

Lena FRENKEN, Münster & Gilbert GREEFRATH, Münster

Analysen eines Testinstruments zum Metawissen über mathematisches Modellieren

Metakognition wird als grundlegende Kompetenz für selbstständiges Arbeiten sowie für mathematische Prozesse angesehen (Sjuts, 2003). Im Speziellen tragen metakognitive Aspekte in Bezug auf eine holistische Modellierungskompetenz von Schülerinnen und Schülern – neben beispielsweise Teilkompetenzen oder motivationalen Komponenten – zu erfolgreichen Lösungsprozessen bei (Kaiser, 2007). Daher wird in dem Projekt „Modi – Modellieren digital“ unter anderem das Ziel verfolgt, die explizite Vermittlung von Metawissen über Modellieren in einer auf holistischen Modellierungsaufgaben basierenden, digitalen Lernumgebung zu untersuchen.

Metawissen über mathematisches Modellieren

Metawissen, also Wissen über auf den Arbeitsprozess einflussreiche Faktoren wie Strategien oder Eigenschaften von Aufgaben, kann als Teil des umfassenden Konzepts Metakognition angesehen werden (Flavell, 1979). In Bezug auf das Metawissen über Modellieren können drei Facetten deklariert werden: Wissen über Eigenschaften von Modellierungsaufgaben, über hilfreiche Strategien und deren Ziele sowie über eigene Fähigkeiten und Fähigkeiten anderer Personen beim Modellieren. In verschiedenen Studien konnte gezeigt werden, dass einzelne dieser Aspekte, wie z. B. das Wissen über den Modellierungskreislauf oder geeignete Strategien, einen positiven Einfluss auf Modellierungsprozesse haben (z.B. Schukajlow, Kolter & Blum, 2015; Stillman & Galbraith, 1998). Auf der Basis durchgeführter Studien zu metakognitiven Aspekten der holistischen Modellierungskompetenz kann weiterhin herausgestellt werden, dass Testinstrumente, mit denen die Förderung solcher Aspekte evaluiert werden kann, entwickelt werden sollten (Vorhölter, 2017). Dies schließt auch die Entwicklung eines Testinstruments zum Metawissen über Modellieren ein.

Forschungsfragen und Methodik

Um die digitale Lernumgebung im Projekt Modi in Hinblick auf die explizite Vermittlung von Metawissen über Modellieren im Prä-Post-Design zu evaluieren, wurden Items konstruiert, die zu zwei verschiedenen Versionen eines Testinstrument zusammengesgeschlossen wurden. Mit den folgenden beiden Forschungsfragen wird das Ziel verfolgt, verschiedene Gütekriterien zu überprüfen und so ein für den Gruppenvergleich geeignetes Testinstrument zu entwickeln.

F1: Ist Metawissen über Modellieren mit Hilfe des entwickelten Testinstruments als latentes Konstrukt messbar?

F2: Welche Schwierigkeiten haben Schülerinnen und Schüler in der Jahrgangsstufe 9 bei der Bearbeitung des entwickelten Testinstruments?

Mit diesen beiden Forschungsfragen geht ein Mixed-Methods-Ansatz einher, der nun kurz vorgestellt wird. Um die Skalierbarkeit des Instruments und die Messbarkeit des Konstrukts zu analysieren, wurde der Test papierbasiert in fünf Klassen an zwei unterschiedlichen Gymnasien in Nordrhein-Westfalen durchgeführt. Insgesamt nahmen 115 Schülerinnen und Schüler an dieser Testung teil, sodass zunächst eine einparametrische Raschanalyse (Rasch, 1960) mit der Software ConQuest (Wu et al., 2007) durchgeführt werden konnte. Forschungsfrage F2 verfolgt darüber hinaus das Ziel Schwierigkeiten während der Testbearbeitung zu extrahieren. Dazu wurden Interviews mit acht Schülerinnen und Schülern aus zwei verschiedenen Klassen anhand eines Leitfadens geführt, welcher aus theoretischen Grundlagen und einer darauf aufbauenden Analyse des Testinstruments hergeleitet wurde. Die an den Interviews teilnehmenden Schülerinnen und Schüler wurden von den Lehrkräften so ausgewählt, dass die befragte Gruppe insgesamt sowohl bezüglich der Mathematikleistung als auch bezüglich der Erstsprache möglichst heterogen war. Die Interviews wurden transkribiert und im Anschluss mit einer inhaltlich-strukturierenden qualitativen Inhaltsanalyse basierend auf deduktiver Kategorienbildung ausgewertet (Mayring, 2010). Das Kategoriensystem ist in zwei Hauptkategorien unterteilt worden, wobei sich Kategorie A auf Schwierigkeiten bezieht, die auf die Gestaltung des Tests zurückzuführen sind, und in Kategorie B Schwierigkeiten, die aufgrund von fehlendem Metawissen über Modellieren aufgetreten sind, kodiert wurden.

Testdesign

Die Struktur des Testinstruments soll an dieser Stelle – bevor die Ergebnisse präsentiert werden – kurz vorgestellt werden. Basierend auf den theoretischen Grundlagen bezüglich allgemeiner Annahmen über Metawissen besteht der Test aus drei Teilen: Strategien, Personen, Aufgaben. Die Erfassung von Wissen über Strategien wurde weiterhin gegliedert, sodass Eigenschaften und Ziele von beim Modellieren hilfreicher Strategien im Combined-Single-Choice-Format und das vorhandene Strategierepertoire im offenen Kurzantwortformat erhoben werden. Das Wissen in den Bereichen Personen und Aufgaben wird ebenfalls im Combined-Single-Choice-Format erhoben, sodass die Ratewahrscheinlichkeit minimiert und die Ökonomie des Tests gewährleistet werden kann. Insgesamt wurden 39 Items in zwei verschiedenen

Testversionen verwendet. Die beiden Testversionen unterscheiden sich lediglich im Teil zur Erfassung des Strategierepertoires, wobei Ankeritems integriert wurden. Neben den Items zum Metawissen über Modellieren wurden außerdem zu Beginn einige personenbezogene Daten wie Erstsprache, Geschlecht, Alter und Mathematiknote erhoben.

Ergebnisse

Die umfassenden Ergebnisse der Raschanalyse wurden zunächst auf Itemebene ausgewertet, indem 12 Items mit einer Trennschärfe unter 0.2 (Kriterium wie bei PISA) aus der Analyse ausgeschlossen wurden (OECD, 2012). Bei der Analyse der übrigen 27 Items wurden neben den Trennschärfeparametern außerdem die Fit-Werte „Mean Square Fit“ (MNSQ) und „Weighted Mean Square Fit“ herangezogen, welche angeben, inwiefern die gegebenen Antworten zu den im Modell berechneten Lösungswahrscheinlichkeiten passen. Die Werte liegen in Bezug auf den MNSQ zwischen 0.82 und 1.13 und in Bezug auf den WMNSQ zwischen 0.94 und 1.1, sodass ein hohes Qualitätslevel gegeben ist (Bond & Fox, 2007; OECD, 2012). Hinsichtlich der Güte des Gesamttests kann weiter von einer zufriedenstellenden EAP/PV-Reliabilität berichtet werden, welche mit 0.64 ausreichend für Gruppenvergleiche ist (Lienert & Raatz, 1998). Nachdem die Kennwerte der Raschanalyse gezeigt haben, dass der Test skalierbar ist, sollten Bearbeitungsschwierigkeiten betrachtet werden, die sich aus den Ergebnissen der Zusammenfassung aus Hauptkategorie A ergeben. Herauszustellen ist jedoch zunächst, dass ein Großteil der genannten Schwierigkeiten darauf zurückzuführen ist, dass Modellierungsaufgaben im bisherigen Mathematikunterricht der beiden Klassen kaum vorkamen und damit einhergehend Metawissen über Modellieren unzureichend vorhanden war. Darüber hinaus wurden die beiden Wörter „Querschnitt“, welches in einer Aufgabe zum Strategierepertoire vorkommt, und „Lösungsplan“, welches im Combined-Single-Choice-Format zu Zielen spezifischer Strategien vorkommt und sogar erläutert wird, teilweise als schwierig deklariert. Das Item zum Querschnitt musste in der Raschanalyse bereits wegen geringer Trennschärfe ausgeschlossen werden. Außerdem wurde genannt, dass die Qualität der in Schwarz-Weiß gedruckten Bilder irritiert hat. Die dritte Problematik bezieht sich auf die Menge an Text in Bezug auf die Items zum Strategierepertoire, welche teilweise laut Angaben der Interviewten dazu geführt hat, dass nicht alles oder nicht genau gelesen wurde.

Diskussion & Ausblick

Insgesamt ist das Testinstrument also für ein Prä-Post-Design mit Gruppenvergleich nutzbar. Die auf Basis der Raschanalyse auszuschließenden Items

können – bis auf eins – mithilfe der qualitativen Studie nicht überarbeitet werden, da keine Zusammenhänge ersichtlich wurden. Darüber hinaus lassen sich die testbedingten Schwierigkeiten lösen, indem der Test in ein digitales Format übertragen wird, sodass Bilder in Farbe und ausreichend groß dargestellt werden können. Außerdem besteht die Möglichkeit, die große Menge an zu lesendem Text bei der Erfassung des Strategierepertoires zu vermeiden, indem die Dialoge als Audiodateien eingebunden werden. Die vertiefende qualitative Analyse hat weiterhin gezeigt, dass ein Großteil der aufgetretenen Schwierigkeiten auf den Mangel an Metawissen und nicht auf konstruktionsbedingte Probleme zurückzuführen sind. Daraus lässt sich für den weiteren Verlauf des Projekts Modi die Wichtigkeit der Vermittlung von Metawissen über Modellieren betonen.

Literatur

- Bond, T. G. & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences* (2. Aufl.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Flavell, J. H. (1979). Metacognition and Cognitive Monitoring. A New Area of Cognitive-Developmental Inquiry. *American Psychologist*, 34(10), 906–911.
- Kaiser, G. (2007). Modelling and Modelling competencies in school. In C. Haines, P. Galbraith, W. Blum & S. Khan (Hrsg.), *Mathematical Modelling (ICTMA 12): Education, Engineering and Economics* (S. 110–119). Chichester, UK: Horwood Publishing.
- Lienert, G. A. & Raatz, U. (1998). *Testaufbau und Testanalyse* (6. Aufl.). Weinheim: Psychologie Verlags Union.
- Mayring, P. (2010). *Qualitative Inhaltsanalyse. Grundlagen und Techniken* (11. Aufl.). Weinheim: Beltz Verlag.
- OECD. (2012). *PISA 2009 technical report*. Paris: OECD.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Nielsen & Lydiche.
- Schukajlow, S., Kolter, J. & Blum, W. (2015). Scaffolding mathematical modelling with a solution plan. *ZDM*, 47(7), 1241–1254. <https://doi.org/10.1007/s11858-015-0707-2>
- Sjuts, J. (2003). Metakognition per didaktisch-sozialem Vertrag. *Journal für Mathematikdidaktik*, 24(1), 18–40.
- Stillman, G. A. & Galbraith, P. L. (1998). Applying mathematics with real world connections: Metacognitive characteristics of secondary students. *Educational Studies in Mathematics*, 36, 157–189.
- Vorhölter, K. (2017). Measuring Metacognitive Modelling Competencies. In G. A. Stillman, W. Blum & G. Kaiser (Hrsg.), *Mathematical Modelling and Applications. Crossing and Researching Boundaries in Mathematics Education* (S. 175–185). <https://doi.org/10.1007/978-3-319-62968-1>
- Wu, M. L., Adams, R. J., Wilson, M. R. & Haldane, S. A. (2007). *ACER ConQuest version 2.0: Generalised item response modelling software*. ACER Press.