

Matthias LEHNER, Kristina REISS, München

Erfassung des Fachwissens von Studierenden im ersten Semester: Einschätzung des kognitiven Anspruchs eines Tests in Einzelinterviews

Kompetenzmessung und das Fachwissen von Lehrkräften

Empirisch validierte Kompetenzstufenmodelle für das mathematische Wissen von Schülerinnen und Schülern gibt es beispielsweise für die Primarstufe (Reiss & Winkelmann, 2009) und die Sekundarstufe (Blum, Roppelt & Müller, 2013). Die mathematischen Inhalte im weiteren Verlauf der Ausbildung sind allerdings deutlich komplexer, sodass es (noch) schwieriger wird, solche Modelle für spätere Ausbildungsabschnitte zu generieren. Dies gilt auch für das Fachwissen von Lehramtsstudierenden zu Beginn des Studiums.

Unser Ziel ist ganz allgemein, geeignete Items für die Messung des Fachwissens von angehenden Studierenden des gymnasialen Mathematiklehramts zu entwickeln. Im vorliegenden Beitrag betrachten wir zunächst die Ausweisung von Anforderungsbereichen in Anlehnung an die Bildungsstandards (vgl. etwa KMK, 2012). Mit Hilfe von schwierigkeitsgenerierenden Merkmalen werden diese Anforderungsbereiche genauer spezifiziert und validiert. Beide Aspekte werden im Folgenden kurz beschrieben.

Anforderungsbereiche

Um bei der Itementwicklung das Leistungsspektrum möglichst breit abzudecken, sind Aufgaben verschiedener Anforderungsbereiche zu erarbeiten (vgl. Klieme et al., 2007). Der erste Schritt der Itementwicklung ist daher die Entwicklung von Anforderungsbereichen. In diesem Beitrag wählen wir dazu exemplarisch den Bereich des Argumentierens. Die Bildungsstandards für die Lehrerbildung konzentrieren sich auf das Benennen konkreter Inhalte, ohne dabei Anforderungsbereiche zu spezifizieren (KMK, 2008). Aus diesem Grund wird im vorliegenden Beitrag eine Extrapolation der Bildungsstandards für die Allgemeine Hochschulreife (KMK, 2012) vorgeschlagen. In diesen Standards konzentriert sich Anforderungsbereich I auf das Wiedergeben von Sachverhalten im bekannten Kontext. Aufgaben im Anforderungsbereich II beinhalten das selbständige Übertragen auf vergleichbare neue Zusammenhänge, und im Anforderungsbereich III steht das Anwenden von Wissen auf neue Problemstellungen sowie die Reflexion des eigenen Vorgehens im Vordergrund (KMK, 2012). In Bezug auf Argumentieren umfasst Anforderungsbereich III insbesondere die Bewertung von Schlüssen in Hinblick auf deren Gültigkeit und Reichweite

In Institut für Mathematik und Informatik Heidelberg (Hrsg.), *Beiträge zum Mathematikunterricht 2016* (S. x–y). Münster: WTM-Verlag

(KMK, 2012). Übertragen auf Studienanfängerinnen und Studienanfänger werden folgende Anforderungsbereiche formuliert:

- Anforderungsbereich I: Die Studierenden können Routineargumentationen wiedergeben und anwenden sowie einfache logische Schlussfolgerungen auf Schulniveau ziehen.
- Anforderungsbereich II: Die Studierenden können einfache logische Schlussfolgerungen ziehen, die über das Schulniveau hinausgehen.
- Anforderungsbereich III: Die Studierenden können Aussagen über die Grenzen der Gültigkeit von Schlüssen treffen.

Diese Anforderungsbereiche sind auf einer normativen Basis entstanden. Aus diesem Grund ist es erstrebenswert, eine empirische Einordnung der Komplexität von Items vorzunehmen, die auf der Grundlage dieser Anforderungsbereiche entstehen, und damit die Anforderungsbereiche zu präzisieren und auch zu validieren.

Komplexität von Aufgaben

Die Kenntnis schwierigkeitsbestimmender Merkmale ermöglicht es, den Anspruch von Aufgaben systematisch einzuschätzen (Cohors-Fresenborg, Sjuts & Sommer, 2004). So wurden etwa die empirisch gefundenen PISA-Schwierigkeiten mit Hilfe der folgenden vier Kategorien der Komplexität von Denkvorgängen von Cohors-Fresenborg et al. (2004) erklärt: Sprachlogische Komplexität, kognitive Komplexität, Formalisierung von Wissen und Formelhandhabung. Die *sprachlogische Komplexität* betrifft Anforderungen beim Identifizieren und Verstehen relevanter Informationen im Aufgabentext vor dessen Übertragung in eine mathematische Beschreibung. *Kognitive Komplexität* spezifiziert das Ausmaß und die Intensität der notwendigen Denkvorgänge und beinhaltet insbesondere die Gleichzeitigkeit sowie das Verketteten von mehreren Denkschritten. Das Kriterium *Formalisierung von Wissen* beschreibt ob oder in welchem Ausmaß formale Repräsentationen verstanden oder erarbeitet werden müssen. *Formelhandhabung* schließlich ist ein Maß dafür, wie viele Termumformungen für die Lösung der Aufgabe notwendig sind.

Diese vier Kriterien zur Komplexität von Denkvorgängen sind nicht auf ein bestimmtes Themengebiet oder eine bestimmte Klassenstufe limitiert. Vielmehr ist es möglich, sie auf verschiedene Aufgabentypen anzuwenden. Dadurch wird es ermöglicht, die Items, die auf Basis der oben spezifizierten Anforderungsniveaus für angehende Studierende formuliert wurden, aus einer anderen Perspektive zu charakterisieren.

Forschungsfrage und Methode

Unsere Studie zielt auf die Frage, ob die Konzeption von Items auf Grundlage der Anforderungsniveaus und Einschätzung der Komplexität der Aufgaben mit Hilfe der vier Kategorien von Cohors-Fresenborg et al. (2004) zur selben Charakterisierung der Komplexität der entwickelten Items führt. Um diese Frage zu beantworten, wurden Items, die auf Grundlage der Anforderungsbereiche für angehende Studierende erarbeitet worden waren, einer Gruppe von $N = 8$ fortgeschrittenen Studierenden (3 weiblich, 5 männlich) vorgelegt, die im Mittel $M = 25.6$ Jahre alt waren ($SD = 1.9$). Sie lösten die Items und bewerteten anschließend deren Komplexität. Dazu wurde den Studierenden ein Bewertungsleitfaden ausgehändigt, auf dem für jede der vier Kategorien von Cohors-Fresenborg et al. (2004) drei Stufen explizit beschrieben wurden. Es war jeweils Stufe 0 das Niveau mit dem geringsten Anspruch, in Stufe 1 und Stufe 2 wird die Aufgabe als anspruchsvoller gewertet. Die Studierenden wurden aufgefordert, jede der sechs vorgelegten Aufgaben mit jedem der vier Kriterien zu bewerten. Die Übereinstimmung der acht Studierenden ist mit Fleiss $\kappa = .29$ nur gering. Die Studierenden stimmen zwar nicht bezüglich der absoluten Bewertung der Aufgaben überein, sehr wohl aber sind sie sich einig, was die Reihenfolge der Aufgaben von der einfachsten bis hin zur schwierigsten betrifft.

Ergebnisse

Wir gehen der Frage nach, ob die Einschätzung der Komplexität der Items zu dem Ergebnis führt, das als Folge der Konzeption der Items auf Grundlage der Anforderungsbereiche zu erwarten ist und berichten dazu die über alle acht Personen gemittelte Einschätzung der Komplexität der Aufgaben:

<i>Item</i>	<i>Komplex. gesamt</i>	<i>Sprachlog. Komplex.</i>	<i>Kognitive Komplex.</i>	<i>Formalismus</i>	<i>Formelhandhabung</i>
1a	0.6	0.3	0.3	0.3	0.0
1b	1.6	0.5	0.6	0.4	0.1
1c	2.5	0.1	1.0	1.0	0.4
2a	3.9	0.3	1.4	1.5	0.8
2b	4.6	0.3	1.5	1.6	1.3
2c	3.3	0.1	1.6	0.9	0.6

Die Items 1a und 2a sind Anforderungsbereich I zuzuordnen, Items 1b und 2b Anforderungsbereich II und 1c sowie 2c Anforderungsbereich III. Es wird innerhalb der Aufgabenblöcke eine steigende Komplexität erwartet.

Die Tabelle zeigt, dass die gesamte Komplexität innerhalb des ersten Aufgabenblockes erwartungsgemäß ansteigt, im zweiten hingegen ist dies nicht der Fall. Eine genauere Betrachtung der Ergebnisse zeigt, dass die sprachlogische Komplexität in allen sechs verwendeten Items ähnlich hoch ist, diese wirkt sich also kaum auf die gesamte Komplexität aus. Die kognitive Komplexität der Aufgaben steigt innerhalb beider Aufgabenblöcke an. Dies deckt sich mit der Erwartung, die auf Grund der Konzeption der Aufgaben mit Hilfe der drei Anforderungsbereiche vorab getroffen wurde. Im zweiten Itemblock ist in Bezug auf Formalisierung von Wissen sowie Formelhandhabung zu beobachten, dass Item 2c weniger komplex ist als die Items 2a und 2b. Dadurch zeigt sich in der Gesamtskala nicht die erwartete Tendenz.

Insgesamt lässt sich feststellen, dass sich die auf Grund der Anspruchsniveaus erwartete Komplexität der Items nicht vollständig in den Einschätzungen der Schwierigkeit zeigt. Vielmehr ist bei der Itementwicklung zu beachten, dass neben der kognitiven Komplexität von Aufgaben weitere Charakteristika der Items deren Schwierigkeit beeinflussen können.

Literatur

- Blum, W., Roppelt, A. & Müller, M. (2013). Kompetenzstufenmodelle für das Fach Mathematik. In H. A. Pant, P. Stanat, U. Schroeders, A. Roppelt, T. Siegle & C. Pöhlmann (Hrsg.), *IQB-Ländervergleich 2012. Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I* (S. 61–73). Münster: Waxmann.
- Cohors-Fresenborg, E., Sjuts, J. & Sommer, N. (2004). Komplexität von Denkvorgängen und Formalisierung von Wissen. In M. Neubrand (Hrsg.), *Mathematische Kompetenzen von Schülerinnen und Schülern in Deutschland: Vertiefende Analysen im Rahmen von PISA 2000*, S. 109-144. VS Verlag: Wiesbaden.
- Klieme, E., Avenarius, H., Blum, W., Döbrich, P., Gruber, H., Prenzel, M., Reiss, K., Riquarts, K., Rost, J., Tenorth, H. E. & Vollmer, H. J. (2003). *Zur Entwicklung nationaler Bildungsstandards. Eine Expertise*. Berlin: BMBF.
- KMK (2008). *Ländergemeinsame inhaltliche Anforderungen für die Fachwissenschaften und Fachdidaktiken in der Lehrerbildung*. Beschluss der Kultusministerkonferenz vom 16.10.2008.
- KMK (2012). *Bildungsstandards im Fach Mathematik für die Allgemeine Hochschulreife*. Beschluss der Kultusministerkonferenz vom 18.10.2012.
- Reiss, K., & Winkelmann, H. (2008). Step by step. Ein Kompetenzstufenmodell für das Fach Mathematik. *Grundschule*, 40(10), 34–37.