

Stefan HOCH, München, Frank REINHOLD, München,
Bernhard WERNER, München, Jürgen RICHTER-GEBERT, München &
Kristina REISS, München

Prädiktion des Lernerfolgs mit Prozessdaten aus einem digitalen Mathematikschulbuch

Einleitung

Forschung zur Unterrichtsqualität stellt eine Reihe an Merkmalen zu Verfügung, um das Unterrichtsangebot positiv zu verändern und dadurch – abhängig von der Nutzung durch Schülerinnen und Schüler – die Wirkung von Unterricht zu verbessern (Helmke, 2009).

Zu diesen Merkmalen zählen unter anderem die Schaffung echter Lernzeit, in der sich Schülerinnen und Schüler aktiv mit dem Lernstoff auseinandersetzen (z. B. Slavin, 1994). Außerdem sollte für Konsolidierung und Sicherung des Lernstoffes ausreichend Raum gegeben sein, hinreichend viel und variationsreich zu üben (z. B. Wellenreuther, 2013). Qualitativ hochwertiger Unterricht zeichnet sich des Weiteren auch dadurch aus, dass in ihm Lernende individuell gefördert werden (Helmke, 2009). Hierbei können Aufgaben unterschiedlicher Anforderungsniveaus oder Lösungshilfen hilfreich sein (z. B. Meyer, 2014). Daneben stellt gutes Feedback ein weiteres Angebot dar, das Schülerinnen und Schüler nutzen können, um Lernziele zu erreichen (Hattie & Timperley, 2007).

Die Erhebung der Nutzung des Unterrichtsangebots durch Schülerinnen und Schüler geschieht traditionell durch Beobachtung oder Videographie (Helmke, 2009). Beide Methoden sind sehr zeitintensiv und können den Unterricht durch die Anwesenheit der Beobachtenden oder der Kameras beeinflussen. Gerade im Mathematikunterricht, in dem das Schulbuch eine wesentliche Rolle spielt (z. B. Valverde et al., 2002), eröffnen sich hier durch digitale Mathematikschulbücher neuartige Möglichkeiten, Daten über die Nutzung des Unterrichtsangebots zu gewinnen: Die digitale Natur eines solchen *E-Books* gestattet es, sämtliche Interaktionen der Nutzerinnen und Nutzer mit dem Mathematikschulbuch zu protokollieren. Aus dieser aufgezeichneten Nutzung von Mathematikschulbüchern durch Schülerinnen und Schüler können im Anschluss Informationen über deren Nutzung des Unterrichtsangebots erschlossen werden. Derartige Nutzungsdaten werden Log- oder Prozessdaten genannt, da sie „geloggt“ werden bzw. den Prozess der Nutzung abbilden (z. B. Goldhammer et al., 2014; Hoch, Reinhold, Werner, Richter-Gebert & Reiss, 2018). Die Erhebung von Prozessdaten geschieht in der Regel vollautomatisch und ohne die Beteiligten zu stören.

Fragestellung

In diesem Beitrag wird dargestellt, wie sich Lernerfolg – als ein Wirkungsziel von Unterricht (Helmke, 2009) – durch Prozessdaten aus einem digitalen Mathematikschulbuch (als Operationalisierung für die Nutzung des Unterrichtsangebots) vorhersagen lässt.

Methode

Insgesamt $N = 151$ Sechstklässlerinnen und Sechstklässler (63 Mädchen, 88 Jungen) nahmen – mit Einverständnis ihrer Erziehungsberechtigten und nach Genehmigung durch das bayerische Kultusministerium – an der vierwöchigen Studie teil. Die Schülerinnen und Schüler entstammten aus sechs Klassen bayerischer Gymnasien, die im Zuge des *ALICE:Bruchrechnen*-Projekts (z. B. Reinhold et al., 2020) mit einem interaktiven Mathematikschulbuch zu Bruchzahlkonzepten (Hoch, Reinhold, Werner, Reiss & Richter-Gebert, 2018) unterrichtet wurden. Dieses E-Book enthält 90 interaktive Komponenten, darunter 75 interaktive Aufgaben. Den Schülerinnen und Schülern stand es dabei prinzipiell frei, wie sie die interaktiven Aufgaben bearbeiteten. In allen interaktiven Aufgaben erhielten sie automatisches, adaptives Feedback auf ihre Aufgabenlösungen, das sofort auf dem Tabletbildschirm angezeigt wurde; in einigen Aufgaben standen ihnen abgestufte Lösungshilfen zur Verfügung, die sie jederzeit im Bearbeitungsprozess aufrufen konnten.

Vor dem Bruchrechnenunterricht wurde das Vorwissen der Schülerinnen und Schüler mit Hilfe eines Pretests (10 Items) erhoben. Der Lernerfolg wurde nach $M = 15.167$ Stunden Unterricht ($SD = 1.329$) durch einen Posttest überprüft (38 Items; s. Reinhold et al., 2020, für Details zu den Tests).

Während der Nutzung des interaktiven Schulbuchs wurden Interaktionen mit den Aufgaben als Prozessdaten aufgezeichnet. Für die Analysen werden Aktivitäten, deren Dauer 0.5 s unterschreitet, nicht beachtet. Nach gängigen Verfahren wurden Ausreißer erkannt und deren Dauer geschätzt (Goldhammer et al., 2014; s. Hoch, Reinhold, Werner, Richter-Gebert & Reiss, 2018). Insgesamt basieren die folgenden Ergebnisse auf 99090 Datenpunkten, aus denen für jede Schülerin und jeden Schüler die folgenden sechs Prozessmaße berechnet wurden: (1) Anzahl an bearbeiteten Aufgaben, (2) Anzahl an genutzten interaktiven Aufgaben in Prozent, (3) Anzahl an selbstreguliert aufgerufenen Lösungshilfen, (4) mittlere Zeit auf Feedback, (5) gesamte Bearbeitungszeit und (6) eine E-Book-interne Lösungsrate.

Für die Auswertung wurden generalisierte lineare Mischmodelle (*GLMMs*) benutzt, um die Lösungswahrscheinlichkeit für Items im Posttest mittels logistischer Regression zu schätzen. *GLMMs* gelten als besonders geeignet für Prädiktionsmodelle auf Prozessdaten (Zheng et al., 2019) und sind in der

Lage, Klassenstrukturen in Datensätzen zu berücksichtigen. Die Berechnungen der GLMMs erfolgten mit dem R-Package *lme4* (Bates et al., 2015).

Ergebnisse & Diskussion

Die Ergebnisse zeigen in Übereinstimmung mit dem Angebots-Nutzungs-Modell, dass die Nutzung des Unterrichtsangebots – hier gemessen durch Prozessdaten aus einem E-Book – durch die Schülerinnen und Schüler prädiktiv für den Lernerfolg im Mathematikunterricht ist. Tabelle 1 zeigt die Schätzungen der vier verschiedenen Prädiktionsmodelle.

	Modell 0		Modell 1		Modell 2		Modell 3	
<i>Fixed Effects</i>	B	SE	B	SE	B	SE	B	SE
<i>Intercept</i>	0.289	0.220	0.338	0.196	0.255	0.195	0.401*	0.197
Vortest	1.720***	0.214	1.500***	0.203	1.516***	0.198	1.164***	0.207
Geschlecht	-0.182	0.123	-0.075	0.119	0.051	0.119	0.099	0.113
Bearbeitete Aufgaben			-0.002	0.031	-0.072	0.041	-0.014	0.041
Bearbeitete Widgets (%)			2.365***	0.499	1.814***	0.522	1.156*	0.519
Aufgerufene Lösungshilfen			-0.015*	0.007	-0.010	0.007	-0.004	0.006
Mittlere Zeit auf Feedback					1.228	0.708	1.362*	0.675
Gesamte Bearbeitungszeit					0.403**	0.143	0.183	0.146
Lösungsrate							2.784***	0.666
<i>Random Effects</i>	Var	SD	Var	SD	Var	SD	Var	SD
Schülerin/Schüler	0.349	0.591	0.294	0.542	0.256	0.506	0.206	0.454
Item	0.905	0.952	0.900	0.949	0.900	0.949	0.901	0.949
Klasse	0.052	0.228	0.003	0.058	0.003	0.056	0.009	0.097
<i>Likelihood Ratio Test</i>	χ^2		χ^2		χ^2		χ^2	
	-		24.126***		11.818**		16.505***	

Anmerkung. 5738 Datenpunkte, 151 Schülerinnen und Schüler; 6 Klassen; 38 Items. Schätzungen in *Log-Odds*. SE = Standardfehler, Var = Varianz, SD = Standardabweichung. Signifikanzniveaus: * $p < .05$, ** $p < .01$, *** $p < .001$.

Tab. 1: Modellschätzungen der vier GLMM-Prädiktionsmodelle.

Die Ergebnisse können dahingehend interpretiert werden, dass es nicht darauf ankommt, möglichst viele Aufgaben im Mathematikunterricht zu rechnen. Vielmehr gibt der signifikant positive Einfluss des Prozentsatzes an bearbeiteten unterschiedlichen interaktiven Aufgaben den Hinweis, dass variationsreiches Üben (vgl. Meyer, 2014) und eine Betrachtung des Lernstoffes aus unterschiedlichen Perspektiven (für den betrachteten Bereich der Bruchrechnung bspw. mit unterschiedlichen Repräsentationen, vgl. Padberg & Wartha, 2017) für den Lernerfolg förderlich sind.

Die Signifikanz des positiven Einflusses der gesamten gemessenen Bearbeitungszeit wird durch die Aufnahme der Lösungsrate im E-Book in das Modell aufgelöst. Der signifikante positive Effekt dieses Leistungsmaßes ist u. a. im Hinblick auf Ergebnisse aus den Learning Analytics (z. B. Zheng et

al., 2019) nicht verwunderlich. Im Gegenzug wird dabei jedoch der positive Effekt der mittleren Zeit, in denen Schülerinnen und Schüler auf Feedback verweilen, signifikant. Dies kann als Indiz für die Wirksamkeit des automatischen Feedbacks – sofern es genutzt wird – gewertet werden und lässt den Schluss zu, dass Feedback beim Arbeiten mit interaktiven Aufgaben als echte Lernzeit verstanden werden sollte.

Insgesamt kann die Nutzung von Prozessdaten als probates Mittel angesehen werden, die Nutzung von Schulbüchern und somit von Teilen des Unterrichtsangebots durch Lernende akkurat im natürlichen Kontext zu erfassen.

Literatur

- Bates, D., Mächler, M., Bolker, B. & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1–48.
- Goldhammer, F., Naumann, J., Stelter, A., Tóth, K., Rölke, H. & Klieme, E. (2014). The time on task effect in reading and problem solving is moderated by task difficulty and skill: Insights from a computer-based large-scale assessment. *Journal of Educational Psychology*, 106(3), 608–626.
- Hattie, J. & Timperley, H. (2007). The Power of Feedback. *Review of educational research*, 77(1), 81–112.
- Helmke, A. (2009). *Unterrichtsqualität und Lehrerprofessionalität: Diagnose, Evaluation und Verbesserung des Unterrichts*. (2. Auflage). Klett/Kallmeyer.
- Hoch, S., Reinhold, F., Werner, B., Reiss, K. & Richter-Gebert, J. (2018). *Bruchrechnen. Bruchzahlen & Bruchteile greifen & begreifen* [Apple iBooks Version] (4. Aufl.). Technische Universität München.
- Hoch, S., Reinhold, F., Werner, B., Richter-Gebert, J. & Reiss, K. (2018). Design and research potential of interactive textbooks: the case of fractions. *ZDM Mathematics Education*, 50, 839–848.
- Meyer, H. (2014). *Was ist guter Unterricht?* (10. Aufl.). Cornelsen Verlag Scriptor.
- Padberg, F. & Wartha, S. (2017). *Didaktik der Bruchrechnung* (5. Aufl.). Springer Spektrum.
- Reinhold, F., Hoch, S., Werner, B., Richter-Gebert, J. & Reiss, K. (2020). Learning fractions with and without educational technology: What matters for high-achieving and low-achieving students? *Learning and Instruction*, 65, 101264.
- Slavin, R. E. (1994). Quality, appropriateness, incentive, and time: A model of instructional effectiveness. *International Journal of Educational Research*, 21(2), 141–157.
- Valverde, G. A., Bianchi, L. J., Wolfe, R. G., Schmidt, W. H. & Houang, R. T. (2002). *According to the Book: Using TIMSS to investigate the translation of policy into practice through the world of textbooks*. Springer Netherlands.
- Wellenreuther, M. (2013). *Lehren und Lernen – aber wie?* (6., vollst. überarb. Aufl.). Schneider Hohengehren.
- Zheng, G., Fancsali, S. E., Ritter, S. & Berman, S. R. (2019). Using Instruction-Embedded Formative Assessment to Predict State Summative Test Scores and Achievement Levels in Mathematics. *Journal of Learning Analytics*, 6(2), 153–174.