

Christine STREIT, Nordwestschweiz

Einsatzmöglichkeiten eines beobachtungsgestützten Paper-and-pencil-Tests zum Erfassen von Schwierigkeiten beim Rechnen

I. Diagnose als schulische Aufgabe

Diagnose und Förderung werden zunehmend als zentrale schulische Aufgaben gesehen. Laut KMK-Beschluss vom 4.12.2003 müssen die Länder Grundsätze zur Diagnose und Förderung von Schülerinnen und Schülern mit besonderen Schwierigkeiten in der Mathematik vorlegen. Die Frage, wie solch ein diagnostischer Auftrag konkret erfüllt werden kann, bleibt allerdings noch weitgehend unbeantwortet.

Standardisierte Testinstrumente wie z.B. ZAREKI, DeMat, OTZ usw. vergleichen die Leistungen einer Schülerin bzw. eines Schülers mit einer normierten Stichprobe. Sie lassen aber keine Rückschlüsse auf individuelle Fehlvorstellungen und die Denkprozesse der Kinder während der Bearbeitung der Aufgaben zu. Dazu sind qualitative Lernstanderhebungen notwendig, die eine prozessorientierte Diagnose ermöglichen und differenzierte Hinweise über den Lernstand eines Kindes in bestimmten Bereichen geben können (vgl. z.B. Schipper 2007). Beobachtungen von materialgebundenen Handlungen sowie der Gestik und Mimik der Kinder in Kombination mit der Methode des lauten Denkens lassen eine differenzierte Analyse der eingesetzten Strategien beim Lösen von Aufgaben zu. Solche qualitativen Instrumente weisen aber auch Nachteile auf: Sie sind sehr zeitintensiv und im Klassenverband nicht durchführbar. Zudem sind sie in der Regel nicht empirisch überprüft, die Analyse und Auswertung bleibt somit hauptsächlich der durchführenden Lehrperson oder sonstigen Fachkraft überlassen, wodurch immer eine gewisse Gefahr der Willkür gegeben ist. Vor allem aber stellt diese Art der Diagnostik hohe Anforderungen an die Lehrperson: Fundierte fachdidaktische Kompetenzen in den zu untersuchenden Inhaltsbereichen sind ebenso Voraussetzung wie das Wissen um qualitative Verfahren und Erfahrung im Führen diagnostischer Gespräche. In Deutschland ist Diagnose aber erst seit kurzem in der Lehrerbildung verankert. Daher ist nicht zu erwarten, dass Lehrkräfte per se über die notwendigen Kompetenzen verfügen.

Um dieser Diskrepanz von Anspruch und Wirklichkeit entgegenwirken zu können, wird aktuell ein beobachtungsgestützter Arithmetiktest (BAT) entwickelt, der als Eingangsdiagnostikum zu Beginn der 3. Klasse eingesetzt werden kann. Parallel dazu wird an einem entsprechenden Diagnosti-

kum für das 5. Schuljahr gearbeitet. Der BAT wird als Gruppentest eingesetzt. Damit ist er zwar zeitaufwändiger als der Bielefelder Rechentest (BIRTE 2), ein computergestütztes Diagnoseverfahren, das ebenfalls versucht diese Lücke zu schließen. Dafür beinhaltet der BAT aber die Möglichkeit der gezielten Beobachtung von typischen Verhaltensweisen, die zählende Rechner häufig zeigen, und die der Computer nicht erfassen kann.

2. Zur Konzeption des Diagnoseinstruments

In Anlehnung an das Berner Screening (Moser Opitz 2007) handelt es sich um ein Diagnoseinstrument, das Schüler und Schülerinnen mit „deutlich unterdurchschnittlichen“ Leistungen im Arithmetikunterricht erfasst. Es soll die Grundlage für eine vertiefte qualitative Diagnose und eine darauf aufbauende gezielte Förderung sein. Damit ist eine rechtssteile Häufigkeitsverteilung angestrebt. So ist eine Differenzierung bei den leistungsschwächeren Schüler/innen möglich.

Die Schüler und Schülerinnen bearbeiten Aufgaben, die nach derzeitigem Stand der Forschung Hinweise auf Schwierigkeiten beim Rechnen liefern können. Im Mittelpunkt stehen folgende Bereiche:

- (1) Zahlverständnis: Zählfertigkeit, Zahlbeziehungen und Stellenwertverständnis
- (2) Rechenfertigkeit und Anwenden von Strategien
- (3) Operationsverständnis

(vgl. z.B. Moser Opitz 2007, Geary 2004, Gerster/Schultz 1998)

Es gilt vor allem zu erfassen, ob die Schülerin bzw. der Schüler über flexible Rechenstrategien verfügt oder ob zählend gerechnet wird. Hier soll ein Beobachtungsbogen weiterführende Erkenntnisse liefern. Ein beigefügtes Testmanual gibt Hinweise zur Durchführung und Auswertung des Tests.

3. Methodisches Vorgehen und erste Ergebnisse

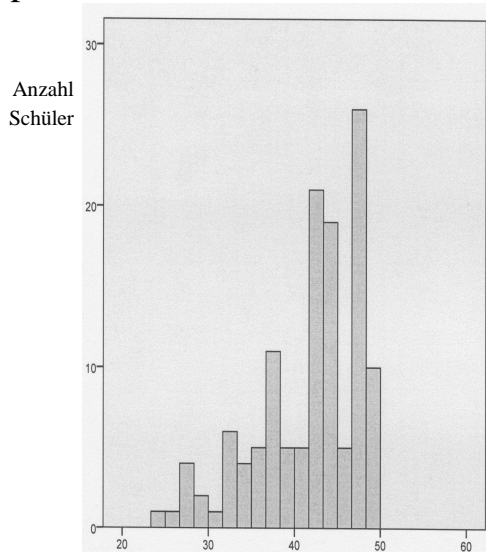
Eine erste Testfassung wurde von 7 Lehrerinnen am Ende der zweiten Klasse erprobt. Dabei wurden die Testaufgaben immer 4 bis 6 Kindern gleichzeitig vorgelegt. Die Lehrerinnen führten den Test nach Anweisung durch. Die Testergebnisse wurden ausgewertet, zusätzlich wurden in gemeinsamen Analysegesprächen die Rückmeldungen und Hinweise der Lehrerinnen gesammelt und die Ergebnisse bei der Überarbeitung der Testaufgaben entsprechend berücksichtigt.

Während der Testdurchführung beobachtete eine Gruppe von 4 Studierenden die Kinder und notierte alle vermeintlichen Auffälligkeiten. In nachfolgenden Gesprächen wurden die Beobachtungen analysiert und kategorisiert. Dieser induktive Zugang wurde durch die Generierung von Kategorien aus theoretischem Vorwissen ergänzt. Zusätzlich wurden qualitative Interviews zur Überprüfung der Kategorien eingesetzt.

Um die Beobachtungsvalidität zu überprüfen, wurde der Test in der jetzigen Fassung dann erneut einer Gruppe von 48 Kindern vorgelegt, bei der immer zwei Studierende parallel den Beobachtungsbogen ausfüllten. Die Beobachterübereinstimmung wurde über Kreuztabellen berechnet. Die Kappa-Werte sind sehr gut (siehe nebenstehende Tabelle; Wirtz & Caspar, 2002: Beurteilerübereinstimmung und Beurteilerreliabilität).

Item	Kappa
B3	1
B4-1	1
B4-2	0,895
B5-1	0,944
B5-2	0,814
B6-A	0,773
B6-B	0,782
B7	0,84
B8-1	1
B8-2	0,891
Median	0,893
Mittelwert	0,8939

In der Pilotierungsphase wurde der Test mit 128 Schülern und Schülerinnen durchgeführt. Die Gruppengröße betrug 4 bzw. 5 Schüler. Es ist geplant, nach der aktuellen Itemanalyse das Instrument zu überarbeiten.



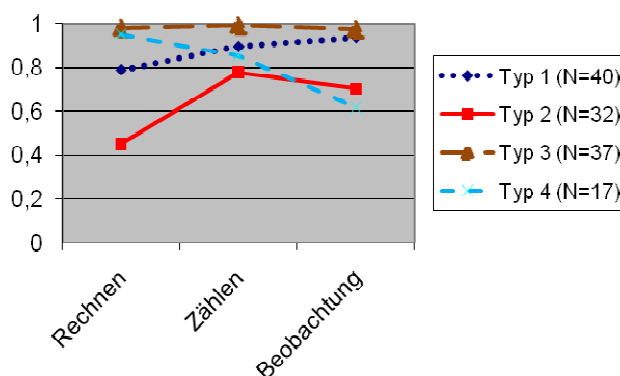
Die Häufigkeitstabelle zeigt keine Normalverteilung. Wie erwartet handelt es sich um eine rechtssteile Verteilung.

<p>n=128 Mittelwert=41 Standardabweichung=6</p>

Um das Risiko zu verringern, dass „Kinder mit Schwierigkeiten“ nicht erfasst werden, und um der rechtssteilen Verteilung zu entsprechen, wurde der Cut-off-Wert auf „Mittelwert minus eine Standardabweichung“ festgelegt. Diese Festlegung konnte durch eine Expertenbefragung (im Sinne der Angoff-Methode) gestützt werden. Die qualitativen Interviews werden mit insgesamt 27 Schüler/innen durchgeführt, deren Ergebnisse entweder in einem der beiden oder in beiden Testteilen im kritischen Bereich lagen. Diesbezüglich liegen noch keine abschließenden Ergebnisse vor.

Daneben wurde durch eine Clusteranalyse eine Gruppe von Schülern ermittelt, deren Testergebnisse ebenfalls durch qualitative Interviews überprüft werden. Diese Gruppe (Typ 4 in der Grafik) zeichnet sich dadurch aus, dass die Kinder im schriftlichen Test akzeptable Ergebnisse zeigten, im Beobachtungsbogen dagegen deutlich auffielen. Hier müssen diagnostische Gespräche zeigen, ob es sich tatsächlich um „zählende Rechner“ handelt, die der schriftliche Test nicht als Risikokinder erfasst hat.

Der Clusteranalyse zugrunde liegen drei reliable Skalen – „Rechnen“ (Cronbachs Alpha .875, 17 Items aus dem Test), „Zahlbeziehungen-Zählen“ (Cronbachs Alpha .853, 6 Items aus dem Test) und „Beobachtungen“ (Cronbachs Alpha .735, 8 Items aus dem Beobachtungsbogen).



4. Ausblick

Obgleich davon ausgegangen wird, dass der BAT Hinweise auf Fehlvorstellungen und vor allem zählendes Rechnen liefern kann, ersetzt er eine vertiefte prozessorientierte Lernstanderhebung nur zum Teil. Der Lehrkraft sollen konkrete Ansatzpunkte für weiterführende Fragen im Sinne einer „Fokussierung des diagnostischen Gesprächs auf zentrale Punkte“ und erste Hinweise für mögliche Förderkonzepte eröffnet werden. Und hierin liegt – neben der Zeitersparnis - auch sein Potential: Es kann auch dazu dienen, dass Lehrer und Lehrerinnen sich gezielter mit der Problematik auseinandersetzen und ihre didaktische und diagnostische Kompetenz in diesem Bereich erweitern.

Literatur

- Geary, D. C. (2004). Mathematics and Learning disabilities. In *Journal of Learning Disabilities*. 37, 4-15
- Gerster, H. D.; Schultz, R. (1998). Schwierigkeiten beim Erwerb mathematischer Konzepte im Anfangsunterricht. Freiburg (PH). Unveröffentlichter Forschungsbericht.
- Moser Opitz (2006). Diagnostik von Mathematikleistungen. In v. Stechow; Hoffmann (Hrsg.) *Sonderpädagogik und PISA*. (S. 279-290) Bad Heilbrunn: Klinkhardt.
- Moser Opitz et.al. (2007). Berner Screening Mathematik 3. Bern Erziehungsdirektion. Download unter: www.erez.be.ch/besmath.
- Schipper, W. (2007). Prozessorientierte Diagnostik von Rechenstörungen. In Lorenz; Schipper (Hrsg) *Hendrik Radatz - Impulse für den Mathematikunterricht* (S. 105 - 116). Braunschweig: Schroedel.