Nicolai VON SCHROEDERS, Erlangen-Nürnberg

Datenanalyse und -kodierung zur Kategorisierung eines Merkmals Rechenschwäche

1. Ausgangslage

BIRTE 2 ist eine computergestützte Diagnostik, die am Institut für Didaktik der Mathematik von Schipper und Wartha entwickelt und im Jahr 2011 veröffentlich wurde. Basierend auf einer Normierungsstichprobe von über 2000 Schülerinnen und Schülern kann der Test eingesetzt werden, um die arithmetischen Kompetenzen von SchülerInnen in der Mitte des 2. Schuljahres zu messen. Anhand von Eingaben zu insgesamt 145 Aufgaben und der Erfassung der Bearbeitungszeiten wird die arithmetische Gesamtleistung eingeschätzt und Hypothesen zu besonderen Auffälligkeiten hinsichtlich verschiedener Fehlertypen formuliert. Unter anderem werden dazu Zählfehler, Zahlendreher, Inverse-Operationsfehler und Ziffernstrategiefehler, die als Indikatoren für Hauptsymptome einer sich entwickelnden Rechenschwäche zu verstehen (SCHIPPER et. al., 2011) sind, mitgezählt und bei Überschreitung einer normativ festgelegten Grenze entsprechende Hinweise formuliert.

Obwohl die Aufgaben nicht an einem Stück bearbeitet werden müssen, bedeutet die große Anzahl an Aufgaben und die damit verbundenen lange Bearbeitungszeit gerade für rechenschwache Kinder eine hohe kognitive Beanspruchung. Die eigentliche Intention von BIRTE 2 ist jedoch nicht, primär die arithmetische Kompetenz zu messen, sondern den Lehrkräften ein Instrument an die Hand zu geben, mit denen potentiell rechenschwache Kinder identifiziert werden können. Intensive Tests mit Kindern, denen anhand vorab durchgeführter Gespräche eine Rechenschwäche attestiert wurde, zeigen, dass diese zwar in BIRTE 2 auffällig werden, die Hypothesen und Ergebnisse aber teilweise nicht so ein deutliches Bild aufzeigen.

2. Grundidee von BIRTE 2 (AS)

Aus diesem Grund ist der Wunsch entstanden, einen BIRTE 2 (AS) Test – das ist vorerst ein Projektname – zu entwickeln, der vor allem den Fokus weg von der arithmetischen Kompetenz hin auf die Messung einer Rechenschwächesymptomatik legt. Dabei soll das Augenmerk viel deutlicher auf Aufgaben gelegt werden, die spezifisch für besondere Fehlertypen sind, und die Aussagen bezüglich einer Rechenschwäche nicht mehr absolut anhand von Grenzwerten festgemacht sondern probabilistisch formuliert werden. Dazu ist es notwendig, die Aufgaben nicht nur für sich isoliert zu be-

In F. Caluori, H. Linneweber-Lammerskitten & C. Streit (Hrsg.), *Beiträge zum Mathematikunterricht 2015*. Münster: WTM-Verlag

trachten, sondern Antwortmuster und Fehlermuster zu identifizieren, die für bestimmte Gruppen von Kindern symptomatisch sind.

3. Kodierung der Daten

In einem ersten Schritt sind dazu sämtliche Lösungseingaben der 2078 SchülerInnen neu kodiert worden. Die Messung einer Rechenschwäche erfordert nun, dass nicht mehr das einzelne Ergebnis erfasst wird, sondern eine dichotome Kodierung der besonderen Fehlertypen vorgenommen wird. Jede Aufgabe wird daraufhin untersucht, ob der Eingabe ein Zählfehler, ein Zahlendreher, ein Inverse-Operationsfehler, ein Ziffernstrategiefehler oder auch eine Kombination der zuvor genannten Fehlertypen zumindest hypothetisch zu Grunde liegen kann.

Zählfehler	Zahlendreher	Ziffernstrategiefehler	Inverse-Operationsfehler
1	1	0	0

Abb1: Kodierung der Fehlertypen je Aufgaben

Für diese Art der Kodierung sind verschiedenste typische Fehler und Fehlstrategien in der Statistik Software R implementiert worden, um programmgesteuert die Eingabedaten auszuwerten. Die folgende Aufzählung zeigt exemplarisch, wie sich Zählfehler oder auch Ziffernstrategiefehler bei Rechenaufgaben manifestieren können.

Zählfehler: +1/-1 Fehler, +10/-10 Fehler, +10/+1, -1/-10 Fehler

Ziffernstrategiefehler bei Aufgaben der Form $z_1e_1 + z_2e_2$:
$z_1+z_2;e_1+e_2$ $e_1+e_2;z_1+z_2$ $z_1+z_2 + e_1+e_2$ ....

Insbesondere die Kombination derartiger Fehler und Fehlstrategien, ermöglicht es, hypothetisch aber zumindest theoriegeleitet verschiedenste Eingaben zu erklären.

75 Zahlendreher im 1. und 2. Summanden (B) $z_1e_1 + e_1$ Fehler (I)(6) +1 Fehler (D)(2) Zahlendreher im Ergebnis (B)
75 Zahlendreher im 1. und 2. Summanden (B) $z_2e_2 + e_2$ Fehler (I)(7) +1 Fehler (D)(2) Zahlendreher im Ergebnis (B)
76 ziffernweise $z_1 + z_2$ und e_1 oder e_2 (I)(11)
79 Kippfehler (G)(2)
80 -1 Fehler (D)(1)
81 richtige Addition 35+46 (R)
81 $z_1+z_2 + e_1+e_2$ Fehler (I)(1) Zahlendreher im Ergebnis (B)
81 ziffernweise ohne Übertrag Fehler (I)(8) +1 Fehler (D)(2) Zahlendreher im Ergebnis (B)
81 Zahlendreher im 1. und 2. Summanden (B) $z_1+z_2 + e_1+e_2$ Fehler (I)(1) Zahlendreher im Ergebnis (B)
82 +1 Fehler (D)(2)

Abb.2: Auszug möglicher Kombinationsfehler für die Aufgaben 35+46

Interessant dabei ist, dass es Aufgaben gibt, bei denen selbst richtige Eingaben sich auch aus der Kombination von Fehlern ergeben können und dass auch verschiedenste Kombinationen zu gleicher Ergebnissen führen.

Für eine Selektion an Aufgaben können somit neben dem Schwierigkeitsgrad (Lösungsanteil) einer Aufgabe auch der Anteil an spezifischen Eingaben oder auch den Anteil spezifischer Eingaben an allen Fehleingaben berücksichtigt werden. Am Beispiel einiger Additionsaufgaben wird deutlich, wie diese Auswahlkriterien sich bei verschiedenen Aufgabentypen (ZE+ZE mit Zehnerübergang, Schnapszahl + Einer mit Zehnerübergang) darlegen.

Zähl-fehler	33+9	57+6	38+40	35+46	53+17	48+37
Fehler-anteil	3,67%	2,78%	2,51%	16,0%	9,35%	8,23%
Lösungs-anteil	76,4%	75,2%	72,8%	57,1%	64,5%	52,3%
Anteil an Fehlern	15,5%	11,2%	9,22%	37,3%	19,6%	17,3%

Abb.3: Fehleranteile und Schwierigkeitsgrad bei Additionsaufgaben

Diese Kriterien alleine reichen aber noch nicht aus, um gute - im Sinne der Identifikation rechenschwacher Kinder - Aufgaben auswählen zu können. Im nächsten Schritt werden deshalb Antwortmuster analysiert.

4. Kategorisierung

Die Kernidee ist, Gruppen von Kinder mit einem annähernd gleichem Antwortverhalten zu identifizieren und die Zugehörigkeit zu diesen Gruppen probabilistisch vornehmen zu können. Als Methode für dichotome Kodierungen stellt die Statistik dafür die Latent Class Analyse (LCA) zur Verfügung (ROST, 2004). Eine Latent Class Analyse für vier Additionsaufgaben hinsichtlich der Symptomatik der Zählfehler liefert als Ergebnis drei Gruppen, deren Antwortwahrscheinlichkeiten die „Schwächen" der Gruppenmitglieder aufzeigen (siehe Abb. 4).

Während die erste Gruppe nur einen Anteil von 1,76 % der Grundgesamtheit präsentiert, verdeutlich das Antwortverhalten bzw. deren Antwortwahrscheinlichkeiten, dass für die Aufgaben 38+40 und 48+37 die Mitglieder mit einer hohen Wahrscheinlichkeit (zu 95,57% und 82,87%) eine Eingabe mit einem Zählfehler abliefern werden. Hingegen für die gleiche Gruppe die Aufgabe 53+17 mit annähernden 100 % keine Zählfehlerproblematik erkennen lässt, was aber nicht gleichbedeutend damit ist, dass diese Aufgabe richtig gelöst wird.

57 + 6

	(0)	(1)
Gruppe 1:	0.7099	0.2901
Gruppe 2:	0.9410	0.0590
Gruppe 3:	0.8556	0.1444

53 + 17

	(0)	(1)
Gruppe 1:	1.0000	0.0000
Gruppe 2:	0.9856	0.0144
Gruppe 3:	0.1184	0.8816

38 + 40

	(0)	(1)
Gruppe 1:	0.0443	0.9557
Gruppe 2:	0.9703	0.0297
Gruppe 3:	0.8263	0.1737

48 + 37

	(0)	(1)
Gruppe 1:	0.1713	0.8287
Gruppe 2:	0.7872	0.2128
Gruppe 3:	0.6747	0.3253

Gruppe 1: 1,76% Gruppe 2: 88,26% Gruppe 3: 9,98%

Abb. 4: LCA für vier Additionsaufgaben, kein Zählfehler (0), Zählfehler (1)

Die Antwortwahrscheinlichkeiten ermöglichen neben der Hervorhebung einzelner Aufgabentypen für jedes beliebige Antwortmuster eine probabilistische Zuordnung zu einer Gruppe. Wird bei einem Kind z.B. bei allen vier Aufgaben ein Zählfehler identifiziert, so liefert die bedingte Gruppenzugehörigkeitswahrscheinlichkeit für die Gruppe 1 einen Wert von 0,0%, für Gruppe 2 0,07% und Gruppe 3 99,92%. Damit ist die Gruppe 3, die insgesamt einen Anteil von 9,98% einnimmt, die Gruppe die hinsichtlich dieser Aufgabenauswahl eine stark ausgeprägte Tendenz für eine Zählfehlerproblematik aufweist.

5. Ausblick

Es gilt nun, mit Hilfe der Latent Class Analyse aussagekräftige und trennscharfe Aufgaben für jeden Fehlertyp zu selektieren, und eine Kodierung der Bearbeitungszeiten vorzunehmen, um auch diesem Indikator rechenschwacher Kinder gerecht zu werden. Die Identifikation bestimmter Aufgabentypkonstellationen kann dann zu neuen Hypothesen für Zusammenhänge sowohl zwischen einzelnen Aufgaben als auch den Fehlertypen führen.

Literatur

Bortz, J., Döring, N. (2006). Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler. Berlin: Springer.

Fischer, G. (1974). Einführung in die Theorie psychologischer Tests. Bern: Huber.

Rost, J.(2004). Lehrbuch Testtheorie und Testkonstruktion. Bern: Huber.

Schipper, W., Wartha, S., von Schroeders, N. (2011). BIRTE 2, Rechentest für das zweite Schuljahr, Handbuch zur Diagnostik. Braunschweig: Bildungshaus Schulbuchverlage.