

Dominik NACCARELLA, Timo LEUDERS, Markus WIRTZ, Freiburg & Regina BRUDER, Darmstadt

## **Empiriegestützte Itemanalyse für die Kompetenzmodellierung funktionalen Denkens mit Graph, Tabelle und Situation**

### **1. Ziele des Projektes**

Im Rahmen des DFG-Schwerpunktprogrammes „Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen“ werden im Projekt HEUREKO<sup>1</sup> Möglichkeiten der mehrdimensionalen Erfassung von Kompetenzen des Problemlösens mit funktionalen Repräsentationen untersucht. Im Mittelpunkt der zweiten Phase stand die Weiterentwicklung des in der ersten Phase entwickelten Instrumentes zur Erfassung von „funktionalem Denken. Dazu wurde eine Itemanalyse durchgeführt, bei der qualitative und quantitative Verfahren miteinander verknüpft wurden, mit dem Ziel der Optimierung und Validierung des Itempools. Hiermit sind die langfristigen Ziele verknüpft, ein für die Schulpraxis nützliches Instrumentarium bereitzustellen, das sowohl zur Diagnostik, als auch zur Bestimmung von Förderbedarf genutzt werden kann.

### **2. Theoretischer Hintergrund**

Funktionales Denken ist durch eine Vielzahl von Aspekten geprägt (Vollrath, 1989). Insbesondere lassen sich Anforderungen im Rahmen von Aufgaben unter dem Aspekt von Grundvorstellungen und als Modellierungstätigkeiten, d.h. als Übersetzung zwischen Situationen und mathematischen Modellen beschreiben. Daneben lassen sich die postulierten kognitiven Prozesse nach der Art der verwendeten Repräsentationen und des Wechsels zwischen ihnen unterscheiden (Swan, 1985; Blum, 2002; Leuders, 2005). Im Projekt wurden die Repräsentationen und der Wechsel zwischen ihnen zur Grundlage von postulierten Dimensionen eines Kompetenzmodells gewählt. Die Grundvorstellungstheorie und die Theorie zum Modellieren bildeten daneben die leitenden Ideen für eine möglichst breite und variantenreiche Konstruktion der einzelnen Items innerhalb der jeweiligen Kompetenzdimensionen. Dies geschah in der Annahme, dass sich Präferenzen und Kompetenzprofile hinsichtlich der externalen Repräsentation vermutlich

---

<sup>1</sup> Diese Veröffentlichung wurde ermöglicht durch eine Sachbeihilfe der Deutschen Forschungsgemeinschaft (Kennzeichen: LE 2335/1-1) im Schwerpunktprogramm „Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen“ (SPP 1293).

leichter psychometrisch erfassen lassen, als internale und komplexe Verarbeitungsprozesse.

### **3. Ausgangspunkt: Ergebnisse der ersten Projektphase**

In der ersten Projektphase (2007-2009) wurde ein Testinstrument zum „funktionalen Denken“ entwickelt und empirisch überprüft. Das Kompetenzstrukturmodell, welches die vier Bereiche:

- Mathematisieren und interpretieren – zwischen situativer und graphischer Repräsentation (**SG**)
- Mathematisieren und interpretieren – zwischen situativer und numerischer Repräsentation (**SN**)
- Verarbeiten innerhalb der graphischen Darstellung (**G**)
- Verarbeiten innerhalb der numerischen Darstellung (**N**),

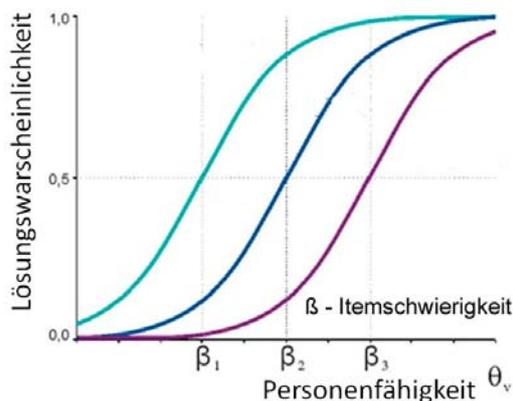
als eigenständige Dimensionen annimmt, wies die beste Datenpassung auf. Leider ist die Qualität der entwickelten Skalen, die EAV/PV-Reliabilität liegt zwischen .52 und .72, für den diagnostischen Einsatz noch nicht ausreichend (siehe Bayerhuber et al, 2010). Trotz der validierten dimensionalen Struktur muss also eine weitere Optimierung und Schärfung der Itemgruppen erfolgen.

### **4. Ziele der empiriegestützten Itemanalyse**

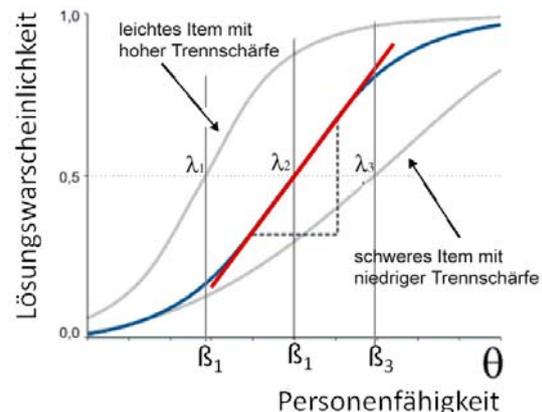
Die grundsätzlichen Ziele der Aufgabenanalyse und Überarbeitung lagen somit in der Erhöhung der Kohärenz der einzelnen Konstrukte. Messtechnisch sollte dies über eine Erhöhung der Reliabilität, also eine Stärkung der internen Konsistenz der einzelnen Skalen erfolgen. Dazu wurden verschiedene Heuristiken zur Aufgabenüberarbeitung und -analyse eingesetzt und miteinander kombiniert. Hierbei wurden in einer empiriegestützten Itemanalyse quantitative und qualitative Verfahren miteinander verknüpft. Der Zweck der einzelnen Analysen ist nicht die Prüfung spezifischer Hypothesen, sondern die Bereitstellung von Informationen als empirische Grundlage für itembasierte Entscheidungen. Es wurden einerseits die Leistungsdaten der ersten Projektphase mithilfe eines mehrdimensionalen zweiparametrischen logistischen Modells reanalysiert. Andererseits wurden ausgewählte Items mittels einer qualitativen Videostudie einer kognitiven Analyse unterzogen. Auf Grundlage dieser beiden Untersuchungen und mit besonderem Augenmerk auf die intendierte Erhöhung der Reliabilität zwischen den Teildimensionen wurden zusätzliche Items entwickelt und pilotiert. Im Folgenden wird ausschließlich über die Ergebnisse der quantitativen Reanalyse der Leistungsdaten berichtet.

## 5. Reanalyse der Daten auf Basis des Birnbaum-Modells

Die Analyse der Daten der ersten Phase erfolgte über ein (mehrdimensionales) Rasch-Modell, in dem sich die einzelnen Items nur in ihrer Schwierigkeit voneinander unterscheiden. Wird die Lösungswahrscheinlichkeit für ein Item in Abhängigkeit von der Personenfähigkeit aufgetragen so sind die einzelnen Itemfunktionen zueinander identisch und nur bzgl. der Abszisse verschoben (vgl. Abb. 1). Es wird somit davon ausgegangen, dass alle Items einer Skala gleich gut zwischen Personen mit hoher und mit niedriger Fähigkeit unterscheiden können - anders ausgedrückt, besitzen alle Items im Rasch-Modell dieselbe Trennschärfe. In zwei-parametrischen IRT-Modellen (Birnbaum-Modell, 2-pl-Modellen) wird diese Einschränkung nicht gemacht. Hier wird für jedes Item ein zusätzlicher Parameter geschätzt, der den (stärksten) Anstieg der Itemfunktion beschreibt (vgl. Abb. 2) (Rost, 2004; Reckase, 2009).



**Abb. 1:** Rasch-Modell parallele Itemfunktionen von drei Items mit versch. Schwierigkeit



**Abb. 2:** Birnbaum-Modell - drei Itemfunktionen mit versch. Schwierigkeit und Trennschärfen

Die über das Birnbaum-Modell aus den Daten ermittelten Trennschärfenkennwerte der Items wurden genutzt, um eine theoretische Itemanalyse durchzuführen. Dabei wurden die Items auf der Basis ihrer Trennschärfen mit Experten analysiert und bzgl. ihrer Skalenpassung diskutiert und insbesondere hinsichtlich der (vermuteten) kognitiven Prozesse die beim Lösen der Items genutzt werden können interpretativ bewertet. So konnten Itemgruppen identifiziert und beschrieben werden, die gemeinsame Merkmale aufwiesen, aber nicht mit dem eigentlich gewünschten Konstrukt harmonieren. Drei besonders problematische Gruppen waren:

- Items mit zu sehr ausgeprägten situativem Kontext innerhalb innermathematischer Dimensionen (N und G)
- „Wissensitems“, die Begriffswissen über Funktionstypen nicht genügend von Übersetzungsfähigkeiten trennen.

- „Kalkül orientierte Items“, bei denen repräsentationsabhängige und funktionale Aspekte durch den Einsatz eines dominanten Kalküls (z. B. Dreisatz) verdrängt wurden. Die Skala schien hier durch modellfremde Kompetenzaspekte dominiert.

Bei der Optimierung des Itempools wurden diese Items hinsichtlich dieser Aspekte überarbeitet bzw. ersetzt.

## **6. Zusammenfassung und Ausblick**

Über die Verbindung der quantitativen Analyse von Itemkonsistenzen mit der qualitativen Analyse der (vermuteten) kognitiven Qualität einzelner Items bzw. Itemgruppen, konnte der Itempool systematisch auf der Basis von empirischen Daten grundlegend überarbeitet werden. Die bisherigen, vorläufigen Analysen der zweiten Erhebung (Herbst 2010) deuten bereits an, dass es gelungen ist das dimensionale Auflösungsvermögen und die Qualität der einzelnen Skalen zu erhöhen.

## **Literatur**

- Bayrhuber, M., Leuders, T., Bruder, R. & Wirtz, M. (2010). Erfassung und Modellierung mathematischer Kompetenz: Aufdeckung kognitiver Strukturen anhand des Wechsels von Darstellungs- und Repräsentationsform. In: Klieme, E. Leutner, D. & Kenk, M.: Kompetenzmodellierung. Zwischenbilanz des DFG-SPP und Perspektiven des Forschungsansatzes. 56. Beiheft der ZfP, Weinheim u. a.: Beltz.
- Blum, W. (2002). On the role of „Grundvorstellungen“ for reality-related proofs – examples and reflections. In: M. A. Mariotti (Ed.), CERME-3 – Proceedings of the Third Conference of the European Society for Research in Mathematics Education. Università di Pisa.
- Leuders, Timo / Prediger, Susanne (2005). Funktioniert's? – Denken in Funktionen. In: PM 2, Jahrgang 47, 1-7.
- Reckase, M.D. (2009). Multidimensional Item Response Theory. Dordrecht u. a.: Springer.
- Rost, J. (1996). Lehrbuch Testtheorie - Testkonstruktion. Bern: Verlag Hans Huber.
- Swan, M. (1985). The Language of Functions and Graphs. Nottingham, UK: Shell Centre for Mathematical Education.
- Vollrath, H.J. (1989). Funktionales Denken. Journal für Mathematikdidaktik, (10), 3-37.