

Julia TELLER, Bärbel BARZEL, Timo LEUDERS, Freiburg

## **Erfassung der Entwicklung diagnostischer Kompetenzen von Lehrkräften - Validität im Mixed-Method-Design**

Validität ist bei quantitativen Untersuchungsdesigns an Kennwerte gebunden, die sich lediglich auf Testergebnisse beziehen. Ein Mixed-Method-Design mit qualitativen Elementen stellt daher generell eine günstige Bedingung für Validität dar. In diesem Beitrag werden erste Auswertungsergebnisse einer Studie zur Förderung diagnostischer Kompetenzen durch eine Lehrkräftefortbildung vorgestellt, die auf qualitativen Auswertungen beruhen und dabei prozessbezogene Validität besonders ins Auge fassen.

### **1. Hintergrund der Studie: Diagnostische Kompetenzen**

Diagnostische Kompetenzen werden als Fähigkeit aufgefasst, adäquate Urteile über Schülerinnen und Schüler zu treffen (Schrader, 2011). Hierzu zählen insbesondere die Fähigkeiten, lern- und leistungsrelevante Merkmale von Schülerinnen und Schülern zu beurteilen und Aufgabenanforderungen adäquat einzuschätzen (Lorenz & Artelt, 2009). In der vorliegenden Studie galt es, diese Facetten durch eine themenspezifische Lehrerfortbildung zu fördern. Zudem wurde ein Online-Fragebogen entwickelt, der Entwicklungsverläufe diagnostischer Kompetenzen aufdecken kann.

Im Rahmen von Large-Scale Studien wie COACTIV und TEDS-M werden diagnostische Kompetenzen als eine Facette von handlungsleitendem Professionswissen bzw. fachdidaktischem Wissen querschnittlich erhoben (Baumert et al., 2008; Döhrmann et al., 2012). Das Ziel der in diesem Beitrag vorgestellten Studie „Förderung diagnostischer Kompetenzen im Bereich Funktionen“ ist hingegen die Untersuchung der Entwicklung diagnostischer Kompetenzen in einem längsschnittlichen Design. Dabei muss die Kompetenzmessung spezifischer angelegt sein als bei oben genannten Ansätzen. Von Interesse sind die folgenden Forschungsfragen: Inwiefern können diagnostische Kompetenzen durch eine unterrichtsbegleitende Fortbildung zum Themenbereich Funktionen gefördert werden? Inwiefern können diese Entwicklungsverläufe valide erfasst werden?

### **2. Mixed-Method-Design der Studie**

Zu diesem Zweck wurde ein quasiexperimentelles Design mit 26 Lehrkräften entwickelt (vgl. Teller, Barzel, & Leuders, 2013). Die Erhebung diagnostischer Kompetenzen vor und nach der Fortbildung geschah in Form eines Online-Fragebogens, der größtenteils offene Antwortformate enthielt, In J. Roth & J. Ames (Hrsg.), *Beiträge zum Mathematikunterricht 2014* (S. 1207–1210). Münster: WTM-Verlag

um Diagnoseprozesse der Teilnehmenden möglichst situationsspezifisch und valide zu erfassen. Die erhobenen Antworten wurden qualitativ auf der Basis der Grounded Theory ausgewertet und von zwei unabhängigen Ratern kodiert.

### **3. Erste Ergebnisse**

Die am Datenmaterial entwickelten Kategorien beschreiben unterschiedliche diagnostische Zugänge zur Beurteilung von Schülerlernständen und erlauben die Darstellung in Form eines Diagnoseprofils. Entwicklungsverläufe der diagnostischen Kompetenzen können nachgewiesen werden, indem vor und nach der Fortbildung entstandene Diagnoseprofile einer Lehrkraft verglichen werden.

Ein Diagnoseprofil einer Lehrkraft bezieht sich auf deren Diagnoseurteile über vier Items hinweg. Es besteht aus 13 Elementen auf fünf Ebenen, die unabhängig voneinander und auch mehrmals erfüllt werden können. Qualitative Unterscheidungen der Diagnoseprofile werden getroffen, indem Diagnoseurteile entweder als allgemein oder konkret (Ebene 1) und korrigierend, beschreibend oder analytisch (Ebene 2) identifiziert werden. Des Weiteren kann sich ein Diagnoseprofil entweder auf ein Defizit oder zwei sowie mehrere Defizite beziehen (Ebene 3). Außerdem kann ein Diagnoseprofil zusätzliche Qualitätsmerkmale beinhalten wie die Anwendung fachdidaktischen Wissens, der Bezug zu Stärken der Schülerlösung, eine Erklärung für Defizite oder der Bezug zu möglichen Problemen des Unterrichts (Ebene 4). Abschließend kann ein Diagnoseurteil auch fehlerhaft oder unverständlich sein (Ebene 5). Situationsspezifische diagnostische Kompetenzen zeigen sich durch unterschiedliche Fähigkeiten (Korrigieren, beschreiben und analysieren) auf Ebene 2.

Auf der Grundlage der qualitativen Datenanalyse konnten bisher drei verschiedene Entwicklungsverläufe diagnostischer Kompetenzen aufgedeckt werden, die bei mehreren Lehrkräften auftreten:

- 1) Eine Lehrkraft diagnostiziert vor der Fortbildung allgemein, konkret, analytisch und bezieht sich auf ein Defizit. Im Anschluss an die Fortbildung wendet die Lehrkraft zusätzlich fachdidaktisches Wissen an.
- 2) Eine Lehrkraft diagnostiziert vor der Fortbildung konkret und beschreibend. Nach der Fortbildung diagnostiziert die Lehrkraft konkret, korrigierend, beschreibend, analytisch und bezieht sich auf ein Defizit der Schülerlösung.

- 3) Eine Lehrkraft diagnostiziert vor der Fortbildung konkret und beschreibend, nach der Fortbildung konkret, analytisch und auf ein Defizit bezogen. Dies stellt eine profundere Diagnose dar, da Förderpotential aus dieser Diagnose abgeleitet werden kann.

#### **4. Zur Frage der Validität**

Aus Sicht der quantitativen Forschung ist Validität durch das Forschungsdesign und den gezielten Ausschluss validitätsgefährdender Aspekte sicherzustellen. Bei der vorliegenden Studie wurden beispielsweise Vorerfahrungen der Probanden in den Bereichen Diagnose und Fortbildung erhoben, damit diese möglichen Störfaktoren ausgeschlossen werden oder in der Auswertung Berücksichtigung finden können. Die direkte Übertragung des Gütekriteriums der Validität aus der quantitativen Forschung ist nicht ausreichend, da sich Validität nicht nur auf Kennwerte eines Tests bezieht sondern auch während des Forschungsprozesses eine Rolle spielt.

Aus der qualitativen Forschung sind Validierungsmaßnahmen bekannt, die sich auf den kompletten Forschungsprozess beziehen (Winter, 2000). Nach Meyrick (2006) stellen Transparenz und Systematik zentrale Charakteristika valider qualitativer Forschung dar. Winter (2000) unterscheidet dabei zwei Typologien von Validität. Zum einen kann ein Forschungsprojekt den Anspruch der Inhalts-, Kriteriums- oder Konstruktvalidität verfolgen und zum anderen interne und externe Validität erfüllen. In dieser Studie wurde die Inhaltsvalidität als zentral zu erfüllendes Gütekriterium angesehen. Zur Instrumententwicklung wurde eine Pilotierung und Expertenbefragung durchgeführt, was nach Messick (1995) zur Inhaltsvalidität beitragen kann. Außerdem wurde bei der Interpretation der Ergebnisse regelmäßig Rückmeldungen anderer Forscherinnen und Forscher eingeholt. Interne Validität als der Grad, zu dem die Ergebnisse der untersuchten Kompetenzfacette zuzuordnen sind, sollte durch die im folgenden vorgestellte Maßnahme gesichert werden (Winter, 2000; Hammersley, 1987). Der Online-Fragebogen enthielt Items zu Vorerfahrungen der Probanden bezüglich Fortbildungen und Diagnose sowie dem Unterrichten von Funktionen. Mit dieser Maßnahme konnte die interne Validität erhöht werden. Externe Validität kann lediglich innerhalb der Theoriegenerierung (Kategorienbildung) bezogen auf die untersuchte Population als erfüllt angesehen werden.

#### **5. Ausblick**

Im weiteren Verlauf der Studie sollen quantitative Datenanalysen durchgeführt werden und dabei unterschiedliche Ausprägungen der Kategorien erfasst werden. Da sich bisherige Auswertungen lediglich auf ein Drittel des

Datensatzes beziehen, werden die Überprüfung und eine mögliche Erweiterung der Entwicklungsverläufe an weiteren Daten vorgenommen. Bei der Interpretation der Ergebnisse werden sowohl validitätsgefährdende Aspekte im Projekt als auch alternative Argumentationsstränge in die Ergebnisdarstellung integriert, um die Ergebnisinterpretation so valide wie möglich zu gestalten (Kane, 2001).

## Literatur

- Baumert, J., Blum, W., Brunner, M., Dubberke, T., Jordan, A., Klusmann U., Krauss, S., Kunter, M., Löwen, K., Neubrand, M., Tsai, Y.-M. (2008). *Professionswissen von Lehrkräften, kognitiv aktivierender Mathematikunterricht und die Entwicklung von mathematischer Kompetenz (COACTIV). Dokumentation der Erhebungsinstrumente*. Materialien aus der Bildungsforschung, 83. Berlin: Max-Planck-Institut für Bildungsforschung.
- Dörmann, M., Kaiser, G., Blömeke, S. (2012). The conceptualisation of mathematics competencies in the international teacher education study TEDS-M. *ZDM Mathematics Education*, 44, 325-340.
- Hammersley, M. (1987). Some notes on the terms 'validity' and 'reliability'. *British Educational Research Journal*, 13 (1), 73-81.
- Kane, M.T. (2001). Current concerns in validity theory. *Journal of Educational Management*, 38 (4), 319-342.
- Lorenz, C., Artelt, C. (2009). Fachspezifität und Stabilität diagnostischer Kompetenz von Grundschullehrkräften in den Fächern Deutsch und Mathematik. *Zeitschrift für Pädagogische Psychologie*, 23 (3-4), 211-222.
- Messick, S. (1995). Validity of Psychological Assessment. Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50 (9), 741-749.
- Meyrick, J. (2006). What is good qualitative research? A first step towards a comprehensive approach to judging rigour/ quality. *Journal of Health Psychology*, 11 (5), 799-808.
- Schrader, F.-W. (2011). Lehrer als Diagnostiker. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (S. 683-698). Münster: Waxmann.
- Teller, J., Barzel, B., & Leuders, T. (2013). Förderung Diagnostischer Kompetenzen von Lehrerinnen und Lehrern im Bereich Funktionales Denken: Eine Interventionsstudie. In G. Greefrath, F. Käpnick, & M. Stein (Hrsg.), *Beiträge zum Mathematikunterricht*, 2 (S. 1002-1005). Münster: WTM Verlag für wissenschaftliche Texte und Medien.
- Winter, G. (2000). A comparative discussion of the notion of 'validity' in qualitative and quantitative research. *The Qualitative Report*, 4 (1&2). Online 20.03.2014: <http://www.nova.edu/ssss/QR/QR4-3/winter.html>