

RÜHL, Lisa & THURM, Daniel  
Siegen

## **KI-Tutoren zur konstruktiven Unterstützung bei Übungsaufgaben**

Large Language Models (LLMs) sind eine Form generativer künstlicher Intelligenz (KI). Auf theoretischer Ebene wurden für den Einsatz von LLMs im Bildungsbereich vielfältige Potenziale hergeleitet (z. B. Kasneci et al., 2023; Ständige Wissenschaftliche Kommission der Kultusministerkonferenz [SWK], 2024). Ein zentrales Potenzial liegt dabei insbesondere in einer individuellen und konstruktiven Unterstützung von Lernenden. Bisher gibt es jedoch kaum konkrete Lernsettings und empirische Studien, die sich gezielt mit dem Einsatz von LLMs zur konstruktiven Unterstützung im Mathematikunterricht befassen. Ausgehend von diesem Desiderat untersucht das in diesem Beitrag vorgestellte Dissertationsprojekt, inwiefern aufgabenspezifische KI-Tutoren Lernende bei der Bearbeitung von Übungsaufgaben konstruktiv unterstützen können.

### **Einsatz von LLMs als Lerntutoren: Potenziale und Herausforderungen im Mathematikunterricht**

Seit der Veröffentlichung von ChatGPT Ende 2022 gewinnen LLMs zunehmend an gesellschaftlicher Bedeutung und entwickeln sich rasant weiter. Auch für den Bildungsbereich werden vielfältige Potenziale in der Nutzung solcher Modelle erwartet (Kasneci et al., 2023; SWK, 2024). Ein zentraler Anwendungsbereich liegt dabei in der konstruktiven Unterstützung von Lernenden durch den Einsatz von LLMs als Lerntutoren.

Die konstruktive Unterstützung von Lernenden im Mathematikunterricht gehört zu den drei Basisdimensionen guten Unterrichts (Praetorius et al., 2018; Ufer et al., 2023). Im Kern umfasst sie die Bereitstellung von angemessenem inhaltlichem, individuellem Feedback sowie emotional-motivationaler Unterstützung. Wesentliche Bestandteile sind die konstruktive Betrachtung von Fehlern, die Berücksichtigung des individuellen Vorwissens und Lerntempos sowie die Förderung eines positiven Lernklimas. Besonders bei der Bearbeitung von Aufgaben zeigt sich die Bedeutung konstruktiver Unterstützung. Lernende stoßen häufig auf Probleme, die sie allein nicht lösen können, wie z. B. das Missverstehen der Aufgabenstellung, das Auftreten von Rechenschwierigkeiten, das Nicht-Weiterkommen oder das Fehlen relevanter Begrifflichkeiten. Angesichts der zeitlichen und personellen Begrenzungen im Unterricht ist es jedoch für Lehrkräfte kaum möglich, allen Lernenden individuelle Unterstützung zu bieten und den Lernprozess aktiv zu begleiten.

In: L. Schick, M. Platz & A. Lambert (Hrsg.),  
Beiträge zum Mathematikunterricht 2025.

Hier können LLMs eine wichtige Rolle spielen, da sie die Möglichkeit bieten konstruktive Unterstützung für Lernende bereitzustellen. Durch ihre Fähigkeit, personalisierte und kontextbezogene Antworten zu generieren, könnten sie gezielt auf individuelle Herausforderungen eingehen. Ihre Flexibilität und Anpassungsfähigkeit eröffnen das Potenzial, durch inhaltliches Feedback, die Klärung von Missverständnissen und die Erklärung von Konzepten als Lerntutoren zu agieren und den Lernprozess effektiv zu begleiten.

Allerdings fehlen bislang umfassende empirische Studien, die untersuchen, inwiefern LLMs tatsächlich konstruktive Unterstützung im Bildungskontext bieten können. Dilling und Herrmann (2024) untersuchten die Nutzung von ChatGPT zur Unterstützung von Lehramtsstudierenden beim Erstellen mathematischer Beweise in der Geometrie und fokussierten dabei u. a. auf die verwendeten Prompts der Studierenden. Mit Blick auf diese zeigt sich, dass häufig der Beweis ohne weitere Angabe von Informationen eingefordert wird. In einigen Fällen haben sich die Studierenden gezielt Feedback zu eigenen Lösungsansätzen von ChatGPT eingeholt. Das Potenzial der Antworten von ChatGPT zur konstruktiven Unterstützung wurde nicht betrachtet.

Eine weitere aktuelle Untersuchung von über 20 KI-Angeboten aus dem deutsch- und englischsprachigen Raum, die u. a. gezielt die konstruktive Unterstützung bei der Bearbeitung von Aufgaben untersucht hat, zeigt erhebliche Defizite auf (Rühl et al., in Veröffentlichung). Die Untersuchung ergab, dass die Unterstützung in vielen Fällen lediglich aus kleinschrittigen Lösungen besteht, die kaum zur Aktivierung der Lernenden beitragen, und nur selten Visualisierungen zur Förderung des Verständnisses enthalten. Dies zeigt, dass die untersuchten Angebote bislang kaum konstruktive Unterstützung für Lernende bei der Aufgabebearbeitung bieten.

Eine vielfältige Möglichkeit, die genannten Herausforderungen zu bewältigen, ist der Einsatz aufgabenspezifischer LLMs als KI-Tutoren zur konstruktiven Unterstützung. Dabei handelt es sich um LLMs, die speziell für eine konkrete Mathematikaufgabe entwickelt wurden. Solche spezialisierten LLMs könnten Lernende effektiver unterstützen als generische Modelle, da sie gezielt auf die spezifischen Anforderungen einer einzelnen Aufgabe zugeschnitten sind. So bietet etwas ChatGPT die Möglichkeit, individuell konfigurierbare GPTs (sogenannte "Custom GPTs") zu erstellen. Durch das Hinterlegen eines aufgabenspezifischen Prompts und relevanter Dateien (z. B. Bilder) können aufgabenspezifische Fehler, typische Fehlvorstellungen und spezifische Unterstützungselemente, wie aufgabenspezifische Sprachmittel und Visualisierungen, durch das LLM berücksichtigt werden.

Wie bei generischen LLMs fehlen bislang jedoch auch für aufgabenspezifische KI-Tutoren empirische Studien, die deren Nutzung durch Lernende im

Mathematikunterricht untersuchen, insbesondere hinsichtlich ihrer Effektivität bei der konstruktiven Unterstützung.

### **Forschungsinteresse**

Im Rahmen des Dissertationsprojekt werden daher aufgabenspezifische KI-Tutoren zur konstruktiven Unterstützung von Lernenden beim Bearbeiten von Übungsaufgaben entwickelt und beforscht. Ziel ist es, durch eine qualitative Studie Erkenntnisse über die Nutzung aufgabenspezifischer KI-Tutoren durch Lernende zu gewinnen und damit verbundene Potenziale sowie mögliche Hürden für konstruktive Unterstützung zu identifizieren. Die übergeordnete Forschungsfrage lautet:

- Wie nutzen Lernende einen aufgabenspezifischen KI-Tutor bei der Bearbeitung von Übungsaufgaben, und welche Potenziale und Hürden lassen sich in Bezug auf konstruktive Unterstützung identifizieren?

Aus dieser übergeordneten Forschungsfrage ergeben sich die folgenden Unterforschungsfragen:

- (Uf1) Wann nutzen die Lernenden den aufgabenspezifischen KI-Tutor? Welche Fragen stellen die Lernenden?
- (Uf2) Inwiefern löst die Interaktion mit dem aufgabenspezifischen KI-Tutor kognitive und metakognitive Prozesse bei den Lernenden aus?

In der geplanten Studie werden KI-Tutoren entwickelt und beforscht, die spezifisch auf drei verschiedene Schulbuchaufgaben zu linearen Gleichungssystemen abgestimmt sind. Für jede Aufgabe wird durch gezieltes Prompting mit Hilfe von Custom GPTs ein aufgabenspezifischer KI-Tutor entwickelt. Dieser Prompt beinhaltet pädagogische, fachliche und fachdidaktische Aspekte. Hierzu wird etwa im Prompt betont, dass die Lernenden sich die Lösung selbst erarbeiten sollen und der aufgabenspezifische KI-Tutor durch gezielte Fragen zum Nachdenken und Reflektieren anregen soll. Weiterhin werden in den Prompt bekannte, relevante Grundvorstellungen zu Variablen, zu Termen und zu Gleichungen integriert. Zudem wird curriculare Vorwissen, das die Lernenden bereits erworben haben sollten (z. B. das Distributivgesetz) angegeben und verschiedene Lösungswege, beispielsweise das Additionsverfahren oder das Gleichsetzungsverfahren, aufgezeigt. Auch bekannte inhaltliche Fehler und mögliche Strategien der Lernenden zur Bearbeitung der Aufgabe werden in den Prompt eingefügt, um den aufgabenspezifischen KI-Tutor für mögliche Fehler und Vorgehensweisen der Lernenden zu sensibilisieren. Außerdem wird eine lernförderliche Visualisierung hinterlegt, die bei Bedarf angezeigt werden soll.

Die Lernenden bearbeiten die Aufgaben in Einzelarbeit und werden aufgefordert, im Bearbeitungsprozess laut zu denken. Während des Bearbeitungsprozesses steht dem Lernenden der entsprechende aufgabenspezifische KI-Tutor als Unterstützung zur Verfügung. Die Bearbeitungsprozesse werden videografiert. Im Anschluss an den Bearbeitungsprozess sind Leitfadeninterviews mit den Lernenden geplant, die ebenfalls videografiert werden.

## Ausblick

Ende November 2024 wurden im Rahmen einer kleinen Pilotstudie eine der drei Aufgaben mit dem zugehörigen KI-Tutor mit drei Lernenden der 9. Klasse erprobt. Auf Grundlage dieser Daten wird nun der aufgabenspezifische KI-Tutor weiterentwickelt. Die videografierten Bearbeitungsprozesse werden genutzt, um präzise Kriterien für die Videoauswertung zu identifizieren. Basierend darauf soll ein qualitatives Analyseschema erarbeitet werden, das zur Beantwortung der Forschungsfragen genutzt werden soll.

Weitere Daten sollen im Frühjahr 2025 erhoben werden, um das Studiendesign sowie die aufgabenspezifischen Angebote final zu optimieren und aktuell noch offene methodische Fragen zu beantworten, bevor die Hauptstudie für den Sommer 2025 geplant ist.

## Literatur

- Dilling, F. & Herrmann, M. (2024). Using large language models to support pre-service teachers reasoning - an exploratory study on ChatGPT as an instrument for creating mathematical proofs in geometry. *Frontiers in Artificial Intelligence*, 7, Artikel 1460337. <https://doi.org/10.3389/frai.2024.1460337>
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., . . . Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, Artikel 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Praetorius, A.-K., Klieme, E., Herbert, B., & Pinger, P. (2018). Generic dimensions of teaching quality: the German framework of Three Basic Dimensions. *ZDM - Mathematics Education*, 50, 407–426. <https://doi.org/10.1007/s11858-018-0918-4>
- Rühl, L., Buchholtz, N., Reinhold, F., Straser, O., & Thurm, D. (in Veröffentlichung). Generative KI-Systeme im Mathematikunterricht.
- Ständige Wissenschaftliche Kommission der Kultusministerkonferenz. (2024). *Large Language Models und ihre Potenziale im Bildungssystem: Impulspapier der Ständigen Wissenschaftlichen Kommission der Kultusministerkonferenz*. <https://doi.org/10.25656/01:28303>
- Ufer, S., Heinze, A., & Lipowsky, F. (2023). Unterrichtsqualität und Instrukionsstrategien. In R. Bruder, A. Büchter, H. Gasteiger, B. Schmidt-Thieme, & H.-G. Weigand (Hrsg.), *Handbuch der Mathematikdidaktik* (2. Auflage, S. 465–492). Springer.