

Timo LEUDERS, Freiburg

## **Validität von Modellierungen mathematischer Kompetenzen**

### **Kompetenzmodellierung**

Die Ergebnisorientierung in der Bildungspolitik und die Forderung nach diagnostischen, der Heterogenität der Schülerschaft Rechnung tragenden Lehrformen, haben das Interesse an Kompetenzmodellierungen gestärkt. Die Fachdidaktiken sind gefragt, fachbezogene Theorien zu Schülerkompetenzen zu konstruieren und empirisch zu fundieren praktikabler Diagnoseinstrumente für die Unterrichtspraxis zu entwickeln.

In dem Überblicksartikel von Leuders (2014) wird ein vergleichender Blick auf einige prototypische Anwendungen der Modellierung mathematischer Kompetenzen geworfen, wie sie in den letzten Jahren vorgeschlagen oder umgesetzt wurden. Ziel ist dabei, theoretisch aufzuzeigen, auf welche Weise die *Validität* solcher Modelle bewertet werden kann und welche Rolle hierbei die fachdidaktische Perspektive spielt. Es wird aufgezeigt, wie eine systematische Bewertung von sechs Validitätsaspekten (inhaltliche, kognitive, strukturelle, generalisierende, externe und konsequentielle Validität) eine differenzierte Einschätzung bestehender Anwendungen von Kompetenzmodellierungen und Hinweise für deren Weiterentwicklung liefern kann.

Im Vortrag wurden einige Beispiele für Kompetenzmodellierungen aus dem genannten Beitrag dargestellt und hinsichtlich ihrer Validität diskutiert. Kompetenzmodellierungen zeichnen sich dadurch aus, dass sie als *theoretischen* Ausgangspunkt einen *Kompetenzbegriff* wählen (Weinert 2001; Klieme 2004) und dass sie *methodisch* auf neuere *psychometrische Modelle* latenter Persönlichkeitsmerkmale zurückgreifen (Rost 2004; Wu u. Adams 2007). Eine Aufgabe der Fachdidaktiken besteht folglich darin, ihr spezifisches Wissen über den Kontext in die Kompetenzforschung einzubringen (Wendt u. Bos 2011; Leuders 2011). Der Begriff *Kompetenzmodellierung* hebt den Prozesscharakter der Konstruktion und Anwendung von Kompetenzmodellen hervor. Diese Auffassung findet man auch im DFG-Schwerpunktprogramm „Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen“ (2007-2013, Klieme u. Leutner 2006), welches sich entlang der Bereiche „theoretische Modelle“, „psychometrische Modelle“, „Messkonzepte und Messverfahren“ und „Nutzung von Diagnostik und Assessment“ strukturiert (vgl. Pellegrino, Chudowsky u. Glaser, 2001).

## Validitätsaspekte

Betrachtet man Kompetenzmodellierung unter der Frage der Validität, so eignet sich in besonderer Weise die Perspektive von Messick (1989, 1995), welche Validität nicht als eine Reihe von Eigenschaften eines Messinstruments sondern als fortwährendem Prozess der argumentativen und empirischen Verteidigung miteinander verbundener Validitätsaspekte ansieht. Die Fachdidaktiken können von dieser differenzierteren Sichtweise auf Validität profitieren. Messick (1995) unterscheidet und beschreibt sechs Validitätsaspekte: (1) Inhaltliche Validität: Curriculare und theoretische Absicherung des modellierten Bereichs (content aspect), (2) Kognitive Validität: Passung der kognitiven Prozesse bei der Kompetenzerfassung zum postulierten theoretischen Kompetenzmodell (substantive aspect), (3) Strukturelle Validität: Passung von theoretischem Kompetenzmodell und gewähltem psychometrischem Messmodell (structural aspect), (4) Verallgemeinerbarkeit: Angemessenheit einer über die Aufgaben- und Personengruppe hinausgehenden Interpretation (generalizability aspect), (5) Externe Validität: Angemessenheit mit Blick auf konvergente, diskriminante und prädiktive Zusammenhänge mit anderen Konstrukten (external aspect), (6) Konsequentielle Validität: Angemessenheit der Nutzung im pädagogischen oder bildungspolitischen Kontext (consequential aspect).

Diese sechs Validitätsaspekte erlauben eine systematische, alle Schritte einer Kompetenzmodellierung durchdringende Validitätsanalyse, die dabei hilft, die für jeden Schritt spezifischen Validitätsbedrohungen zu erkennen und zu bewältigen.

**(1) Inhaltliche Validität (content aspect):** Am Beginn einer Kompetenzmodellierung steht die Setzung eines inhaltlichen Rahmens, dessen Breite und Auflösungsvermögen erheblich variieren kann: von der gesamten schulischen Domäne Mathematik bis hinunter zur Beschreibung bestimmter Einzelfähigkeiten, wie z.B. der Addition im Zehneraum.

Bei der Modellierungen breiter und mittlerer Kompetenzbereiche fußt die Absicherung der inhaltlichen Validität oft auf einer Ableitung aus normativen Beschreibungen eines Kompetenzbereiches, was man auch als curriculare Validität bezeichnen kann. Bei mittlerer Breite kann die von Newell u. Simon (1972) beschriebene cognitive task analysis (CTA) besonders affin zum Kompetenzkonzept, da sich Kompetenzen sich ja qua Definition über typische Anforderungssituationen definieren (Weinert 2001). Dabei besteht oft ein Spannungsverhältnis zwischen deskriptiven und normativen Zielsetzungen, die mit einer Kompetenzmodellierung vorgenommen werden: Sollen beispielsweise die Kompetenzmodelle, die Bildungsstandards zugrunde

liegen, das normalerweise Erreichbare beschreiben oder sollen sie neue curriculare Impulse setzen? Bei der Modellierung engerer Kompetenzbereiche können Theorien über die jeweils relevanten Kognitionen zu Rate gezogen werden (Embretson 1994; Mislevy 1996; Rupp u. Mislevy 2007). Dies kann sowohl theoretisch als auch empirisch durchgeführt werden, was in der Fachdidaktik in etwa den beiden komplementären Ansätzen der stoffdidaktischen Analyse durch Expertinnen und Experten einerseits und der empirischen Analyse individueller Lernprozesse in Fallstudien andererseits entspricht.

**(2) Kognitive Validität (substantive aspect):** Die Brücke zwischen theoretischen Kompetenzmodellen und ihrer empirischen Erfassung bildet die Konkretisierung der Theorieelemente durch Aufgabensituationen. Eine mangelnde Passung zwischen den theoretischen Konstrukten und den tatsächlich ablaufenden Kognitionen, die durch die spezifische Operationalisierung angeregt werden, kann eine ernste Validitätsbedrohung darstellen. Der Weg von den allgemeinen Situationen, welche als konstitutiv für einen Kompetenzbereich angesehen werden, bis hin zu den konkreten Erfassungssituationen gleicht einem mehrschrittigen Übersetzungsvorgang, bei dem jedesmal die inhaltliche Bedeutung beeinträchtigt werden kann.

Argumente zur kognitiven Validität können theoretischer oder empirischer Natur sein. Sie können über das Urteil von Expertinnen und Experten generiert werden (Rubio et al. 2003) oder durch eine Untersuchung der Bearbeitungsprozesse unter testnahen Bedingungen in so genannten „cognitive labs“ (Snow u. Lohman 1989) erfolgen.

**(3) Strukturelle Validität (structural aspect):** Zur psychometrischen Beschreibung von Kompetenzen kann auf eine Vielzahl von Messmodellen zurückgegriffen werden, welche durch eine probabilistische Modellierung latenter Fähigkeiten dem nicht-deterministischen Charakter menschlicher Leistungen Rechnung tragen (Rost 2004; DiBello et al. 2007). Eine Vielzahl von Modellvarianten ergibt sich durch (i) Art und Zahl zu schätzender Parameter, (ii) Art und Zahl der latenten Variablen und (iii) strukturelle Eigenschaften (z.B. Dimensionalität). Die große Vielfalt dieser Modelle erlaubt es, solche auszuwählen, die das Verhalten der Probanden auf eine möglichst passende Weise beschreiben und damit reliable und valide Messungen ermöglichen. Wenn eine solche Passung jedoch nicht exploratorisch durch Anpassung hinreichend vieler Parameter geschehen soll, braucht es a priori eine strukturelle Korrespondenz zwischen gewähltem psychometrischem Modell und dem zu modellierenden Kompetenzbereich. Hartig (2008) zählt zu Kriterien für eine strukturell valide Modellwahl die angemessene Wahl der Art der Personenvariablen (kontinuierlich/kategorial),

der dimensionalen Struktur oder der (nicht)kompensatorischen Modellierungsansätze.

**(4) Verallgemeinerbarkeit (generalizability aspect) und**

**(5) Externe Validität (external aspect)** wurden im Vortrag nicht behandelt – s. aber Leuders (2014).

**(6) Konsequentielle Validität (consequential aspect):** Die Nutzung von Kompetenzmodellierungen für das Bildungsmonitoring auf Systemebene, die Rückmeldung von Leistungsdaten auf Klassen- und Schulebene mit dem Ziel der Schul- und Unterrichtsentwicklung, die Professionalisierung im Sinne der Förderung diagnostischer Kompetenzen von Lehrkräften oder die Unterstützung individueller Diagnose zur Vorbereitung pädagogischer Förderentscheidungen bedarf jeweils angepasster Formate (Pellegrino et al. 2001, S.2: „Often a single assessment is used for multiple purposes; in general, however, the more purposes a single assessment aims to serve, the more each purpose will be compromised. For instance, many state tests are used for both individual and program assessment purposes. This is not necessarily a problem, as long as assessment designers and users recognize the compromises and trade-offs such use entails.“

### **Beispielmodellierungen**

Im Vortrag und in Leuders (2014) wurde eine Reihe von Kompetenzmodellierungen für den Mathematikunterricht der letzten Jahrzehnte vorgestellt und hinsichtlich der genannten Validitätsaspekte diskutiert:

*Typ 1: Eindimensionale Kompetenzskalen* für breite curricularen Bereichen mit eine kontinuierliche Fähigkeitsvariable zum direkten Vergleich zwischen Individuen oder Gruppen. *Typ 2: Mehrdimensionale Kompetenzskalen* für einen curricular engeren Bereich durch mehrere kontinuierliche Fähigkeitsvariablen zur Identifikation von unterschiedlichen Fähigkeitsprofilen. *Typ 3: Kompetenzstufenmodelle* mit theoretisch a priori definierten Stufen sowie *Typ 4: Kategoriale Kompetenzstrukturmodelle* mit einem engen Fähigkeitsbereich auf der Basis kognitiver Theorien und kategorialer Fähigkeitsvariablen zur Rückmeldung über die Zugehörigkeit zu einem bestimmten Kompetenzprofil („kognitive Diagnosemodelle“)

Die kritische Darstellung von Vorzügen aber auch typischen Validitätsbedrohungen, sowie eine ausführliche Literaturliste findet man bei

Leuders, T. (2014). Modellierungen mathematischer Kompetenzen – Kriterien für eine Validitätsprüfung aus fachdidaktischer Sicht. *Journal für Mathematik-Didaktik*, 35(1), 7-48. doi: 10.1007/s13138-013-0060-3