

SCHAAF, Julius; ROLFES, Tobias; NAGY, Gabriel & HEINZE, Aiso
Frankfurt a. M., Frankfurt a. M., Kiel, Kiel

Zusammenhang zwischen der Nutzung intelligenter tutorieller Systeme (ITS) und dem Lernzuwachs in Mathematik in der Mittelstufe

Einleitung

In den Schulen werden digitale Medien in Form digitaler Mathematikwerkzeuge (z.B. grafikfähige Taschenrechner, dynamische Mathematiksysteme) bereits seit längerem eingesetzt und sind auch fester Bestandteil der Bildungsstandards (KMK, 2022). Insbesondere durch die Corona-Pandemie hat das Mathematiklernen mit digitalen Medien weiter an Bedeutung gewonnen.

Nach Hattie (2023) ist die durchschnittliche Effektgröße des Einsatzes von Technologie im Unterricht klein bis mittel ($d = 0,34$). Dabei wurde jedoch eine relativ große Streuung der Effekte in den Studien festgestellt. Diese große Variabilität deutet darauf hin, dass der Einsatz von Technologie nicht per se lernförderlich ist. Vielmehr scheint entscheidend zu sein, wie digitale Medien in den Unterricht integriert werden. So zeigten die Hattie-Studien zum Beispiel, dass die Nutzung digitaler Medien in Zweiertteams zu einem höheren Lernzuwachs führte als in Einzel- oder Gruppenarbeit.

Ein möglicherweise bedeutsamer Faktor für die Lernwirksamkeit ist der Grad der Adaptivität der digitalen Medien. Gerade in der Corona-Pandemie rückten digitale Systeme in den Fokus, die individuell auf Schüler*innen reagieren und dadurch den selbstgesteuerten Lernprozess effizient unterstützen können. Eine Kategorie digitaler Medien, die dieses versucht, sind sogenannte *intelligente tutorielle Systeme* (ITS), welche eine Weiterentwicklung der *Computer Aided Instruction* (CAI) darstellen. Diese adaptiven ITS besitzen oft eine breite Basis an bereitgestellten Inhalten mit direktem Feedback. Die Studienlage zur Lernwirksamkeit von ITS in Mathematik ist uneindeutig (Steenberg-Hu & Cooper, 2013, Kulik & Fletcher, 2016). Daher soll im Rahmen dieser Studie der Einfluss der Nutzung eines ITS (konkret: Bettermarks) auf den Lernzuwachs in Mathematik von Schüler*innen untersucht werden.

Forschungsfragen

1. Existiert ein Zusammenhang zwischen der durchschnittlichen Nutzungshäufigkeit des ITS in einzelnen Klassen und dem Lernzuwachs von Schüler*innen dieser Klassen in Mathematik?
2. Existiert ein Zusammenhang zwischen der individuellen Nutzungshäufigkeit des ITS relativ zum Klassenmittelwert und dem Lernzuwachs von Schüler*innen dieser Klassen in Mathematik?

In: P. Ebers, F. Rösken, B. Barzel, A. Büchter, F. Schacht & P. Scherer (Hrsg.),
Beiträge zum Mathematikunterricht 2024.

57. Jahrestagung der Gesellschaft für Didaktik der Mathematik. WTM.
<https://doi.org/10.37626/GA9783959872782.0>

1261

Stichprobe

Im Längsschnitt über ca. acht Monate liegen die Daten von 942 Testpersonen (489 weiblich, 445 männlich und 8 ohne Angabe) aus 57 Klassen der 7. und 8. Klasse aus Schleswig-Holstein vor. Von den Testpersonen besuchten 771 ein Gymnasium und 165 eine Gemeinschaftsschule. Das durchschnittliche Alter der Stichprobe betrug 12,8 Jahre ($SD = 0,7$ Jahre).

Design

An zwei Messzeitpunkten (Prättest: September bis November 2021, Posttest: Juni bis Juli 2022) wurde ein computerbasierter Fragebogen administriert. Da der genaue Testzeitpunkt von den Lehrkräften eigenständig innerhalb vorab definierter Zeitfenster festgelegt wurde, variierten die Zeitabstände von Prä- und Posttest zwischen den Klassen und betragen im Mittel 253 Tage (Median). Zu beiden Messzeitpunkten wurden die Mathematikleistungen und weitere Merkmale gemessen (siehe Messinstrumente). Darüber hinaus wurde während des Schuljahres die Aktivität der einzelnen Schüler*innen im ITS gespeichert.

Die Schüler*innen bearbeiteten in dem ITS sogenannte *Worksheets* (Arbeitsblätter). Ein Worksheet besteht aus einer Anzahl ähnlicher Aufgaben zu einem gemeinsamen Thema. Bei jeder Bearbeitung eines Worksheets wurde erfasst, aus wie vielen Aufgaben es besteht, wie viele davon korrekt gelöst wurden und wann die Bearbeitung erfolgte.

Die Nutzung des ITS war für die beteiligten Schulen kostenlos und wurde vom Bildungsministerium finanziert. Es gab keine Vorgaben, in welchem Ausmaß, zu welchen Themengebieten, zu welchen Zeitpunkten (während oder außerhalb der Unterrichtszeit) oder mit welcher Methode das ITS genutzt werden sollte. Somit entschieden die Lehrkräfte und die Schüler*innen nach ihrem Ermessen über die Art und Weise der ITS-Nutzung.

Messinstrumente

Zur Messung der Mathematikleistung wurden 60 Items verwendet. Bei den administrierten Items handelte es sich größtenteils um Items aus der internationalen *Trends in Mathematics and Science Study* (TIMSS) für die Jahrgangsstufe 8. Die Items wurden in einem Multi-Matrix-Design administriert, sodass jeder Testperson jeweils 20 Items im Prä- und Posttest vorgelegt wurden.

Um eine valide Messung der Mathematikleistung zu ermöglichen, deckten die Items die Unterrichtsinhalte der Jahrgangsstufen 7 und 8 der vier Leitideen *Zahl und Operation*, *Raum und Form*, *Größen und Messen* und *Struk-*

turen und funktionaler Zusammenhang entsprechend der aktuellen Bildungsstandards der KMK ab (KMK, 2022). Aufgrund der limitierten Testzeit wurde die Leitidee *Daten und Zufall* nicht in die Testung aufgenommen. Neben den Mathematikleistungen wurden noch weitere Schüler*innenmerkmale (z.B. End- und Halbjahresnoten in sprachlichen und naturwissenschaftlichen Fächern, kulturelles Kapital des Elternhauses, Freude und Interesse an Mathematik, gewissenhafte Arbeitshaltung in Mathematik) erfasst.

Analyse

Die Testergebnisse wurden mit Hilfe eines Rasch-Modells skaliert. Die Itemanalyse zeigte, dass alle eingesetzten Items eine akzeptable Passung aufwiesen (INFIT zwischen 0,87 und 1,15, Trennschärfe über 0,24). Auch wiesen nur drei der 60 Items substanzielles differentielles Itemfunktionieren (DIF) zwischen Prä- und Posttest auf. Insgesamt kann davon ausgegangen werden, dass die Mathematikleistungen in ausreichendem Maße valide und reliabel gemessen wurden. Auf der Grundlage des Rasch-Modells sowie eines Hintergrundmodells wurden für die Mehrebenenanalysen für jede Person Fähigkeitswerte in Form von *Plausible Values* berechnet (EAP/PV-Reliabilität: 0,84).

Ergebnisse

Der durchschnittliche Lernzuwachs der Stichprobe betrug von Prä- zu Posttest 0,35 Standardabweichungen. Um den Zusammenhang zwischen der Nutzungshäufigkeit des ITS und dem Lernzuwachs der Schüler*innen zu ermitteln, wurde eine Mehrebenenanalyse (Ebene 1: Testpersonen, Ebene 2: Klasse) durchgeführt. Dafür wurde die ITS-Nutzungshäufigkeit als die Anzahl von einer Testperson zwischen den beiden Messzeitpunkten bearbeiteten unterschiedlichen Worksheets operationalisiert.

Darauf aufbauend wurde in der Mehrebenenanalyse der Einfluss verschiedener Prädiktoren auf die Mathematikleistung im Posttest untersucht. Dabei fand sich erwartungskonform ein positiver Einfluss des Prätests sowie der Schulform (Testpersonen des Gymnasiums schnitten im Posttest signifikant besser ab als Testpersonen von Gemeinschaftsschulen). Allerdings wurde weder auf Klassen- noch auf Individualebene ein signifikanter Zusammenhang zwischen der Nutzungshäufigkeit des ITS und dem Lernzuwachs der Schüler*innen gefunden. Das heißt, der Lernzuwachs war unabhängig davon, wie häufig in den Klassen im Schnitt das ITS genutzt wurde, aber auch unabhängig davon, ob eine Testperson in einer Klasse seltener oder häufiger das ITS nutzte als der Durchschnitt dieser Klasse.

Diskussion

Die Mehrebenenanalyse zeigte, dass die Häufigkeit der ITS-Nutzung weder

auf Klassen- noch auf Individualebene einen signifikanten Effekt auf die Mathematikleistung hatte. Das ITS ersetzt vermutlich Aufgaben und Materialien (Arbeitsbögen, Schulbücher etc.), die anscheinend ähnliche Effektivität für den Lernzuwachs aufweisen. Aus theoretischer Perspektive existieren sowohl Gründe, die für eine höhere Lerneffektivität sprechen (direktes & adaptives Feedback, Aufzeigen von Wissenslücken) als auch solche, die eine niedrigere vermuten lassen (weniger Klassendiskussion über Aufgaben, weniger kollaboratives Lernen, Beschränkung der Aufgabenformate). Eine mögliche Interpretation der vorliegenden Ergebnisse ist, dass sich diese positiven und negativen Aspekte bei der ITS-Nutzung in der Studie kompensiert haben.

Auf Grund des Studiendesigns konnten nur korrelative Zusammenhänge untersucht werden, kausale Effekte können allenfalls theoretisch fundiert vermutet werden. Leider liegen wenig Informationen darüber vor, in welcher Art und Weise (z.B. Sozialform) die Arbeit mit dem ITS im Unterricht erfolgt ist und ob bzw. in welchem Ausmaß die bearbeiteten Aufgaben von der Lehrkraft besprochen wurden.

Fazit

In unserer Studie konnte kein Zusammenhang zwischen der Nutzung von ITS und der Mathematikleistung gefunden werden. Allerdings sind auch keine nachteiligen Effekte aufgrund einer intensiven ITS-Nutzung feststellbar. In weiteren Analysen soll untersucht werden, ob Inhaltsgebiete oder Aufgabeneigenschaften zu identifizieren sind, die mit positiven Effekten der ITS-Nutzung zusammenhängen.

Literaturverzeichnis

- KMK (Sekretariat der Ständigen Konferenz der Kultusminister der Länder der Bundesrepublik Deutschland). (Hrsg.). (2015). *Bildungsstandards im Fach Mathematik für die Allgemeine Hochschulreife (Beschluss der Kultusministerkonferenz vom 15.10.2004 i.d.F 26.2022)*.
- Steenberg-Hu, S., & Cooper, H. (2013). A meta-analysis of the effectiveness of intelligent tutoring systems on K–12 students' mathematical learning. *Journal of Educational Psychology*, Vol. 105(4), 970–987 VS. <https://doi.org/10.1037/a0032447>
- Hattie, J. (2023). *Visible learning: The sequel: A synthesis of over 2,100 meta-analyses relating to achievement*. Routledge. <https://doi.org/10.4324/9781003380542>
- Kulik, J. A., & Fletcher, J. D. (2016). Effectiveness of intelligent tutoring systems: A meta-analytic review. *Review of Educational Research*, 86(1), 42–78. <https://doi.org/10.3102/0034654315581420>