

Stefan KRAUSS<sup>1</sup>, Jürgen BAUMERT<sup>1</sup>, Werner BLUM<sup>2</sup>, Michael NEU-  
BRAND<sup>3</sup>, Alexander JORDAN<sup>2</sup>, Martin BRUNNER<sup>1</sup>, Mareike KUNTER<sup>1</sup>  
und Katrin LÖWEN<sup>1</sup> <sup>1</sup> Berlin, <sup>2</sup> Kassel, <sup>3</sup> Oldenburg

## **Die Konstruktion eines Tests zum fachlichen und zum fachdi- daktischen Wissen von Mathematiklehrkräften**

### **1. Zur Konzeption des Tests**

Das Fachwissen und das fachdidaktische Wissen sind aus fachspezifischer Sicht elementare Bausteine des Professionswissens von Lehrkräften (bei Shulman, 1986, als *content knowledge* und als *pedagogical content knowledge* konzeptualisiert). Im Rahmen des DFG-Schwerpunktprogramms BIQUA (Bildungsqualität von Schule) wurde im Projekt COACTIV (Berlin/Kassel/Oldenburg) ein Test entwickelt, der das Fachwissen und das fachdidaktische Wissen von Mathematiklehrkräften misst (ausführlichere Informationen zur Gesamtrationale der COACTIV-Studie sowie eine Beschreibung weiterer im Rahmen dieser Studie entwickelten Instrumente finden sich im Beitrag von M. Neubrand in diesem Band).

Eine empirische Untersuchung dieser beiden Wissensbereiche erfordert die Entwicklung von Skalen direkt „aus dem Fach heraus“ (im Gegensatz zu den mittlerweile empirisch bewährten, aber fachunspezifischen Unterrichtsskalen, siehe z.B. Gruehn, 2000). Die in COACTIV erfolgte Konstruktion entsprechender Items für Mathematiklehrkräfte ist wissenschaftliches Neuland und soll in diesem Beitrag ausführlich beschrieben werden. Insbesondere soll dabei auf die Konzeptualisierung des fachdidaktischen Wissens, das maßgeblich zur Gestaltung fachlich gehaltvoller Lernumgebungen beiträgt (Blum, 1999), eingegangen werden.

Unterrichten ist die Verhandlung fachlicher Inhalte mit Schülern. In Anlehnung an diese theoretischen Eckpfeiler von Mathematikunterricht wurden in COACTIV drei Subtests für fachdidaktisches Wissen entwickelt. Dabei steht im Folgenden (1) exemplarisch für das *didaktische Wissen über Inhalte*, (2) exemplarisch für das *didaktische Wissen über Schüler* und (3) exemplarisch für das *didaktische Wissen über Verhandlung*:

#### Subtest (1): Wissen über multiple Lösbarkeit von Mathematikaufgaben

Aufgaben stellen im Mathematikunterricht eine zentrale Trägerform mathematischer Inhalte dar. Das Potential von Aufgaben für die Unterstützung von Lernprozessen kann u.a. über deren multiple Lösbarkeit erschlossen werden (für eine empirische Bestätigung der lernförderlichen Wirkung multipler Lösungswege siehe z.B. Große, 2005). Zur Operationalisierung des didaktischen Wissens über Inhalte wurden vier (auch für Schüler ge-

eignete) Mathematikaufgaben gewählt, jeweils mit der Instruktion für die Lehrkraft, möglichst viele substantiell verschiedene Lösungswege anzugeben (siehe Beispielitem „Nachbarzahlen“ im Beitrag von Neubrand).

#### Subtest (2): Didaktisches Wissen über fachliche Schülerkognitionen

Um Unterricht adaptiv gestalten zu können, muss eine Lehrkraft über Kenntnisse zu typischen inhaltlichen Schülerkognitionen verfügen. Nach Matz (1982) offenbaren vor allem Probleme und Fehler das implizite Wissen des Problemlösers und machen kognitive Prozesse so oftmals überhaupt erst erkennbar. Um Schülerfehler und typische Schwierigkeiten als eine didaktische Chance für verständnisvolles Lernen nutzbar zu machen, muss eine Lehrkraft Schülerfehler erkennen, konzeptuell einordnen und analysieren können. Zur Operationalisierung des didaktischen Wissens über Schüler wurden acht Unterrichtssituationen konstruiert, in denen Schülerfehler oder Schwierigkeiten erkannt und/oder analysiert werden mussten (siehe Beispielitem „Hoch Null“ im Beitrag von Neubrand).

#### Subtest (3): Didaktisches Wissen über lokale Vermittlungsstrategien

Die Wissenskonstruktion des Schülers kann oft nur mit instruktionaler Anleitung gelingen. Zur Operationalisierung des didaktischen Wissens über Verhandlung wurden acht Situationen aus dem Mathematikunterricht konstruiert, in denen die unmittelbare Unterstützung lokaler Verständnisprozesse erforderlich war (siehe Beispielitem „Minus mal minus“ im Beitrag von Neubrand). Da profundes Wissen über mathematische Repräsentationen die Verfügbarkeit eines großen Repertoires zum Erklären mathematischer Sachverhalte bedeutet, wurde dabei ein Schwerpunkt auf solches Wissen gesetzt.

Das fachdidaktische Wissen wurde somit mit insgesamt 24 Items operationalisiert. In Abgrenzung zum fachdidaktischen Wissen wurde auch ein Test zum Fachwissen von Mathematiklehrkräften (4) konstruiert:

#### (4) Fachwissen

Das Fachwissen wurde in COACTIV als vertieftes Hintergrundwissen über den Schulstoff gefasst (in Anlehnung an Shulman, 1986). In diesem Test wurden den Lehrkräften insgesamt dreizehn Aufgaben (mit vergleichsweise hohem Schwierigkeitsniveau, aber explizit *kein* ausgewiesenes Universitätswissen) vorgelegt. Eine Beispielaufgabe dafür ist:

„Ist  $2^{1024} - 1$  eine Primzahl?“<sup>1</sup>

---

<sup>1</sup> Nein, denn es gilt:  $a^2 - b^2 = (a - b)(a + b)$ . Demnach lässt sich  $2^{1024} - 1$  zerlegen in  $(2^{512} - 1)(2^{512} + 1)$ .

(Dieses Item wurde der Pilotierungsphase entnommen).

Insgesamt wurde das Fachwissen und das fachdidaktische Wissen der Mathematiklehrkräfte also anhand von 37 Items erfasst. Neu (im Vergleich zu den meist fachunspezifischen Unterrichtsskalen) ist dabei nicht nur die dezidiert fachspezifische Ausrichtung, ein entscheidender Schritt ist vor allem auch die *direkte Testung* dieser Konstrukte (falls überhaupt, wird das Fachwissen von Lehrkräften üblicherweise distal, z.B. über Examensnoten, „erhoben“). Da die Erfassung der fachlichen Kompetenzen von Schülern durch direkte Leistungstests selbstverständlich sind, ist es erstaunlich, dass es sich bei der direkten Erfassung des fachbezogenen Professionswissens von Lehrkräften weitestgehend um Forschungsneuland handelt.

## 2. Zur Testkonstruktion

Im Folgenden soll das Vorgehen bei der Testkonstruktion und der Auswertung geschildert werden: Das erste Problem war, Items mit geeignetem Schwierigkeitsgrad zu finden, da Items, die von fast allen Lehrkräften bzw. von kaum einer Lehrkraft gelöst werden können, vermieden werden mussten. Mithilfe mehrerer Pilotierungsrunden konnten Items gefunden werden, die prinzipiell für Lehrkräfte aller Schulformen lösbar waren (und die auch eine hohe „face-validity“ hatten, d.h., auch die Lehrkräfte selbst waren der Meinung, dass durch diese Items relevantes fachliches bzw. fachdidaktisches Wissen getestet wird). Ein weiteres Problem war, dass es für fachdidaktische Items oft keine eindeutig normativ richtigen Antworten gibt: Welche Analyse eines Schülerfehlers z.B. ist „richtig“? Welche Erklärung eines mathematischen Sachverhalts ist „didaktisch geeignet“? Unter Einbeziehung von Fachdidaktikern *und* von Lehrkräften wurde ein Kategorienschema entwickelt, das es den Testauswertern ermöglichen sollte, ihre Beurteilungen objektiven Kriterien anzupassen (Kriterien für eine „didaktisch geeignete Lehrerantwort“ waren beispielsweise eine sehr hohe Übereinstimmung unter den Testkonstruktoren bzw. eine in der Literatur zu findende empirische Bestätigung einer „didaktisch angemessenen“ Vermittlungsstrategie)<sup>2</sup>. Die Durchführung des Tests geschah unter Aufsicht eines Testleiters (in den meisten Fällen am Nachmittag des PISA-Testtages). Bei der Auswertung wurde jede Lehrerantwort von zwei unabhängigen Ratern beurteilt, dabei konnte eine befriedigende Übereinstimmung erreicht werden (die Interraterreliabilität lag in der Regel bei über .80).

---

<sup>2</sup> Die Schwierigkeit, „didaktisch richtige“ Antworten zu identifizieren, ist auch einer der Gründe, warum weitere in COACTIV implementierte Paradigmen, die eindeutig der Fachdidaktik zuzuordnen sind (wie zum Beispiel die Auswahl und die Anordnung von Aufgaben bzw. die Einschätzung von zur Lösung von Mathematikaufgaben benötigten Kompetenzen; siehe Beitrag von Neubrand) nicht in den Fachdidaktik-*Test* integriert wurden (für die in COACTIV implementierten Videosequenzen geben die Lehrerantworten z.B. eher didaktische *Orientierungen* als didaktisches *Wissen* wieder).

Für die einzelnen Skalen ergaben sich folgende Reliabilitäten: Für die Subskala „Aufgaben“ ergab sich ein Cronbach  $\alpha$  von .55 (4 Items), für die Subskala „Schülerkognitionen“ ein  $\alpha$  von .54 (8 Items), und für die Subskala „Vermittlung“ ein  $\alpha$  von .64 (12 Items). Für die aus diesen Items gebildete Gesamtskala „fachdidaktisches Wissen“ ergab sich ein  $\alpha$  von .78 (24 Items) und für die Skala „Fachwissen“ ein  $\alpha$  von .83 (13 Items).

Unter psychometrischen Gesichtspunkten ist es also gelungen, sowohl eine reliable Testskala für das fachdidaktische Wissen als auch eine für das Fachwissen zu konstruieren. Die Subtests des fachdidaktischen Wissens dagegen müssen mit Vorsicht interpretiert werden.

### **3. Erste deskriptive Ergebnisse**

Abschließend sollen noch einige erste Ergebnisse auf deskriptiver Ebene berichtet werden: Weder für das Fachwissen noch für das fachdidaktische Wissen ergaben sich signifikante Unterschiede bezüglich Alter oder Geschlecht der Lehrkräfte. Im Vergleich zu den anderen Schulformen haben Gymnasiallehrkräfte jedoch einen signifikanten Vorsprung in beiden Wissensbereichen (der beim Fachwissen noch deutlicher ausgeprägt ist). Erste Ergebnisse zur Struktur des Wissens sind: Die fachdidaktischen Subfacetten korrelieren untereinander sehr hoch (eine geeignete Erklärung eines mathematischen Sachverhaltes erfordert z.B. meist auch didaktisches Wissen über Inhalte bzw. Wissen über Schülerkognitionen). Die Korrelation zwischen Fachwissen und fachdidaktischem Wissen (als Gesamtskala) ist ebenfalls hoch, aber differentiell bezüglich der Schulformen: Sie fällt für Gymnasiallehrkräfte noch einmal höher aus als für Lehrkräfte anderer Schulformen.

In zukünftigen Analysen wird es darum gehen, gemeinsame Modelle mit weiteren in COACTIV erhobenen Skalen zu entwickeln und anhand der Daten zu überprüfen. Von besonderem Interesse wird dabei natürlich der Zusammenhang der beiden hier vorgestellten Wissensbereiche mit dem in der PISA Studie 03/04 erhobenen Leistungszuwachs der Schüler sein.

### **Literatur**

- Blum, W. (1999). Unterrichtsqualität am Beispiel Mathematik – was kann das bedeuten, wie ist das zu verbessern? In: *SEMINAR - Lehrerbildung und Schule*, 4, S 8-16.
- Gruehn, S. (2000). Unterricht und Lernen. Schüler als Quellen der Unterrichtsbeschreibung. Waxmann.
- Große, C.S. (2005). Lernen mit multiplen Lösungswegen. Waxmann.
- Matz, M. (1982). *Towards a process model for high school algebra errors*. In: Sleeman, D. & Brown, J.S.: *Intelligent Tutoring Systems*, S. 25-50. Academic Press.
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching, *Educational researcher*, 15 (2), S. 3-14.