

BARTON, Daniel & KATTER, Valentin
Osnabrück, Bielefeld

"Wo liegt mein Fehler?" Einsatz von generativer KI zur Fehlerreflexion. Erste Zwischenergebnisse des Projekts KI-BEBA.

Reflexionen über eigene Fehler haben sich als eine effektive Methode erwiesen, um nachhaltiges Lernen zu fördern (Heemsoth & Heinze, 2015). Insbesondere das Nachdenken über die Ursachen von Fehlern ermöglicht es Lernenden, eigene kognitive Modelle zu überprüfen und zu erweitern, was sowohl das prozedurale als auch das konzeptionelle Wissen verbessert. Gleichzeitig zeigt sich, dass gerade leistungsschwächere Schüler*innen häufig zusätzliche Unterstützung benötigen, um Fehler effektiv reflektieren zu können. Hier bietet sich der Einsatz von LLMs als digitale individuelle Lernassistenz an, die gezielte und differenzierte Hilfestellungen bieten kann. Aufgrund der rasanten Entwicklungen im Bereich der künstlichen Intelligenz gibt es bisher allerdings kaum Erfahrungen und Forschungsergebnisse darüber, ob LLMs geeignet sind individuelles Feedback für Lernende zu generieren und wie dies effektiv im Unterricht eingesetzt werden kann.

In diesem Beitrag wird die BEBA-Strategie untersucht, ein strukturierter Ansatz zur Fehleranalyse, der sich im Mathematikunterricht bewährt hat (ebd.). Ziel der Untersuchung ist es, Bedingungen einer Lernumgebung zu identifizieren, die einen effektiven Einsatz von ChatGPT im Mathematikunterricht ermöglichen, so dass Schüler*innen individuelle und differenzierte Rückmeldungen zu ihren gemachten Fehlern erhalten und daraus lernen können. Dabei dient ChatGPT als digitaler Lernassistent, der es den Lernenden ermöglicht, ihre eigenen Denk- und Lösungswege kritisch zu reflektieren und gezielt zu verbessern.

Design-Research

Die Ergebnisse der Untersuchung sollen dazu beitragen, praxisnahe Einsatzmöglichkeiten von LLMs im Mathematikunterricht zu entwickeln, um Lehrkräfte und Schüler*innen gleichermaßen zu unterstützen. Aus diesem Grund wurde für die vorliegende Studie ein Design-Research Ansatz ausgewählt.

Design-Research ist ein Forschungsansatz, der sich durch die enge Verknüpfung von Forschung und Praxis auszeichnet. Ziel dieses Ansatzes ist es, praxisrelevante Probleme aus dem Bildungsbereich zu adressieren und gleichzeitig fundierte theoretische Erkenntnisse zu generieren. Im Fokus stehen dabei die iterative Entwicklung, Erprobung und Verbesserung von didaktischen Ansätzen oder Materialien.

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

Den methodischen Rahmen dieser Studie bildet das Dortmunder Modell des Design-Research, welches sich durch vier Schritte auszeichnet: (1) Lerngegenstände spezifizieren und strukturieren, (2) Design (weiter-)entwickeln, (3) Design-Experimente durchführen und auswerten und (4) lokale Theorien zu Lehr-Lernprozessen (weiter-)entwickeln (Prediger et al. 2012). Dabei folgen diese Schritte nicht zwangsläufig strikt nacheinander, sondern gehen ineinander über. Auch der Startpunkt eines solchen Zyklus kann variieren.

Lerngegenstände spezifizieren und strukturieren

In dem Forschungsprojekt geht es um methodische Überlegungen zum Einsatz von KI im Mathematikunterricht. Aufgrund der potentiellen Möglichkeit von ChatGPT individuelles Feedback für Schüler*innen bereitzustellen wurden Lernumgebungen untersucht, in denen ChatGPT als digitaler Lernassistent eingesetzt wird. Grundlage dafür ist der Ansatz von Brüggemann (2024), die die BEBA-Strategie zusammen mit ChatGPT bereits erprobt hat. Hierbei reflektierten Schüler*innen ihre Fehler aus Mathematikklausuren und durchliefen dabei die vier Schritte der BEBA-Strategie: (1) Fehler beschreiben, (2) den Denkprozess erklären, (3) Fehler berichtigen und (4) eine ähnliche Aufgabe erstellen und lösen.

Design weiterentwickeln

Bereits in der Erprobung durch Brüggemann (2024) wurden einige Kritikpunkte am ursprünglichen Design deutlich. Die Schüler*innen bemängelten, dass ChatGPT zu viel Text generierte, was die Nutzung der Rückmeldungen aufwändig macht und die Bearbeitung unnötig in die Länge ziehen kann. Zudem kam es gelegentlich zu Missverständnissen oder Fehlern in den Antworten von ChatGPT. Um diesen Herausforderungen bereits im ersten Zyklus der Studie zu begegnen, wurden drei Anpassungen vorgenommen. (1) Schüler*innen erhielten eine gezielte Einführung in die Nutzung von ChatGPT zur Unterstützung bei mathematischen Problemen. (2) ChatGPT wurde instruiert, sich wie eine Lehrkraft zu verhalten, die keine direkten Lösungen vorgibt, sondern Schülerinnen und Schüler dabei unterstützt, eigenständig zur richtigen Lösung zu gelangen. (3) Um sicherzustellen, dass ChatGPT die BEBA-Strategie effektiv unterstützen kann, wurden entsprechende Informationen hinterlegt.

Designexperiment durchführen und auswerten

Es nahmen insgesamt 25 Schüler*innen einer zehnten Klasse eines Gymnasiums in NRW teil. Die Schüler*innen schrieben zunächst eine reguläre Klassenarbeit zum Thema quadratische Gleichungen. In der darauffolgenden Stunde wurde das ChatGPT-Tool von fobizz vorgestellt, und die

Schüler*innen lösten drei Aufgaben zu quadratischen Gleichungen mithilfe von ChatGPT. In der dritten Stunde wurde die Klassenarbeit zurückgegeben, und die Schüler*innen hatten die Möglichkeit, die von ihnen gemachten Fehler mit der BEBA-Strategie unter Zuhilfenahme von ChatGPT zu bearbeiten. Die Schüler*innen wurden dazu angehalten, eine Aufgabe auszuwählen und ein Arbeitsblatt zur BEBA-Strategie auszufüllen (vgl. Brüggemann, 2024). Anschließend wurde die Unterrichtssequenz mit einem Fragebogen evaluiert, der das Unterstützungsangebot, die Vertrauenswürdigkeit, das mathematische Selbstkonzept und die Technikaffinität der Schüler*innen erfasste.

Zentral für die Auswertung der Nutzung sind zudem die Chatprotokolle, die während der Anwendung von ChatGPT erstellt wurden. In einer qualitativen Inhaltsanalyse wurden zunächst die Prompts der Schüler*innen kodiert. Diese Kodierung ermöglichte es, die Nutzung von ChatGPT durch die Schüler*innen systematisch zu analysieren und Unterschiede in der Qualität und Art der Interaktion mit der KI sichtbar zu machen.

Lokale Theorien zu Lehr-Lernprozessen entwickeln

Die Auswertung des Fragebogens zeigt, dass die Schüler*innen die Vertrauenswürdigkeit und das Unterstützungsangebot von ChatGPT insgesamt positiv bewerteten. Die Befunde deuten darauf hin, dass die Schüler*innen ChatGPT als hilfreiches Werkzeug wahrnahmen.

Die Analyse der Chatprotokolle zeigt, dass die Schüler*innen ChatGPT - wie von ihnen verlangt - überwiegend zur Fehleranalyse nutzten. Dabei griffen sie häufig auf den zuvor ausgeteilten Prompt "Was habe ich falsch gemacht?" zurück. Jedoch beschränkten sich die Prompts meist darauf, die Aufgabe und ihre fehlerhafte Lösung zu nennen, ohne den Lösungsweg bereitzustellen, was eine präzise Fehleranalyse erheblich erschwerte.

In den Fällen, in denen ChatGPT ausreichend Informationen hatte, um den Fehler zu erkennen, war das generierte Feedback oftmals nicht hilfreich. Statt den spezifischen Fehler herauszustellen, reagierte ChatGPT in den meisten Fällen mit der Bereitstellung einer Musterlösung zur Aufgabe (vgl. Abb. 1). Selbst bei offensichtlichen Fehlern konnte ChatGPT die spezifischen Probleme häufig nicht identifizieren und adressieren.

Darüber hinaus wurde erneut die Problematik deutlich, dass die Rückmeldungen von ChatGPT zu lange und zu komplex waren.

<p>Du Ich habe bei der Umformung der Gleichung $-\frac{1}{2}x^2 - \frac{1}{2}x - 1 = 9,5$ einen Fehler gemacht. Ich habe herausbekommen, dass es keine Lösung für "X" gibt. Wo liegt mein Fehler.</p> <p>GPT-4o-mini Um den Fehler bei der Umformung der Gleichung zu finden, lass uns die Gleichung Schritt für Schritt betrachten: Die ursprüngliche Gleichung lautet: $-\frac{1}{2}x^2 - \frac{1}{2}x - 1 = 9,5$ Zuerst bringen wir alle Terme auf eine Seite der Gleichung, indem wir 9,5 auf die linke Seite subtrahieren: [...]</p>	<p>Du Erkläre mir genauer, wo mein Fehler liegt.</p> <p>GPT-4o-mini Natürlich! Lass uns die Umformung der Gleichung Schritt für Schritt durchgehen und genau auf den Punkt eingehen, an dem der Fehler aufgetreten ist. Die ursprüngliche Gleichung lautet: $-\frac{1}{2}x^2 - \frac{1}{2}x - 1 = 9,5$ Schritt 1: Subtrahiere 9,5 von beiden Seiten Um die Gleichung zu vereinfachen, subtrahieren wir 9,5 von beiden Seiten: [...]</p>
---	--

Abb. 1: Auszüge aus Chatprotokoll

Vorläufiges Fazit des ersten Zyklus

Die Ergebnisse der Studie zeigen, dass es notwendig ist, Schüler*innen gezielt darin zu schulen, wie sie ChatGPT präzise Anweisungen geben können, um hilfreiches Feedback zu erhalten, da sie ohne entsprechende Vorbereitung oft ungenaue oder unzureichende Eingaben machen. Zudem wird deutlich, dass ChatGPT selbst optimiert werden muss, um kürzere, weniger formale und besser verständliche Antworten zu generieren, die den spezifischen Bedürfnissen der Schüler*innen besser gerecht werden.

Ausblick

Ausgehend von den Erkenntnissen aus dem ersten Zyklus lassen sich folgende Adaptionen im Hinblick auf den zweiten Zyklus vornehmen:

- **Optimierung von ChatGPT:** Durch ein gezieltes Vorprompten können die Rahmenbedingungen der Antworten definiert werden (kürzere Antworten, weniger Formalismus), was zu einer präziseren Rückmeldung führen soll.
- **Schulung der Schüler*innen:** Eine gezielte Einführung in den Umgang mit ChatGPT ist notwendig, damit die Schüler*innen lernen, präzise Anweisungen zu formulieren und den Output von ChatGPT kritisch zu reflektieren. Dadurch wird sichergestellt, dass die generierten Antworten den gewünschten Nutzen bringen.

Literatur

- Brüggemann, J. (2024). BEBA-Strategie trifft ChatGPT: Reflexion und Korrektur eigener Fehler motivieren. *mathematik lehren* (244), 43–48.
- Heemsoth, T. & Heinze, A. (2016). Secondary School Students Learning From Reflections on the Rationale Behind Self-Made Errors: A Field Experiment. *The Journal of Experimental Education*, 84(1), 98–118.
- Prediger, S., Link, M., Hinz, R., Hußmann, S., Thiele, J. & Ralle, B. (2012). Lehr-Lernprozesse initiieren und erforschen: Fachdidaktische Entwicklungsforschung im Dortmunder Modell. *MNU*, 65(8), 452–457.