

Sigrid Blömeke

Lehrerbildung – Lehrerhandeln – Schülerleistungen

Perspektiven nationaler und internationaler empirischer Bildungsforschung

Antrittsvorlesung vom 10. Dezember 2003

Humboldt-Universität zu Berlin

2005

(= Öffentliche Vorlesungen; 139)

Die Lehrerausbildung ist in vielen Ländern hinsichtlich ihrer Gestaltung und Leistungsfähigkeit Kritik ausgesetzt. In Deutschland wird ihr vorgeworfen, sie sei praxisfern, in ihrer ersten Phase beliebig und in ihrer zweiten Phase mehr eine Meisterlehre als eine theoriegeleitete Ausbildung, sie dauere zu lange, sei nicht koordiniert und kenne keine geteilten Standards. Empirisch geprüft wurde ihre Wirksamkeit allerdings bisher nicht. Erst eine „externe“ Entwicklung führte dazu, dass entsprechende Forderungen derzeit zunehmen: Die Ergebnisse der internationalen Leistungsvergleiche zu den Kompetenzen von Kindern und Jugendlichen legten Defizite offen, die den Blick unter anderem auf die Lehrerausbildung als mögliche Ursache richten. Je konkreter das Vorhaben einer Wirksamkeitsprüfung allerdings diskutiert wird ist, desto stärker zeigt sich, wie schwierig sie zu realisieren ist. Im Folgenden werden daher zunächst methodologische Fragen der Erfassung von „Wirkungen“ der Lehrerausbildung diskutiert (1). Anschließend erfolgt ein Resümee des derzeitigen Forschungsstands zu den Lernvoraussetzungen, mit denen zukünftige Lehrerinnen und Lehrer in ihre Ausbildung eintreten, und zu den Wirkungen, die die Lehrerausbildung entfaltet (2). Abschließend werden Aufgaben der empirischen Lehrerausbildungsforschung für die Zukunft herausgearbeitet (3) und diese in ihren Chancen und Grenzen reflektiert (4).¹

1 Methodologische Probleme empirischer Forschung zur Lehrerausbildung

Eine Untersuchung der Wirksamkeit von Lehrerausbildung ist nur sinnvoll, wenn man davon ausgeht, dass Lehrerhandeln Schülerleistungen beeinflusst und dass das Lehrerhandeln mindestens zum Teil von der Qualität der Lehrerausbildung abhängt. Im Hinblick auf den ersten Punkt gibt es sowohl im deutschsprachigen (vgl. z.B. Weinert & Helmke 1997, Helmke & Jäger 2002, Ditton 2002, Baumert, Watermann & Schümer 2003) als auch im angelsächsischen Raum Studien (vgl. als klassische Studie Coleman et al. 1966 sowie z.B. Murnane & Phillips 1981; Monk & King 1994; Rowan, Chiang & Miller 1997; Wright, Horn & Sanders 1997, Wenglinsky 2002, Wayne & Youngs 2003), die den genannten Zusammenhang mindestens zum Teil belegen. Insbesondere zum zweiten Punkt gibt es allerdings widersprüchliche Befunde. Insofern besteht eine zentrale Herausforderung an empirische Forschung zur Lehrerausbildung darin zu klären, wie deren Wirksamkeit modelliert werden kann. Herrmann (2003, S. 625) führt hierzu aus: „Institutionen und Personen, Faktoren und Prozesse, Ziele, Erwartungen und Interessen, Realbedingungen und Beteiligungsvoraussetzungen bilden bei Bildungsprozessen fast unauflösbare Ursachen-, Bedingungs- und Wirkungszusammenhänge.“

Vor diesem Hintergrund kann nicht verwundern, dass die Wahl der Variablen umstritten ist. Sie hängt insbesondere von der Modellierung des Zusammenhangs von Lehrerausbildung, Lehrerwissen, Lehrerhandeln und Schülerleistungen ab. So stellt sich zum ersten die Frage, von welchen Faktoren der Lehrerausbildung eine Wirksamkeit auf Lehrerwissen angenommen wird: Sind es die Inhalte der Lehrangebote, ihre didaktisch-methodische Präsentation, der Grad der Strukturierung der Ausbildung, die Abfolge ihrer Phasen, der Ort, die Verknüpfung oder die Länge der Phasen? Zum zweiten ist es grundlegend zu fragen, wodurch Lehrerhandeln beeinflusst wird und welche Rolle das Wissen in diesem Zusammenhang spielt: Lässt sich hier ein ungebrochener Zusammenhang herstellen? Die Forschungslage spricht eher dagegen (Gruber & Renkl 2000, Mandl & Gerstenmaier 2001, Klinzing 2002, Blömeke, Eichler & Müller 2003). Zum dritten ist die Auswahl des Kriteriums für „Wirksamkeit“ der Lehrerausbildung zentral: Können dies die Leistungen von Schülerinnen und Schülern sein? Selbst wer einen solchen Ansatz vom Anspruch her als sinnvoll einschätzt, muss zugeben, dass es sich um eine lange Wirkungskette handelt, deren einzelne Verknüpfungen empirisch ungeklärt sind und die methodisch nur schwer zu bewältigen ist.

Sind diese drei Problemkreise schon in sich kontrovers, ist die Modellierung noch weit umstrittener, nimmt man die zahlreichen normativen Entscheidungen in den Blick, die bei der Definition der zu erhebenden Konstrukte zu fällen sind (Blömeke 2002, Blömeke, Herzig & Tulodziecki 2005): Impli-

zieren „Schülerleistungen“ nur kognitive oder auch affektiv-motivationale und vielleicht sogar sozial-moralische Dimensionen? Wird die Erfassung von „Lehrerhandeln“ auf Unterrichten beschränkt oder umfasst es auch Beurteilen und ggf. sogar Erziehen und Mitwirken an der Schulentwicklung als Untersuchungsdimensionen? Impliziert „Lehrerwissen“ nur kognitiv-fachliche Dimensionen oder auch Einstellungen, Selbstwirksamkeitserwartungen, berufunspezifische Persönlichkeitsmerkmale wie psychische Stabilität und Kontaktbereitschaft und sogar das Berufsethos? Es ist kaum anzunehmen, dass sich zu diesen Fragen einheitliche Positionen herauskristallisieren (vgl. entsprechend Oser 2004, S. 189ff., der vier unerschiedliche Typen an Wirksamkeitsvorstellungen unterscheidet).

Die Wahl der Untersuchungsmethode stellt eine weitere Herausforderung an Studien zur Wirksamkeit der Lehrerausbildung dar. Eine Durchsicht der Literatur zeigt, dass es keine einheitliche Methodologie gibt (Kane, Sandretto & Heath 2002, S. 180ff.). Darüber hinaus werden unterschiedliche Instrumente und innerhalb dieser unterschiedliche Items verwendet, so dass die vergleichsweise geringe Konsistenz in den Ergebnissen der vorliegenden Studien möglicherweise auch auf diese Uneinheitlichkeit zurückzuführen ist. Insgesamt „müssen die Methoden, mit denen Wirkungen der Lehrerausbildung in der Vergangenheit untersucht wurden, kritisch diskutiert werden (Wideen, Mayer-Smith & Moon 1998, S. 161ff.; Chatterij 2002, S. 367): In der Regel handelt es sich nur um qualitative Fallstudien. Ihre Ergebnisse können daher weniger als empirisch geprüfte Aussagen genommen werden denn als Hypothesen, die noch einer Überprüfung in *large scale*-Studien bedürfen. Bei Studien mit größeren Fallzahlen handelt es sich oft nur um Re-Analysen von Surveys, die andere Fragestellungen verfolgten, so dass eine präzise und konsistente Bearbeitung von Fragen zur Lehrerausbildung schwierig ist. (Feld-)experimentelle Studien und Langzeitstudien finden sich so gut wie nicht. Schülerleistungen werden in der Regel nicht erhoben – und wenn, dann werden sie (vor allem in älteren Untersuchungen) nur selten auf der Basis standardisierter Tests festgestellt. Eine erhebliche Forschungslücke ist darüber hinaus festzustellen, was die Bedeutsamkeit der fachwissenschaftlichen Ausbildung angeht; die überwiegende Zahl der Studien fokussiert auf pädagogische *beliefs*.

Für die Bundesrepublik Deutschland lässt sich ein generelles Defizit an empirischer Forschung zur Lehrerausbildung feststellen. Zahlreichen Veröffentlichungen mit normativ-konzeptioneller Orientierung stehen nur punktuelle empirische Studien gegenüber, die zudem in der Regel lediglich auf Befragungen und Dokumentenanalysen beruhen (vgl. zum Überblick Merzyn 2002). Schlee (1992, S. 559) formuliert daher als Fazit in Bezug auf die Forschung bis 1990, „dass zu den konzeptionellen Kernfragen und Brennpunkten der Lehrerausbildungsreform keine systematische, theoriebezogene und experimentelle Forschung durchgeführt wurde“. Zum selben Ergebnis kommen Döbrich et al. (2003, S. 3; vgl. auch Fried 1998, Oelkers 2002, Schaefers 2002) in Bezug auf die Gegenwart: „Verfügbar ist eine große Fülle von thematisch und regional begrenzten Einzeluntersuchungen, die aber oft kein Urteil darüber erlauben, inwiefern die gewonnenen Erkenntnisse generalisierbar sind.“ Larcher und Oelkers (2004, S. 129) gehen sogar so weit zu sagen, „wenn es eine Krise in der Lehrerausbildung gibt, dann ist es wesentlich eine Krise der fehlenden Daten“. Im deutschsprachigen Raum kann lediglich die Schweiz auf Ansätze empirischer Forschung zur Lehrerausbildung verweisen (Oser & Oelkers 2001), womit die noch vor einigen Jahren von Criblez und Wild-Näf (1998) auch für dieses Land konstatierten Defizite angegangen wurden.

Mit der Gewinnung von Daten allein ist die Frage der Wirksamkeit von Lehrerausbildung zudem noch nicht beantwortet. Dies machen die widersprüchlichen *Interpretationen* der wenigen vorliegenden Befunde deutlich. Welche Interpretationsperspektive eingenommen wird, hängt dabei unter anderem von der bildungspolitischen Position ab (s. dazu im Einzelnen Abschnitt 2.3).

2 Kernergebnisse der empirischen Forschung zur Lehrerausbildung

2.1 Befunde zur Bedeutsamkeit von Lernvoraussetzungen

Die Wirksamkeit der Lehrerausbildung kann nicht angemessen beurteilt werden, wenn nicht zuvor geklärt wird, mit welchen Lernvoraussetzungen zukünftige Lehrerinnen und Lehrer ihre Ausbildung aufnehmen. Diese lassen sich hinsichtlich zweier Dimensionen unterscheiden: zum einen im Hinblick auf sozio-kulturelle Merkmale und zum anderen im Hinblick auf Vorannahmen, Einstellungen bzw. Erwartungen (mangels konsistenter Begrifflichkeit in der Forschung im Folgenden mit dem in anglo-amerikanischen Studien verbreiteten Begriff *beliefs* bezeichnet) sowie Berufswahlmotive.

Während zu den sozio-kulturellen Merkmalen der deutschen Lehramtsstudierenden – mit Ausnahme historischer Untersuchungen (vgl. z.B. Kaelble 1978, Bölling 1983) – kaum aktuelle und detaillierte repräsentative Informationen vorliegen, lässt sich in Bezug auf Nordamerika festhalten: „The typical candidate for teacher education in the United States and Canada is a White, Anglo-Saxon, lower- or middle-class female who has grown up in a suburban or rural area. She is monolingual in English, has traveled very little beyond a 100-mile radius of her home, and has attended a local college or university close to her home.“ (Wideen, Mayer-Smith & Moon 1998, S. 141)

Mehr Studien lassen sich im deutschsprachigen Raum zu den Berufswahlmotiven finden, die im Kern auf die beiden folgenden Aussagen reduziert werden können: Lehramtsstudierende nehmen ihr Studium in der Absicht auf, später als Lehrerin bzw. Lehrer zu arbeiten (Terhart et al. 1994, Ulich 2000, Drechsel 2001), und sie haben dafür in erster Linie pädagogische oder fachliche Motive (vgl. als Überblick Schaefers 2002, S. 67f.). Die Validität dieser Erkenntnisse kann allerdings trotz mehrfacher Replikationen insofern in Frage gestellt werden, als es sich jeweils nur um nachträgliche Befragungen handelt, die teilweise nach vielen Jahren im Berufs stattfanden.

Im Hinblick auf die *beliefs* lassen sich international übereinstimmend folgende zentrale Ergebnisse der empirischen Forschung festhalten (Koch 1972; Zeichner 1986; Clark 1988; Kagan 1992; Pajares 1992; Block & Hazelip 1995; Richardson 1996; Bullough 1997; Wideen, Mayer-Smith & Moon 1998; Richardson & Placier 2001; Kane, Sandretto & Heath 2002):

- Lehramtsstudierende treten mit Vorstellungen über schul- und unterrichtsrelevante Aspekte in die Lehrerausbildung ein.
- Diese *beliefs* sind erfahrungsbasiert, so dass Lortie (1975) von der Schulzeit als „apprenticeship of observation“ spricht.
- *Beliefs* sind weitgehend veränderungsresistent, so dass es im Laufe der Lehrerausbildung eher selten zu grundlegenden Veränderungen kommt.
- Ihre Funktionsweise in der Ausbildung kann als „Filter“ beschrieben werden. Es werden überwiegend nur solche Informationen aus der Lehrerausbildung aufgenommen, die sich in das vorhandene System an Überzeugungen einpassen lassen.
- Erfolg versprechend scheinen Lehrprozesse zu sein, die aktiv an die vorhandenen *beliefs* anknüpfen und diese so schrittweise verändern.

In inhaltlicher Hinsicht ist in Bezug auf die *beliefs* festzuhalten, dass die Lehramtsstudierenden sehr optimistisch sind, was ihre zukünftigen Fähigkeiten zu unterrichten angeht (vgl. Kagan 1992, S. 140). Sie meinen zu wissen, wie Unterricht auszusehen hat (i.e. eine Lehrperson überliefert Wissen frontal vor der Klasse an die Schülerinnen und Schüler, die dieses rezeptiv aufnehmen), und sie meinen, nur noch ein bestimmtes Methodenrepertoire erlernen zu müssen, um beruflich beginnen zu können. Damit ist die Annahme verbunden, dass die Persönlichkeit einer Lehrkraft von höherer Bedeutung sei als ihre kognitiven Fähigkeiten oder ihr pädagogisches bzw. fachliches Wissen (Wideen, Mayer-Smith & Moon 1998, S. 142f.). Als besonders wichtig sehen es die Studierenden an, ihre Klassen „im Griff“ zu haben. Gleichzeitig – und in einem gewissen Spannungsverhältnis hierzu – ist ihr Fokus auf soziale Faktoren des Unterrichtens ausgerichtet, indem sie ein verständnisvolles und sorgendes Verhältnis zu ihren Schülerinnen und Schülern als Ziel angeben. Um die (begrenzte) Wirkung der Lehrerausbildung auf anfängliche Einstellungen, Interessen und Werthaltungen der Lehramtsstudierenden deutlich zu machen, kann auf die bereits 30 Jahre alte,

immer noch maßgebliche, da bisher einzige Längsschnittuntersuchung im deutschsprachigen Raum von Koch (1972) Bezug genommen werden. Eingesetzt wurden Skalen zu schul- und unterrichtsbezogenen Einstellungen und Überzeugungen: Bedeutsamkeit von „Anlage versus Umwelt“ und „Druck versus Zug“ als Bedingungen sowie von „Allgemeinbildung versus Spezialisierung“ als Ziel von Lernen, Verständnis des eigenen Berufs als „Berufung versus Job“ bzw. als „Pädagoge versus Fachwissenschaftler“ und schließlich „negative ReformEinstellung versus Veränderungsbereitschaft“. Zwei Kernergebnisse dieser Studie lauten: (1) Lehramtsstudierende unterscheiden sich in ihren Einstellungen zu Lernen und Lehren weder im Vergleich zu anderen Abiturienten noch im Vergleich zu Studienanfängern in anderen Studiengängen. (2) Im Prozess der Lehrerausbildung lassen sich Veränderungen gegenüber den zu Beginn des Studiums vorhandenen Einstellungen, Interessen und Werthaltungen feststellen. Diese werden aber mit dem Eintritt in die Schulpraxis während des Referendariats wieder rückgängig gemacht („Konstanzer Wanne“). Von der Tendenz her gilt, dass vor dem Studium eher konservative Einstellungen vorhanden sind, während des Studiums liberalere Auffassungen stärkere Zustimmung finden (i.e. stärkere Zuschreibung von Umweltfaktoren und Unterstützung als Erfolgsfaktoren für Lernen sowie größere Veränderungsbereitschaft), bevor im Referendariat wieder zu konservativeren Positionen zurückgekehrt wird. Eine *follow up*-Untersuchung zeigt, dass sich die für das Referendariat festgestellte Tendenz zu konservativen pädagogischen Einstellungen, Interessen und Werthaltungen in den ersten Jahren des Berufslebens weiter fortsetzt (Dann, Müller-Fohrbrodt & Cloetta 1981).

Dass angehende Lehrerinnen und Lehrer in der Ausbildung erworbene Einstellungen wieder aufgeben, stellen 20 Jahre später auch Brouwer und ten Brinke (1995a, b) in einer Längsschnittuntersuchung fest, die sie in den Niederlanden zur Erhebung der Wirkungen eines neuen Lehrerausbildungsprogramms durchgeführt haben. In Bezug auf die Ursache dieses Einstellungswandels kommen die Autoren zu dem Ergebnis, dass sich die neu in eine Schule kommenden Lehrerinnen und Lehrer gezwungen sehen, gemäß den dort üblichen Handlungsroutinen und gegen ihre eigenen Einstellungen zu handeln und dass diese sich dann sukzessive den Einstellungen der übrigen Lehrerinnen und Lehrer angleichen.

Vor diesem Hintergrund wird verständlich, wenn Lortie (1975) – wie Koch (1972) und Richardson (1996) – die Lehrerausbildung im Vergleich zur eigenen Schulerfahrung und der späteren beruflichen Sozialisation als „low impact enterprise“ bezeichnet. In Bezug auf die Ausbildung von Mathematiklehrerinnen und -lehrern nehmen Ball, Lubienski und Mewborn (2001, S. 437) eine ähnliche Einschätzung vor. Was immer Lehrpersonen im Laufe ihrer Ausbildung an Einstellungen erworben haben, erscheint „washed out by school experience“ (Zeichner & Tabachnik 1981).

2.2 Befunde zur Wirksamkeit der Lehrerausbildung

Im Folgenden wird auf zentrale Ergebnisse so genannter „Programmevaluationen“ eingegangen, in denen die Lehrerausbildung als Ganze oder zumindest substanzielle Teile von ihr untersucht werden (vgl. im Einzelnen Blömeke 2004a). Gegenstand der Untersuchung ist also nicht die individuelle Lehrperson, ihre Lernvoraussetzungen und deren Veränderung im Prozess der Lehrerausbildung, sondern in *large scale*-Studien und Meta-Analysen werden mit Hilfe multipler Regression Merkmale des Berufsstandes zu Schülerleistungen oder Lehrerhandeln in Beziehung gesetzt. Um die Effekte der Lehrerausbildung abschätzen zu können, werden bekannte Einflüsse (z.B. finanzielle Ausstattung der Schulen, Dauer der beruflichen Erfahrung) konstant gehalten. Während sich entsprechende Studien in der Bundesrepublik Deutschland nicht finden lassen, stellt diese Art von Studien in den USA das Kernverständnis von „Evaluation der Lehrerausbildung“ dar. Drei große Reviews, die sich zum Teil auf dieselbe Datenlage beziehen, fassen den aktuellen Forschungsstand zusammen (Abell Foundation 2001a, b; Darling-Hammond 2000; Wilson, Floden & Ferrini-Mundy 2001). Auf ihrer Basis lässt sich das folgende Resümee ziehen:

- Zunächst zur Frage, was sich in Bezug auf den Abschluss eines Lehramtsstudiengangs im Vergleich zu Notausbildungen bzw. gänzlich fehlender Lehrerausbildung feststellen lässt. Unter das Merkmal „Lehramtsstudiengang“ fallen dabei sowohl universitäre Studien in Erziehungswissenschaft als auch in der Fachwissenschaft, in der unterrichtet wird. Obwohl von diesem Merkmal am ehesten signifikante Wirkungen auf Schülerleistungen bzw. Lehrerhandeln erwartet werden könnten, sind die Ergebnisse der Studien aus den letzten 50 Jahren nicht durchgängig positiv. Auf der Ebene eines Bachelor-Abschlusses lässt sich zwar überwiegend folgern, dass Lehrpersonen mit einem Lehramtsabschluss höhere Schülerleistungen bzw. positivere Einschätzungen ihres Handelns erreichen als Lehrpersonen ohne einen solchen Abschluss. Einer Minderheit an Studien zufolge ist ein solcher Abschluss allerdings bedeutungslos. In beiden Fällen macht es im Übrigen keinen Unterschied, um welche Art des Bachelors – ob in einer Fachwissenschaft oder in Erziehungswissenschaft – es sich handelt.
- Dieses widersprüchliche Ergebnis spiegelt sich auf der Ebene der lehramtsspezifischen Masterabschlüsse wieder. Den meisten Studien zufolge erreichen Lehrpersonen mit einem Master höhere Schülerleistungen oder ihr Handeln wird besser beurteilt als das von Lehrpersonen, die nur einen Bachelor besitzen. Es liegen aber auch Studien vor, die einen solchen Unterschied nicht feststellen können.
- In Bezug auf die Wirksamkeit einzelner Komponenten der Lehrerausbildung soll zunächst die Frage nach der Wirksamkeit der fachwissenschaftlichen Ausbildung gestellt werden. Diese wird in Form von Tests des Lehrwissens, durch Zählungen belegter Fachkurse oder in Form von akademischen Abschlüssen operationalisiert. Ein durchgängig positiver Zusammenhang zwischen Fachwissen und Schülerleistungen bzw. Lehrerhandeln ist nicht vorhanden. Zwar spricht erneut eine Mehrheit der Studien für die Wirksamkeit einer umfassenden fachwissenschaftlichen Ausbildung, gleichzeitig liegen allerdings Studien vor, die nicht-signifikante oder sogar negative Zusammenhänge feststellen. Um diese Befunde angemessen interpretieren zu können, kann die Langzeitstudie von Monk und King (1994) herangezogen werden. In dieser wird von einem „Deckeneffekt“ ausgegangen: Um hohe Schülerleistungen zu erreichen, ist ein Mindestmaß an Fachwissen notwendig. Über eine gewisse Schwelle hinaus ist jedoch kein zusätzlicher positiver Effekt mehr festzustellen. Die mathematikspezifische Lehrerausbildungsforschung analysiert dieses Phänomen noch tiefer (Ball & Bass 2000; Ball, Lubienski & Mewborn 2001): Die aus dem höheren fachwissenschaftlichen Niveau folgende Art der kognitiven Strukturierung von Inhalten macht zum einen eine Elementarisierung schwieriger, die für die Vermittlung der Inhalte an Schüler notwendig ist. Zum anderen ist die Vertiefung fachwissenschaftlicher Inhalte in der Lehrerausbildung in der Regel mit einer stärkeren Orientierung der Dozenten an konventionellen Lehrmethoden verbunden, was offensichtlich – möglicherweise im Sinne von erfahrungsbasiertem Lernen bzw. Modelllernen – zu einer stärkeren eigenen Orientierung der zukünftigen Lehrpersonen an solchen Lehrmethoden führt.
- Die Frage nach der Wirksamkeit lässt sich auch in Bezug auf die fachdidaktische Ausbildung stellen. Dieser Bestandteil der Lehrerausbildung spielt im angelsächsischen Raum allerdings eine geringe Rolle. Den wenigen „fachdidaktisch“ ausgerichteten Studien lässt sich – im Unterschied zu den bisher angesprochenen Merkmalen – ein konsistent positiver Zusammenhang mit Schülerleistungen bzw. Lehrerhandeln entnehmen. Das Verständnis von „Fachdidaktik“ ist in den USA allerdings vergleichsweise eng. Vor dem Hintergrund eines pragmatischen Konzepts von literacy werden vor allem fachspezifische unterrichtsmethodische und lernpsychologische Fragen behandelt, während in Deutschland zum einen auch normative Fragen wie die Ziele von Fachunterricht und sein Beitrag zur Allgemeinbildung sowie seine Geschichte und die außerschulische Rolle des Fachs thematisiert und zum

anderen die zugrunde liegenden fachlichen Inhalte erarbeitet werden. Letzteres wird deutlich an Seminartiteln wie „Lyrik im Unterricht (mit fachdidaktischem Anteil)“.

- In Bezug auf die erziehungswissenschaftliche Ausbildung sind die Ergebnisse erneut widersprüchlich. Einzelne Studien sprechen gegen ihre Wirksamkeit, in der Mehrheit sind aber positive Einflüsse auf Schülerleistungen bzw. Lehrerhandeln festgestellt worden. Zwei Aspekte sind in diesem Zusammenhang herauszustellen: So sind die positiven Effekte vermutlich nur bei einem Zusammenspiel von umfangreicher erziehungswissenschaftlicher und fachlicher Ausbildung zu verzeichnen. Darüber hinaus ist das Verständnis von Erziehungswissenschaft – wie im Fall der Fachdidaktik – in der angelsächsischen Lehrerbildung deutlich enger als in Mitteleuropa. Es gibt in der Regel nur Angebote, die unmittelbar auf den Lernprozess der Schülerinnen und Schüler bzw. auf das methodische Handeln der Lehrpersonen bezogen sind. Damit fehlen zum einen selbst Themen, die in Deutschland zum Kernbereich der Schulpädagogik gehören (z.B. Schultheorie), vor allem aber alle Angebote aus den Bereichen der Allgemeinen, Historischen und Vergleichenden Erziehungswissenschaft.
- Neben der Wirkung dieser Komponenten von Lehrerbildung wurde das Vorhandensein berufunspezifischer Fähigkeiten der Lehrpersonen untersucht. In Bezug auf ihr intellektuelles Niveau müssen widersprüchliche Zusammenhänge mit Schülerleistungen bzw. Lehrerhandeln festgehalten werden. Die verbalen Fähigkeiten einer Lehrperson sind dagegen durchgehend signifikant positiv mit diesen assoziiert.

2.3 Interpretation der vorliegenden Befunde

Angesichts der Widersprüchlichkeit der meisten Ergebnisse kann nicht überraschen, dass deren Interpretation ebenfalls widersprüchlich ausfällt. Seit einigen Jahren wird in den USA eine intensive Auseinandersetzung über die Bedeutung von gesetzlichen Regelungen der formalen Anforderungen an die Ausbildung von Lehrerinnen und Lehrern geführt. Auf der einen Seite stehen diejenigen, die sich für eine weitgehende Freigabe aller Vorschriften aussprechen. Letztlich solle der Markt das Problem regeln, indem die besten Lehrerinnen und Lehrer – ob mit oder ohne Ausbildung – über Angebot und Nachfrage gefunden würden. Diese Gruppe der so genannten „Deregulierer“ – vorrangig vertreten durch die Abell Foundation – berücksichtigt in ihren Reviews nur Ergebnisse, die auf der Basis von *large scale*-Studien entstanden sind und Drittvariablen wie soziale Herkunft etc. kontrollieren, die zudem die Wirksamkeit der Lehrerbildung anhand von Schülerleistungen messen und die schließlich in Zeitschriften mit externer Blindbegutachtung erschienen sind. Ergebnisse qualitativer Studien und quantitativer Studien mit kleineren Fallzahlen sowie von Studien, die sich nur auf Befragungen von Schulleitern stützen, das Lehrerwissen testen oder Lehrerhandeln beobachten, sind somit ausgeschlossen. Auf der anderen Seite steht die Gruppe derjenigen, die sich für eine Reform der Lehrerbildung in dem Sinne einsetzt, dass für alle Staaten eine verbindliche Mindestnorm an zu belegenden Inhalten (vor allem in der Kombination von Fachwissenschaft und Erziehungswissenschaft sowie angemessener schulpraktischer Ausbildungsanteile) sowie zu erbringenden Abschlüssen (Master) gefordert wird. Prominenteste Vertreterin der so genannten „Professionalisierer“ ist Darling-Hammond, die für ihre Reviews eingeschränkt auch Ergebnisse von Doktorarbeiten, qualitativen Studien und Beiträgen in Zeitschriften mit redaktionsinternem Review-Verfahren akzeptiert.

Darling-Hammond (2000, S. 1) zieht angesichts der Tendenz in der Mehrheit der Studien die Schlussfolgerung „that measures of teacher preparation and certification are by far the strongest correlates of student achievement in reading and mathematics, both before and after controlling for student poverty and language status“. Sie spricht sich daher für ein stärkeres staatliches Engagement aus, was Regelungen von Mindestanforderungen an die Lehrerbildung angeht.

Im Review der Abell Foundation wird ein Großteil der Studien, die positive Wirkungen der Lehrerausbildung belegen, nicht akzeptiert, da sie die zuvor festgelegten methodischen Anforderungen verfehlen. Die Stiftung kommt daher zu einer Schlussfolgerung, die der von Darling-Hammond entgegen gesetzt ist. In ihrem Hauptbericht lautet die Kernaussage, dass sich keine hinreichenden Belege dafür finden ließen, Personen mit abgeschlossener Lehrerausbildung würden höhere Schülerleistungen erzielen als Personen ohne Ausbildung (Abell Foundation 2001a, S. iii). Wenn überhaupt ein stabiler Prädiktor gefunden werden könne, dann nur außerhalb der Lehrerausbildung, bei den berufunspezifischen verbalen Fähigkeiten. Die Folgerungen der Abell Foundation lauten (ebd., S. vii f.): Abschaffung der Regularien im Hinblick auf Qualifikationen von Lehrpersonen, stattdessen lediglich die Forderung eines Bachelors (und zwar ohne nähere Spezifizierung der Fachrichtung, wie es in den USA in Bezug auf Privatschulen gilt; Walsh 2001, S. 6). Diese Formalie soll mit einem Einstellungstest kombiniert werden, der verbale Fähigkeiten misst sowie – bei Elementarschullehrern – ihre Fähigkeit, Lesen zu unterrichten, bzw. – bei Sekundarschullehrern – ihr Fachwissen. Wenn es überhaupt lehrerausbildende Bemühungen geben müsse, dann sollten diese weniger auf erziehungswissenschaftliche Qualifikationen zukünftiger Lehrerinnen und Lehrer als auf die Förderung ihrer allgemeinen Sprach- und Lesefertigkeiten gerichtet sein. Insgesamt solle – zu Lasten der Grundausbildung – stärker die Berufseingangsphase unterstützt werden (Abell Foundation 2001a, b).

Diese Unterschiede in der Interpretation der Ergebnisse gehen vor allem auf normative Unterschiede bei der Beantwortung der Frage zurück, was als Kriterium hoher Qualität von Lehrerausbildung akzeptiert wird. Standardisierte Erhebungen von Schülerleistungen können als das am weitesten reichende Kriterium angesehen werden, sind bessere Schülerleistungen doch das „ultimate goal of teacher education“ (Cochran-Smith 2001, S. 530). Allerdings beginnen die Kontroversen spätestens dann, wenn dieses Kriterium spezifiziert werden soll. Cochran-Smith (ebd.): „If we attempt to describe the relationship between teacher learning and professional practice, attempt to explain what we mean by teacher learning and student learning, attempt to elaborate the theoretical bases and consequences of the kind of student learning we are trying to account for [...], we uncover differences, some of which represent deep philosophical and political divides.“ So sieht beispielsweise von Prondczynski (2001, S. 94) schon im Heranziehen des Kriteriums der *teaching excellence* eine Reduktion der Lehrerausbildung „auf eine kurzfristige Praxis“. Hier spiegelt sich das oben bereits angesprochene Problem der Normativität vieler Entscheidungen im Rahmen empirischer Bildungsforschung.

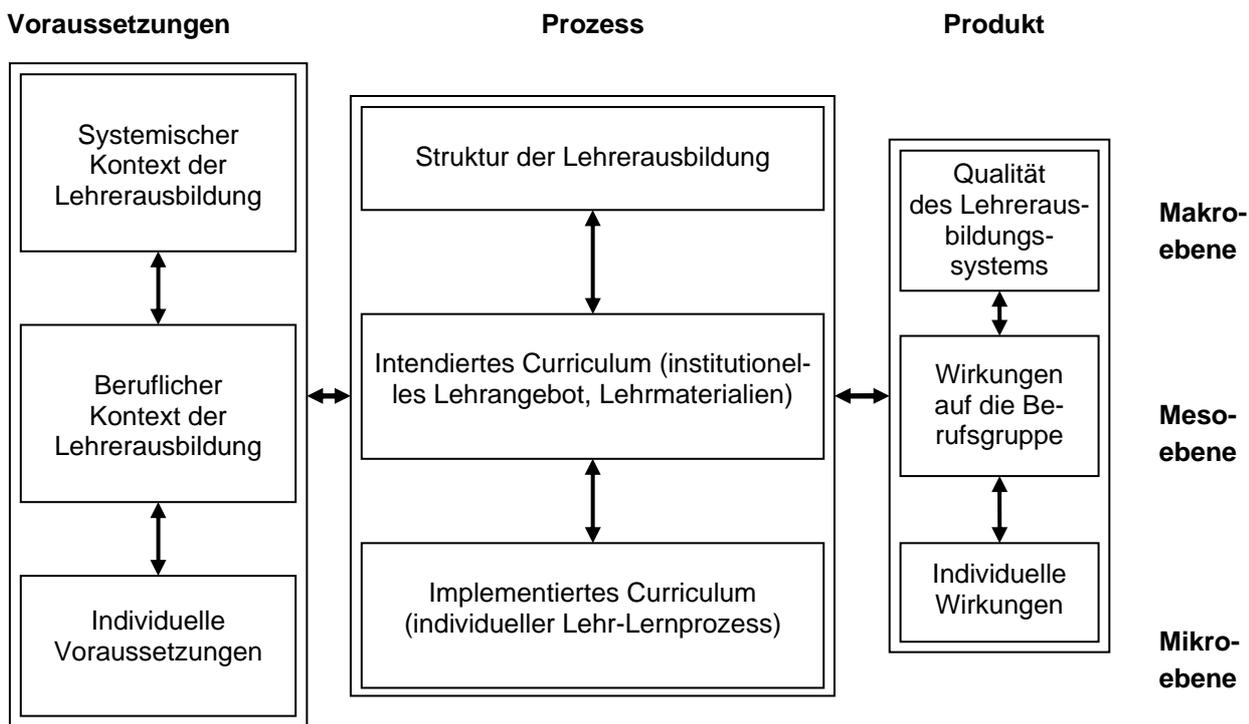
2.4 Zwischenfazit

Die Ausgangsfrage, ob eine hohe Qualität von Lehrerhandeln durch die Lehrerausbildung erreicht wird oder ob die Qualität eher Resultat von Persönlichkeitseigenschaften („Talent“) bzw. von langjähriger Berufspraxis („Übung“) ist – diese Frage kann auf der Basis der vorhandenen Befunde nicht beantwortet werden. Darüber hinaus werden in den dargestellten Reviews normative Unterschiede zur Frage deutlich, was als Kriterium hoher Qualität von Lehrerausbildung akzeptiert wird. Aus den bislang vorliegenden Studien lässt sich auch nicht ableiten, wie Lehrerausbildung im Detail strukturell und inhaltlich optimalerweise auszusehen hat. Wie viel Fachwissenschaft, Fachdidaktik und Erziehungswissenschaft werden mindestens benötigt? Sind Unterschiede nach Schulformen zu machen? Und wie sollen die jeweiligen Inhalte konkret aussehen? Eine Operationalisierung allein über Abschlüsse und Kurszählungen geht möglicherweise am Kern der Problematik von Lehrerausbildung vorbei. Dass sich diese Fragen bisher auf empirischer Basis nicht beantworten lassen, müssen allerdings sowohl die Verfechter der bestehenden Lehrerausbildung zur Kenntnis nehmen als auch ihre Gegner.

3 Aufgaben empirischer Forschung zur Lehrerausbildung

Vor diesem Hintergrund stellt sich die Aufgabe, die Durchführung einer quantitativen Studie zur Lehrerausbildung mit fachspezifischem Fokus bei gleichzeitiger Berücksichtigung der komplexen Zusammenhänge von erziehungswissenschaftlicher, fachdidaktischer, fachwissenschaftlicher und praktischer Ausbildung zu entwickeln, an einer kleineren Fallzahl von Befragten zu erproben und durchzuführen, um die Basis für eine international-repräsentative Erhebung zu legen. Entsprechende Vorhaben werden im Folgenden kurz skizziert² und anschließend in ihren Chancen und Grenzen reflektiert.

Die *International Association for the Evaluation of Educational Achievement* (IEA) wird in den Jahren 2005 bis 2009 in ca. 30 Ländern unter dem Titel „Learning to Teach Mathematics – Teacher Education and Development Study (TEDS)“ eine internationale Vergleichsstudie zur Wirksamkeit der Lehrerausbildung durchführen (Tatto 2004). In Deutschland hat sich eine Forschergruppe gebildet, die die Beteiligung der Bundesrepublik Deutschland, Österreichs und der Schweiz an dieser Studie anstrebt.³ Seit 2003 findet in acht Ländern (Bulgarien, Deutschland, England, Italien, Mexiko, Südkorea, Taiwan und USA) eine Vorstudie statt.⁴ Am Beispiel der Lehrerausbildung für die Sekundarstufe I im Unterrichtsfach Mathematik werden im Rahmen von Fallstudien Instrumente entwickelt, erprobt und auf ihre Einsetzbarkeit unter engen finanziellen Rahmenbedingungen geprüft. Eine Reihe an Datenerhebungen hat bereits stattgefunden (vgl. z.B. Blömeke 2004b, d). Beiden Studien liegt folgendes Mehrebenenmodell zur Wirksamkeit der Lehrerausbildung zugrunde (Blömeke 2003):



Grafik 1: Mehrebenenmodell der Wirksamkeit von Lehrerausbildung

Bei der Konzeptionierung solcher Studien besteht immer ein Zwiespalt zwischen Notwendigkeiten aus theoretischer Sicht und methodischer Machbarkeit. Mit dieser Modellierung greifen P-TEDS und TEDS zwei Traditionen der empirischen Bildungsforschung auf, in denen bedeutsame Zu-

sammenhänge herausgearbeitet wurden: die Wechselwirkung zwischen dem Lehr-Lernprozess und den Lernergebnissen als Produkt sowie die Wechselwirkung zwischen Lernvoraussetzungen und Lehr-Lernprozess. Darüber hinaus kommen neben dem Verlauf der Lehrerausbildung die unterschiedlichen gesellschaftlichen Ebenen, die an ihrer Gestaltung beteiligt sind, in den Blick.

Auf der Makroebene knüpfen die Studien mit ihrem internationalen Vergleich der Lehrerausbildungssysteme und des Lehrerberufs an die Vorarbeiten von Eurydice (2001, 2002, 2003, 2004) und der OECD (2004a, b) an. Sie gehen gleichzeitig darüber hinaus, indem eine systematische Verknüpfung der gewonnenen Erkenntnisse untereinander erfolgt und so eine Systemtypologie unter Berücksichtigung der kulturellen Kontexte entwickelt werden soll (Blömeke 2004b, d). Untersuchungsdimensionen sind die Ziele der Lehrerausbildung, die Leitideen ihrer Institutionalisierung, der gesellschaftliche Stellenwert von Lehrpersonen, ihre Arbeitsbedingungen, Grundorientierungen in Gesellschaft und Bildungssystem und die Kosten der Lehrerausbildung. In Bezug auf den Untersuchungsgegenstand sind Profile der Ausbildungswege von Mathematiklehrern zu zeichnen sowie näher zu bestimmen, was sich hinter Begriffen wie „Lehrerausbildungsinstitutionen“, „Lehrerausbildnern“ und „zukünftigen Lehrpersonen“ als den zentralen Untersuchungsebenen verbirgt.

Auf der Mesoebene untersuchen die Studien das intendierte und das implementierte Curriculum der Lehrerausbildung im Hinblick auf quantitative und qualitative Charakteristika. Dabei werden zum einen die Institutionen als Ganze, aber auch die Lehrerbildner mit ihrem Wissen, ihren Einstellungen und ihren demographischen Merkmalen in den Blick genommen. Was machen die einzelnen Universitäten und Studienseminare aus den staatlichen Vorgaben? Welche Inhalte werden konkret gelehrt? Wie sehen die Praxiserfahrungen der Auszubildenden aus? Trotz staatlicher Rahmenvorgaben können die Unterschiede im Detail groß sein. Als Beispiel kann auf die Vernetzung der Lehrveranstaltungen untereinander verwiesen werden: Diese kann innerhalb der Fächer, innerhalb der Fachdidaktiken und innerhalb der Erziehungswissenschaft geschehen oder querschnittlich, indem eine fachwissenschaftliche Lehrveranstaltung mit einer darauf ausgerichteten fachdidaktischen und einer erziehungswissenschaftlichen Lehrveranstaltung gekoppelt wird. Trotz eines quantitativ gleichen Lehrangebots sollten die beiden Richtungen mit einem qualitativen Unterschied einhergehen. Um Wirkungen von Lehrerausbildung angemessen beurteilen zu können, ist es wichtig, auch solche qualitativen Unterschiede zu erfassen. In Bezug auf den Prozess der Lehrerausbildung gilt es zudem, eine Perspektive in den Blick zu nehmen, die in einzelnen Professionstheorien als das „versteckte“ Curriculum bezeichnet wird. Es gibt Hinweise, dass es möglicherweise nicht so sehr ausschlaggebend ist, welche Inhalte dargeboten werden, sondern in welcher Form dies geschieht: ob zum Beispiel vorwiegend strukturiert anhand der Systematik der zugrunde liegenden akademischen Disziplin oder fallbasiert anhand von typischen beruflichen Situationen, ob zum Beispiel vorwiegend in Form von Präsentationen oder über aktive Aneignung durch die Studierenden.

Die Ergebnisse der Untersuchung auf der Mesoebene werden mit den Ergebnissen auf der Makroebene (Typen von Lehrerausbildungssystemen) in Beziehung gesetzt. Da an vielen Orten in den kommenden Jahren noch die traditionellen grundständigen Lehrerausbildungssysteme bestehen werden, während sich parallel Bachelor- und Masterstudiengänge etablieren, wird auch eine Evaluation dieser beiden Konzepte im Rahmen der Studien möglich.

Auf der Mikroebene werden Wissen, Einstellungen und Persönlichkeitsmerkmale (sowie ggf. Handeln) der zukünftigen Lehrpersonen untersucht und zu den Ergebnissen auf der Mesoebene (Qualität und Quantität des Lehrerausbildungs-Curriculums) in Beziehung gesetzt. Getestet werden mathematisches, mathematikdidaktisches und erziehungswissenschaftliches Wissen sowie entsprechende Einstellungen, um das gesamte Spektrum an Lehrerwissen zu erfassen, das in der Lehrerausbildung vermittelt werden soll. Auf allen Dimensionen und Subdimensionen erfolgt eine Fokussierung auf so genannte „big ideas“ im Sinne von empirisch oder theoretisch identifizierten fundamentalen Merkmalen der Lehrerausbildung, von denen eine enge Verbindung zu Lehrerwissen, Lehrerhandeln und Schülerleistungen erwartet wird. Dabei sollen nicht nur Studierende und

Referendare erfasst werden, sondern auch Lehrpersonen mit bis zu fünf Jahren Berufserfahrung. Dahinter steht die Absicht, konkurrierenden Professionstheorien gerecht zu werden. In einigen Ländern lassen sich Traditionslinien feststellen, denen zufolge ein Lehrer seinen Beruf erst in dem Moment wirklich lernt, in dem er zu unterrichten anfängt.

Die Instrumente müssen im Detail noch entwickelt werden. Das grundsätzliche Format sieht Dokumentenanalysen, Fragebögen und Tests vor. Da Lehrerhandeln eine stark situationsbestimmte Aktivität ist, wurde als Grundprinzip der Tests festgelegt, eine Unterrichtssituation darzustellen, gefolgt von offenen und geschlossenen Fragen: Was würden Sie tun? Wie würden Sie dies tun? Warum würden Sie dies tun? Darüber hinaus muss es möglich sein, professionelle *Entwicklung* abzubilden. Erste Entwürfe liegen vor (angefertigt von Mathematikern und Mathematikdidaktikern aus Australien, Belgien, Bulgarien, Japan und den USA sowie von Erziehungswissenschaftlern aus Australien, Deutschland, England, Italien und den USA).

Zum Untersuchungsdesign ist festzuhalten, dass die Untersuchungen in einem Quasi-Längsschnitt stattfinden werden. Auf Mikroebene werden Studienanfänger, Referendare zu Beginn und am Ende der zweiten Phase sowie junge Lehrer bis zum fünften Berufsjahr als Kohorten behandelt, bei denen die Erhebung zum selben Zeitpunkt stattfindet. Offen ist, ob auch Lehramtsstudierende am Ende ihres ersten Ausbildungsabschnitts befragt werden können. Zumindest in Deutschland mit seiner individuellen Ausbildungsorganisation, in der sich die Studierenden das Studium selbst organisieren und einteilen, dürfte es schwierig sein, diese Gruppe zu erfassen. Auf der Mesoebene werden zum einen Institutionen (Universitäten und Studienseminare) und zum anderen Personen (Lehrerausbildner der ersten und zweiten Ausbildungsphase) gezogen.

4 Chancen und Grenzen empirischer Forschung zur Lehrerausbildung

Abschließend werden die Chancen und Grenzen von empirischer Lehrerausbildungsforschung, wie sie hier dargestellt wurde, reflektiert. Zunächst gehe ich auf ihre Chancen ein, und zwar in bildungstheoretischer, methodologischer und entwicklungsorientierter Hinsicht.

4.1 Chancen empirischer Forschung zur Lehrerausbildung

Seit rund 250 Jahren wird in Deutschland eine intensive bildungstheoretische Diskussion über die Lehrerausbildung geführt. Dabei wird nicht nur über legitime Ziele diskutiert, es werden auch Wege zur Erreichung dieser Ziele entfaltet (Blömeke 2004c, Blömeke 2005, Blömeke & Paine 2005). Es stellt sich aber die Frage, worin eigentlich die Überzeugung begründet liegt, dass die jeweils entwickelten Modelle tragen. Von Prondczynski (2001, S. 98) spricht zu Recht von einem „Teufelskreis von Reform, Kritik, erneuter Reform und Kritik“. Diesen Teufelskreis wenigstens ansatzweise zu durchbrechen, indem zumindest die von Larcher und Oelkers so bezeichnete „Krise der Daten“ (s.o.) beendet wird, ist eine der Hauptchancen der skizzierten Vergleichsstudien.

Ein solches Vorhaben ist dabei auch aus bildungstheoretischer Perspektive interessant. Zu empirischen Erkenntnissen über die Lehrerausbildung kann man nicht nur in technokratischer Weise kommen. Auf der Schulebene ist es ein Unterschied, ob Schülerleistungen wissens- oder kompetenzorientiert erfasst werden. Es ist ein Unterschied, ob man Faktoren zum Unterrichtsklima und zur moralischen Entwicklung berücksichtigt (weil man sie aus sozialen Gründen für bedeutsam hält) oder ob man sie unberücksichtigt lässt (weil sie ggf. wenig zu kognitiven Schülerleistungen beitragen). Entsprechend ist es auf der Ebene der Lehrerausbildung ein Unterschied, ob man Persönlichkeitsmerkmale der zukünftigen Lehrpersonen mit aufnimmt oder sich nur auf die Erhebung ihres Fachwissens konzentriert, ob man den Lehr-Lernprozess in der Ausbildung einbezieht oder sich auf das Produkt konzentriert, ob man dieses Produkt über die Kognitionen der Lehrpersonen und ihr Handeln oder über die Leistungen ihrer Schülerinnen und Schüler definiert. Dahinter stehen unterschiedliche bildungstheoretische Ansprüche an Schule und Lehrerausbildung. Diese An-

sprüche lassen sich durchaus benennen – und damit in Variablen umsetzen. Auf diese Weise bildungstheoretisch fundiert, kann die Studie dann auch einen Beitrag zur erziehungswissenschaftlichen Theoriebildung leisten.

In diesem Zusammenhang stellt schon allein die für empirische Studien notwendige Modellbildung eine bildungstheoretische Chance dar. Erst ein Modell mit seiner systematischen Definition von gesellschaftlichen Ebenen, Stadien von Bildungsprozessen sowie Einflussfaktoren und Wirkungen macht abstrakte Bildungsfragen diskutierbar. Auf dieser Basis kann dann der rein funktionale Lesebegriff der PISA-Studien kritisiert werden (Kämper-van den Boogaart 2003, 2004, Spinner 2003). Und aus einer Tradition, die sich nicht vornehmlich an der Psychologie orientiert, kann dann der Beitrag der Unterrichtsfächer zur Allgemeinbildung kritisch und anders reflektiert werden, als dies in den PISA-Studien der Fall ist (Ladenthin 2003, Ofenbach 2003, Sacher 2003). Vergleichbare Diskussionen regen hoffentlich auch die anstehenden Studien zur Lehrerausbildung an. Das Problem der bisher eher wenig empirisch orientierten Erziehungswissenschaft und Fachdidaktiken ist nämlich, dass ihre Reflexionen zwar bedeutsam sind, dass in der Vergangenheit aber häufig Operationalisierungen ausblieben, die es ermöglichen, einen plausiblen Lösungsweg tatsächlich daraufhin zu überprüfen, ob er das leistet, was seine Verfechter anstreben. Benner (2002) formuliert in diesem Sinne zu Recht aus der Perspektive der Bildungstheorie: „Zwar lassen sich Bildungstheorie und Empirie nicht in ein Kontinuum aufheben. Die Spannungen zwischen bildungstheoretischen und empirischen Fragen wechselseitig fruchtbar zu machen, ist dagegen ein Unternehmen, das für beide Seiten lohnend sein könnte.“

Ergänzend zu den angeführten bildungstheoretischen Chancen kann ein methodologischer Mehrwert von internationalen Vergleichsstudien zur Lehrerausbildung festgestellt werden. In Bezug auf die Lehrerausbildung handelt es sich um den ersten Versuch, eine mehrebenenanalytische Herangehensweise umzusetzen. Am Ende zu Indikatoren für Qualität in der Lehrerausbildung zu kommen, die theoretisch und empirisch fundiert sind, wäre ein wichtiger Fortschritt. Darüber hinaus würden Universitäten und Studienseminaren Instrumente bereitgestellt, die sie adaptieren und so für eigene Evaluationszwecke nutzen können.

Neben diesen theoretischen und methodologischen Chancen besteht eine andere wichtige Perspektive der Studie darin, dass aus ihr fundierte Ansatzpunkte für eine Weiterentwicklung der Lehrerausbildung gewonnen werden können – insbesondere aufgrund ihrer internationalen Orientierung. Dass in Deutschland überhaupt wieder über das Bildungssystem diskutiert wird, dass andere Institutionalisierungsformen als die bestehenden in den Sinn kommen, ist eine Folge internationaler Vergleiche. Internationalität öffnet den Blick dafür, dass Systemstrukturen und Lehr-Lernprozesse auch ganz anders organisiert werden können, als es vor Ort der Fall ist. Im globalen Maßstab kann man die unterschiedlichen Institutionalisierungen der Lehrerausbildung mit einer quasi-experimentellen Situation vergleichen. Die Variation zwischen den Ländern ermöglicht einen Vergleich ihrer relativen Wirksamkeit (vgl. entsprechend Terhart 2004, S. 50, in Bezug auf die Chancen eines Bundesländer-Vergleichs). Man ist also noch nicht einmal gezwungen, normativ ein Kriterium für eine „gute“ Lehrerausbildung zu setzen. In einer ersten Annäherung würde es auch genügen festzustellen: Verglichen mit „xy“, wird unter diesen und jenen Umständen bei „yz“ mehr oder weniger erreicht.

4.2 Grenzen empirischer Forschung zur Lehrerausbildung

Von Prondczynski kritisiert, dass das Kriterium der Qualität von Lehrerhandeln Lehrerausbildung „auf eine kurzfristige Praxis“ reduziert (s.o.). Abgesehen davon, dass hier mit „Qualität“ vermutlich ein eingeschränktes Begriffsverständnis assoziiert wird, spricht von Prondczynski dennoch ein Grundsatzproblem an. Angesichts der normativen Ausrichtung von Bildungsprozessen und ihrer nur indirekten Messmöglichkeiten ist eines der gravierendsten Probleme empirischer Forschung, dass selbst bei komplexer Untersuchungsanlage „Bildung“ – je nach Begriffsverständnis – gar

nicht, mindestens aber nur schwer messbar ist. Das gilt vor allem, wenn Bildung nicht allein auf kognitive Leistungen reduziert werden soll. Aber selbst bei einer differenzierten Berücksichtigung affektiv-motivationaler Komponenten und berufunspezifischer Persönlichkeitsmerkmale wird man in jedem Fall den Anspruch an empirische Forschung zur Lehrerausbildung einschränken müssen. Der Technologisierbarkeit von Bildungsprozessen sind enge Grenzen gesetzt. Schülerleistungen zum alleinigen Maßstab für die Qualität von Lehrerausbildung zu nehmen, würde eine Linearität pädagogischer Prozesse unterstellen, die nicht vorhanden ist. Ein Ingenieur kann an der Beschaffenheit seines Werkes gemessen werden. Pädagogen haben es aber mit nicht-trivialen, autonomen Subjekten zu tun. Sie sind mit dem Paradox konfrontiert, etwas von außen zu wollen, was nur von den Subjekten selbst hervorgebracht werden kann. Schülerleistungen können daher immer nur ein Maßstab unter mehreren sein.

Bescheidenheit ist vermutlich auch angebracht, was die praktischen Folgen empirischer Studien angeht. Schon in den 50er Jahren wurde in den USA in Bezug auf das Schulsystem eine Mängel­liste aufgestellt, die eine analoge Beschreibung der gegenwärtig festgestellten Probleme sein könnte. Als Konsequenz wurde unter anderem in Bezug auf die Lehrerausbildung die Einführung von Standards gefordert, die überprüft werden sollten. Obwohl diese und andere Reformen durchgeführt wurden, wiederholen sich heute Kritik und Forderungen. Oelkers (2002a, S. 3) führt daher aus: „Warum fünfzig Jahre später immer noch keine adäquate Problemlösung vorliegt, bleibt unklar und wird zumeist gar nicht erst diskutiert, weil die historische Dimension fehlt. Die Frage nach der Wiederholung des Neuen stellt sich erst gar nicht, so dass bestimmte Ideen immer neu ins Spiel gebracht werden können, ohne sich auf Erfahrungen beziehen zu müssen.“

Vor diesem Hintergrund stellt sich die Frage, wem mit Studien wie den vorgestellten in *praktischer* Hinsicht „gedient“ wird. Folgt man dem dem Habermas'schen Verständnis von empirischer Forschung als „Orientierungsforschung“ im Interesse von Information und Aufklärung heißt das für die Bildungsforschung, dass ihre Erkenntnisse langfristig auch in der Schule oder anderen Bildungsrealitäten ankommen müssen. Empirische Studien können Hinweise liefern und die Aufmerksamkeit auf bedeutsame Einflussfaktoren lenken. Diese können dann wieder Ausgangspunkt für neue Überlegungen und Untersuchungen zur Verbesserung der Praxis sein. Die PISA-Studien haben beispielsweise Erkenntnisse zu Defiziten unseres Schulsystems und gleichzeitig Anhaltspunkte für Ursachen geliefert. Im nächsten Schritt muss es nun darum gehen, diese unter dem Blickpunkt, was verändert werden kann und soll, in ein Konzept umzusetzen, das dann erprobt und wiederum evaluiert werden muss.

Mit unmittelbarer Anwendung von empirischen Erkenntnissen in der Praxis darf das nicht verwechselt werden. Zu meinen, aus empirischen Studien direkt Handlungsanleitungen ableiten zu können, ist ein Trugschluss. Politikberatung, wie es der Begriff der „Evaluation“ in einem zentralen Verständnis nahe legt, ist daher nur schwer möglich – abgesehen von dem generellen Problem der Vermischung von Wertfreiheit im Begründungszusammenhang und legitimatorischen Interessen. Diese Skepsis hat gerade etwas mit dem Gegenstand von Erziehungswissenschaft zu tun, den Bildungsprozessen. In nur wenigen Ländern ist die Bildungsdebatte so stark politisiert, geradezu von politischen „Grabenkämpfen“ durchzogen wie in Deutschland. Das hängt unter anderem damit zusammen, dass zwar immer die Autonomie der Bildung betont wird, dass diese ihr implizit aber nicht zugestanden wird. Durch Bildungsreformen soll immer gleich die Gesellschaft verändert werden – oder es wird immer gleich die Gesellschaft in ihrem Bestand bedroht gesehen. Unter diesen Voraussetzungen zu erwarten, dass die Bildungsdebatte mit empirischen Erkenntnissen rational fundiert werden könne, ist erst einmal nur eine Hoffnung. Darüber hinaus bleiben die bei empirischer Forschung gewonnenen „harten“ Fakten immer interpretationsbedürftig. Cochran-Smith (2001, S. 532) formuliert hierzu: „No matter what conclusions are ultimately drawn about the evidence, they will depend to a great extent on complex and prior decisions about what counts and how it counts as evidence, decisions that are based on values, goals, and priorities. In that sense, the debate will never be settled solely through empirical means.“

Hinzu kommt, dass Bildungspolitik eine andere Rationalität zugrunde liegt als Wissenschaft. Politik ist notwendigerweise normativ orientiert und interessengeleitet. Wissenschaftliche Erkenntnisse sollten vor diesem Hintergrund zwar berücksichtigt werden, müssen aber nicht unbedingt handlungsleitend sein – ganz abgesehen von dem Problem, dass es „die“ Wissenschaft auch nicht gibt. Die Nach-PISA-Diskussion hat gezeigt, wie unterschiedlich die empirischen Ergebnisse in der Erziehungswissenschaft interpretiert werden und wie unterschiedlich wissenschaftliche Empfehlungen und politische Entscheidungen ausfallen können (vgl. in Bezug auf die Schulebene beispielsweise die Unterschiede zwischen Böttcher 2003, Brügelmann 2003, Buhlmann, Wolff & Klieme 2003, Klieme 2004, KMK 2003 a, b, c und Peters 2004 sowie in Bezug auf die Ebene der Lehrerbildung beispielsweise die Unterschiede zwischen Blömeke 2004e, Brandenburg & Keller 2002, KMK 2004, Oser 1997a, b, 2004 und Terhart 2002 sowie zwischen Döbrich et al. 2003 und KMK 2003d). Diese Phänomene sind Hinweise darauf, dass man Bildungsfragen unter verschiedenen normativen Perspektiven betrachten kann.

Vor diesem Hintergrund soll abschließend mit Bezug auf die aktuelle Diskussion um Bildungsstandards für Schule und Lehrerbildung auf mögliche gesellschaftliche Folgen von internationalen Vergleichsstudien zur Lehrerbildung eingegangen werden. Seit PISA ist eine latente Tendenz zu registrieren, von einer *input*-Steuerung des Bildungssystems – sei es Schule, sei es Lehrerbildung – abzuweichen und *output*-Erwartungen zu entwickeln. Ihre Erreichung soll dann durch Tests überprüft werden. Ohne die Sinnhaftigkeit dieser Überlegungen grundsätzlich in Frage stellen zu wollen, soll auf drei Problemkreise aufmerksam gemacht werden, die bisher möglicherweise nicht hinreichend diskutiert werden.

Zum ersten ist die vergleichsweise triviale Feststellung noch einmal herauszustellen, dass sich in Tests nur messbare Leistungen erfassen lassen. So wichtig es ist sicherzustellen, dass alle Kinder und Jugendlichen sowie Studierende und Referendare bestimmte Kenntnisse, Fertigkeiten und Fähigkeiten erwerben, so sehr würde eine ausschließliche Fokussierung von Schule und Lehrerbildung darauf eine Reduzierung bisheriger Bildungsansprüche bedeuten. Die Entwicklung in den USA, wo sich mehr und mehr Schulen im Interesse besserer Testleistungen ihrer Schülerinnen und Schüler auf den Unterricht in Mathematik und Englisch konzentrieren, sollte ein warnendes Beispiel sein. Statt einer vollständigen Umsteuerung auf eine Orientierung an *outputs* sowie deren (finanziell) folgenreichen Behandlung scheint im Interesse eines weiten Bildungsbegriffs eine Beibehaltung von – ggf. entschlackten – Curricula in Ergänzung zu Tests wichtig zu sein.

Ein zweites Problem ist, dass für die Entwicklung von Tests viel Spezialwissen benötigt wird. Dies ist der Anlass dafür, dass die Kultusministerkonferenz an der Humboldt-Universität ein An-Institut fördert, das diese Aufgabe für den Schulbereich übernehmen wird. Demokratietheoretisch liegt die Gestaltung und Kontrolle des Bildungssystems aber bei Öffentlichkeit und Politik. Angesichts der Tatsache, dass mit der Entwicklung von Tests zahlreiche normative Entscheidungen verbunden sind, wie oben ausgeführt wurde, wird es für die Legitimation des An-Instituts zentral sein, eine angemessene Balance zwischen wissenschaftlicher Fundierung und öffentlicher Diskussion über zugrunde liegende Konzepte zu finden.

In diesem Zusammenhang stellt sich drittens das Problem der Uneinheitlichkeit von „Wissenschaft“. Die nationalen und internationalen Vergleichsstudien zu Schule und Lehrerbildung dürfen nicht damit einhergehen, Vielfalt in der Erziehungswissenschaft und in den Fachdidaktiken zu reduzieren. Dass dies aber ggf. eine reale Gefahr ist, wird schon jetzt an den TIMS-, PISA- und IGLU-Studien deutlich: Eine andere Wahl als das dahinter stehende, sehr einheitliche Bildungsverständnis zu akzeptieren, besteht für Öffentlichkeit und Politik kaum. In einer pluralistischen Demokratie ist es aber nicht nur legitim, sondern sogar wünschenswert, dass unterschiedliche Auffassungen von Bildung existieren und sich gegenseitig herausfordern. Von herausragender Bedeutung wird es daher sein, dass sich in den universitären Disziplinen weiterhin unterschiedliche Bildungsverständnisse widerspiegeln. Auch in der Wissenschaft sind Monopole nicht unbedingt der beste Weg zur Erkenntnis.

Literatur

- Abell Foundation (2001a): Teacher Certification Reconsidered. Stumbling for Quality. Baltimore: Abell <http://www.abell.org/pubsitems/ed_cert_1101.pdf > [30.09.2003]
- Abell Foundation (2001b): Teacher Certification Reconsidered. Stumbling for Quality – Appendix. Review of Research Teacher Certification and Effective Teaching. Baltimore: Abell <http://www.abell.org/pubsitems/ed_cert_appendix_1101.pdf> [30.09.2003]
- Ball, D. L. & Bass, H. (2000): Interweaving content and pedagogy in teaching and learning to teach. Knowing and Using Mathematics. In: Boaler, J. (Hrsg.): Multiple Perspectives on the Teaching and Learning of Mathematics. Westport: Ablex, S. 83-104 <<http://www-personal.umich.edu/~dball/BallBassInterweavingContent.pdf> > [30.09.2003]
- Ball, D. L., Lubienski, S. T. & Mewborn, D. S. (2001): Research on Teaching Mathematics. The Unsolved Problem of Teachers' Mathematical Knowledge. In: Richardson, V. (Hrsg.): Handbook of Research on Teaching. Washington: American Educational Research Association (4. Aufl.), S. 433-456
- Baumert, J., Watermann, R. & Schümer, G. (2003): Disparitäten der Bildungsbeteiligung und des Kompetenzerwerbs. Ein institutionelles und individuelles Mediationsmodell. In: Zeitschrift für Erziehungswissenschaft 6 (1), S. 46-72
- Benner, D. (2002): Die Struktur der Allgemeinbildung im Kerncurriculum moderner Bildungssysteme. Ein Vorschlag zur bildungstheoretischen Rahmung von PISA. In: Zeitschrift für Pädagogik 48 (1), S. 68-90
- Block, J. H. & Hazelp, K. (1995): Teachers' beliefs and belief systems. In: Anderson, L. W. (Hrsg.): International encyclopaedia of teaching and teacher education. Oxford: Elsevier Science Ltd. (2. Aufl.), S. 25-28
- Blömeke, S. & Paine, L. (2005): Benefits and problems of doing empirical research work on teacher education on an international level (eingereicht)
- Blömeke, S. (2002): Universität und Lehrerbildung. Bad Heilbrunn/Obb.: Klinkhardt
- Blömeke, S. (2003): Learning to teach mathematics and science – A hierarchical linear model for case studies in six countries. Paper presented at the IEA TEDS Expert Panel Meeting (15-18 December 2003, Hamburg)
- Blömeke, S. (2004a): Empirische Forschung zur Wirksamkeit der Lehrerbildung. In: Blömeke, S., Reinhold, P., Tulodziecki, G. & Wildt, J. (Hrsg.) (2004): Handbuch Lehrerbildung. Bad Heilbrunn/ Braunschweig: Klinkhardt/ Westermann, S. 59-91
- Blömeke, S. (2004b): Grundlegende Strukturmerkmale und kulturelle Kontexte der Lehrerbildung im internationalen Vergleich (eingereicht)
- Blömeke, S. (2004c): How Do Elementary and Middle School Teachers "come to know"? The Case of German Teacher Education. Paper presented at the Education Policy Forum "What Teachers Need to Know" (Washington, D.C., 2004, June 2-3)
- Blömeke, S. (2004d): How efficient is the teacher education of mathematics teachers? Typology of cultural context factors and their connection to student achievement. Proposal submitted for the 15th ICMI study „The professional education and development of teachers of mathematics“ in Brazil 2005
- Blömeke, S. (2004e): Standards für die Lehrerbildung im Bereich der Berufswissenschaften. Vorlage für die Grundsätze-Arbeitsgruppe des Vizepräsidenten für Studium und Lehre der Humboldt-Universität zu Berlin (Stand: 24.10.2004)
- Blömeke, S. (2005): German teacher-education reform at the background of global tendencies in teacher education and accountability. Erscheint in: Tatto, M. T. (Hrsg.): Teacher Education Reform and Accountability Systems in a Global Time. A Comparative Analysis. Oxford (im Druck)
- Blömeke, S., Eichler, D. & Müller, Ch. (2003): Rekonstruktion kognitiver Strukturen von Lehrpersonen als Herausforderung für die empirische Unterrichtsforschung. In: Unterrichtswissenschaft 31 (2), S. 103-121
- Blömeke, S., Herzig, B. & Tulodziecki, G. (2005): Gestaltung von Schule. Eine Einführung in die Schulpädagogik. Bad Heilbrunn/Obb.: Klinkhardt (im Druck)
- Bölling, R. (1983): Sozialgeschichte der deutschen Lehrer. Ein Überblick von 1800 bis zur Gegenwart. Göttingen: Vandenhoeck und Ruprecht
- Böttcher, W. (2003): Bildung, Standards, Kerncurricula. Ein Versuch, einige Missverständnisse auszuräumen. In: Die Deutsche Schule (Weinheim) 95 (2), S. 152-165
- Brandenberg, M. & Keller, B. (2002): Entwicklung von Standards. Methodisches Vorgehen und Prozesserfahrungen. *journal für lehrerinnen- und lehrerbildung* 2 (2002) 1, S. 61-69
- Brouwer, N. & ten Brinke, St. (1995a): Der Einfluss integrativer Lehrerbildung auf die Unterrichtskompetenz (I). In: *Empirische Pädagogik* 9 (1), S. 3-31
- Brouwer, N. & ten Brinke, St. (1995b): Der Einfluss integrativer Lehrerbildung auf die Unterrichtskompetenz (II). In: *Empirische Pädagogik* 9 (3), S. 289-330
- Brügelmann, H. (2003): Das kurze Gedächtnis der großen Reformer. Anmerkungen zum Beitrag von Wolfgang Böttcher in diesem Heft. In: Die Deutsche Schule (Weinheim) 95 (2), S. 168-171

- Buhlman, E., Wolff, K. & Klieme, E. (2003): Zur Entwicklung nationaler Bildungsstandards. Eine Expertise. Berlin <http://www.bmbf.de/pub/zur_entwicklung_nationaler_bildungsstandards.pdf> [30.09.2003]
- Bullough, R.V. (1997): Becoming a teacher. Self and the social location of teacher education. In: Biddle, B. J., Good, T. L. & Goodson, I. F. (Hrsg.): International handbook of teachers and teaching. Dordrecht: Kluwer Academic Publishers, S. 87-148
- Chatterji, M. (2002): Models and Methods for Examining Standards-Based Reforms and Accountability Initiatives. Have the Tools of Inquiry Answered Pressing Questions on Improving Schools? In: Review of Educational Research 72 (3), S. 345-386
- Clark, C. (1988): Asking the right questions about teacher preparation. Contributions of research on teacher thinking. In: Educational Researcher 17 (2), S. 5-12
- Cochran-Smith, M. (2001): The outcomes question in teacher education. In: Teaching and teacher education 17, S. 527-546
- Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J. M., Mood, A., Weinfeld, F. D. & York, R.L. (1966): Equality of educational opportunity. Washington: Government Printing Office
- Criblez, L. & Wild-Näf, M. (1998): Lehrerbildungsforschung in der Schweiz. Forschungsdesiderata zur Stellung der Erziehungswissenschaften in der Lehrerbildung. In: Zeitschrift für Pädagogik (Weinheim) 44 (1), S. 21-39
- Dann, H.-D., Müller-Fohrbrod, G. & Cloetta, B. (1981): Sozialisation junger Lehrer im Beruf. „Praxisschock“ drei Jahre später. In: Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie 13 (3), S. 251-262
- Darling-Hammond, L. (2000): Teacher Quality and Student Achievement. A Review of State Policy Evidence. In: Education Policy Analysis Archives 8 (1) <<http://epaa.asu.edu/epaa/v8n1/>> [27.05.2003]
- Ditton, H. (2002): Unterrichtsqualität – Konzeption, methodische Überlegungen und Perspektiven. In: Unterrichtswissenschaft 30 (3), S. 197-212
- Döbrich, P., Klemm, K., Knauss, G. & Lange, H. (2003): Ausbildung, Einstellung und Förderung von Lehrerinnen und Lehrern (OECD-Lehrerstudie). Ergänzende Hinweise zu dem Nationalen Hintergrundbericht (CBR) für die Bundesrepublik Deutschland, Bericht für die Kultusministerkonferenz
- Drechsel, B. (2001): Subjektive Lernbegriffe und Interesse am Thema Lernen bei angehenden Lehrpersonen. Münster u.a.: Waxmann (= Internationale Hochschulschriften; 363)
- Eurydice (2001): Schlüsselthemen im Bildungsbereich in Europa. Bd. 3: Der Lehrerberuf in Europa: Profil, Tendenzen und Anliegen. Bericht I: Lehrerausbildung und Maßnahmen für den Übergang in das Berufsleben – Allgemein bildender Sekundarbereich I <<http://www.eurydice.org/Documents/KeyTopics3/de/KT3R1DE00.pdf>> [17.09.2004]
- Eurydice (2002): Schlüsselthemen im Bildungsbereich in Europa. Bd. 3: Der Lehrerberuf in Europa: Profil, Tendenzen und Anliegen. Bericht II: Angebot und Nachfrage – Allgemein bildender Sekundarbereich I <<http://www.eurydice.org/Documents/KeyTopics3/de/FrameSet2.htm>> [17.09.2004]
- Eurydice (2003): Schlüsselthemen im Bildungsbereich in Europa. Bd. 3: Der Lehrerberuf in Europa: Profil, Tendenzen und Anliegen. Bericht III: Beschäftigungsbedingungen und Gehälter – Allgemeinbildender Sekundarbereich I <<http://www.eurydice.org/Documents/KeyTopics3/de/FrameSet3.htm>> [17.09.2004]
- Eurydice (2004): Key Topics in Education in Europe. Bd. 3: The Teaching Profession in Europe: Profile, Trends und Concerns. Report IV: Keeping Teaching Attractive for the 21st Century – General Lower Secondary Education <<http://www.eurydice.org/Documents/KeyTopics3/en/FrameSet4.htm>> [17.09.2004]
- Fried, L. (1998): Zwischen Wissenschaftsorientierung und Orientierung an der Berufspraxis – Bilanz der Lehrerbildungsforschung. In: Empirische Pädagogik 12 (1), S. 49-90
- Gruber, H. & Renkl, A. (2000): Die Kluft zwischen Wissen und Handeln. Das Problem des trägen Wissens. In: Neuweg, G. H. (Hrsg.): Wissen, Können, Reflexion. Ausgewählte Bestimmungen. Innsbruck: Studienverlag, S. 155-174
- Helmke, A. & Jäger, R. S. (Hrsg.) (2002): Das Projekt MARKUS. Mathematik-Gesamterhebung Rheinland-Pfalz. Kompetenzen, Unterrichtsmerkmale, Schulkontext. Landau: VEP (= VEP-Aktuell; 2)
- Herrmann, U. (2003): „Bildungsstandards“ – Erwartungen und Bedingungen, Grenzen und Chancen. In: Zeitschrift für Pädagogik 49 (5), S. 625-639
- Kaelble, H. (1978): Probleme der Modernisierung in Deutschland. Sozialhistorische Studien zum 19. und 20. Jahrhundert. Opladen: Westdeutscher Verlag
- Kagan, D. M. (1992): Professional Growth Among Preservice and Beginning Teachers. In: Review of Educational Research 62 (2), S. 129-169
- Kämper-van den Boogart, M. (2003): Lesekompetenzen – Hauptsache flexibel. Zu einer Parallele zwischen Literaturdidaktik und empirischer Lesepsychologie. In: Abraham, U./ Bremerich-Vos, A./ Frederking, V./ Wieler, P. (Hrsg.): Deutschdidaktik und Deutschunterricht nach PISA. Freiburg/Br.: Fillibach, S. 26-46
- Kämper-van den Boogart, M. (2004): PISA und die Interpretationsrituale des Deutschunterrichts. In: Kämper-van den Boogart, M. (Hrsg.): Deutschunterricht nach der PISA-Studie. Reaktionen der Deutschdidaktik. Frankfurt/M. u.a.: Peter Lang, S. 59-81
- Kane, R., Sandretto, S. & Heath, Ch. (2002): Telling Half the Story. A Critical Review of Research on the Teaching Beliefs and Practices of University Academics. In: Review of Educational Research 72 (2), S. 177-228
- Klieme, E. (2004): Was sind Kompetenzen und wie lassen sie sich messen? In: Pädagogik 56 (6), S. 10-13

- Klinzing, H. G. (2002): Wie effektiv ist Microteaching? Ein Überblick über fünfundreißig Jahre Forschung. In: Zeitschrift für Pädagogik 48 (2), S. 194-214
- Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (KMK 2003a): Bildungsstandards im Fach Deutsch für den Mittleren Schulabschluss. Bonn: KMK
- Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (KMK 2003b): Bildungsstandards für die erste Fremdsprache (Englisch/ Französisch) für den Mittleren Schulabschluss. Bonn: KMK
- Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (KMK 2003c): Bildungsstandards im Fach Mathematik für den Mittleren Schulabschluss. Bonn: KMK
- Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (KMK 2003d): Ausbildung, Einstellung und Förderung von Lehrerinnen und Lehrern. Nationaler Hintergrundbericht (CBR) für die Bundesrepublik Deutschland (Ms.)
- Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (KMK 2004): Standards für die Lehrerbildung: Bildungswissenschaften. Entwurf v. 06.10.2004
- Koch, J.-J. (1972): Lehrer-Studium und Beruf. Einstellungswandel in den beiden Phasen der Ausbildung. Ulm: Süddeutsche Verlagsgesellschaft
- Ladenthin, V. (2003): PISA – Recht und Grenzen einer globalen empirischen Studie. Eine bildungstheoretische Betrachtung. Vierteljahrschrift für Wissenschaftliche Pädagogik, 79 (3), S. 354-375
- Larcher, S. & Oelkers, J. (2004): Deutsche Lehrerbildung im internationalen Vergleich. In: Blömeke, S., Reinhold, P., Tulodziecki, G. & Wildt, J. (Hrsg.): Handbuch Lehrerbildung. Bad Heilbrunn/ Braunschweig: Klinkhardt/ Westermann, S. 128-150
- Lortie, D. (1975): School teacher. A sociological study. Chicago: University of Chicago Press
- Mandl, H. & Gerstenmaier, J. (Hrsg.) (2001): Die Kluft zwischen Wissen und Handeln. Empirische und theoretische Lösungsansätze. Göttingen u.a.: Verlag für Psychologie
- Merzyn, G. (2002): Stimmen zur Lehrerbildung. Überblick über die Diskussion zur Gymnasiallehrerbildung, basierend vor allem auf Stellungnahmen von Wissenschafts- und Bildungsgremien sowie auf Erfahrungen von Referendaren und Lehrern. Hohengehren: Schneider
- Monk, D. H. & King, J. A. (1994): Multilevel teacher resource effects in pupil performance in secondary mathematics and science. The case of teacher subject matter preparation. In: Ehrenberg, R. G. (Hrsg.): Choices and consequences. Contemporary policy issues in education. Ithaca: ILR Press, S. 29-58
- Murnane, R. J. & Phillips, B. R. (1981): What do effective teachers of inner-city children have in common? In: Social Science Research 10, 83-100
- OECD (2004a): Education at a Glance. OECD Indicators 2004. Paris: OECD
- Organisation for Economic Co-operation and Development, Directorate for Education, Education and Training Policy Division (OECD 2004b): Attracting, Developing and Retaining Effective Teachers. Synthesis Report: First Draft. Paris: OECD
- Oelkers, J. (2002a): Die Entwicklung der Lehrerbildung als internationales Problem? Vortrag im Rahmen der Ringvorlesung der Universität Erfurt „Herausforderung der Bildungsgesellschaft“ am 25. Juni 2002 in der Michaeliskirche in Erfurt. <<http://www.paed.unizh.ch/ap/ErfurtLAB.rtf>> [30.09.2003]
- Oelkers, J. (2002b): Befunde und Fragen zur Wirksamkeit der Lehrerbildung. In: Daschner, P./ Drews, U./ Durdel, A. (Hrsg.): Kursbuch Referendariat. Weinheim/ Basel: Beltz 4. Aufl., S. 152-174
- Ofenbach, B. (2003): Zur Differenz von pädagogisch und ökonomisch motivierter Leistungsbeurteilung. Pädagogische Rundschau 57 (2003) 4, S. 391-398
- Oser, F. (1997a): Standards in der Lehrerbildung. Teil I: Berufliche Kompetenzen, die hohen Qualitätsmerkmalen entsprechen. In: Beiträge zur Lehrerbildung 15, S. 26-37
- Oser, F. (1997b): Standards in der Lehrerbildung. Teil II: Wie werden Standards in der schweizerischen Lehrerbildung erworben? Erste empirische Ergebnisse. In: Beiträge zur Lehrerbildung 15, S. 210-228
- Oser, F. (2004): Standardbasierte Evaluation der Lehrerbildung. In: Blömeke, S., Reinhold, P., Tulodziecki, G. & Wildt, J. (Hrsg.): Handbuch Lehrerbildung. Bad Heilbrunn/ Braunschweig: Klinkhardt/ Westermann, S. 184-206
- Oser, F. & Oelkers, J. (Hrsg.) (2001): Die Wirksamkeit der Lehrerbildungssysteme. Von der Allrounderbildung zur Ausbildung professioneller Standards. Chur/ Zürich: Rüegger (= Nationales Forschungsprogramm; 33)
- Pajares, F. (1992): Teachers' beliefs and educational research. Cleaning up a messy construct. In: Review of Educational Research 62 (3), S. 307-332
- Peters, J. (2004): Qualitätssicherung im Deutschunterricht. Probleme praktischer Umsetzung der KMK-Bildungsstandards. In: Pädagogik 56 (2004) 6, S. 14-17
- Prondczynski, A. von (2001): Evaluation der Lehrerbildung in den USA. Geschichte, Methoden, Befunde. In: Keiner, E. (Hrsg.): Evaluation (in) der Erziehungswissenschaft. Weinheim/ Basel: Beltz (= Beiträge zur Theorie und Geschichte der Erziehungswissenschaft; 23), S. 91-140
- Richardson, V. & Placier, P.: Teacher Change. In: Richardson, V. (Hrsg.): Handbook of Research on Teaching. Washington: American Educational Research Association (4. Aufl.), S. 905-947
- Richardson, V. (1996): The role of attitudes and beliefs in learning to teach. In: Sikula, J., Buttery, T. & Guyton, E. (Hrsg.): Handbook of Research on Teacher Education. New York: Macmillan (2. Aufl.), S. 102-119

- Rowan, B., Chiang, F. S. & Miller, R. L. (1997): Using research on employees' performance to study the effect of teachers on student performance. In: *Sociology of Education* 70 (4), S. 256-284
- Sacher, W. (2003): Schulleistungsdiagnose – pädagogisch oder nach dem Modell PISA? *Pädagogische Rundschau* 57 (2003) 4, S. 399-417
- Schaefers, Ch. (2002): Forschung zur Lehrerbildung in Deutschland – eine bilanzierende Übersicht der neueren empirischen Studien. In: *Schweizerische Zeitschrift für Bildungswissenschaften* 24 (1), S. 65-88
- Schlee, J. (1992): Empirische Forschung zur Lehrerbildung. In: Ingenkamp, K., Jäger, R. S., Petillon, H. & Wolf, B. (Hrsg.): *Empirische Pädagogik 1970-1990. Eine Bestandsaufnahme der Forschung in der Bundesrepublik Deutschland*. Weinheim: Deutscher Studien Verlag, S. 558-565
- Spinner, K. H. (2003): Standards für einen kreativen Deutschunterricht? *Deutschunterricht* 56 (5), S. 37-41
- Tatto, M. T., Schwille, J., Schmidt, W., Ingvarson, L., Beavis, A., Cogan, L., Ferrini-Mundy, J., Foy, P., Paine, L., Plank, D. & Rodrigues, E. (2004): *Conceptual Framework and Plan for IEA First Teacher Education Study in Mathematics (TEDS-M)*. o.O.: IEA 2. Aufl.
- Terhart, E. (2002): *Standards für die Lehrerbildung. Eine Expertise für die Kultusministerkonferenz*. Münster: Westfälische Wilhelms-Universität/ Institut für Schulpädagogik und Allgemeine Didaktik
- Terhart, E. (2004): *Struktur und Organisation der Lehrerbildung in Deutschland*. In: Blömeke, S., Reinhold, P., Tulodziecki, G. & Wildt, J. (Hrsg.) (2004): *Handbuch Lehrerbildung*. Bad Heilbrunn/ Braunschweig: Klinkhardt/ Westermann, S. 37-59
- Terhart, E., Czerwenka, K., Ehrich, K., Jordan, F. & Schmidt, H. J. (1994): *Berufsbiographien von Lehrern und Lehrerinnen*. Frankfurt/M.: Lang
- Ulich, K. (2000): *Lehrer/innen-Ausbildung im Urteil der Betroffenen. Ergebnisse und Folgerungen*. In: *Die Deutsche Schule* 88 (1), S. 81-97
- Walsh, K. (2001): *Teacher Certification Reconsidered. Stumbling for Quality – A Rejoinder*. Baltimore: Abell Foundation <http://www.abell.org/pubsitems/ed_cert_rejoinder_1101.pdf> [30.09.2003]
- Wayne, A. J. & Youngs, P. (2003): *Teacher Characteristics and Student Achievement Gains*. In: *Review of Educational Research* 73 (1), S. 89-122
- Weinert, F. E. & Helmke, A. (1997): *Entwicklung im Grundschulalter*. Weinheim: Psychologie Verlags Union
- Wenglinsky, H. (2002): *How Schools Matter. The Link Between Teacher Classroom Practices and Student Academic Performance*. In: *Education Policy Analysis Archives* 10 (12) <<http://epaa.asu.edu/epaa/v10n12/>> [07.11.2004]
- Wideen, M., Mayer-Smith, J. & Moon, B. (1998): *A Critical Analysis of the Research on Learning to Teach. Making the Case for an Ecological Perspective on Inquiry*. In: *Review of Educational Research* 68 (2), S. 130-178
- Wilson, S. M., Floden, R. E. & Ferrini-Mundy, J. (2001): *Teacher Preparation Research. Current Knowledge, Gaps, and Recommendations*. Washington: Center for the Study of Teaching and Policy <<http://depts.washington.edu/ctpmail/PDFs/TeacherPrep-WFFM-02-2001.pdf>> [30.09.2003]
- Wright, S. P., Horn, S. P. & Sanders, W. L. (1997): *Teacher and classroom context effects on student achievement. Implications for teacher evaluation*. In: *Journal of Personnel Evaluation in Education* 11, S. 57-67
- Zeichner, K. & Tabachnik, B. R. (1981): *Are the Effects of University Teacher Education Washed Out by School Experience?* In: *Journal of Teacher Education* 32, S. 7-11
- Zeichner, K. M. (1986): *Lehrersozialisation und Lehrerbildung. Forschungsstand und Perspektiven*. In: *Bildung und Erziehung* 39 (3), S. 263-278

¹ Aktualisierte und für die Drucklegung aufbereitete Fassung der Antrittsvorlesung. Für kritische Hinweise danke ich Hans-Günther Bracht, Christiane Buchholtz, Dana Eichler, Anja Felbrich, Christiane Müller, Sebastian Hacke, Annegret Hilligus, Peter Reinhold, Heike Schaumburg, Peter Strutzberg und Gerhard Tulodziecki.

² Der Ruf an die Humboldt-Universität zu Berlin eröffnete der Verfasserin die Möglichkeit, die in der empirischen Lehrerbildungsforschung bestehenden Forschungsdefizite anzugehen. Die enge Anbindung der HU an die *International Association for the Evaluation of Educational Achievement* (IEA, deren deutscher Repräsentant Rainer Lehmann ist, Professor für Empirische Bildungsforschung und Methodenlehre im Institut für Erziehungswissenschaften der HU), das hervorragende Umfeld für interdisziplinäre empirische Forschung und die zahlreichen internationalen Kontakte sind ideale Voraussetzungen für einen internationalen Vergleich der Lehrerbildung. Entsprechend wurde die Verfasserin in das internationale Expertenpanel berufen, das die im Folgenden dargestellte IEA-Studie vorbereitet. Sie erhielt zudem den Zuschlag für die Organisation und Durchführung der Vorstudie P-TEDS in der Bundesrepublik Deutschland.

³ Die Forschergruppe besteht derzeit (Stand: 10.11.2004) aus: Sigrid Blömeke/ Anja Felbrich (HU Berlin, Erziehungswissenschaft), Ewald Terhart (WWU Münster, Erziehungswissenschaft), Eckard Klieme/ Hermann Josef Abs (DIPF Frankfurt/M., Empirische Bildungsforschung), Rainer Lehmann (HU Berlin, Empirische Bildungsforschung), Gabriele Kaiser (U Hamburg, Didaktik der Mathematik), Wolfgang Schulz (HU Berlin, Didaktik der Mathematik), Hermann Lange (Hamburg, Bildungsplanung/ -ökonomie), Fritz Oser (Fribourg/CH, Pädagogische Psychologie), Johannes Mayr (Linz/A, Pädagogische Psychologie)

⁴ Die Vorstudie wird von der US-amerikanischen National Science Foundation unter dem Stichwort „P-TEDS: Learning to teach mathematics“ (REC-0231886) und von der Alexander von Humboldt-Stiftung sowie vom Bundesministerium für Bildung und Forschung im Rahmen des TransCoop-Programms gefördert (DEU/1117152).