

Digitales Testen am Beispiel des proportionalen Schließens auf dem Prüfstand

Umfassender Evidenz zufolge sagen Fähigkeiten zum proportionalen Schließen Lernerfolg in der Bruchrechnung vorher (Schadl, 2020). Bruchrechnungsfähigkeiten präzisieren wiederum schulischen und beruflichen Erfolg (Siegler et al., 2012). In der Literatur wird der Zusammenhang zwischen den Fähigkeiten zum proportionalen Schließen und Bruchrechnungsfähigkeiten häufig mit Prädiktionsmodellen beschrieben (z.B. Hansen et al., 2015), welche keine differenzierte Beschreibung über ein „mehr ist besser“ hinaus erlauben. In jüngster Zeit beschreibt Schadl (2020) diese Zusammenhänge auf Grundlage von IRT-skalierten Stufenmodellen über die Linearität hinaus. Diese Befunde basieren jedoch auf papierbasierten Testungen, sodass sie im praktischen Schulkontext bislang noch kaum nutzbar sind. Hinderlich ist hier vor allem die hohe Material- und Auswertungsintensität. Diesbezüglich bieten digitale Tests hohes Potential. Gleichzeitig wird ein solcher Unterricht als lernwirksam betrachtet, in dem Lerngelegenheiten adaptiv auf Lern- und Leistungsentwicklungen abgestimmt sind. Daher erweist sich im Speziellen die Diagnostik von Lernverläufen als zentraler, aber auch herausfordernder Bestandteil von Lehrkräfteprofession. Fuchs (2004) postuliert im Kontext von Lernverlaufsdagnostik, zuerst die psychometrische Qualität der Instrumente zu prüfen, bevor Leistungsentwicklungen und in einer weiteren Phase die praktische Nutzbarkeit durch Lehrkräfte untersucht werden. Für die Validierung von Lernverlaufsskalierungen gelten IRT-basierte Raschskalierungen als gängiges Verfahren (Lobato & Walters, 2017). Im Bruchrechnungskontext fehlen bisher digitale Tests für Lernverlaufsmessungen. Der vorliegende Beitrag fokussiert daher am Beispiel eines digitalen Tests zum proportionalen Schließen zunächst einmal die Prüfung der psychometrischen Qualität.

Fähigkeiten zum proportionalen Schließen werden häufig mit missing value tasks erfasst (van Dooren et al., 2009). Diese erfordern bei drei gegebenen Größen die Bestimmung einer vierten Größe, wie der folgende Einkaufskontext zeigt (Schadl, 2020; van Dooren et al., 2009): In einem Schreibwarenladen kosten 6 Packungen mit Stiften 12 €. Eine Lehrerin kauft 24 solche Packungen ein. Wie viel € muss sie bezahlen? Schreibe Deine Rechnung auf. Solche missing value tasks beschreiben entweder proportionale (vgl. Beispiel) oder nicht proportionale Mengenrelationen. Um den missing value task zu lösen, können die Einkaufs- (6) und Kostenmengen (12) verglichen werden (externales Verhältnis, $6 \cdot \underline{2} = 12$). Erfolgt der Vergleich innerhalb einer Menge (Kostenmenge, 6 und 24), dann liegt der Fokus auf dem internalen Verhältnis ($6 \cdot \underline{4} = 24$). In dem Beispiel sind das externale (2) und internale

(4) Verhältnis aus dem natürlichen Zahlbereich gewählt (numerische Struktur: NN). Es ist empirisch dokumentiert, dass missing value tasks für einige Lernende schwieriger werden, wenn diese Verhältnisse aus dem rationalen Zahlbereich gewählt sind (z.B. Schadl, 2020). Auf einer Datenbasis von 784 Lernenden der Klassenstufen 4 bis 6 beschreibt Schadl (2020) ein entsprechendes IRT-skaliertes Stufenmodell. Auf Niveau 1 beschreibt das Modell den Einkaufskontext mit mindestens einem natürlichen Verhältnis. Auf Niveau 2 beschreibt das Modell missing value tasks zu verschiedenen Kontexten mit zwei natürlichen (NN) oder einem natürlichen externalen und einem rationalen internalen Verhältnis (NQ). Auf Niveau 3 werden missing value tasks zu Kontexten mit einem rationalen externalen und natürlichen internalen Verhältnis (QN) beschrieben. Niveau 4 umfasst missing value tasks zu Kontexten mit zwei rationalen Verhältnissen (QQ). Auf Grundlage dieser Stufenbeschreibungen fokussiert dieser Beitrag folgende Forschungsfrage: *Inwiefern lassen sich mithilfe eines adaptierten digitalen Testinstrumentes die Fähigkeiten zum proportionalen Schließen mit ähnlichen Niveaus beschreiben wie in dem früheren Papersetting?*

Methodik

Die Fähigkeiten zum proportionalen Schließen wurden in der laufenden Studie bisher bei $n=93$ gymnasialen Lernenden der Klassenstufe 5 mit der Onlineplattform Levumi erfasst (Gebhardt et al., 2016). Die Lernenden bearbeiteten 14 Aufgaben (proportionale missing value tasks) in ca. 30 Minuten (Instrument in Anlehnung an Schadl, 2020). Analog zu dem Papersetting wurden die numerischen Strukturen NN, NQ, QN und QQ sowie die Kontexte systematisch variiert. Während die Testaufgaben im Papersetting offen gestellt waren („Schreibe Deine Rechnung auf“), wurde für das digitale Setting ein geschlossenes MC-Format gewählt, bei dem drei fiktive Schülerlösungen angegeben waren, ebenso wie die Antwortoption, dass keine der vorgegebenen Schülerlösungen korrekt ist. Dabei wurden die fiktiven Schülerlösungen systematisch variiert. Insgesamt umfasste das Instrument 4 Testaufgaben zum Einkaufskontext sowie 10 zu anderen Kontexten (z.B. Sport, Mischverhältnisse). Zwei dieser Testaufgaben wurden als Distraktoren eingesetzt, bei denen alle fiktiven Schülerlösungen falsch waren. Jede numerische Struktur wurde mit drei Testaufgaben berücksichtigt, wobei sich jeweils eine Testaufgabe auf einen Einkaufskontext und zwei auf andere Kontexte bezogen. Hierbei wurden die Zahltripel der missing value tasks als Vielfache von 10 (Vz) oder keine Vielfachen von 10 (kVz) gewählt. Die vorläufige Datenanalyse erfolgte mithilfe des dichotomen Raschmodells, die Niveaustufen wurden mit der Bookmark-Methode bestimmt (Mitzel et al., 2001).

Ergebnisse

Es zeigten sich auf Basis der bisherigen Analysen akzeptable bis gute Fit-Indizes (WLE-Reliabilität > 0.65 ; $0.80 \leq \text{Item Infits} \leq 1.09$, Details s. Schadl & Lindmeier, accepted). Abbildung 1 zeigt die Wright-Map des vierstufigen Modells zur Beschreibung der Fähigkeiten zum proportionalen Schließen. Dieses Modell beschreibt auf Niveau 1 Einkaufskontexte mit einem internen oder externalen natürlichen Verhältnis. Weniger vertraute Einkaufskontexte und verschiedene Kontexte (z.B. Mischverhältnisse) mit den numerischen Strukturen NN und NQ werden im Modell auf Niveau 2 beschrieben. Niveau 3 beschreibt mit verschiedenen Kontexten zur Struktur QN die Fähigkeiten zu einem *flexiblen* proportionalen Schließen. Niveau 4 beschreibt ein proportionales Schließen mit rationalen Verhältnissen (QQ). Insgesamt zeigten sich damit ähnliche Niveaus wie im Papersetting von Schadl (2020).

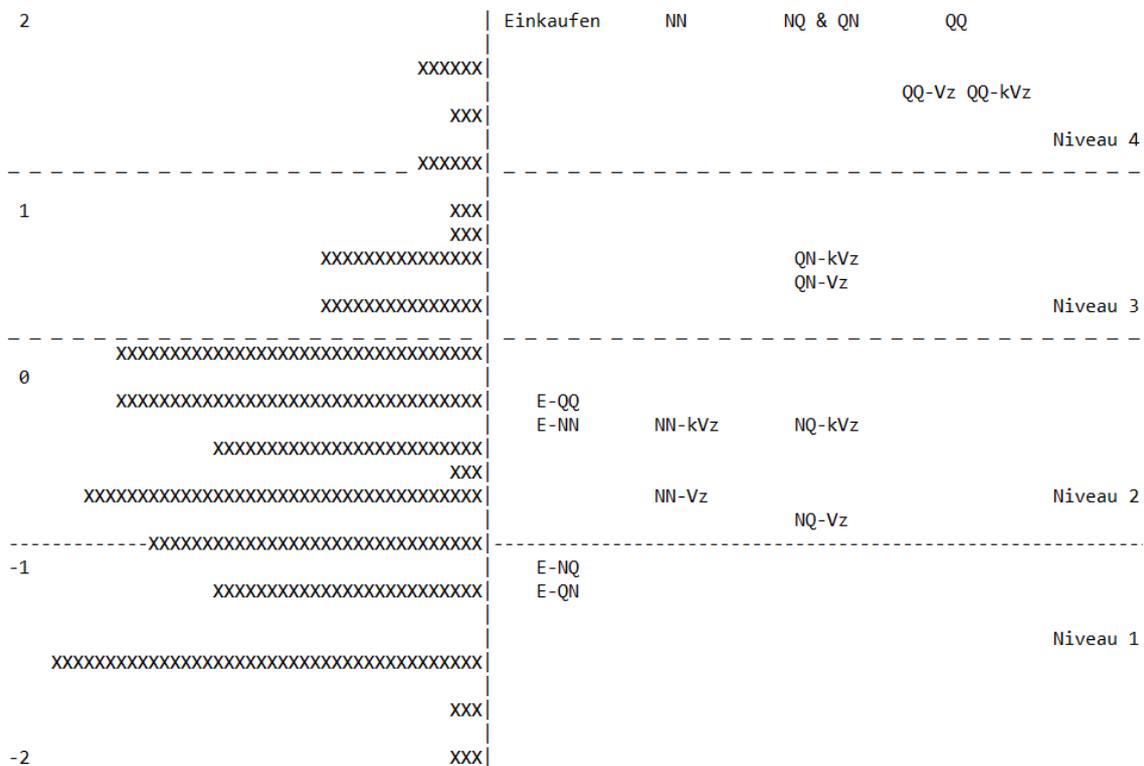


Abb. 1: Wright-Map auf Basis der vorläufigen Analysen zu den Fähigkeiten zum proportionalen Schließen. Präfix „E“ kennzeichnet Einkaufskontexte.

Diskussion und Ausblick

Mit dem entwickelten digitalen Testformat lassen sich in der laufenden Studie gemäß den ersten Analysen ähnliche Niveaus wie in dem papierbasierten Format (Schadl, 2020) beschreiben. Während auf den unteren Niveaus ein proportionales Schließen in Einkaufskontexten (NN, NQ, QN, QQ) und Kontexten mit den Strukturen NN und NQ beschrieben wird, beschreibt das

Modell auf höheren Niveaus ein proportionales Schließen in verschiedenen Kontexten zu den Strukturen QN und QQ. Ob Einkaufskontexte ähnlich wie bei Schadl (2020) ein separates Niveau beschreiben sollten, ist auf Grundlage des vorliegenden Datensatzes noch unklar. Zahlentripel in missing value tasks als (keine) Vielfache von 10 zu wählen, scheint besonders auf niedrigeren Niveaus für den Schwierigkeitsgrad relevant zu sein. So erweisen sich hier missing value tasks mit Vz als Zahlentripel für einen Großteil der Lernenden einfacher als solche mit kVz. Insgesamt bedarf es weiterer Evidenz, um die Ergebnisse auf einer breiteren Datengrundlage abzusichern. Anschließende Forschungsbemühungen könnten die von Fuchs (2004) postulierten Phasen 2 und 3 im Kontext von Lernverlaufsmessungen adressieren. Diese weiteren Forschungsbemühungen erscheinen insofern bedeutend, als dass die Fähigkeiten zum proportionalen Schließen als zentraler Prädiktor für verschiedene Bruchrechenfähigkeiten auch für späteren Lernerfolg im Fach Mathematik relevant erscheinen. Langfristig gesehen könnte das vorgelegte digitale Testinstrument damit einen Beitrag leisten, erfolgreiches individuelles Mathematiklernen im Bruchrechenkontext zu unterstützen.

Literatur

- Fuchs, L. (2004). The past, present, and future of curriculum-based measurement research. *School Psychology Review*, 33(2), 188–192.
- Gebhardt, M., Diehl, K. & Mühling, A. (2016). Online Lernverlaufsmessung für alle SchülerInnen in inklusiven Klassen. *Zeitschrift für Heilpädagogik*, 66, 444–453.
- Hansen, N., Jordan, N., Fernandez, E., Siegler, R., Fuchs, L., Gersten, R. & Micklos, D. (2015). General and math-specific predictors of sixth-graders' knowledge of fractions. *Cognitive Development*, 35, 34–49.
- Lobato, J. & Walters, C. (2017). A taxonomy of approaches to learning trajectories and progressions. In J. Cai (Ed.), *Compendium for research in mathematics education* (pp. 74–101). NCTM.
- Mitzel, H., Lewis, D., Patz, R. & Green, D. (2001). The bookmark procedure: Psychological perspectives. In G. Cizek (Ed.), *Setting performance standards: Concepts, methods, and perspectives* (pp. 249–281). Lawrence Erlbaum Associates.
- Schadl, C. (2020). *Individuelle Lernvoraussetzungen für den Erwerb des Bruchzahlkonzepts. Strukturanalysen und Untersuchung der längsschnittlichen Prädiktivität*. Waxmann.
- Schadl, C. & Lindmeier, A. (accepted). Modelling proportional reasoning skills in levels within a digital setting. In C. Fernández, S. Llinares, A. Gutiérrez & N. Planas (Eds.). *Proceedings of the PME45*.
- Siegler, R., Duncan, G., Davis-Kean, P., Duckworth, K., Claessens, A., Engel, M., Susperreguy, M. & Chen, M. (2012). Early predictors of high school mathematics achievement. *Psychological Science*, 23(7), 691–697.
- Van Dooren, W., De Bock, D., Evers, M. & Verschaffel, L. (2009). Students' overuse of proportionality on missing-value problems: How numbers may change solutions. *Journal for Research in Mathematics Education*, 40(2), 187–211.