[Quelle: Kraler, Christian & Schratz, Michael (Hrsg.) (2007), Ausbildungsqualität und Kompetenz im Lehrerberuf. Wien: LIT (= Österreichische Beiträge zur Bildungsforschung; 4), S. 191-208]

Sigrid Blömeke

Messung der professionellen Kompetenz zukünftiger Lehrpersonen

Standards empirischer Lehrerausbildungsforschung, bildungstheoretische Herausforderungen und exemplarische Ergebnisse einer Studie in Deutschland

Vor dem Hintergrund einer Einschätzung des Forschungsstands empirischer Lehrerausbildungsforschung anhand von 10 Standards, die von der Autorin in Anlehnung an übergreifende Qualitätsmerkmale empirischer Bildungsforschung entwickelt worden sind, werden die Erkenntnisse aus Studien resümiert, die den Standards genügen. Anschließend werden zwei Forschungsprojekte vorgestellt, in denen erstmals versucht wird, die professionelle Kompetenz angehender Lehrpersonen auf dem avisierten methodischen Niveau zu messen.

1 Standards für die empirische Lehrerausbildungsforschung

Nachdem der Erziehungswissenschaft und den Fachdidaktiken lange Zeit vorgeworfen wurde, sich der empirischen Überprüfung ihrer Postulate zu entziehen, hat sich die Unterrichtsforschung in den letzten 20 Jahren in so raschem Tempo entwickelt, dass ihr heute eine hohe Qualität bescheinigt werden kann. In Bezug auf die Lehrerausbildungsforschung ist dies noch keineswegs der Fall. Die folgenden 10 Standards sollen einen Maßstab darstellen, an dem die Qualität künftiger Vorhaben gemessen werden kann. Zentral ist:

1. dass *anwendungsorientiert* versucht wird, Antworten auf drängende Fragen der Lehrerausbildung zu finden, ohne die gleichzeitig notwendige intensive *Grundlagenforschung* zu vernachlässigen,
2. dass umfangreiche *large-scale assessments* durchgeführt, aber auch vertiefende *Fallstudien* vorgelegt werden,
3. dass hypothesen*prüfend* vorgegangen wird und gleichzeitig hypothesen*generierende* Studien existieren,
4. dass die Studien stärker auf *nachhaltige* Ergebnisse, denn auf kurzfristige Wirkungen bedacht sind,
5. dass in ihnen *unterschiedliche Messkriterien* berücksichtigt werden,
6. dass sie auch bei Erhebungen auf der Individualebene *institutionenübergreifend* angelegt sind,
7. dass Studien das *komplexe Wirkungsgeflecht* von individuellen, institutionellen und nationalen Einflussfaktoren erfassen, auch in internationalen Vergleiche,

8. dass die Studien *anschlussfähig* an zentrale Theorien der Bildungsforschung sind,
9. dass sie *kumulativ* auf vorliegenden Erkenntnissen aus anderen Studien aufbauen
10. und dass sie von einem *interdisziplinären Team* an fachkundigen Wissenschaftlern ausgeführt werden, die die verschiedenen notwendigen Spezialisierungen abdecken.

Derzeit muss zu jedem der zehn Standards festgehalten werden, dass die Lehrerausbildungsforschung erhebliche Defizite aufweist (vgl. im Detail auch Blömeke, 2004; Cochran-Smith & Zeichner, 2005):

- Die Lehrerausbildungsforschung ist stark von aktuellen Fragen geleitet, die zu einem schnellen Wechsel von Forschungsschwerpunkten führen, sodass kein kumulativer Aufbau einer gesicherten Erkenntnisbasis zur Lehrerausbildung stattfindet.
- Die Lehrerausbildungsforschung ist überwiegend evaluativ angelegt. Es fehlen sowohl systematische, wissenschaftlich fundierte standardisierte Studien als auch anspruchsvolle qualitative Studien.
- Die Lehrerausbildungsforschung ist häufig auf die eigene Region, wenn nicht sogar auf die eigene Ausbildungsinstitution bezogen, an der die jeweiligen Forscher/innen tätig sind. Wenn es sich um eine neue Ausbildungsidee handelt, wird die Forschung häufig sogar von *den* Kolleg/innen durchgeführt, die die Idee selbst entwickelt haben. Damit ist man bei der Frage, ob solche Studien eigentlich scheitern können.
- Und es stellt sich die Frage, ob die in der Regel gewählte Analyseeinheit der Studierenden eigentlich die richtige ist. Maßnahmen der Lehrerausbildung werden gruppenweise von den Lehrerausbildner/innen einer Institution übermittelt. Wenn die Zuordnung auch nicht so eindeutig ist wie in der Unterrichtsforschung, erscheint es doch angemessener, die Ausbildungsinstitution als korrekte Analyseeinheit anzusehen (mit allen Konsequenzen: Wahl angemessen definierter Vergleichsgruppen, hinreichende Anzahl für Mehrebenenanlaysen etc.).
- In vielen Projekten wird die Forschung mit lokal entworfenen Untersuchungsinstrumenten durchgeführt. Für andere Zwecke sind sie damit kaum brauchbar (Stichwort: Anschlussfähigkeit). Zudem können die Instrumente den Verdacht unkontrollierter *ad hoc*-Entwicklung manchmal nur schwer ausräumen.
- Die fehlende Anschlussfähigkeit auf der Ebene von Instrumenten spiegelt häufig eine fehlende Anschlussfähigkeit in theoretischer Hinsicht. Manche Studie baut auf einem individuell zusammengesetzten Theorienkonglomerat auf, das den Transfer in andere Studien erschwert.
- Die fehlende Anschlussfähigkeit geht zum Teil so weit, dass eine Idee nicht einmal im eigenen Arbeitszusammenhang mit Anschlussprojekten weiterverfolgt oder vertieft wird. Die vielen punktuellen Dissertationen,

die in der Erziehungswissenschaft produziert werden, zeugen davon. Langzeitstudien finden sich sowieso so gut wie nie.

- Und schließlich: Die Studien werden in der Regel von einer begrenzten Anzahl an Wissenschaftler/innen durchgeführt, sodass ein Defizit an Breite in der Kompetenz festzustellen ist. Lehrerausbildungsforschung kann aufgrund ihrer Komplexität nur im angemessenen Zusammenwirken von Erziehungswissenschaftler/innen, Psycholog/innen, Fachdidaktiker/innen und Fachwissenschaftler/innen durchgeführt werden. Aus meiner Sicht ist es schlicht nicht hinreichend, wenn Erziehungswissenschaftler/innen oder Psycholog/innen allein zum Beispiel Analysen zur Lehrqualität in der Lehrerausbildung durchführen. Das bedeutet eine Unterschätzung fachinhaltlicher Expertise. Umgekehrt gilt, dass die Fachdidaktiken allein die Reichweite ihrer Erkenntnisse in der Regel nicht einschätzen können, da ihnen die Erfahrung mit übergreifenden Fragestellungen und der Blick über das eigene Fach hinaus fehlen.

Viele dieser Kritikpunkte sind grundsätzlicher Art und hängen mit der tradierten Arbeitsweise von Erziehungswissenschaftler/innen zusammen – und trifft in Teilen selbst noch die Unterrichtsforschung (National Research Council, 2002; Helmke, 2004; Confrey, 2006; What Works Clearinghouse <http://www.whatworks.ed.gov>). Der Erziehungswissenschaft fällt es offensichtlich schwer, in Abgrenzung zu Psychologie und Soziologie einen eigenständigen, genuin bildungstheoretisch fundierten empirischen Forschungszugang zu entwickeln.

2 Bildungstheoretische Herausforderungen

Es fehlt ein grundsätzliches Modell, das die Wirkungsfaktoren der Lehrerausbildung bildungstheoretisch fundiert erfasst und das auch die zentrale abhängige Variable im Detail modelliert, die professionelle Kompetenz von Lehrerinnen und Lehrern. Die überzeugendste Konzeptualisierung von Kompetenz stammt derzeit von Weinert (1999, 2001). Weinert differenziert Kompetenz in die „bei Individuen verfügbaren oder durch sie erlernbaren kognitiven Fähigkeiten und Fertigkeiten, um bestimmte Probleme zu lösen, sowie die damit verbundenen motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten, um die Problemlösungen in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können“.

Forschungsmethodisch ist dieser Anspruch des Abrufens und Vernetzens von komplexen Wissensbeständen in berufstypischen Situationen eine Herausforderung. Oser (1997a, b) plädiert daher für ein besonders Item-Format in der Lehrerausbildungsforschung. Um das situationsbezogene Zusammenspiel von Wissen und Überzeugungen zu erfassen und damit der Gefahr zu entgegen, lediglich „träges“ systematisches, disziplinorientiertes Wissen zu messen, sollen Tests aus

solchen berufstypischen Situationen bestehen. Solche echten Kompetenztests sind selbst in der Unterrichtsforschung außerhalb von Großstudien wie PISA selten. In der Lehrerausbildungsforschung gibt es sie so gut wie nicht.
Eine solche Perspektive ist zudem nicht unumstritten. Kritisiert wird daran beispielsweise, dass sie Lehrerausbildung technologisch „auf eine kurzsichtige Praxis“ reduziere (Prondczynski, 2001). Angesichts der normativen Ausrichtung von Bildungsprozessen und ihrer nur indirekten Messmöglichkeiten sei „Bildung“ und insofern auch Lehrerausbildung nicht messbar. Das gelte vor allem, wenn man Bildung nicht auf kognitive Leistungen reduzieren wolle.
Mit dieser Kritik ist das Verhältnis von Bildung und Kompetenz angesprochen sowie die Frage der Messmöglichkeiten:
Was in empirischen Studien erfasst wird, ist Performanz, also aktuelles Verhalten. Dieses Verhalten wird als Indikator für Kompetenz genommen. Einigkeit dürfte in der Einschätzung bestehen, dass Bildung bzw. Lehrerausbildung deutlich darüber hinaus geht. Gleichzeitig bestreitet niemand, dass Kompetenz eine notwendige, wenn auch nicht hinreichende Voraussetzung für Bildung ist. Diese Position kann weiter radikalisiert werden. Heid (2004, S. 460) hat vor einiger Zeit ausgeführt: „Es kann doch nicht sinnvoll sein, von einer Bildung zu reden, eine Bildung zu fordern und eine Bildung fördern zu wollen, die man für unfassbar hält. Woher weiß man, durch welche Handlungen die Entwicklung eines Menschen gefördert oder behindert wird, die wir ‚Bildung’ nennen? Und woher weiß man, ob und in welchem Maß Aktivitäten erfolgreich waren oder sind, die wir ‚Bildung’ […] zu nennen pflegen?“
Messprobleme gibt es in allen Disziplinen, selbst in der Physik sind Konstrukte wie Kraft oder Energie nicht unmittelbar beobachtbar. Um noch einmal Bezug auf Heid zu nehmen (ebd., S. 461): „Es ist Sache forschungsmethodischer Exzellenz und bildungspraktischer Diagnostik, so gut wie eben möglich zu sichern, dass das Verhalten oder die Performanz sich als Ausdruck jener Kompetenz sichern lässt, die wir als ‚Bildung’ bezeichnen oder die uns als Bildung interessiert.“
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, ob der Begriff nur kognitive oder auch affektiv-motivationale Dimensionen umfasst. 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 Einstellungen und Persönlichkeitsmerkmale der zukünftigen Lehrpersonen mit aufnimmt oder sich nur auf die Erhebung ihres Fachwissens konzentriert, ob die Erfassung von Lehrerhandeln auf Unterrichten und Diagnostizieren beschränkt wird oder auch Erziehen und Mitwirken an der Schulentwicklung umfasst, ob man den Entwicklungsprozess während der Lehrerausbildung 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 und schließlich ob man versucht, auch das Berufsethos einzubeziehen oder nicht. Hinter den unterschiedlichen Konzepten stehen unterschiedliche bildungstheoretische Ansprüche an Schule und Lehrerausbildung. Diese Ansprüche lassen sich durchweg benennen und in Variablen umsetzen.
Auf diese Weise bildungstheoretisch fundiert, könnten Studien zur Lehrerausbildung dann auch wesentliche Beiträge zur erziehungswissenschaftlichen Theoriebildung leisten. Benner (2002) formuliert in diesem Sinne im Anschluss an Adorno und Brezinka: „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."

3 IEA's Teacher Education and Development Study: Learning to teach mathematics (TEDS-M)

Die kommende IEA-Studie zur Lehrerausbildung, TEDS-M, ist die erste IEA-Studie im tertiären Bereich und die erste international-vergleichende *large-scale* Studie zur Lehrerausbildung (siehe zur Konzeption der Studie im Detail Tatto et al., 2004). Sie wird auf internationaler Ebene von einem Projektteam gemanagt, das aus Wissenschaftler/innen der Michigan State University (USA), des *Australian Cuncil for Educational Research* (Australien) und Statistics Canada besteht. Von TEDS-M können große Fortschritte im Bereich der Lehrerausbildungsforschung erwartet werden. Rund 20 Länder haben zur Zeit Verträge mit der IEA abgeschlossen, die ihre Teilnahme an TEDS-M bestätigen. Darunter befinden sich Deutschland, England, Italien, Norwegen, Polen, Russland und die Schweiz sowie die Philippinen, Singapur, Taiwan und Thailand, Chile, Mexiko und die USA sowie Botswana. Deutsche Projektleiter/innen sind die Autorin, Rainer Lehmann (ebenfalls Humboldt-Universität zu Berlin) und Gabriele Kaiser (Universität Hamburg).
In TEDS-M werden angehende Mathematiklehrer/innen der Primarstufe und der Sekundarstufe I untersucht, die sich im letzten Jahr ihrer Ausbildung befinden. In Deutschland werden drei Routen untersucht: die Ausbildung für die Grundschule, auch wenn Mathematik nicht Unterrichtsfach ist. Als Generalisten unterrichten Grundschullehrer *per definitionem* auch Mathematik. Auf der Ebene der Sekundarstufe I untersuchen wir einerseits zukünftige Haupt- und Realschullehrer (angehende Mathematiklehrer/innen mit einer Lehrbefähigung bis zur Klasse 10) und andererseits zukünftige Mathematiklehrer für das Gymnasium und die Gesamtschule (Lehrbefähigung bis zur Klasse 13).
TEDS-M geht drei zentralen Fragekomplexen nach:

1. Welche S*truktur* und welches N*iveau* professioneller Kompetenz weisen angehende Mathematiklehrpersonen im letzten Jahr ihrer Ausbildung auf?

Von welchen individuellen Bedingungen hängen die Struktur und das Niveau ab? Und welche Unterschiede zeigen sich zwischen den Ländern?

2. Wie sind die institutionellen Lernbedingungen in der Mathematik-Lehrerausbildung ausgeprägt? Welche Unterschiede zeigen sich auf internationaler Ebene? Und welche dieser Rahmenbedingungen sind mit welcher individuellen Kompetenz verbunden?

<table>
<tr><td rowspan="6">Natio-
nale
Ebene</td><td>Level
III</td><td colspan="6">Gesellschaftssystem</td></tr>
<tr><td></td><td colspan="2">Allgemeiner Entwicklungsstand</td><td colspan="2">Status des Lehrerberufs</td><td>Status von Mathematik</td></tr>
<tr><td>Level
II</td><td colspan="6">Bildungssystem</td></tr>
<tr><td></td><td colspan="2">Steuerung und Kontrolle</td><td colspan="2">Ziele der Schule</td><td>Arbeitsbedingungen im Lehrerberuf</td></tr>
<tr><td>Level
I</td><td colspan="6">Lehrerausbildungssystem</td></tr>
<tr><td></td><td>Ziele/ Standards</td><td>Ausbildungs-
komponenten</td><td>Kosten pro Absol-
vent/in</td><td>Institutionalisierung</td><td>Eingangs-
selektivität</td></tr>
<tr><td rowspan="5">Institu-
tionelle
Ebene</td><td>Level
II</td><td colspan="6">Institutionell intendiertes Curriculum</td></tr>
<tr><td></td><td>Ziele und Inhalte</td><td>Lehr-
Lernmethoden</td><td>Kontrolle und
Steuerung</td><td>Beratung und Unter-
stützung</td><td>Selektivität</td></tr>
<tr><td>Level
I</td><td colspan="3">Lehrerausbildner/innen</td><td colspan="3">Implementiertes Curriculum</td></tr>
<tr><td></td><td>Wissen</td><td>Beliefs</td><td>Demographi-
sche Daten</td><td>Ziele und
Inhalte</td><td>Lehr-Lern-
methoden</td><td>Selektivität</td></tr>
<tr><td></td><td colspan="2">Ziele und Inhalte</td><td>Lehr-Lernmethoden</td><td>Kontrolle,
Steuerung</td><td>Beratung</td><td>Studierenden-
schaft</td></tr>
<tr><td rowspan="5">Indivi-
duelle
Ebene</td><td rowspan="3">Level
II</td><td colspan="3">Lernvoraussetzungen</td><td colspan="3">Nutzung des Lehrangebots</td></tr>
<tr><td>Wissen</td><td colspan="2">Beliefs</td><td colspan="2">Inhalte</td><td>Lehr-Lernmethoden</td></tr>
<tr><td>Persönlichkeitsmerkmale</td><td colspan="2">Demographisches</td><td>Investierte Lernzeit</td><td>Lernstrategien</td><td>Affektive
Komponenten</td></tr>
<tr><td>Level
I</td><td colspan="6">Erworbene professionelle Kompetenz</td></tr>
<tr><td></td><td colspan="2">Professionelles Wissen</td><td colspan="2">Professionelle beliefs</td><td>Persönlichkeitsmerkmale</td><td>Demographisches</td></tr>
</table>

3. Wie stellen sich die nationalen Rahmenbedingungen der Mathematiklehrerausbildung dar? Welche Unterschiede zeigen sich auf internationaler Ebene? Und welche dieser Rahmenbedingungen sind mit welchen institutionellen Rahmenbedingungen und welcher Kompetenz verbunden?

Im Rahmen einer Vorstudie – der Prepatory Teacher Education and Develoment Study (P-TEDS) – haben wir ein Modell entwickelt, das diese globalen Fragen konkretisiert und auf dessen Basis die Instrumententwicklung stattfand. Dieselben Untersuchungsdimensionen werden auch in TEDS-M angesprochen (siehe Darstellung 1).
Die Dimensionen werden entsprechend ihrer Komplexität mit einem ganzen Bündel an Instrumenten untersucht:

- Auf der Ebene des Individuums wird ein Kompetenztest eingesetzt, der das mathematische, mathematikdidaktische und pädagogisch-psychologische Wissen der angehenden Mathematiklehrer/innen sowohl systematisch als auch situationsorientiert erfasst. Darüber hinaus spielen Überzeugungen eine wichtige Rolle.
- Ebenfalls auf der Ebene des Individuums erfolgen Befragungen zu Lerngelegenheiten in der Mathematiklehrerausbildung.
- Auf der Ebene der Institutionen erfolgen dann eine parallelisierte Befragung der Lehrerausbildner/innen sowie eine Curriculum-Analyse.
- Ebenfalls auf der Ebene der Institution finden eine Expertenbefragung zu Ausbildungsbedingungen und eine Kostenanalyse statt. Diese Kostenanalyse umfasst zwei Elemente, und zwar Kosten der Mathematiklehrerausbildung (als Indikator für Ausbildungsbedingungen) und Lehrergehälter (als Indikator für Arbeitsbedingungen).
- Auf der nationalen Ebene findet eine Curriculum-Analyse der schulischen Ausbildung in Mathematik statt – mit der doppelten Zielsetzung, einen Referenzrahmen für die Interpretation der Lerngelegenheiten in der Lehrerausbildung zu gewinnen und einen Indikator für das Vorwissen der Studierenden zu haben.
- Und schließlich wird ein *Policy Report* zu den allgemeinen Rahmenbedingungen erstellt.

Die effektive Stichprobengröße auf der individuellen Ebene der zukünftigen Mathematiklehrer/innen beträgt 400 pro Sample. Wegen Klumpungseffekten wird die reale Zahl der Befragten aber deutlich höher liegen.
Da auch Aussagen über institutionenbezogene Zusammenhänge zwischen Ausbildungsmerkmalen und Kompetenzausprägungen gemacht werden sollen, wird zum einen eine hinreichende Anzahl an Institutionen im Sample benötigt und zum anderen eine hinreichend große Stichprobe innerhalb der Institutionen. Die Mindestzahl an Institutionen liegt pro Sample bei 50. Innerhalb der Institutionen sollen – falls vorhanden – 30 Ausbildner/innen befragt.

Wenn wirklich Kompetenz getestet und nicht nur Wissen erfasst werden soll, ist ein zentraler Schritt die Definition dessen, was „berufstypische Anforderungen“ sind. Die folgende Synthese stellt die aus P-TEDS stammende Basis der Diskussionen in TEDS-M dar:

Teacher tasks	**Situations**
A: Choice of themes, methods; sequencing of learning processes	1. Selecting and justifying content of instruction 2. Designing and evaluating of lessons
B: Assessment of student achievement; counselling of students/ parents	1. Diagnosing student achievement, learning processes, misconceptions, preconditions 2. Assessing students 3. Counselling students and parents 4. Dealing with errors, giving feedback
C: Support of students' social, moral, emotional development	1. Establishing teacher-student relationship 2. Foster the development of morals and values 3. Dealing with student risks 4. Prevention of, coping with discipline problems
D: School development	1. Initiating, facilitating cooperation 2. Understanding of school evaluation
E: Professional ethics	1. Accepting the responsibility of a teacher

Aufgrund unterschiedlicher Akzentsetzungen in den Ländern und normativ unterschiedlicher Bewertungen haben sich die Länder darauf geeinigt, dass A und B (in einer deutschen Begrifflichkeit etwa „Unterrichten“ und „Diagnostizieren“; siehe KMK, 2004) internationale Kernaufgaben darstellen und dass zum gegenwärtigen Zeitpunkt nur diese methodisch sinnvoll testbar sind.
Vor diesem Hintergrund einer relativ engen Definition professioneller Kompetenz, wurde entwickelt, welche Wissens- und Überzeugungsdimensionen beim Handeln von Mathematiklehrern in beruflichen Situationen zum Tragen kommen:

Mathematics Knowledge (TIMSS 2003)
4 Content Domains: number, algebra, probability, geometry.
3 Cognitive Domains: knowing, applying, reasoning.
Mathematics Pedagogy Knowledge
Mathematical Curricular Knowledge [Pre-active]
Knowledge of planning for mathematics teaching [Pre-active]
Knowledge of enacting mathematics [Interactive]
General Pedagogy
Knowledge of students

Classroom management and assessment
Beliefs/Attitudes/Perspectives
Nature of mathematics, and of teaching & learning mathematics
Personality
Self efficacy

TEDS-M lässt sich also wie folgt zusammenfassen:

(1) Die Wirksamkeit der Mathematiklehrerausbildung wird auf der Ebene des mathematischen, mathematikdidaktischen und pädagogisch-psychologischen Wissens, der Überzeugungen und der Persönlichkeitsmerkmale festgemacht, die zusammen professionelle Kompetenz ausmachen.

(2) Struktur und Niveau professioneller Kompetenz werden durch Faktoren beeinflusst, die auf verschiedenen Ebenen liegen. Basis der Studie ist daher ein Mehrebenenmodell, das zwischen nationalen, institutionellen und individuellen Bedingungen sowie individuellen Wirkungen unterscheidet.

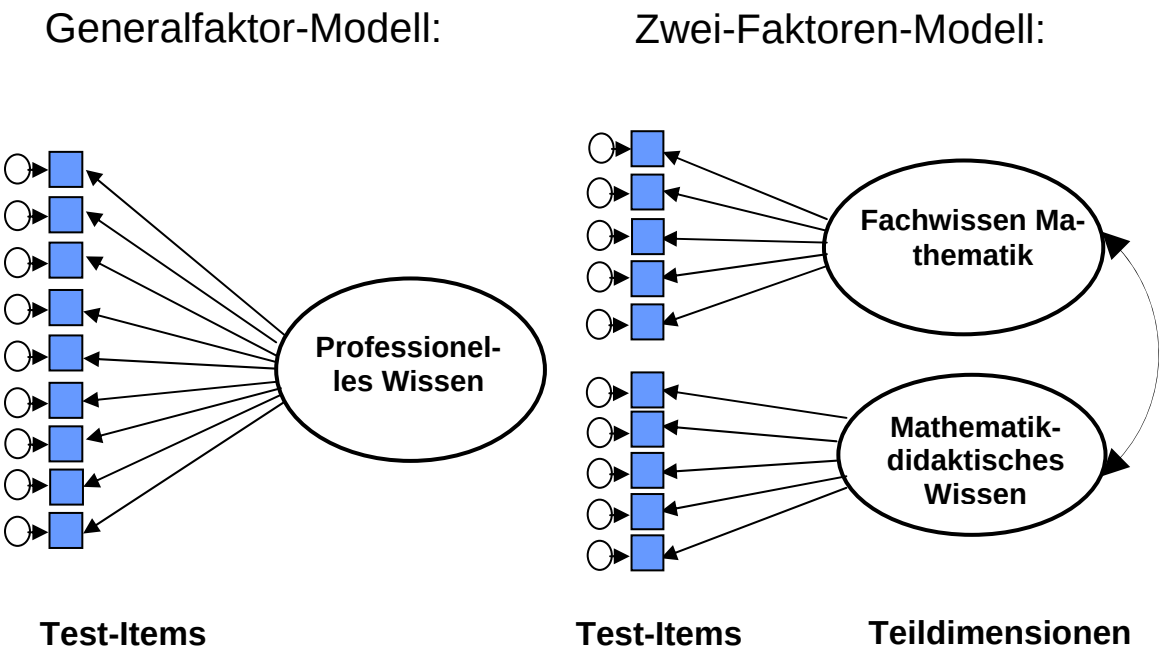
(3) Die Erfassung des Wissens erfolgt über situationsbezogene Items, die zentrale berufliche Anforderungen von Mathematiklehrer/innen repräsentieren. Gleichzeitig wird es rein wissensorientierte Items geben, sodass eine Prüfung von Zusammenhängen zwischen deklarativem Wissen in Mathematik, Mathematikdidaktik und Pädagogik-Psychologie, entsprechenden *beliefs* und handlungsbezogenen (prozeduralen) Ausformungen des Wissens möglich ist, wie sie Kompetenz kennzeichnen.

Die TEDS-M Hauptstudie ist für Frühjahr 2008 angesetzt. Im Dezember 2009 werden die Ergebnisse der Studie veröffentlicht.

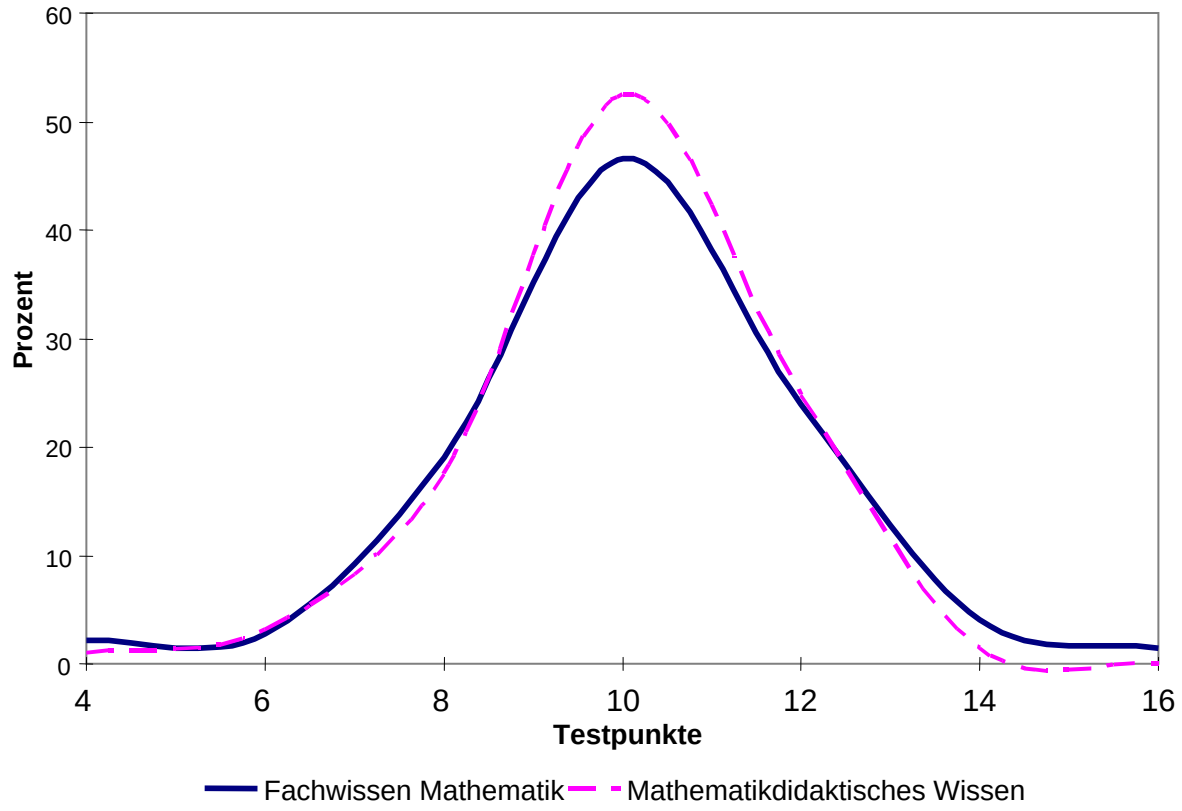
4 Erste Ergebnisse aus P-TEDS

P-TEDS ist eine Studie in sechs Ländern, und zwar in Deutschland, Bulgarien, Mexiko, USA, Südkorea und Taiwan. Prinzipiell sind die Instrumente und die abzudeckenden Fragestellungen dieselben wie in TEDS-M. Vom Design her ist bei P-TEDS aber zusätzlich eine quasi-längsschnittliche Modellierung von Kompetenzentwicklung möglich, weil drei Kohorten untersucht wurden: Studienanfänger/innen, Hauptstudiumsstudierende und Referendar/innen. Gleichzeitig war P-TEDS auf die Sekundarstufe I beschränkt, um die Komplexität der Studie bewältigen zu können. Hier wurden zwei Routen in den Blick genommen, nämlich angehende Mathematik-Grund-, Haupt- und Realschullehrer/innen mit einer Lehrbefähigung für die Klassen 1 bis 10 sowie Mathematiklehrpersonen an Gymnasien und Gesamtschulen mit einer Lehrbefähigung für die Klassen 5 bis 13. Die Datenerhebung für P-TEDS fand von April bis Juni 2006 statt. Ei-

ne Veröffentlichung der detaillierten Ergebnisse erfolgt im Herbst 2007 (Blömeke, Kaiser & Lehmann, in Vorbereitung).
Insgesamt nahmen in Deutschland 878 Studierende und Referendare aus 4 Universitäten und 22 Studienseminaren an der Studie teil. Um die Zweiphasigkeit der deutschen Lehrerausbildung zu bewältigen, wurden zunächst vier Universitäten ausgewählt und dann alle umliegenden Studienseminare in die Stichprobe aufgenommen. Aus anderen Untersuchungen wissen wir, dass die Mehrheit der Studierenden die Ausbildung in einem nahe gelegenen Studienseminar fortsetzt.
Der Test dauerte 90 Minuten und wurde unter Aufsicht durchgeführt. Die Studierenden und Referendare wurden in 20 Lehrveranstaltungen der ersten und in 22 Veranstaltungen der zweiten Ausbildungsphase getestet. Je nach Standort und Kohorte konnten wir eine Rücklaufquote zwischen 85 und 100% erreichen, im Mittel 96,7%. In der dritten Kohorte, bei den Referendar/innen, konnte eine Ausschöpfungsquote von 80 Prozent erreicht werden, d.h. 80% aller formal im Dienst befindlichen Referendar/innen in den 22 Institutionen (je nach Standort zwischen 72 und 100 Prozent). In Bezug auf die Hauptstudiums-Studierenden an der Universität liegt die Ausschöpfungsquote zu niedrig, um diese Kohorte in Vergleiche einbeziehen zu können. Sie bleibt daher außen vor.
Unsere erste Frage war: Stellt die Kompetenz, Mathematik zu unterrichten, ein einheitliches Konstrukt dar oder lassen sich mehrere Kompetenzdimensionen separieren? Diese Frage ist deswegen so interessant, weil die Modellierungen in P-TEDS zu Diskussionen darüber geführt haben, was eigentlich eine Fachdidaktik ist. Ist Mathematikdidaktik überhaupt ohne Mathematik denkbar? Die Grafik veranschaulicht diese Frage:



Wir prüfen diese beiden Modellannahmen mit Hilfe ein- und mehrdimensionaler Raschmodelle. Den Informationsindizes lässt sich entnehmen, dass das zweidimensionale Modell im Vergleich zur absoluten Höhe der Abweichungen eine zwar geringe, aber signifikante Modellverbesserung ergibt. Die WLE-Reliabilität der beiden Subskalen beträgt .83 bei Mathematikdidaktik und .74 bei Mathematik, die nur mit etwa der Hälfte der Items gemessen wurde. Die latente Korrelation zwischen Mathematik und Mathematikdidaktik beträgt .81. Damit kann man die Schlussfolgerung ziehen, dass der Gesamttestwert prinzipiell ein erschöpfendes Maß für die Beurteilung des fachbezogenen Wissens angehender Mathematiklehrerinnen und -lehrer sowie für die Analyse von Mittelwertdifferenzen darstellen würde. Differenzielle Befunde in den Subtests sind aber möglich und bieten ggf. weitere Aufschlüsse (siehe in diesem Sinne auch die Entscheidung in TIMSS III; Klieme, 2000: 74).
Blickt man auf die Verteilung der Leistungen in den beiden Dimensionen, ist auffällig, wie ähnlich diese verläuft. Beide Teilkompetenzen sind annähernd normalverteilt. Die Mathematik-Skala wurde auf einen Mittelwert von 100 und eine Standardabweichung von 20 transformiert. Für die Mathematikdidaktik-Skala ergibt sich daraus ein Mittelwert von knapp über 100, allerdings eine deutlich geringere Standardabweichung.



Detailliertere Analysen liegen derzeit zur begrenzt vor. Aus diesen wird deutlich, dass in den beiden Kohorten der Studienanfänger/innen und der Referen-

dar/innen sowohl in Mathematik als auch in Mathematikdidaktik substanzielle Unterschiede zwischen den Studiengängen und Ausbildungsorten festzustellen sind. Diese liegen jeweils bei rund 1 Standardabweichung. Alle drei Faktoren – Studiengang, Kohorte und Standort – klären signifikant Varianz auf, wobei der Studiengang der stärkste Faktor ist.
Über die kognitiven Aspekte professioneller Kompetenz hinaus werden in P-TEDS auch umfassend *beliefs* erhoben, und zwar sowohl fachgebunden als auch fächerübergreifend. Ein Beispiel sind epistemologische Überzeugungen zur Natur der Mathematik. Hier haben wir das Instrument von Grigutsch, Raatz und Törner (1998) verwendet, und zwar in einer anhand der Faktorladungen gekürzten Form.
Detaillierte Analysen stehen noch aus. Es ist aber bereits deutlich geworden, dass sich die bekannten vier Faktoren Formalismus, Schemaorientierung, Prozess und Anwendung auch für unsere Stichprobe extrahieren lassen. Die Partialkorrelationen dieser vier Faktoren wiesen ebenfalls eine ähnliche Struktur auf, indem Formalismus und Schematismus sowie Prozess und Anwendung positiv korrelieren. Diese Struktur kann also als stabil angesehen werden. Anders als bei Grigutsch, Raatz und Törner (1998) lässt sich allerdings keine negative Korrelation zwischen Schemaorientierung und Formalismus nachweisen – im Gegenteil korrelieren Formalismus und Anwendung positiv. Hier werden sich Analysen anschließen müssen, um Ursachen für die Unterschiede auf die Spur kommen zu können.

5 Diskussion

Für P-TEDS wurde ein Leistungstest entwickelt, mit dem international-vergleichend die fachbezogenen kognitiven Komponenten professioneller Kompetenz angehender Mathematiklehrerinnen und -lehrer der Sekundarstufe I erfasst werden können. Der Item-Entwicklung lag ein bildungstheoretisch fundiertes Modell professioneller Kompetenz zugrunde.
Empirische Überprüfungen zeigen, dass das fachbezogene Wissen von Studierenden und Referendaren aus zwei unterschiedlichen Subdimensionen besteht, die latent zu .81 korrelieren. Im deutschen Raum verfolgt nur eine weitere Arbeitsgruppe eine ähnliche Forschungsrichtung, wie sie hier dargestellt ist. Im Rahmen des Projekts „Kognitiv-aktivierender Unterricht“ (COAKTIV) wird von Baumert, Blum und Neubrand mit Hilfe eines Bezugs zu PISA 2003 untersucht, worin die Expertise von im Beruf stehenden Mathematik-Lehrkräften besteht und in welchem Zusammenhang diese mit der Leistungsentwicklung der Schüler steht (Brunner et al., 2006; Brunner et al., im Druck).[1]

[1] Wir danken Jürgen Baumert und Kolleg/innen, die uns bereits vor Erscheinen eine Reihe von Aufsätzen zur Verfügung gestellt haben.

In der Prüfung der Strukturannahmen kommen beide Studien zu vergleichbaren Ergebnissen: Grundsätzlich lassen sich mathematisches und mathematikdidaktisches Wissen reliabel empirisch erfassen (COAKTIV: α = .83 bzw. .78; P-TEDS: WLE-Reliabilität .74 und .83). Dabei handelt es sich zwar um eng miteinander verknüpfte, aber dennoch unterscheidbare und in sich raschhomogene Konstrukte (COAKTIV: latente messfehlerbereinigte Korrelation auf der Basis einer Konfirmatorischen Faktorenanalyse = .79; P-TEDS latente messfehlerbereinigte Korrelation aufgrund einer zweidimensionalen Raschmodellierung = .81). Vor diesem Hintergrund können die Strukturannahmen zum professionellen Wissen von Mathematiklehrpersonen als stabil angesehen werden. Sie wurden auf der Basis einer vergleichbaren Konzeption mit zwei unterschiedlichen Tests an zwei unterschiedlichen Stichproben (angehende Lehrpersonen, praktizierende Lehrpersonen) gewonnen.

Dass P-TEDS und TEDS-M in einen internationalen Rahmen eingebettet sind, hat seinen Preis – gleichzeitig ist es mit erheblichen Vorteilen verbunden. Die Unterschiede zwischen den teilnehmenden Ländern sind beträchtlich, wenn man sich die Struktur der Lehrerausbildung ansieht. Dies stellt vor allem für Deutschland mit seiner Zweiphasigkeit der Lehrerausbildung ein Problem dar. Andererseits zeigt sich, dass es offensichtlich einen gemeinsamen Kern dessen gibt, was von Lehrer/innen erwartet wird. Welche Anforderungen zu bewältigen sind, ist erstaunlich unumstritten. Der Vorteil internationaler Studien ist, dass sie das *Know How* zur Lehrerausbildungsforschung aus vielen Ländern zusammenbringen. Die Testung von Lehrerkompetenzen erfordert ein komplexes Herangehen, da immer fachliche und didaktische Fragestellungen hineinspielen – und hierfür Items zu entwickeln ist anspruchsvoll. Gerade in Deutschland bestand bisher wenig Erfahrung mit einem empirischen Zugang zur Lehrerausbildung. Hier konnte auf internationale Expertise zurückgegriffen werden. Schon die Konzeptualisierung von P-TEDS und TEDS-M hat Fortschritte in der Theoriebildung mit sich gebracht, die durch die kommenden empirischen Erkenntnisse weitergeführt werden können.

Literatur

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

Blömeke, S. (2004). Empirische Befunde zur Wirksamkeit der Lehrerbildung. In: Blömeke, S., Reinhold, P, Tulodziecki, G. & Wildt, J. (Hrsg.), Handbuch Lehrerbildung (pp. 59-91). Bad Heilbrunn/ Braunschweig: Klinkhardt/ Westermann.

Blömeke, S. Kaiser, G. & Lehmann, R. (Hrsg.) (in Vorbereitung): Kompetenzmessung bei angehenden Lehrerinnen und Lehrern. Ergebnisse einer empirischen Studie zum Wissen, zu den Überzeugungen und zu den Lerngelegenheiten von Mathematik-Studierenden und -Referendaren. Münster: Waxmann.

Brunner, M., Kunter, M., Krauss, S., Baumert, J., Blum, W., Neubrand, M., Dubberke, T., Jordan, A., Klusmann, U., & Tsai, Y.-M. (im Druck). Welche Zusammenhänge bestehen

zwischen dem fachspezifischen Professionswissen von Mathematiklehrkräften und ihrer Ausbildung sowie beruflichen Fortbildung? Zeitschrift für Erziehungswissenschaft.

Brunner, M., Kunter, M., Krauss, S., Klusmann, U., Baumert, J., Blum, W., Neubrand, M., Dubberke, T., Jordan, A., Löwen, K., & Tsai, Y.-M. (2006): Die professionelle Kompetenz von Mathematiklehrkräften. Konzeptualisierung, Erfassung und Bedeutung für den Unterricht. In: BIQUA 2006.

Cochran-Smith, M. & Zeichner, K. M. (Eds.) (2005). Studying Teacher Education. The Report of the AERA Panel on Research and Teacher Education. Mahwah, N.J.: Lawrence Erlbaum.

Confrey, J. (2006). Comparing and Contrasting the National Research Council Report *On Evaluating Curricular Effectiveness* With the What Works Clearinghouse Approach. In: Educational Evaluation and Policy Analysis 28 (3), pp. 195-213.

Grigutsch, S., Raatz, U., & Törner, G. (1998). Einstellungen gegenüber Mathematik bei Mathematiklehrern. Journal für Mathematik-Didaktik, 19, 3-45.

Heid, H. (2004): Bildung als Gegenstand empirischer Forschung. In: Vierteljahrsschrift für wissenschaftliche Pädagogik 80 (4), S. 456-477.

Helmke, A. (2004). Unterrichtsqualität: Erfassen, Bewerten, Verbessern (3. Aufl.). Seelze: Kallmeyersche Verlagsbuchhandlung.

Klieme, E. (2000). Fachleistungen im voruniversitären Mathematik- und Physikunterricht. In Baumert, J., Bos, W., & Lehmann, R. (Eds.) (2000). TIMSS/III: Dritte Internationale Mathematik- und Naturwissenschaftsstudie – Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn (pp. 57-128). Opladen: Leske + Budrich.

(KMK) Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (2004): Standards für die Lehrerbildung: Bildungswissenschaften. Beschluss der Kultusministerkonferenz v. 16.10.2004.

National Research Council (2002). Scientific Research in Education. Washington, DC: The National Academies Press.

Oser, F. (1997a). Standards in der Lehrerbildung. Teil 1: Berufliche Kompetenzen, die hohen Qualitätsmerkmalen entsprechen. Beiträge zur Lehrerbildung, 1997 (1), 26-37.

Oser, F. (1997b). Standards in der Lehrerbildung. Teil 2: Wie werden Standards in der schweizerischen Lehrerbildung erworben? Erste empirische Ergebnisse. In: Beiträge zur Lehrerbildung, 1997 (2), 210-228.

Tatto, M. T., Swille, J., Senk, Sh., Schmidt, W. H., Ingvarson, L., Rowley, G. & Peck, R. (2004). IEA Teacher Education Study in Mathematics (TEDS-M). Conceptual Framework. Unpublished manuscript.

Weinert, F. E. (1999). Konzepte der Kompetenz. Gutachten zum OECD-Projekt "Definition and Selection of Competencies: Theoretical and Conceptual Foundations (DeSeCo)". Neuchatel, Schweiz: Bundesamt für Statistik.

Weinert, F. E. (2001). Concept of Competence: A Conceptual Clarification. In D. S. Rychen & L. H. Salganik (Eds.), Defining and Selecting Key Competencies (pp. 45-66). Göttingen: Hogrefe.