

Inwiefern sind Kompetenzmodelle dazu geeignet kognitive Prozesse von Lernenden zu beschreiben?

Timo Leuders · Beate Sodian

Zusammenfassung: Die zunehmende Verwendung von Kompetenzmodellierungen in den Domänen schulischen Wissens wirft die Frage auf, inwieweit Kompetenzmodelle geeignet sind, auch kognitive Prozesse abzubilden. Der Beitrag analysiert in dieser Hinsicht aktuelle Entwicklungen im Bereich der Kompetenzskalen, Kompetenzstufenmodelle, Kompetenzstrukturmodelle und kognitiven Diagnosemodelle und bewertet ihren Einsatz zur Erfassung von Lernprozessen in der fachdidaktischen und entwicklungspsychologischen Forschung.

Schlüsselwörter: Kompetenzmodelle · Kognitive Prozesse · Kompetenzentwicklung

To what extent can competence models describe cognitive processes?

Abstract: The increasing use of competence models for modelling domain specific knowledge in school raises the question to what extent such models can represent cognitive processes. The article analyses the recent development of models for competence scales, competence stages, competence structures and so-called cognitive diagnostic models and taxes their value in the context of subject specific educational sciences and developmental psychology.

Keywords: Cognitive processes · Competence development · Competence models

© Springer Fachmedien Wiesbaden 2013

Prof. Dr. T. Leuders (✉)
Institut für Mathematische Bildung, Pädagogische Hochschule Freiburg,
Kunzenweg 21, 79117 Freiburg, Deutschland
E-Mail: leuders@ph-freiburg.de

Prof. Dr. B. Sodian (✉)
Entwicklungspsychologie und Pädagogische Psychologie, Ludwig-Maximilians-Universität
München, Leopoldstr. 13, 80802 München, Deutschland
E-Mail: Sodian@psy.lmu.de

Einhergehend mit der fortschreitenden empirischen Fundierung der Fachdidaktiken sowie mit den Bemühungen um eine empiriebasierte Schulentwicklung, mehren sich auch die Versuche der psychometrischen Modellierung von Kompetenzen in den verschiedenen Domänen schulischen Wissens (vgl. Leuders 2011). Dabei stellt sich die Frage, inwieweit solche Kompetenzmodelle – bei zunehmender Verfeinerung – in der Lage sind, kognitive Prozesse der Lernenden zu beschreiben. Im Folgenden möchten wir an verschiedenen Beispielen die Frage beleuchten, inwieweit bestehende Versuche der Kompetenzmodellierung kognitive Prozesse abzubilden instande sind.

1 Kompetenzskalen und post-hoc-Kompetenzstufen

Das wohl gängigste und einflussreichste Verfahren für die Modellierung von Kompetenzen ist die in Large-Scale-Assessments wie PISA und TIMSS verwendete Erfassung von domänenspezifischen Kompetenzen über eindimensionale, kontinuierliche Leistungsskalen (vgl. Baumert et al. 2000; PISA-Konsortium Deutschland 2007) in Form von Rasch-Modellen. Die hierbei postulierten, operationalisierten und empirisch überprüften Konstrukte modellieren beispielsweise „mathematische Grundbildung“ (*mathematical literacy*) oder „Lesekompetenz“ (*reading literacy*). Die genannten Studien beziehen sich schon in ihrer Anlage nicht primär auf kognitive Prozesse. Ihre Aufgabe ist die Bereitstellung eines eindimensionalen Index, der einen (am normativen Rahmen) gemessenen Vergleich der Schülerleistungen zwischen Bildungssystemen erlaubt.

Um die so gemessenen Leistungsunterschiede interpretierbar zu machen, kann man auf die besondere Eigenschaft der Rasch-Skala zurückgreifen: Da die latente Leistungsvariable der Proband(inn)en und die Schwierigkeitsvariable der Items auf einer gemeinsamen Skala liegen, kann man – wie es üblicherweise praktiziert wird – versuchen, die Fähigkeiten der Proband(inn)en durch die Anforderungen der von ihnen potenziell einigermmaßen zuverlässig bewältigbaren Aufgaben zu beschreiben. Auf diese Weise entstand das Verfahren des „scale anchoring“ (Beaton und Allen 1992), welches zum Standardverfahren für die inhaltliche Interpretation von Rasch-Skalen geworden ist: Die empirisch geprüfte Skala wird post-hoc auf der Basis einer inhaltlichen Analyse der Anforderungen in Stufen unterteilt und den Proband(inn)en, deren Fähigkeitsvariable in einem entsprechenden Intervall liegt, wird die den dort lokalisierten Aufgaben entsprechende Kompetenz zugeschrieben. Die Validität dieses Verfahrens stand aber von Beginn an in der kritischen Diskussion. Die Kritik lässt sich folgendermaßen zusammenfassen: 1) Es besteht Unklarheit über die Stabilität und daher die Validität der durch Schwellenwerte spezifizierten Stufen in Abhängigkeit von der Variabilität der gewählten Items oder der befragten Expert(inn)en. 2) Die Interpretation von Aufgabenmerkmalen wird qua Identifikation auf der Rasch-Skala auf Personenmerkmale übertragen. Bei diesem Verfahren hängt der Rückschluss auf kognitive Prozesse stark von post-hoc Interpretationen ab; es liegt kein vorab formuliertes Modell zugrunde, das einer empirischen Überprüfung unterzogen werden könnte. 3) Auch gegenüber der Anwendung einer solchen „interpretierten Skala“ als Feedbackinstrument für die zentrale Leistungsmessung (wie z. B. bei Reiss und Winkelmann 2008) besteht Skepsis: Die Nützlichkeit einer vereinfachenden Einordnung von Schülerinnen und Schülern in eine Kompetenzstufe scheint zurzeit weder

für die Selbsteinschätzung der Lernenden noch für die Unterrichtsentwicklung durch die Lehrpersonen eine tragende Funktion zu entwickeln (vgl. Schulz 2010). Schließlich gibt es kein allgemein anerkanntes Rationale, nach dem entschieden werden kann, ob empirische Abweichungen von einer solchen Skala entweder zu einer Entfernung der entsprechenden Items oder zur Postulierung nicht erfasster kognitiver Prozesse und somit zu einer Erweiterung des Modells führen sollten.

Ein Vorschlag zur empiriebasierten Generierung von Kompetenzstufen wurde von Hartig (vgl. 2007) im Rahmen der DESI-Studie entwickelt. Anzahl und Breite der Kompetenzniveaus werden in diesem Ansatz nicht willkürlich festgelegt, sondern nach empirischen Kriterien bestimmt. Für die Definition von Niveaus sind drei separate Analyseschritte notwendig: Zunächst werden die empirischen Aufgabenschwierigkeiten anhand einer Rasch-Analyse ermittelt. Dann werden den Aufgaben schwierigkeitsbestimmende Merkmale zugeordnet und die empirischen Aufgabenschwierigkeiten werden dann mittels einer Regressionsanalyse auf diese Merkmale (bzw. eine Auswahl solcher Merkmale, die insgesamt eine hinreichende Varianz der Schwierigkeit erklären können) zurückgeführt. Die Grenzen zwischen den Kompetenzstufen schließlich findet man bei den Schwierigkeitswerten, die sich rechnerisch aus den Regressionskoeffizienten bei Vorliegen verschiedener Kombinationen der Merkmale ergeben. Dieses Vorgehen stellt das Auffinden von Kompetenzstufen auf eine rationale Basis; bislang fehlen aber systematische Analysen dazu, wie plausibel, stabil und valide die so gewonnenen Stufen sind. Immerhin lässt sich durch Auswahl geeigneter kognitionsnaher Merkmale von Aufgaben – sofern sich solche finden lassen – eine höhere kognitive Plausibilität der Stufen erreichen.

2 Kompetenzstruktur- und Kompetenzentwicklungsmodelle

Kognitionspsychologische Forschung zur begrifflichen Entwicklung und zum begrifflichen Wandel (*conceptual change*) hat für verschiedene Domänen theoretische und empirische Grundlagen für Kompetenzstruktur- bzw. entwicklungsmodelle geliefert (vgl. z.B. Vosniadou und Verschaffel 2004). So sind zum Beispiel Entwicklungsverläufe der begrifflichen Differenzierung von Gewicht und Dichte oder der Restrukturierung kosmologischer Begriffe im Grundschulalter gut erforscht. Auf dieser Basis können Kompetenzstrukturmodelle entwickelt werden, die unter anderem Annahmen über das gemeinsame Auftreten von verschiedenen Fehlvorstellungen enthalten, sowie Kompetenzentwicklungsmodelle, die Annahmen über distinkte Schritte oder Stufen auf dem Weg zu einem wissenschaftlich adäquaten Verständnis einer Domäne enthalten. Bisher stehen Bemühungen, theoriegeleitet psychometrische Modelle der Kompetenzentwicklung zu erstellen und zu überprüfen, allerdings noch ganz am Anfang.

Ein Beispiel für diesen Ansatz sind die Arbeiten von Wilson und Sloane (2000; siehe auch Wilson 2005). Sie untersuchen Schülerfähigkeiten zunächst inhaltlich und an konkreten Aufgaben bzw. auf Basis bestehender Theorien zum Begriffsaufbau. Die qualitative Analyse von kognitiven Prozessen ermöglicht eine Ausdifferenzierung der theoretischen Modelle und liefert Indikatoren für die Operationalisierung eines validen Ratingverfahrens der Schülerantworten. Nach Wilson (2005) können geeignet konstruierte reflexive,

offenere Aufgaben eingesetzt werden, um Lernverläufe zu erfassen (z. B. „Warum sinken oder schwimmen Dinge? Erkläre so ausführlich wie möglich!“; vgl. Kennedy und Draney 2007). Diese Aufgaben können organisch in den Unterrichtsprozess eingebettet und im formativen Sinne für die Planung des Unterrichts gezielt genutzt werden (*embedded assessment*). Zugleich führt die lernbegleitende, durch die Lehrkräfte vorgenommene Analyse von Schülerantworten auf diese offeneren Aufgaben zu der Zuweisung einer Einzelleistung zu einer inhaltlich definierten Fähigkeitsstufe (im genannten Beispiel auf die Erklärung des Auftriebs auf verschiedenen abgrenzbaren Verstehensniveaus:

1. Nicht anschlussfähige Misskonzepte,
2. Bezug auf *eine* bedeutsame Variable: Masse *oder* Volumen,
3. Bedeutung beider Variablen,
4. Bedeutung des Quotienten = Dichte,
5. Verständnis der relativen Dichte.

Auf der Basis solcher Ratingdaten kann nun eine polytome Skala zur empirischen Absicherung der Lernfortschritte entwickelt werden. Das hier skizzierte Verfahren zeichnet sich dadurch aus, dass die Leistungsskala *a priori* auf der kognitiven Interpretation einer Schülerleistung basiert (nämlich der theoriegeleiteten Entwicklung sogenannter *construct maps*) und nicht erst *a posteriori* eine solche Interpretation zu einer theoretisch wenig durchgeformten Itemgruppe zuordnet.

Ein Versuch zur theoriegeleiteten Erstellung eines Struktur- und Entwicklungsmodells für naturwissenschaftliche Kompetenz in der Grundschule wird in einem Teilprojekt des DFG-Schwerpunktprogramms unternommen (Hardy et al. 2010; Kleickmann et al. 2011; Koerber et al. 2011). Die Einbeziehung zweier Kompetenzdimensionen (naturwissenschaftliches Wissen und Wissen über Naturwissenschaften) und mehrerer Inhaltsbereiche sowie die Anforderungen an eine psychometrische Modellierung erzwingen allerdings eine Generalisierung über domänenspezifische Entwicklungsverläufe hinweg, sodass heuristisch ein allgemeines Modell des Übergangs von einem naiven Verständnis über eine Ebene der „Zwischenvorstellungen“ hin zu einem wissenschaftlich integrierten Verständnis erstellt wurde. Die Realisierung verschiedener Aufgabenformate sowie die Einbeziehung von Kontrollvariablen (z. B. Leseverständnis) erlauben es zu kontrollieren, dass die Schwierigkeitsunterschiede zwischen den Items tatsächlich auf die angenommenen Kompetenzstufen zurückzuführen sind.

3 Mehrdimensionale Kompetenzstrukturmodelle

Erkennt man an, dass die beim Versuch der Konstruktion eindimensionaler Skalen auftretenden Diskrepanzen ein Indiz für differenziertere Kompetenzstrukturen sind, so legt dies den Versuch nahe, solche Strukturen auch theoretisch zu modellieren und empirisch zu überprüfen. Einen solchen Ansatz verfolgt ein weiteres Projekt aus dem Schwerpunktprogramm (Bayrhuber et al. 2010; Wirtz et al. 2012). Hier wird innerhalb eines engen Fokus auf den Bereich des elementaren Problemlösens mit Funktionen untersucht, inwiefern unterschiedliche externe Repräsentationen von Funktionen (Graph, Tabelle, Situation) und die Übersetzungsprozesse zwischen ihnen in einem Kompetenzstrukturmodell

abbildbar sind. In der Tat stellt sich heraus, dass ein vierdimensionales Raschmodell, das die verschiedenen Übersetzungen zwischen Situation, Tabelle und Graph modelliert, die empirischen Daten einer Querschnittserhebung in siebten und achten Klassen im Vergleich mit „gröberen“ Modellen niedriger Dimensionen am besten wiedergibt. Die hier gewählten Kompetenzdimensionen repräsentieren kognitive Prozesse (z. B. Übersetzen zwischen Eigenschaften funktionaler Graphen und der durch sie repräsentierten Situationen) und liegen offenbar bei Probandengruppen in unterschiedlicher Ausprägung vor. Dennoch ist festzustellen, dass trotz des relativ engen Fokus immer noch ganze Bündel kognitiver Prozesse untrennbar voneinander erfasst werden. So kann zum Beispiel anhand der Operationalisierung nicht festgestellt werden, ob eine Übersetzung zwischen zwei Repräsentationen in die eine oder die andere Richtung verläuft.

Eine dimensionale Analyse eines fokussierteren Kompetenzbereiches versuchen Hammann et al. (vgl. 2007) mit der „experimentellen Kompetenz“ im Bereich Biologie. Hier zeigt sich jedoch, dass die empirisch tragfähigen Dimensionen eher die unterschiedlichen Aufgabentypen widerspiegeln (Multiple Choice, Experiment, Offene Frage) als die inhaltlichen Fähigkeitskomponenten (Suche im Hypothesenraum, Testen von Hypothesen, Analyse von Evidenzen). Die genauere Betrachtung der hier beschriebenen Kompetenzdimensionen lässt zudem Zweifel aufkommen, ob diese Dimensionen als voneinander trennbare kognitive Prozesse aufzufassen sind.

4 Kognitive Diagnosemodelle

Kognitive Diagnosemodelle postulieren a priori den Zusammenhang von gewissen kognitiv interpretierbaren Teilprozessen bei der Bearbeitung einer Aufgabe, insbesondere können sie kompensatorische und nicht-kompensatorische Aspekte des logischen Zusammenhangs der untersuchten Fähigkeit abbilden (vgl. Hartig 2008; Rupp und Mislevy 2007).

Nicht-kompensatorische Modelle gehen davon aus, dass bestimmte Teilprozesse zur Lösung einer Aufgabe notwendig sind und nicht durch Stärken in einem anderen Teilprozess kompensiert werden können. Die Erfahrungen mit solchen Modellen sind allerdings noch begrenzt. Die Ergebnisse hängen stark von der Validität des postulierten Modells ab, wie am folgenden Beispiel nach Kunina-Habenicht et al. (vgl. 2009) diskutiert werden soll. Hier wurden zur Erfassung der arithmetischen Kompetenzen von Grundschülerinnen und -schülern Items zu den Grundrechenarten über eine sogenannte *Q-Matrix* von vier latenten Variablen (Addition, Subtraktion, Multiplikation und Divisionskompetenz) in Beziehung gesetzt. Die Ergebnisse zeigen, dass ein solches Modell empirisch nicht trägt und bestenfalls zwischen „Punktrechnen“- und „Strichrechnen“-Kompetenz unterschieden werden kann. Dabei bleibt es unklar, ob diese aufgedeckte Stufung nicht allein durch die curriculare Chronologie induziert ist. Die von der fachdidaktischen Forschung der letzten Jahrzehnte zutage geförderten Analysen von Rechenstrategien von Schülerinnen und Schülern (vgl. z. B. Verschaffel et al. 2007) legen nahe, dass eine rein fachlich-systematische Trennung der Grundrechenarten nicht kognitiv valide ist, sondern dass es eher einer Ausdifferenzierung typischer Rechenstrategien (z. B. Subtrahieren durch Ergänzen, Rechnen bis zum nächsten Zehner usw.) bedarf.

Trotz dieser kritischen Anmerkungen sehen wir künftig in der Zusammenarbeit von psychometrischer Methodenentwicklung und fachdidaktischer Erforschung kognitiver Prozesse eine Chance für die Entwicklung empirisch tragfähiger Kompetenzstrukturmodelle. Insbesondere scheint es angeraten, die Entwicklung von Kompetenzmodellen stärker mit empirischen Analysen auf Individual- und Klassenebene zu vernetzen, also etwa mit der Klassifikation typischer Fehlvorstellungen in einem Gegenstandsbereich oder mit der Erfassung unterschiedlicher curricularer Rahmenbedingungen der getesteten Schülergruppen.

5 Fazit

Aus der Bewertung bisheriger Ansätze gewinnen wir die folgenden Kriterien für künftige Versuche der Modellierung kognitiver Prozesse durch psychometrische Kompetenzmodelle: Die Konstruktion des Kompetenzmodelles sollte sich von Anfang an auf empirisch gestützte Theorien zu kognitiven Prozessen beziehen. Die Validität des Modells und seiner Operationalisierungen sollte zudem durch Einbindung in geeignete kognitionspsychologische und fachdidaktische Analysen gesichert werden (z. B. durch qualitative Analysen aus Fallstudien).

Diese Kriterien beschreiben Minimalanforderungen an Kompetenzmodelle als Voraussetzung für eine kognitive Interpretation. Aber selbst bei empirisch bestätigten Kompetenzmodellen ist Vorsicht geboten vor allzu einfachen Rückschlüssen von Kovariationsstrukturen interindividueller Unterschiede auf die Dynamiken mentaler Prozesse (vgl. Renkl 2012). Kompetenzmodellierungen können kognitionswissenschaftliche Studien ergänzen, aber nicht ersetzen.

Danksagung: Diese Veröffentlichung wurde ermöglicht durch Sachbeihilfen der Deutschen Forschungsgemeinschaft (Kennz.: LE 2335/1-2 und SO 213/29-2) im Schwerpunktprogramm „Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen“ (SPP 1293).

Literatur

- Baumert, J., Bos, W., & Lehmann, R. (2000). *TIMSS/III: Dritte Internationale Mathematik- und Naturwissenschaftsstudie – Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn*. Opladen: Leske + Budrich.
- Bayrhuber, M., Leuders, T., Bruder, R., & Wirtz, M. (2010). Repräsentationswechsel beim Umgang mit Funktionen – Identifikation von Kompetenzprofilen auf der Basis eines Kompetenzstrukturmodells. In E. Klieme, D. Leutner & M. Kenk (Hrsg.), *Kompetenzmodellierung Eine aktuelle Zwischenbilanz des DFG-Schwerpunktprogramms* (56. Beiheft der Zeitschrift für Pädagogik, S. 28–39). Weinheim: Beltz.
- Beaton, E., & Allen, N. (1992). Interpreting scales through scale anchoring. *Journal of Educational Statistics*, 17, 191–204.
- Hammann, M., Phan, T. H., & Bayrhuber, H. (2007). Experimentieren als Problemlösen: Lässt sich das SDDS-Modell nutzen, um unterschiedliche Kompetenzen beim Experimentieren zu messen. In I. Gogolin, M. Prenzel & H.-H. Krüger (Hrsg.), *Kompetenzdiagnostik* (8. Sonderheft der Zeitschrift für Erziehungswissenschaft, S. 33–49). Wiesbaden: VS Verlag für Sozialwissenschaften.

- Hardy, I., Kleickmann, T., Koerber, S., Mayer, D., Möller, K., Pollmeier, J., Schwippert, K., & Sodian, B. (2010). Die Modellierung naturwissenschaftlicher Kompetenz im Grundschulalter. In E. Klieme, D. Leutner & M. Kenk (Hrsg.), *Kompetenzmodellierung Eine aktuelle Zwischenbilanz des DFG-Schwerpunktprogramms* (56. Beiheft der Zeitschrift für Pädagogik, S. 115–125). Weinheim: Beltz.
- Hartig, J. (2007). Skalierung und Definition von Kompetenzniveaus. In B. Beck & E. Klieme (Hrsg.), *Sprachliche Kompetenzen: Konzepte und Messung – DESI-Studie* (S. 72–82). Weinheim: Beltz.
- Hartig, J. (2008). Psychometric models for the assessment of competencies. In J. Hartig, E. Klieme, & D. Leutner (Hrsg.), *Assessment of competencies in educational contexts* (S. 69–90). Göttingen: Hogrefe.
- Kennedy, C. A., & Draney, K. (2007). Interpreting and using multidimensional performance data to improve learning in science. In X. Liu & W. Boone (Hrsg.), *Applications of Rasch measurement to science education* (S. 247–284). Maple Grove: JAM Press.
- Kleickmann, T., Hardy, I., Pollmeier, J., & Möller, K. (2011). Zur Struktur naturwissenschaftlichen Wissens von Grundschulkindern. Eine personen- und variablenzentrierte Analyse. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 43, 200–212.
- Koerber, S., Sodian, B., Kropf, N., Mayer, D., & Schwippert, K. (2011). Die Entwicklung des wissenschaftlichen Denkens im Grundschulalter: Theorieverständnis, Experimentierstrategien, Dateninterpretation. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 43, 16–21.
- Kunina-Habenicht, O., Rupp, A. A., & Wilhelm, O. (2009). A practical illustration of multidimensional diagnostic skills profiling: Comparing results from confirmatory factor analysis and diagnostic classification models. *Studies in Educational Evaluation*, 35, 64–70.
- Leuders, T. (2011). Kompetenzorientierung – eine Chance für die Weiterentwicklung des Mathematikunterrichts? In K. Eilerts, A. Hilligus, G. Kaiser & P. Bender (Hrsg.), *Kompetenzorientierung in Schule und Lehrerbildung. Perspektiven der bildungspolitischen Diskussion, der Bildungsforschung und der Mathematik-Didaktik* (S. 287–306). Münster: LIT.
- PISA-Konsortium Deutschland. (Hrsg.). (2007). *PISA 2006. Die Ergebnisse der dritten internationalen Vergleichsstudie*. Münster: Waxmann.
- Reiss, K., & Winkelmann, H. (2008). Step by step: Ein Kompetenzstufenmodell für das Fach Mathematik. *Grundschule*, 40(10), 18–21.
- Renkl, A. (2012). Modellierung von Kompetenzen oder von interindividuellen Kompetenzunterschieden: Ein unterschätzter Unterschied? *Psychologische Rundschau*, 63, 50–53.
- Rupp, A. A., & Mislevy, R. J. (2007). Cognitive foundations of structured item response theory models. In J. Leighton & M. Gierl (Hrsg.), *Cognitive diagnostic assessment in education: Theory and applications* (S. 205–241). Cambridge: Cambridge University Press.
- Schulz, A. (2010). *Ergebnisorientierung als Chance für den Mathematikunterricht? Innovationsprozesse qualitativ und quantitativ erfassen*. München: Herbert Utz.
- Verschaffel, L., Greer, B., & De Corte, E. (2007). Whole number concepts and operations. In F. Lester (Hrsg.), *Second handbook of research on mathematics teaching and learning* (2. Aufl., S. 557–628). Charlotte: Information Age Publishing.
- Vosniadou, S., & Verschaffel, L. (2004). Extending the conceptual change approach to mathematics learning and teaching. *Learning and Instruction*, 14, 445–451.
- Wilson, M. (2005). *Constructing measures: An item response modeling approach*. Mahwah: Erlbaum.
- Wilson, M., & Sloane, K. (2000). From principles to practice: An embedded assessment system. *Applied Measurement in Education*, 13, 181–208.
- Wirtz, M., Naccarella, D., Henning, J., Kröhne, U., Bruder, R., & Leuders, T. (2012). *Komparative Analyse von Kompetenzstrukturmodellen im Bereich funktionaler Zusammenhänge. Vortrag gehalten auf dem 48. Kongress der Deutschen Gesellschaft für Psychologie (DGPs)*, Bielefeld.