wir sie mit unserer Studie begonnen haben, reicht in die Gestaltung der Lehrerbildung hinein. Uns sind keine wissenschaftlich geprüften Kriterien zur Analyse und Bewertung von schriftlichen Unterrichtsplanungen bekannt, die gerade in der zweiten Ausbildungsphase nicht nur prüfungsrelevant, sondern in Form von Staatsexamensnoten beim Einstieg in den Beruf sogar auch karriereentscheidend sind. Es ist die Aufgabe der (empirischen) Lehrerbildungsforschung, geeignete Verfahren zu entwickeln, die die Beurteilung kompetenter Planung von Unterricht von zukünftigen Lehrkräften wissenschaftlich absichern helfen und in der Aufgabe der Beurteilung einen engen Anschluss zur praktischen Ausbildung ermöglichen.

Anmerkungen

- 1 Von rund 20% der in diese Studie einbezogenen Lehramtsanwärterinnen und Lehramtsanwärter fehlen die demografischen Kennwerte. Anhand der hier berichteten Verteilungen nach Geschlecht, Alter oder Abitur- sowie Staatsexamensnote, welche wir über die gültigen Werte erhalten, vermuten wir jedoch nicht, dass eine Verzerrung der Stichprobe vorliegt, da die Verteilungen jener der in der Studie untersuchten Population im Wesentlichen entsprechen.
- 2 Analog zu diesem Item wurde ein weiteres Kriterium definiert (Item 221), das erfasste, ob die Aufgabenstellung, die die Schülerinnen und Schüler zur Phase der „Erarbeitung“ erhalten werden, in Hinblick auf motivationale Unterschiede bei den Schülerinnen und Schülern differenziert. Da jedoch nur zwei Planungen (<1%) dieses Kriterium erfüllten, wurde es für die weiteren Analysen ausgeschlossen.

- 3 Für die höchst sorgfältig durchgeführten Kodierarbeiten danken wir herzlich Klaus-Peter Hachmann und Mona Zimmermann.
- 4 Die in unserem Fall verwendete – für IRT-Anwendungen – relativ kleine Stichprobe stellt ferner eine üblicherweise empfohlene Modellierung von möglicherweise voneinander abhängigen Items als *partial-credit* (Masters 1982) in Frage, sodass wir allein dichotome Items in die Skalierung einbeziehen. Ein Ausschluss betreffender Items würde wiederum erheblichen Informationsverlust nach sich ziehen und damit einen vergleichbar größeren Nachteil mit sich bringen. Eine paarweise Prüfung der Zellenbesetzungen verdeutlicht jedoch, dass jeweils alle Zellen besetzt sind und lediglich bei den Item-Paaren 10/11 sowie 21/22 eine leicht erhöhte Abhängigkeit vorliegt (Kappa-Werte 0,34 und 0,38). Dieses skalierungstechnische Problem wird im Diskussionsteil aufgegriffen und stellt unseres Erachtens ein Desiderat für zukünftige Skalierungen an größeren Stichproben von Unterrichtsplanungen dar, welche dann mit entsprechenden *partial-credit*-Modellierungen arbeiten sollten.
- 5 Der Verzicht auf eine latente Modellierung hat auch lediglich zur Folge, dass die Zusammenhänge zwischen den Konstrukten unterschätzt werden; angesichts unserer Forschungsfrage handelt es sich somit um ein konservatives Vorgehen.

Open Access Dieser Artikel unterliegt den Bedingungen der Creative Commons Attribution License. Dadurch sind die Nutzung, Verteilung und Reproduktion erlaubt, sofern der/die Originalautor/en und die Quelle angegeben sind.

Literatur

- 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.
- Arnold, K. -H., & Koch-Priewe, B. (2010). Traditionen der Unterrichtsplanung in Deutschland. *Bildung und Erziehung*, 63(4), 401–416.
- Aufschnaiter, C. v., & Blömeke, S. (2010). Professionelle Kompetenz von (angehenden) Lehrkräften erfassen – Desiderata. *Zeitschrift für Didaktik der Naturwissenschaften*, 16, 361–367.
- Baer, M., Kocher, M., Wyss, C., Guldemann, T., Larcher, S., & Dörr, G. (2011). Lehrerbildung und Praxiserfahrung im ersten Berufsjahr und ihre Wirkung auf die Unterrichts Kompetenzen von Studierenden und jungen Lehrpersonen im Berufseinstieg. *Zeitschrift für Erziehungswissenschaft*, 14(1), 85–117.
- Baumert, J., & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 9(4), 469–520.
- Baumert, J., Kunter, M., Brunner, M., Krauss, S., Blum, W., & Neubrand, M. (2004). Mathematikunterricht aus Sicht der PISA-Schülerinnen und -Schüler und ihrer Lehrkräfte. In M. Prenzel, J. Baumert, W. Blum, R. Lehmann, D. Leutner, M. Neubrand, R. Pekrun, H.-G. Rolff, J. Rost, & U. Schiefele (Hrsg.), *PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland. Ergebnisse des zweiten internationalen Vergleichs* (S. 314–354). Münster: Waxmann.
- Berliner, D. C. (2004). Describing the behavior and documenting the accomplishments of expert teachers. *Bulletin of Science, Technology & Society*, 24, 200–212.
- Beyerbach, B. A. (1988). Developing a technical vocabulary on teacher planning: Preservice teachers' concept maps. *Teaching and Teacher Education*, 4(4), 339–347.
- Blankertz, H. (1969). *Theorien und Modelle der Didaktik*. München: Juventa.
- Blömeke, S., Risse, J., Müller, C., Eichler, D., & Schulz, W. (2006). Analyse der Qualität von Aufgaben aus didaktischer und fachlicher Sicht. Ein allgemeines Modell und seine exemplarische Umsetzung im Unterrichtsfach Mathematik. *Unterrichtswissenschaft*, 34(4), 330–357.
- Blumschein, P. (2014). Einblicke. In P. Blumschein (Hrsg.), *Lernaufgaben – Didaktische Forschungsperspektiven*. Bad Heilbrunn: Klinkhardt.
- Bönsch, M. (2004). *Intelligente Unterrichtsstrukturen*. Hohengehren: Schneider.
- Borko, H., & Livingston, C. (1989). Cognition and improvisation: Differences in mathematics instruction by expert and novice teachers. *American Educational Research Journal*, 26, 473–498.
- Bremerich-Vos, A. (2008). Benjamin S. Bloom (und andere) revisited. In A. Bremerich-Vos, D. Granzer, & O. Köller (Hrsg.), *Lernstandsbestimmung im Fach Deutsch: Gute Aufgaben für den Unterricht* (S. 29–49). Weinheim: Beltz.

- Bromme, R. (1981). *Das Denken von Lehrern bei der Unterrichtsvorbereitung. Eine empirische Untersuchung zu kognitiven Prozessen von Mathematiklehrern*. Weinheim: Beltz.
- Bromme, R. (1992). *Der Lehrer als Experte: zur Psychologie des professionellen Wissens*. Bern: Huber.
- Bruder, R. (2008). Vielseitig mit Aufgaben arbeiten. Mathematische Kompetenzen nachhaltig entwickeln und sichern. In R. Bruder, T. Leuders, & A. Büchter (Hrsg.), *Mathematikunterricht entwickeln. Bausteine für kompetenzorientiertes Unterrichten* (S. 18–52). Berlin: Cornelsen-Scriptor.
- Buchholtz, C., Dohmen, D., Köller, M., & Tschackert, K. (2012). *Evaluation des Berliner Vorbereitungsdienstes 2010–2012. Evaluationsbericht im Auftrag der Senatsverwaltung für Bildung, Jugend und Wissenschaft des Landes Berlin*. Hrsg. v. Forschungsinstitut für Bildungs- und Sozialökonomie, Berlin.
- Clark, C. M., & Peterson, P. L. (1986). Teachers' thought processes. In M. Wittrock (Hrsg.), *Handbook of research on teaching* (S. 255–296). New York: Macmillan.
- Cook, N. J. (1992). Modeling human expertise in expert systems. In R. Hoffman (Hrsg.), *The psychology of expertise* (S. 29–60). New York: Springer.
- von Davier, A. A., Carstensen, C. H., & von Davier, M. (2006). Linking competencies in horizontal, vertical, and longitudinal settings and measuring growth. In J. Hartig, E. Klieme, & D. Leutner (Hrsg.), *Assessment of competencies in educational contexts* (S. 53–80). Göttingen: Hogrefe.
- De Jong, T., & Ferguson-Hessler, M. G. M. (1996). Types and qualities of knowledge. *Educational Psychologist*, 31(2), 105–113.
- Deutsches PISA-Konsortium. (Hrsg.) (2000). *Schülerleistungen im internationalen Vergleich. Eine neue Rahmenkonzeption für die Erfassung von Wissen und Fähigkeiten*. Berlin: MPI.
- Döhrmann, M., Kaiser, G., & Blömeke, S. (2010). Messung des mathematischen und mathematikdidaktischen Wissens: Theoretischer Rahmen und Teststruktur. In S. Blömeke, G. Kaiser, & R. Lehmann (Hrsg.), *TEDS-M 2008. Professionelle Kompetenz und Lerngelegenheiten angehender Primarstufenlehrkräfte im internationalen Vergleich* (S. 169–194). Münster: Waxmann.
- Ericsson, K. A. (1996). The acquisition of expert performance: An introduction to some of the issues. In K. A. Ericsson (Hrsg.), *The road to excellence: The acquisition of expert performance in the arts and science, sports, and games* (S. 1–50). Mahwah: Lawrence Erlbaum.
- Fauth, B., Decristan, J., Rieser, S., Klieme, E., & Büttner, G. (2014). Student ratings of teaching quality in primary school: Dimensions and prediction of student outcomes. *Learning and Instruction*, 29, 1–9.
- Fend, H. (2008). *Schule gestalten. Systemsteuerung, Schulentwicklung und Unterrichtsqualität*. Wiesbaden: Springer VS.
- Haag, L., & Streber, D. (2010). Unterrichtsvorbereitung bei Lehrern – mit System? *Lehrerbildung auf dem Prüfstand*, 3, 107–117.
- Haas, A. (1998). *Unterrichtsplanung im Alltag von Lehrerinnen und Lehrern. Eine empirische Untersuchung zum Planungshandeln von Hauptschul-, Realschul- und Gymnasiallehrern*. Regensburg: S. Roederer.
- Hammerness, K., Darling-Hammond, L., Bransford, J., Berliner, D., Cochran-Smith, M., McDonald, M., & Zeichner, K. (2005). How teachers learn and develop. In L. Darling-Hammond & J. Bransford (Hrsg.), *Preparing teachers for a changing world: What teachers should learn and be able to do* (S. 358–389). San Francisco: Jossey-Bass.
- Hattie, J. (2009). *Visible learning. A synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.
- Heckhausen, H. (1989). *Motivation und Handeln* (2. Aufl.). Berlin: Springer.
- Heimann, P. (1965). Didaktik 1965. In P. Heimann, G. Otto, & W. Schulz (Hrsg.), *Unterricht. Analyse und Planung* (S. 7–12). Hannover: Schroedel.
- Heimann, P., Otto, G., & Schulz, W. (Hrsg.) (1965). *Unterricht – Analyse und Planung*. Hannover: Schroedel.
- Helmke, A. (2010). *Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts*. Seelze: Kallmeyer.
- Hiebert, J., Gallimore, R., Garnier, H., Givvin, K. B., Hollingsworth, H., & Jacobs, J. (2003). *Teaching mathematics in seven countries: Results from the TIMSS 1999 video study*. NCEES (2003–013). Washington DC: U.S. Department of Education/National Center for Education Statistics.
- Hill, J., Yinger, R. J., & Robbins, D. (1983). Instructional planning in a laboratory preschool. *The Elementary School Journal*, 83(3), 182–193.
- Hopperditzel, A., & Arnold, K.-H. (2008). Förderung von Kompetenz: Die Perspektive der Lehr-Lern-Forschung. In K.-H. Arnold, O. Graumann, & A. Rakhkochkine (Hrsg.), *Handbuch Förderung* (S. 36–44). Weinheim: Beltz.

- Horstkemper, M. (2004). Erziehungswissenschaftliche Ausbildung. In S. Blömeke, P. Reinhold, G. Tulodziecki, & J. Wildt (Hrsg.), *Handbuch Lehrerbildung* (S. 461–476). Bad Heilbrunn: Klinkhardt.
- Housner, L. D., & Griffey, D. C. (1985). Teacher cognition: Differences in planning and interactive decision-making between experienced and in-experienced teachers. *Research Quarterly for Exercise and Sport*, 56, 45–53.
- Jacobs, C. L., Martin, S. N., & Otieno, T. C. (2008). A science lesson plan analysis instrument for formative and summative program evaluation of a teacher education program. *Science Education*, 92(6), 1096–1126.
- Jank, W., & Meyer, H. (2002). *Didaktische Modelle* (5. Aufl.). Berlin: Cornelsen.
- Johnson, K. (2003). *Designing language tasks*. Hamshire: Palgrave Macmillan.
- Kagan, D. M. (1992). Professional growth among preservice and beginning teachers. *Review of Educational Research*, 62(2), 129–169.
- Klafki, W. (1975). *Studien zur Bildungstheorie und Didaktik. Durch ein kritisches Vorwort ergänzte Aufl.* Weinheim: Beltz.
- Kleinknecht, M., Bohl, T., Maier, U., & Metz, K. (2011). Aufgaben und Unterrichtsplanung. In K. H. Arnold, T. Bohl, S. Hopmann, B. Hudson, E. Kiel, H. Kiper, B. Koch-Priewe, K. Reussner, N. Seel, & K. Zierer (Hrsg.), *Jahrbuch für Allgemeine Didaktik* (S. 59–75). Baltmannsweiler: Schneider Verlag.
- Kleinknecht, M., Bohl, T., Maier, U., & Metz, K. (Hrsg.) (2013). *Lern- und Leistungsaufgaben im Unterricht. Fächerübergreifende Kriterien zur Auswahl und Analyse*. Bad Heilbrunn: Klinkhardt.
- Klieme, E., & Leutner, D. (2006). Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen. *Zeitschrift für Pädagogik*, 52, 876–903.
- KMK = Kultusministerkonferenz. (2004). *Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. Standards für die Lehrerbildung: Bildungswissenschaften*. Beschluss der Kultusministerkonferenz vom 16.12.2004.
- Koch-Priewe, B. (1997). Grundlegung einer Didaktik der Lehrerbildung. In M. Bayer, U. Carle, & J. Wildt (Hrsg.), *Brennpunkt: Lehrerbildung* (S. 139–163). Opladen: Leske + Budrich.
- König, J. (Hrsg.) (2012). *Teachers' pedagogical beliefs. Definition and operationalisation – connections to knowledge and performance – development and change*. Münster: Waxmann.
- König, J., & Blömeke, S. (2009). Pädagogisches Wissen von angehenden Lehrkräften: Erfassung und Struktur von Ergebnissen der fachübergreifenden Lehrerausbildung. *Zeitschrift für Erziehungswissenschaft*, 12(3), 499–527.
- König, J., & Blömeke, S. (2010). *Pädagogisches Unterrichtswissen (PUW). Dokumentation der Kurzfassung des TEDS-M-Testinstruments zur Kompetenzmessung in der ersten Phase der Lehrerausbildung*. Berlin: Humboldt-Universität.
- Krumm, V. (1985). Anmerkungen zur Rolle der Aufgabe in Didaktik, Unterricht und Unterrichtsforschung. *Unterrichtswissenschaft*, 13(2), 102–115.
- Kuhnke-Lerch, I., & Bruder, R. (2010). Wie analysieren Lehramtsstudierende Unterrichtsentwürfe? – Vorstellung einer Studie zur Erfassung von Lehrpersonen im Lehramtsstudium. In Gesellschaft für Didaktik der Mathematik (Hrsg.), *Beiträge zum Mathematikunterricht* (S. 517–520). Münster: WTM Verlag.
- Kunze, I., & Solzbacher, C. (Hrsg.) (2008). *Individuelle Förderung in der Sekundarstufe I und II*. Baltmannsweiler: Schneider.
- Leisen, J. (2006). Aufgabenkultur im mathematisch-naturwissenschaftlichen Unterricht. *Der mathematische und naturwissenschaftliche Unterricht (MNU)*, 59(5), 260–266.
- Lipowsky, F., Kastens, C., Lotz, M., & Faust, G. (2011). Aufgabenbezogene Differenzierung und Entwicklung des verbalen Selbstkonzepts im Anfangsunterricht. *Zeitschrift für Pädagogik*, 57(6), 868–884.
- Maier, U., Kleinknecht, M., Metz, K., & Bohl, T. (2010). Ein allgemeindidaktisches Kategoriensystem zur Analyse des kognitiven Potenzials von Aufgaben. *Beiträge zur Lehrerbildung*, 28(1), 84–96.
- Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, 47(2), 149–174.
- McCutcheon, G. (1980). How do elementary school teachers plan? The nature of planning and influences on it. *Elementary School Journal*, 81, 4–23.
- Morine-Dershimer, G. (1993). Tracing conceptual change in preservice teachers. *Teaching and Teacher Education*, 9(1), 15–26.
- Müller, K., Gartmeier, M., & Prenzel, M. (2013). Kompetenzorientierter Unterricht im Kontext nationaler Bildungsstandards. *Bildung und Erziehung*, 66, 127–144.
- Muthén, B. O., & Muthén, L. K. (1998–2006). *Mplus* (Version 4.2) [Computer software]. Los Angeles.

- Neubrand, M., Jordan, A., Krauss, S., Blum, W., & Löwen, K. (2011). Aufgaben im COACTIV-Projekt: Einblicke in das Potenzial für kognitive Aktivierung im Mathematikunterricht. In Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S., & Neubrand, M. (Hrsg.), *Professionelle Kompetenzen von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV*. Münster: Waxmann.
- Peterßen, W. H. (2003). *Lehreraufgabe Unterrichtsplanung. Das Weingartener Planungsmodell*. München: Oldenbourg.
- Peterson, P. L., Marx, R. W., & Clark, C. M. (1978). Teacher planning, teacher behavior, and student achievement. *American Educational Research Journal*, 15(3), 417–432.
- Plöger, W. (2008). *Unterrichtsplanung*. Köln: Kölner Studien Verlag.
- Ramm, G., Prenzel, M., Baumert, J., Blum, W., Lehmann, R., Leutner, D., Neubrand, M., Pekrun, R., Rolff, H.-G., Rost, J., & Schiefele, U. (2006). *PISA 2003: Dokumentation der Erhebungsinstrumente*. Münster: Waxmann.
- Rost, J. (1996). *Lehrbuch Testtheorie, Testkonstruktion*. Bern: Huber.
- Roth, W. -M., Lawless, D., & Masciotra, D. (2001). Spielraum and teaching. *Curriculum Inquiry*, 31(2), 183–207.
- Schmit, S., Peters, S., & Kiper, H. (2014). Wissenserwerb durch Lernaufgaben. In P. Blumschein (Hrsg.), *Lernaufgaben – Didaktische Forschungsperspektiven*. Bad Heilbrunn: Klinkhardt.
- Schrader, F.-W., & Helmke, A. (2008). Unter der Lupe: Wie in deutschen Klassenzimmern differenziert wird. *Klexer. Magazin für die Grundschule*, 10(20), 3–5.
- Schwarzer, R., & Steinhagen, K. (Hrsg.) (1975). *Adaptiver Unterricht. Zur Wechselwirkung von Schülermerkmalen und Unterrichtsmethoden*. München: Kösel.
- Seel, A. (1997). Von der Unterrichtsplanung zum konkreten Lehrerhandeln – Eine Untersuchung zum Zusammenhang von Planung und Durchführung von Unterricht bei Hauptschullehrerstudentinnen. *Unterrichtswissenschaft*, 25(3), 257–273.
- Seidel, T., Prenzel, M., Duit, R., & Lehrke, M. (Hrsg.) (2003). *Technischer Bericht zur Videostudie „Lehr-Lern-Prozesse im Physikunterricht“*. Kiel: IPN.
- Shavelson, R. J. (1987). Planning. In M. J. Dunkin (Hrsg.), *International encyclopedia of teaching and teacher education* (S. 483–486). Oxford: Pergamon Press.
- Smith, T. W., & Strahan, D. (2001). Toward a prototype of expertise in teaching. A descriptive case study. *Journal of Teacher Education*, 55(4), 357–371.
- Stanat, P., & Christensen, G. (2006). *Where immigrant students succeed—a comparative review of performance and engagement in PISA 2003*. Paris: OECD.
- Steyer, R., Eid, M., & Schwenkmezger, P. (1997). Modeling true intraindividual change: True change as a latent variable. *Methods of Psychological Research Online*, 2(1), 21–33.
- Stigler, J. W., & Hiebert, J. (1999). *The teaching gap: Best ideas from the world's teachers for improving education in the classroom*. New York: Free Press.
- Tulodziecki, G., Herzig, B., & Blömeke, S. (2004). *Gestaltung von Unterricht. Eine Einführung in die Didaktik*. Bad Heilbrunn: Klinkhardt.
- Weinert, F. E. (1997). *Psychologie des Unterrichts und der Schule. Bd. 3: Enzyklopädie der Psychologie, Themenbereich D (Serie I)*. Göttingen: Hogrefe.
- Westermann, D. A. (1991). Expert and novice teacher decision making. *Journal of Teacher Education*, 42(4), 292–305.
- Wu, M. L., Adams, R. J., & Wilson, M. R. (1997). *ConQuest: Multi-aspect test software* [computer program]. Camberwell: Australian Council for Educational Research.
- Wygotski, L. S. (1978). *Mind in society. The development of higher psychological processes*. Cambridge: University Press.
- Yinger, R. J. (1977). A study of teacher planning: Description and theory development using ethnographic and information processing methods. Unpublished doctoral dissertation, Michigan State University.
- Zahorik, J. A. (1975). Teachers' planning models. *Educational Leadership*, 33, 134–139.