

Verbindung der Ergebnisse von sprachgewandt und KiDiT® Bereich Sprache

Den Schulen und Kindergärten des Kantons Zürich werden in Zukunft zwei Instrumente zur Beurteilung der Sprachkompetenzen von Kindern im Vorschulalter zur Verfügung gestellt. *sprachgewandt* ist ein diagnostisches Testinstrumentarium zur Erfassung der sprachlichen Fähigkeiten von Kindern im Vorschulalter. KiDiT® ist ein Instrument zur Dokumentation von Beobachtungen, unter anderem im Bereich «Sprache».

Die beiden Instrumente wurden im Januar und Februar 2011 in einer repräsentativen Stichprobe von rund 500 Kindern im Alter von 4 bis 7 Jahren eingesetzt. Das bedeutet, dass die Kinder im Bereich Sprache einerseits von einer externen Person mit *sprachgewandt* getestet und andererseits von ihrer Lehrperson mit KiDiT® beobachtet wurden. Die Daten wurden anschliessend vom Institut für Bildungsevaluation skaliert und ausgewertet.

Der gleichzeitige Einsatz der beiden Instrumente wurde dazu genutzt, einen Bezug zwischen den Ergebnissen der beiden Instrumente herzustellen. Dazu wurden in einem ersten Schritt die Rohdaten des Tests *sprachgewandt* nach dem Rasch-Modell für dichotome Items, einem Modell der probabilistischen Testtheorie, skaliert. Das Rasch-Modell basiert auf der Annahme, dass die Wahrscheinlichkeit der Lösung einer Aufgabe ausschliesslich von der Fähigkeit der untersuchten Person und der Schwierigkeit der Aufgabe abhängt. Dementsprechend wird die Wahrscheinlichkeit, eine Aufgabe lösen zu können, als Funktion der Fähigkeit und der Aufgabenschwierigkeit berechnet (Rost, 2004).

Aufgrund der Beziehung zwischen Aufgabenschwierigkeit und Fähigkeit lassen sich die Aufgaben und die Fähigkeiten auf der gleichen Skala abbilden. Die Beziehung zwischen der Aufgabenschwierigkeit und der Fähigkeit ermöglicht es somit, aufgrund des Testergebnisses (gelöste Aufgaben) Aussagen darüber zu machen, über welche Fähigkeiten ein Kind verfügt.

Für eine förderorientierte Interpretation der Testergebnisse wurden Aufgaben mit ähnlicher Schwierigkeit zu Niveaus zusammengefasst. Für jedes Niveau wurden mit Hilfe der Aufgabenmerkmale Fähigkeiten umschrieben, über die ein Kind verfügt, das aufgrund seiner Testergebnisse dem Niveau zugeordnet wird. Die einzelnen Niveaus unterscheiden sich in den Anforderungen, die mit den entsprechenden Aufgaben gestellt werden, und bauen aufeinander auf. Das bedeutet, dass ein Kind in einem bestimmten Niveau immer auch über die Fähigkeiten aller tieferen Niveaus verfügt.

In einem zweiten Schritt wurden die Rohdaten des Tests *sprachgewandt* und die Rohdaten der Beobachtungen mit KiDiT® gemeinsam nach dem Rasch-Modell skaliert. Aufgrund der mehrstufigen Antwortskala von KiDiT kam das sogenannte Partial Credit Model für polytome Items zur Anwendung. Dieses Modell wurde ursprünglich für Aufgaben mit teilrichtigen Lösungen (partial credit) entwickelt. Das Modell gibt – bezogen auf die Anwendung des Beobachtungsbogens – die Wahrscheinlichkeit für ein Kind mit einer bestimmten Fähigkeit an, bei einem Item mit einer bestimmten Kategorie beurteilt zu werden. Während bei einer Aufgabe, die richtig oder falsch gelöst werden kann, ein Schwierigkeitsparameter berechnet wird, lassen sich bei einer mehrstufigen Antwortskala mehrere Schwellenparameter berechnen. Die Schwellenparameter können als Schwierigkeit interpretiert werden, eine bestimmte Schwelle zu überschreiten beziehungsweise mit einer nächsthöheren Kategorie beurteilt zu werden (Bühner, 2006, S. 330; Wright & Masters, 1982, S. 40f.).

Für die gleichzeitige Skalierung der Rohdaten aus dem Test *sprachgewandt* und den Beobachtungen mit KiDiT® wurden die im ersten Schritt berechneten Schwierigkeitsparameter der Testaufgaben importiert und fixiert. Dadurch wurden die Schwellenparameter der Beobachtungen mit KiDiT® auf der Skala der Aufgabenschwierigkeiten des Tests *sprachgewandt* abgebildet und es wird somit ein Bezug zwischen gemessenen und beobachteten Fähigkeiten her-

gestellt. Mit der Verwendung der gleichen Skala wird sichtbar, wie die mit dem Test erfassten Fähigkeiten mit den durch KiDiT® beobachteten Fähigkeiten zusammenhängen.

Die Beobachtungsergebnisse wurden inhaltlich an den Niveaus des Tests ausgerichtet. Mit den Beobachtungsergebnissen kann deshalb nachgewiesen werden, wann die Sprachfähigkeiten altersentsprechend sind und kein Förderbedarf besteht (grüner Bereich), wann die Beobachtungsergebnisse auf eine nicht altersadäquate Entwicklung der Sprachfähigkeit hinweisen und Förderbedarf grundsätzlich angezeigt ist (roter Bereich) und wann die Beobachtungsergebnisse zu keinen klaren Hinweisen über den Förderbedarf führen (gelber Bereich).

Die gemeinsame Anwendung eines «objektiven» Tests und einer «subjektiven» Beobachtung hat unter anderem den Vorteil, dass die beiden Quellen unabhängig zustande kommen. Führen beide Instrumente zum gleichen Ergebnis, dann ist eine zuverlässige Interpretation doppelt abgesichert. Führen die Instrumente zu unterschiedlichen Ergebnissen, dann braucht es eine differenzierte Abklärung und eine entsprechend vorsichtige Interpretation der Ergebnisse.

Urs Moser & Nicole Bayer

Literatur

Rost, J. (2004). *Lehrbuch Testtheorie – Testkonstruktion*. Bern: Verlag Hans Huber.

Bühner, M. (2006). *Einführung in die Test- und Fragebogenkonstruktion*. München: Pearson Studium.

Wright, B.D. & Masters, G.N. (1982). *Rating Scale Analysis. Rasch Measurement*. Chicago: Mesa Press.