Andreas SCHULZ, Timo LEUDERS, Freiburg

Entwicklung und Validierung eines kognitiven Diagnosemodells zur Eingangsdiagnose und -förderung in Klasse 5 – Teilmodell zu Schriftlichen Rechenverfahren

Die empirische Modellierung mathematischer Kompetenzen beruht auf der kohärenten Abstimmung vier grundlegender Arbeitsschritte: (1) Entwicklung eines theoretischen Modells, (2) Beschreibung durch ein psychometrisches Modell, (3) Operationalisierung mit geeigneten Messkonzepten und Messverfahren und (4) Nutzung im Rahmen von Diagnostik und Assessment. Vor dem Hintergrund des im schulischen Kontext wichtigen Anliegens, kriteriale individualdiagnostische Information bereit zu stellen, spielen kognitive Diagnosemodelle (Rupp & Mislevy, 2007) eine zunehmend bedeutsame Rolle (Leuders, 2014): Diese fokussieren (1) auf einen engen Kompetenzbereich, welcher über kognitive Theorien und Modelle definiert wird. Mittlerweile stehen (2) eine Vielzahl psychometrischer Modelle für die Anwendung zur Verfügung, die in Verbindung mit (3) validen Repräsentationen von Kognitionen, bspw. auf Grundlage von eingesetzten Aufgaben, (4) individuell unterschiedliche Merkmalsausprägungen auch auf kategorialen latenten Variablen rückmelden. Eine solche empirische Informationsgrundlage, welche in Passung zu einem kognitiven Modell über einen eng abgegrenzten Kompetenzbereich steht, eignet sich für die Anbahnung spezifischer Fördermaßnahmen.

Die kohärente Abstimmung dieser vier Arbeitsschritte für die Implementierung einer Eingangsdiagnose in Klasse 5, die in Baden-Württemberg landesweit und schulartenübergreifend zum Einsatz kommen soll, wird am Beispiel der schriftlichen Rechenverfahren mit Daten aus einer Pilotierungsstudie berichtet. In Entwicklung sind ebenfalls kognitive Diagnosemodelle für Operations- und Zahlverständnis sowie halbschriftliches/flexibles Rechnen.

Schriftliche Rechenverfahren sind Algorithmen, die sich der Stellenwertstruktur und mathematischer Verknüpfungsgesetze (Kommutativ-, Assoziativ- und Distributivgesetz) bedienen. Ihre Effektivität beruht darauf, dass nicht mehr komplette Zahlen im Blick behalten und miteinander verrechnet werden müssen, sondern dass regelgeleitete Operationen mit Ziffern zum Ergebnis führen. Aus Sicht der cognitive-load-Theory (vgl. Van Merrienboer & Sweller, 2005) können Rechenverfahren auch als domänenspezifische Schemata bezeichnet werden, welche im Langzeitgedächtnis gespeichert werden und welche es erlauben, Informationen so zu verarbeiten, dass kognitive Ressourcen im Arbeitsgedächtnis für weitergehende Aufga-

In J. Roth & J. Ames (Hrsg.), *Beiträge zum Mathematikunterricht 2014* (S. 1115–1118). Münster: WTM-Verlag

ben frei werden. Schriftliche Rechenverfahren ermöglichen Schülerinnen und Schülern zudem einen Einblick in die Bedeutung von Algorithmen als eine der zentralen Leitideen der Mathematik, vermitteln Rechensicherheit und vereinfachen stark das Rechnen mit großen Zahlen (Padberg & Benz, 2011).

Mehrdimensionales Stufenmodell zu schriftlichen Rechenverfahren

Bestehende Diagnosekonzepte und empirische Studien zu schriftlichen Rechenverfahren konzentrieren sich überwiegend auf die rechenartspezifische Erfassung und Analyse typischer Schülerfehler (bspw. Gerster, 1982/2012). Im Gegensatz hierzu war unser Ziel, ein rechenartenübergreifendes Stufenmodell zu entwickeln, welches Kompetenzen von Lernenden über Hierarchien in den Teilanforderungen schriftlicher Verfahren beschreibt. Für ein landesweites Screening-Instrument bietet dies den Vorteil, dass gut verständliche Rückmeldungen über den Leistungsstand von Einzelschülern und Schülergruppen generiert werden können. Hierbei kann ohne den vergleichsweise hohen Analyseaufwand bei der Erfassung von Fehlertypen dennoch zielgerichtet auf differenzierte Fördermaßnahmen auf unterschiedlichen Niveaustufen hingewiesen werden. Aus Sicht der Grundlagenforschung ergänzt die Studie den Forschungsstand zum Einfluss schwierigkeitsgenerierender Faktoren bei den schriftlichen Rechenverfahren.

<table>
<tr><th>Rechenartenübergreifende Stufenbeschreibung</th><th>Beispiele Subtraktion</th><th>Stufeninhärente Schwierigkeiten (Bsp. Subtraktion)</th></tr>
<tr><td>Stufe 1:
Einstellige Prozedur</td><td>8078 - 4018</td><td rowspan="4">- Unterschiedliche Stellenzahl
- Anzahl der Stellen
- Null im Minuend oder Subtrahend oder beides
- Null im Ergebnis (gleiche Ziffern) inkl. Wegfall der ersten Stelle
- Art des Übertrags (auf 0, 9, leere Stelle, mehrfach) ab Stufe 3</td></tr>
<tr><td>Stufe 2: Einschrittig-mehrstellige Prozedur</td><td>7037 - 5821</td></tr>
<tr><td>Stufe 3: Mehrschrittig-mehrstellige Prozedur</td><td>3018 - 74</td></tr>
<tr><td>Stufe 4:
Komplexe Prozedur</td><td>8765
- 4241
- 52</td></tr>
</table>

Auf Stufe 1 (einstellige Prozedur) sind alle Aufgaben durch rein ziffernweises Rechnen lösbar, es werden kein Übertrag (Subtraktion), Behaltensziffer (Multiplikation) oder Teilquotientenbildung mit Rest (Division) notwendig. Auf Stufe 2 (einschrittig-mehrstellig) erfordern die Aufgaben einen Übertrag, eine Behaltensziffer oder einen Teilquotienten mit Rest, wobei diese, im Unterschied zu Stufe 3 (mehrschrittig-mehrstellig), auf

Stufe 2 noch nicht mit zusätzlichen Schwierigkeiten im Übertrag etc. (Subtr.: Übertrag auf die 0, 9 oder leere Stelle, erneuter Übertrag) verbunden sind. Stufe 4 (komplexe Prozeduren) beinhaltet Aufgaben mit zweifachem Minuend oder zweistelligem Multiplikator bzw. Divisor. Die Variation der Aufgaben innerhalb einer Stufe ergibt sich aus den stufeninhärenten Schwierigkeiten (Subtr.: vgl. Tabelle). Das Stufenmodell postuliert, dass die Schwierigkeit einer Aufgabe vorrangig vom Stufenmerkmal bestimmt wird und nachrangig von den stufeninhärenten Schwierigkeiten (Subtr.: u.a. Rechnen mit der Null, ungleiche Stellenzahl, Anzahl der Stellen). Derart wurden Stufenmodelle für alle vier Rechenarten entwickelt.

Die Pilotierung erfolgte in 16 Klassen (GS: 2; WRS: 4; RS: 5; GY: 5; N = 403). Die relativen Lösungshäufigkeiten pro Verfahren gemittelt zeigten (Add: 0,94; Subtr: 0,79; Mult: 0,7; Div: 0,65), dass die schriftliche Addition im 4. bis 6 Schuljahr keine Schwierigkeiten mehr bereitet, sodass nur die 58 Items zu den drei anderen Rechenverfahren analysiert wurden.

Überprüfung der Dimensionalität

Die explorative Faktorenanalyse wies auf eine eindeutig dreidimensionale Struktur hin, die sich in der konfirmatorischen Faktorenanalyse im Modellvergleich deutlich bestätigte (Kriterium AIC/BIC minimal; 1-dim.: 16120/16352; 3-dim.: 14901/15161). In die Analysen gingen 57 Items mit einem Fit-Residual-Wert im jeweils eindimensionalen Raschmodell kleiner als 2,5 ein (vgl. RUMM2030-Manual; Sub/Mult/Div: 26/17/14 Items). Die latenten Korrelationen sind moderat: Sub-Mult: 0,525; Sub-Div: 0,459; Mult-Div: 0,599, die Reliabilitäten sehr gut (PSI für Gesamtstichprobe, in Klammern Cronbach-Alpha für Klasse 5 als Zielstichprobe): Sub: 0,765 (0,823); Mult: 0,829 (0,871); Div: 0,845 (0,929). Damit stellen die drei Rechenverfahren nicht nur konzeptuell unterschiedliche, sondern auch messmethodisch klar voneinander abgrenzbare Fähigkeitsbereiche dar.

Überprüfung der Stufen der Items / der Anforderungen

Die Hypothese, dass die Schwierigkeit der Items (Logitskala der Raschmodellierung) pro Rechenverfahren vorrangig über deren a-priori beschriebene Stufenzugehörigkeit erklärt werden kann, wurde mittels Regression der Dummy-codierten Stufenzugehörigkeiten (Stufe 1/2/3) auf die Itemschwierigkeit überprüft:

$$\mathit{Itemschwierigkeit} = k + b_2(\mathit{Stufe}2) + b_3(\mathit{Stufe}3) + \mathrm{b}_4(\mathit{Stufe}4)$$

Für die Subtr/ Mult/ Div ergab sich eine erklärte Varianz R^2 von 0,828/ 0,868/ 0,898. Nach einer leichten Präzisierung der Stufenmerkmale (nur noch Items mit 4-5 Stellenwerten & Gegebensein der Stufe 3-

Anforderungen auch für Stufe 4 Items) für Subtr./ Mult. vergrößerte sich deren erklärte Varianz nochmals auf 0,934/ 0,949. Die Itemschwierigkeit wird demnach zum allergrößten Teil von den originären Stufenmerkmalen und nicht von den stufeninhärenten Schwierigkeiten bestimmt.

Bestimmung von Kompetenzstufen auf Personenebene

Gängiges Vorgehen zur Bestimmung von Kompetenzstufen ist es, für das schwierigste Item einer Stufe eine Lösungswahrscheinlichkeit von 0,65 (bzw.0,5) zu fordern und diesen Itemschwierigkeitswert als Schwellenwert auf der Personen-Fähigkeitsskala zu verwenden. Hierbei entscheidet jedoch nur ein einzelnes Item über die Lage des Schwellenwertes, dabei wird die inhaltliche Klammer um die kohärenten Itemmerkmale der Stufe nicht berücksichtigt. Daher wurde als Kriterium für die Schwellenwerte verwendet, dass im arithmetischen Mittel alle Items einer Stufe (anstelle nur des schwierigsten Items) mit einer bestimmten erwarteten Wahrscheinlichkeit gelöst werden. Für die Lösungswahrscheinlichkeiten 0,5/ 0,6/ 0,65/ 0,7/ 0,8 wurde vergleichend aus der Datenmatrix ermittelt, welcher Anteil der Aufgaben jeweils entsprechend stufenkonform gelöst bzw. nicht gelöst wurde. Bei geforderten Lösungswahrscheinlichkeiten von wahlweise 0,6 oder 0,65 werden ca. 90% der entsprechenden Items stufenkonform beantwortet, sodass eine verlässliche Rückmeldung pro Schüler ermöglicht wird.

In allen Schulformen identifizierte die gesonderte Auswertung für Klasse 5 in den drei Rechenverfahren substantielle Schülergruppen auf den Kompetenzstufen 0-2, was bedeutet, dass diesen Schülerinnen und Schülern bei Aufgaben mit besonderen Anforderungen im Übertrag (bzw. bei der Behaltensziffer oder im Teilquotient) systematische Fehler unterlaufen. Erst ab Kompetenzstufe 3 ist davon auszugehen, dass die Rechenverfahren hinreichend sicher beherrscht werden und keine besondere Förderung/ Wiederholung in Klasse 5 notwendig ist.

Literatur

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 Univ. Pr.

Leuders, T. (2014). Modellierungen mathematischer Kompetenzen – Kriterien für eine Validitätsprüfung aus fachdidaktischer Sicht. *Journal für Mathematikdidaktik.*

Padberg, F. & Benz, C. (2011). *Didaktik der Arithmetik*. Heidelberg: Springer.

Van Merrienboer, J., & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Ed. Psy. Review, 17,* 147–177.

Gerster, D.(1982/ 2012). *Schülerfehler bei schriftlichen Rechenverfahren. Diagnose und Therapie*. Münster: WTM.