

HEER, Benedikt Tobias  
Siegen

## **Mathematisches Problemlösen mithilfe eines KI-basierten Tutors**

Seit der Veröffentlichung des von OpenAI entwickelten Chatbots ChatGPT 3.5 am 30. November 2022 sind Large Language Models (LLMs) ein integraler Bestandteil des alltäglichen Lebens. Diese Entwicklung macht dabei auch nicht vor dem Bildungswesen halt. Beispielhaft kann hierfür die aktuelle Handlungsempfehlung zum Umgang mit Künstlicher Intelligenz (KI) der Kultusministerkonferenz gesehen werden (KMK, 2024). Mit den Veränderungen gehen für den Mathematikunterricht indes auch große Chancen einher (Dilling et al., 2024 & Schorcht et al., 2024). In dieser Studie soll mit dem Problemlösen ein Kernelement mathematischen Arbeitens (Rott et al., 2023) im Hinblick auf die Verwendung von LLMs mit einem Fallstudienansatz untersucht werden.

Zu Beginn soll kurz die durchgeführte Fallstudie skizziert werden. Dabei ist zu nennen, dass insgesamt vier Lerngruppen mit jeweils drei bis vier Teilnehmenden der Jahrgangsstufe Q2 eines Gymnasiums in NRW untersucht wurden. Zwei der Gruppen bestanden dabei ausschließlich aus Lernenden eines Leistungskurses, während die anderen beiden Gruppen exklusiv aus Schüler\*innen eines Grundkurses gebildet wurden. Alle Beteiligten weisen nach eigener Aussage sowohl im privaten, als auch im schulischen Bereich schon viel Erfahrung im Umgang mit LLMs auf, haben aber noch nie eine Einführung in Promptingtechniken erhalten. Der Versuchsablauf war bei allen identisch: In einem Zeitraum von 75 Minuten sollten die Lerngruppen eine Problemlöseaufgabe zur Geometrie der Mittelstufe bearbeiten (<https://uni-siegen.sciebo.de/s/MkweHWRqE2S7Dce>). Dabei konnten diese freiwillig auf einen mit der Aufgabenstellung vorgeprompteten CustomGPT zurückgreifen. Im Anschluss wurden die Lernenden hinsichtlich ihrer Erfahrungen und den Eindrücken aus der vorangegangenen Aufgabe interviewt. Sowohl das Interview, als auch die Arbeitsphase wurden videographiert und bilden in Kombination mit den Chatverläufen und den Notizen der Schüler\*innen die Basis für die Auswertung. Im Zentrum steht dabei die Frage, wann und wie die Lernenden die KI im Lösungsprozess einbinden.

Zuerst soll im Vordergrund stehen, wie der Problemlöseprozess ausgestaltet ist und in welchen Phasen die KI verwendet wurde. Dafür wurde als Grundlage das deskriptive Modell von Rott et al. (2021) hinzugezogen und die Bearbeitungsphase entsprechend kategorisiert. An dieser Stelle bietet sich ein Blick in Abbildung 1 und 2 an. Diese stehen exemplarisch für zwei der untersuchten Gruppen und decken sowohl den Leistungskurs (Gruppen 1 & 2),

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

als auch den Grundkurs (Gruppen 3 & 4) ab. Generell lassen sich in allen vier Prozessen ähnliche Muster bezüglich des Einsatzes der KI erkennen: So ist als erstes ein ähnlicher Aufbau zu nennen. Wie in den beiden Abbildungen erkennbar ist, findet sich zu Beginn eine Konfrontation mit dem Problem, welche durchweg ohne das Hinzuziehen der KI abläuft. Ähnlich verhält es sich mit der anschließenden Analyse der Aufgabenstellung. Einzig Gruppe 4 greift an dieser Stelle schon auf die Unterstützung des Chatbots zurück, um zu überprüfen, ob die Behauptung generell zutreffend ist. Im Anschluss differenziert sich dann das Bild: Die Schüler\*innen greifen zumeist dann auf das technische Hilfsmittel zurück, wenn sie selbst keine Idee für ihr weiteres Vorgehen haben. Bei Gruppe 1 geschieht dies allerdings erst nach 35 Minuten. Zuvor lässt sich eine ausgeprägte Ablehnung gegenüber der KI ausmachen:

"Das macht mich ein bisschen fertig. [...] Das gibt einfach meinem Ego [...] so auch nicht den Raum, das zu nutzen. Weil ich fühle mich jetzt so, als [...] könnte ich das Problem nicht ohne lösen."

Erst nach längerer Zeit wird die KI seitens der Schüler\*innen zur Überprüfung des eigenen Ansatzes verwendet (Abb. 1). Die Lernenden greifen den Lösungsvorschlag auf und implementieren diesen schrittweise. Die Software kommt einzig zur Erklärung eines geometrischen Begriffs zum Einsatz.

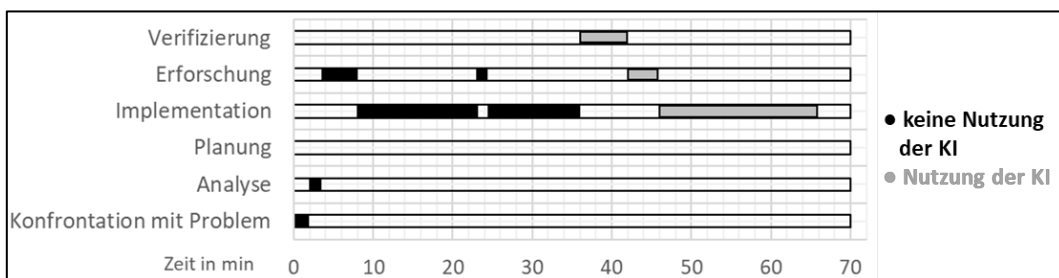


Abb. 1: Phasen des Problemlöseprozesses von Gruppe 1

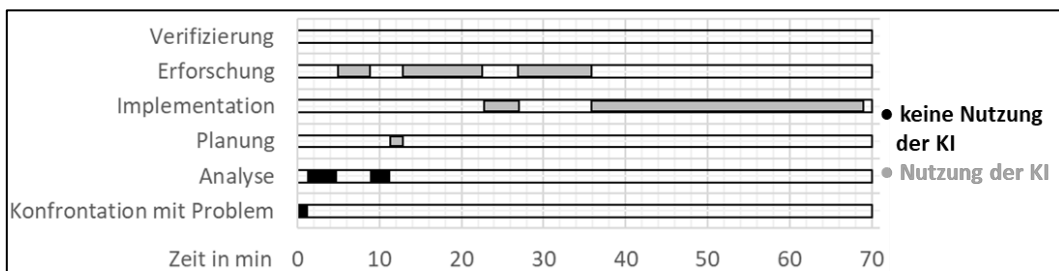


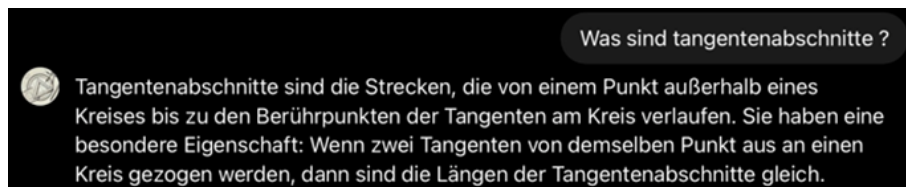
Abb. 2: Phasen des Problemlöseprozesses von Gruppe 3

Die andere Gruppe des Leistungskurses nutzt den Chatbot noch weniger. Im Gegensatz zur vorigen Gruppe wird der Lösungshinweis allerdings viel früher eingefordert. Auch hier findet sich im Anschluss nur eine Frage zur Erklärung der "Winkelhalbierenden".

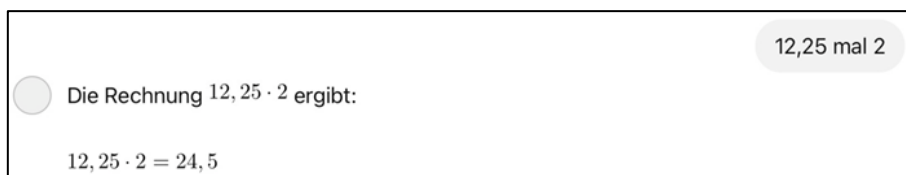
Anders sieht es bei den restlichen Proband\*innen aus. Wie in Abbildung 2 zu erkennen ist, nimmt die Arbeit mit der KI bedeutend mehr Raum ein. Hier

ist diese nahezu omnipräsent. Die Schüler\*innen fordern schnell eine Anleitung zur Lösung der Aufgabe an und orientieren sich stets daran.

In einem zweiten Schritt sollen nun einzelne besonders interessante Nutzungsweisen näher betrachtet werden. Dabei lassen sich grundsätzlich zwei verschiedene Muster hinsichtlich der Einsatzformen differenzieren. Als erstes muss die Rolle als Werkzeugersatz definiert werden. Diese findet sich – in unterschiedlich starker Ausprägung – bei allen Beteiligten. Omnipräsent ist die Verwendung der KI als Lehrbuch, um sich mathematische Begriffe wie die "Tangentenabschnitte" erklären zu lassen (vgl. Abb. 3). Eine Gruppe geht dabei sogar noch weiter und ersetzt den eigentlich vorliegenden Taschenrechner durch die KI (vgl. Abb. 4). Dem Chatbot wird somit schnell die Rolle als "Universalwerkzeug" zuteil.

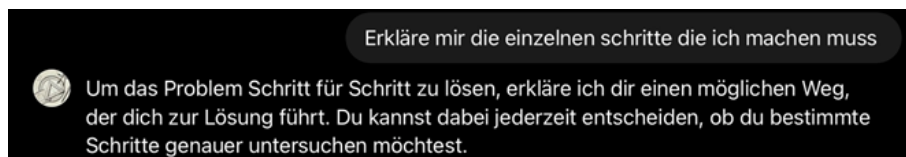


**Abb. 3:** Ausschnitt aus dem Chatverlauf von Gruppe 2



**Abb. 4:** Ausschnitt aus dem Chatverlauf von Gruppe 4

Des Weiteren ist die Rolle eines Tutors zu finden. Alle Gruppen nutzen einen Prompt ähnlich zu dem in Abbildung 5. Dort steht im Fokus, eine Anleitung zur Lösung des Problems zu erhalten. Diese Rolle als Initiator des eigentlichen Problemlöseprozesses würde andernfalls der Lehrkraft zufallen.



**Abb. 5:** Ausschnitt aus dem Chatverlauf der dritten Gruppe

Allen ist dabei gemein, dass nur Zero-Shot-Prompts (Schorcht et al., 2024) verwendet werden. Dies kann als Ausdruck der fehlenden Einführung in das Prompt Engineering gesehen werden.

Abschließend gilt es metareflexiv die Aussagen der Lernende zu betrachten. Hier lassen sich teilweise stark divergente Perspektiven auf diese finden. So gibt es einzelne Schüler\*innen, die die KI aufgrund der sofortigen Verfügbarkeit als Ergänzung für die Lehrkraft und Ersatz für das Schulbuch sehen.

"Du kannst effizienter deine Aufgaben rechnen. [...] Beziehungsweise du musst nicht im Buch nachblättern. Das dauert halt ewig."

Diese positive Einstellung teilen allerdings nicht alle Befragten. So kritisiert eine Gruppe ihre eigenen fehlenden Promptingkenntnisse als Hindernis für einen gezielten Einsatz der KI.

"Ich würde auf jeden Fall die Aufklärung für wichtig befinden, also dass man halt generell den vernünftigen Umgang mit der KI kennenlernt und auch versteht, [...] was man machen muss, damit man zu seinem Ziel kommt."

Die erste Gruppe ist sehr kritisch gegenüber LLMs. Dies klingt dabei schon im zuvor zitierten Ausschnitt an. Sie sehen das Ergebnis nicht mehr als eigene Leistung und nennen im Kontrast zu einer realen Lehrkraft auch klare pädagogische Defizite. Ihre Kritik ist aber fundamentaler und stellt die KI-Nutzung in der Mathematik generell in Frage.

"Sprachliche Sachen, alles schick, aber es wird überbewertet für Recherche und logische Gedankengänge. Und das als KI zu bezeichnen, ist meiner Meinung nach grenzwertig. Intelligenz würde ich philosophisch gesehen anders definieren."

Abschließend bleibt zu resümieren, dass in dieser Fallstudie die KI seitens der Lernenden unterschiedlich stark zum Einsatz kam. Dies manifestiert sich dabei in den divergenten Bewertungen des KI-Gebrauchs der Gruppen. Es muss indes angemerkt werden, dass die fehlenden Promptingkenntnisse einen besonders ergiebigen Einsatz der Technologie verhindert haben. Für eine weiterführende Forschung wäre sicherlich eine vorangestellte Schulung zum Prompt Engineering trotz der Vorerfahrungen der Schüler\*innen sinnvoll.

## Literatur

- Dilling, F., Herrmann, M., Müller, J., Pielsticker, F., & Witzke, I. (2024). Initiating interaction with and about ChatGPT – an exploratory study on the angle sum in triangles. In E. Faggiano, A. Clark-Wilson, M. Tabach, & H.-G. Weigand (Hrsg.), *Proceedings of the 17th ERME topic conference MEDA 4* (S. 145–152). University of Bari Aldo Moro.
- Kultusministerkonferenz (2024). *Handlungsempfehlung für die Bildungsverwaltung zum Umgang mit Künstlicher Intelligenz in schulischen Bildungsprozessen*. [https://www.kmk.org/fileadmin/veroeffentlichungen\\_beschluesse/2024/2024\\_10\\_10-Handlungsempfehlung-KI.pdf](https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2024/2024_10_10-Handlungsempfehlung-KI.pdf)
- Rott, B., Bruder, R., Heinrich, F., Bauer, C. (2023). Problemlösen lernen. In: Bruder, R., Büchter, A., Gasteiger, H., Schmidt-Thieme, B., Weigand, HG. (Hrsg.) *Handbuch der Mathematikdidaktik*. Springer Spektrum. [https://doi.org/10.1007/978-3-662-66604-3\\_10](https://doi.org/10.1007/978-3-662-66604-3_10)
- Rott, B., Specht, B. & Knipping, C. (2021). A descriptive phase model of problem-solving processes. *ZDM Mathematics Education*, 53, 737–752. <https://doi.org/10.1007/s11858-021-01244-3>
- Schorcht, S., Buchholtz, N. & Baumanns, L. (2024). Prompt the problem - investigating the mathematics educational quality of AI-supported problem solving by comparing prompt techniques. *Frontiers in Education*, 9, Artikel 1386075. <https://doi.org/10.3389/feduc.2024.1386075>