

KI als Lernbegleiter

Förderung von Autonomie durch einen Chatbot in der MatheArena-App

Thomas Benesch¹, Eva-Maria Infanger², Corinna Hörmann³

DOI: <https://doi.org/10.53349/re-source.2026.i2.a1562>

Zusammenfassung

Digitale Werkzeuge führen zu einer tiefgreifenden Veränderung von Lehr- und Lernprozessen. Dies führt zu dem nachfolgenden Beitrag über einen auf pädagogischer Basis entwickelten KI-Chatbot und dessen Integration in die mobile Mathematik-Lernplattform *MatheArena*. Individuelles Lernen soll durch eine dynamische individuelle Unterstützung gefördert werden, die auf die spezifischen Kontextbedingungen eingeht und so die Motivation der Lernenden fördert. Der Chatbot ergänzt adaptive Aufgabenvorschläge der App, die auf einem Elo-Algorithmus basieren und den Schwierigkeitsgrad laufend anpassen. Didaktisch orientiert sich der Ansatz am sokratischen Tutoring: Statt fertige Lösungen zu liefern, begleitet der Chatbot Lernende durch Fragen, Hinweise und Unterstützung bei eigenständigem Problemlösen. Technisch basiert das System auf dem Large Language Model *Gemini Flash*, das hinsichtlich pädagogischer Steuerung, Themenfokus und Datenschutz optimiert wurde. Der Beitrag diskutiert, wie ein solcher Chatbot sinnvoll in die Didaktik eingebettet werden kann, um mathematisches Denken und Problemlösekompetenzen in digitalen Lernumgebungen gezielt zu fördern.

Stichwörter: Mobile Learning, Chatbot, Adaptive Lernsysteme, Autonomes Lernen, KI in der Bildung

1 Einleitung

In den letzten 25 Jahren haben die PISA-Ergebnisse in vielen europäischen Ländern einen stetigen Rückgang gezeigt (OECD, 2025), was die anhaltenden Debatten über die Wirksamkeit,

¹ MatheArena FlexCo, Engersdorf 30, 4921 Hohenzell

E-Mail: thomas@mathearena.com

² MatheArena FlexCo, Engersdorf 30, 4921 Hohenzell

E-Mail: infanger.eva-maria@gmx.at

³ Universität Salzburg, Jakob-Haringer-Straße 2, 5020 Salzburg

E-Mail: corinna.hoermann@plus.ac.at

Relevanz und Anpassungsfähigkeit der derzeitigen Bildungssysteme belebt. In diesem Zusammenhang hat sich die Digitalisierung sowohl als Herausforderung als auch als möglicher Hebel für Transformation erwiesen. Die zunehmende Präsenz digitaler Werkzeuge in Bildungseinrichtungen eröffnet neue Möglichkeiten für Engagement und personalisiertes Lernen, birgt aber auch erhebliche Risiken.

Eine herausragende Entwicklung ist die rasche Einführung von Technologien, die auf Künstlicher Intelligenz (KI) basieren, wie z. B. Claude, Gemini oder ChatGPT. Eine Möglichkeit, KI im Bildungsbereich einzusetzen, sind Chatbots. Das Ziel eines Chatbots ist es, Interaktion mit Benutzer*innen über verschiedene Kommunikationskanäle (textuell, auditiv) zu erleichtern, um Befehle, Abfragen oder Anfragen zu verarbeiten und entsprechende Antworten zu liefern. Er ist so konzipiert, dass er einen intelligenten menschlichen Dialog simuliert und eine Erfahrung bietet, die der Interaktion mit einer realen Person unter Verwendung natürlicher Sprache möglichst ähnlich ist (Moral-Sánchez et al., 2023; Cheng et al., 2022). Während diese Werkzeuge vielversprechend sind, um das Lernen und die Kommunikation zu verbessern, bergen sie auch potenzielle Gefahren, darunter die Auseinandersetzung mit Fehlinformationen, Manipulation oder anderen Formen von Cyber-Risiken. Jüngste Erkenntnisse deuten darauf hin, dass 94% der österreichischen Jugendlichen KI-Chatbot nutzen, wobei ChatGPT am häufigsten genannt wird (Saferinternet.at, 2026). Dies unterstreicht den hohen Wert der Implementierung geschlossener, gut regulierter Chatbot-Systeme in Bildungskontexten.

Als Reaktion auf diese Entwicklungen hat *MatheArena* vor Kurzem einen eigenen KI-gestützten Chatbot integriert, der sich auch an die aktuelle mathematische Aufgabe der Nutzer*innen anpassen kann. Im Gegensatz zu Open-Access-Systemen ist diese Version auf die Bildungsziele der App zugeschnitten und bietet den Lernenden geführte Unterstützung in einer sicheren Umgebung, die auf den Lehrplan abgestimmt ist. *MatheArena* ist eine mobile Lernanwendung, die entwickelt wurde, um das Mathematiklernen in einer Vielzahl von Bildungsumgebungen und Altersgruppen zu unterstützen und zu verbessern. Mit lernerzentriertem Design und dem Fokus auf Offline-Zugänglichkeit bietet die Anwendung eine flexible und ansprechende Plattform für Lernende von der Primarstufe bis zur Matura (Klasse 12) sowie für erwachsene Lerner, die sich im lebenslangen Lernen engagieren. Die Anwendung ist für die individuelle Nutzung auf Smartphones und Tablets optimiert, was sie zu einem wertvollen Werkzeug für das selbstgesteuerte Lernen außerhalb des Klassenzimmers macht. Zudem wird adaptives Lernen unterstützt, indem Inhalte und Feedback auf den Fortschritt und die Bedürfnisse der Lernenden zugeschnitten werden.

2 Theoretischer Hintergrund

2.1 Individualisiertes Lernen und adaptive Systeme

Individualisiertes Lernen zählt zu den zentralen Herausforderungen zeitgemäßer Bildung. In heterogenen Lerngruppen variieren Vorwissen, Lerngeschwindigkeit und Motivation erheb-

lich – ein einheitliches Lernangebot kann diesen Unterschieden strukturell nicht gerecht werden. Die Ressourcen der einzelnen Lehrkraft im Klassenkontext reichen für den Differenzierungsbedarf gleichzeitig nicht aus. Auch für selbstbestimmtes Lernen sind individualisierte Zugänge hoch relevant, um die Motivation hochzuhalten und einen nachhaltigen Lernerfolg zu ermöglichen.

2.1.1 Adaptive Lernumgebungen

Adaptive Lernumgebungen begegnen dieser Herausforderung, indem sie Lerninhalte, Aufgabenschwierigkeiten und Rückmeldungen dynamisch an den individuellen Lernstand anpassen und eine vermittelnde Zwischeninstanz in der Lehr-Lern-Situation zur Verfügung stellen, wo Antworten abgegeben, Fragen gestellt, Feedback eingeholt und formative Diagnosen erhoben werden können.

Gallagher et al. (2020) beschreiben adaptive Lernumgebungen als Systeme, die kontinuierlich Leistungsdaten erheben, auswerten und auf deren Grundlage den weiteren Lernpfad modifizieren. Im Unterschied zu statischen Lernmaterialien reagieren solche Systeme auf das tatsächliche Verhalten der Lernenden: Fehlerquoten, Bearbeitungszeiten und Lösungsmuster fließen in die Steuerung des Lernangebots ein. Metaanalytische Befunde belegen die grundsätzliche Wirksamkeit dieses Ansatzes: Ma et al. (2014) zeigen in einer Metaanalyse über 107 Studien, dass intelligente Tutorsysteme gegenüber konventionellem Unterricht zu deutlichen Leistungsverbesserungen führen, mit einer mittleren Effektstärke von $d = 0.66$.

Allerdings ist bei der Interpretation solcher Befunde Vorsicht geboten. Die von Ma et al. (2014) untersuchten Systeme sind überwiegend regelbasierte *Intelligent Tutoring Systems* (ITS) mit explizitem Schüler-, Domänen- und Tutormodell – eine Architektur, die sich grundlegend von generativen Sprachmodellen wie Gemini Flash unterscheidet. Lin et al. (2023) zeigen in ihrer systematischen Überprüfung des Forschungsfeldes, dass KI-gestützte Tutorsysteme der neueren Generation zwar erhebliches Potenzial für personalisiertes Lernen aufweisen, dass aber robuste empirische Evidenz zur Lernwirksamkeit LLM-basierter Ansätze im schulischen Kontext noch aussteht. Awang et al. (2025) bestätigen diesen Befund in ihrer aktuellen Überblicksarbeit zu KI im Mathematikunterricht: Trotz wachsender Verbreitung KI-gestützter Lernwerkzeuge fehlen belastbare Langzeitstudien, die deren spezifische Wirksamkeit im Fach Mathematik empirisch absichern. Für den KI-Chatbot der *MatheArena*-App bedeutet dies, dass Wirksamkeitsannahmen mit entsprechender Zurückhaltung formuliert bzw. erst aufgezeigt werden müssen: Das System versteht sich ausdrücklich nicht als Ersatz für bewährte ITS-Architekturen oder für menschliche Tutoring-Interaktionen, sondern als niedrighschwellige Ergänzung zum adaptiven Kernsystem der App – eine Unterstützungsebene, die dort einsetzt, wo algorithmische Aufgabenauswahl allein nicht ausreicht.

2.1.2 Der Elo-Algorithmus als Grundlage adaptiver Aufgabensteuerung

Ein verbreitetes Verfahren zur Modellierung von Leistungsniveaus in adaptiven Systemen ist der Elo-Algorithmus. Ursprünglich aus dem Schachsport stammend, wurde das Modell für

Bildungskontexte adaptiert. Pelánek (2016) zeigt in einer Analyse verschiedener Anwendungsszenarien, dass der Elo-Algorithmus für adaptive Lernsysteme gut geeignet ist, sofern ausreichend Aufgabenbearbeitungen für eine stabile Schätzung vorliegen. In der *MatheArena*-App bildet der Elo-Algorithmus die technische Grundlage für die Aufgabenauswahl: Jede Aufgabe besitzt einen Schwierigkeitswert, und jede*r Schüler*in wird durch einen entsprechenden Kompetenzwert beschrieben. Nach jeder gelösten oder nicht gelösten Aufgabe werden beide Werte aktualisiert – gelingt die Lösung einer schwierigen Aufgabe, steigt der Kompetenzwert stärker als bei einer leichten; bei Misserfolg sinkt er entsprechend. Auf diese Weise entsteht ein kontinuierlicher Kalibrierungsprozess, der den Schwierigkeitsgrad der nächsten Aufgabe auf einem für das Lernen förderlichen Niveau hält – konzeptuell verankert im Konstrukt der „optimalen Herausforderung“ (Csikszentmihalyi, 1990).

Gleichwohl sind die Grenzen dieses Ansatzes nicht zu übergehen. Der Elo-Algorithmus modelliert Kompetenz als eindimensionalen Wert und bildet weder die innere Wissensstruktur der Lernenden noch die aufgabenspezifischen Teilkompetenzen ab. Pelánek (2016) weist zudem darauf hin, dass der Algorithmus bei stark variierenden Aufgabentypen und kleinen Stichproben an Genauigkeit verliert. Lernzuwächse, die schnell und sprunghaft erfolgen – etwa nach einem klärenden Unterrichtsgespräch – werden vom Algorithmus nur verzögert erfasst. Hinzu kommen kontextuelle Faktoren wie Tagesform oder emotionaler Zustand, die im Modell unberücksichtigt bleiben. Infanger et al. (2022) greifen diese Limitation im Kontext von *MatheArena* direkt auf und schlagen zeitinhomogene Markov-Ketten als alternatives Modellierungsverfahren vor, das Leistungsveränderungen über die Zeit differenzierter abbilden kann als der statische Elo-Ansatz – etwa indem es unterschiedliche Übergangswahrscheinlichkeiten für verschiedene Aufgabenschwierigkeiten und Lernphasen berücksichtigt. Dieser Ansatz verweist auf ein aktives Weiterentwicklungsfeld der zugrunde liegenden Kompetenzmodellierung, das für zukünftige Versionen der App relevant sein könnte. Weitere Alternativen wie das „Bayesian Knowledge Tracing“ (Corbett & Anderson, 1994) oder „Deep Knowledge Tracing“ (Piech et al., 2015) bieten ebenfalls differenziertere Möglichkeiten, sind jedoch mit höherem technischem Aufwand verbunden. Unter der Prämisse von bildungsgerechter, leistbarer und leicht zugänglicher Lernsoftware stellt der Elo-Algorithmus damit einen praktikablen Ausgangspunkt dar: robust, transparent und für den Schulkontext gut handhabbar – unter der Bedingung, dass seine Vereinfachungen bewusst in Kauf genommen und, wie von Infanger et al. (2022) angeregt, perspektivisch weiterentwickelt werden. Der Chatbot ergänzt diese algorithmische Grundlage um eine qualitative Dimension, die der Elo-Mechanismus strukturell nicht leisten kann: die prozessbegleitende, dialogische Unterstützung beim Verstehen.

2.1.3 Relevanz für das Mathematiklernen

Mathematik ist ein kumulatives Fach: Neue Inhalte bauen systematisch auf zuvor erworbenen Konzepten auf. Unangenehmer Nebeneffekt dieses nachhaltigen Lernansatzes sind Lücken im Vorwissen. Sie pflanzen sich sukzessive fort und können langfristig zu einem Rückstand führen,

der durch frontalen Unterricht kaum mehr aufgeholt werden kann. Aktuelle PISA-Ergebnisse verdeutlichen, dass mathematische Grundkompetenzen in vielen Bildungssystemen nach wie vor eine zentrale Herausforderung darstellen (OECD, 2026). Adaptive Systeme sind in diesem Kontext besonders wertvoll, weil sie nicht nur den aktuellen Leistungsstand abbilden, sondern auch gezielte Übungsangebote für Bereiche bereitstellen, in denen individuelle Schwächen identifiziert wurden.

Die *MatheArena*-App differenziert dabei nach Entwicklungsstufe: Die Variante *Junior* richtet sich an Lernende der Sekundarstufe I, in der grundlegende algebraische und geometrische Konzepte aufgebaut werden; die Variante *Classic* adressiert Lernende der Sekundarstufe II, bei denen abstrakteres Denken und komplexere Problemlösestrategien im Vordergrund stehen. Benesch et al. (2023) dokumentieren in einem Fallbeispiel zum Einsatz von *MatheArena Junior* in einer ersten Klasse der Sekundarstufe I, dass die App eine sinnvolle Ergänzung zum Regelunterricht darstellen kann, wenn sie didaktisch eingebettet wird. Diese Differenzierung ist lernpsychologisch bedeutsam: Während Lernende der Sekundarstufe I stärker von konkreten Handlungshinweisen und engmaschigem Feedback profitieren, benötigen Lernende der Sekundarstufe II zunehmend Unterstützung bei der Entwicklung metakognitiver Strategien (Schneider & Artelt, 2010). Hinzu kommt der sogenannte „Expertise Reversal Effect“ (Kal-yuga et al., 2003): Unterstützungsmaßnahmen, die für Lernende mit geringem Vorwissen förderlich sind, können für fortgeschrittenere Lernende redundant oder sogar hinderlich werden, da sie kognitive Kapazität binden, die für tieferes Verarbeiten genutzt werden könnte. Der Chatbot trägt dieser Unterschiedlichkeit Rechnung, indem er Hinweisdichte, Abstraktionsgrad und Kommunikationsweise an die jeweilige Zielgruppe anpasst.

2.2 Pädagogische Grundlagen

2.2.1 Drei Theorien – ein integratives Rahmenmodell

Die pädagogische Fundierung des Chatbot-Konzepts stützt sich auf drei einander ergänzende Theorietraditionen: die „Selbstbestimmungstheorie“ nach Deci und Ryan (SDT), Vygotskys „Konzept der Zone der proximalen Entwicklung“ (ZPD) und die „Cognitive Load Theory“ nach Sweller (CLT). Diese drei Perspektiven sind nicht unabhängig, sondern greifen im konkreten Interaktionsmoment ineinander: Die ZPD definiert den Lernraum, in dem die Aufgabe angesiedelt sein muss, damit nachhaltiger Lernfortschritt möglich wird. Die CLT beschreibt die Ausgestaltung des Userflows sowie die kognitive Architektur, innerhalb derer Scaffolding-Maßnahmen des Chatbots wirksam sein können. Die SDT erklärt, unter welchen motivationalen Bedingungen Lernende bereit sind, sich auf diesen Prozess einzulassen und wie die bewusste Implementierung der *MatheArena* im Unterrichtsgeschehen ideal orchestriert werden kann, um den Lernprozess bestmöglich zu begleiten.

Kommt es in einer konkreten Interaktion zu einem Konflikt zwischen diesen Dimensionen – etwa wenn ein*e Schüler*in kognitiv überlastet ist (CLT), gleichzeitig aber das Autonomieerleben durch zu direktive Hinweise oder emotional belastete Beziehung zur Lehrkraft gefähr-

det würde (SDT) – bietet der Chatbot eine alternative Anlaufstelle mit der klaren Priorität: Zunächst wird die kognitive Belastung durch Rückfragen benannt, erfasst und durch Beschreibungen und Anleitungsfragen reduziert, um überhaupt eine Grundlage für produktives Denken herzustellen; erst dann wird der Grad der Offenheit der Chatbot-Fragen schrittweise erhöht. Ohne kognitive Kapazität ist Autonomie im Lernprozess nicht erfahrbar.

2.2.2 Self-Determination Theory (Deci & Ryan)

Die Selbstbestimmungstheorie (Deci & Ryan, 2000) postuliert drei grundlegende psychologische Bedürfnisse, deren Erfüllung intrinsische Motivation begünstigt: Autonomie, Kompetenzerleben und soziale Eingebundenheit. Im Kontext der *MatheArena*-App kommt der SDT eine doppelte Funktion zu: Sie fundiert zum einen das Design des Chatbots selbst, zum anderen begründet sie die Einbettung der App als ergänzendes Unterrichtsmittel, das die Autonomie der Lernenden durch vielfältige Entscheidungsmöglichkeiten stärkt, ohne dabei die zwischenmenschliche soziale Interaktion zu ersetzen.

Autonomie bezeichnet das Erleben, das eigene Handeln selbst bestimmen zu können. Die App stärkt dieses Erleben, indem sie Lernenden vielfältige Entscheidungsmöglichkeiten eröffnet: Sie können selbst wählen, wann sie üben, welche Themenbereiche sie vertiefen und in welchem Ausmaß sie den Chatbot um Unterstützung bitten. Diese Wahlfreiheit ist eine qualitative Erweiterung gegenüber dem klassischen Unterrichtssetting, in dem Aufgabenauswahl und Tempo weitgehend durch die Lehrkraft bestimmt werden – eine zentrale und daher rare Ressource für die Einzelperson im Klassensetting. Li (2025) zeigt in einer aktuellen Untersuchung zum Einsatz KI-gestützter Chatbots im Mathematikunterricht, dass wahrgenommene Autonomie im Umgang mit dem System einen signifikanten positiven Effekt auf das Lerninteresse ausübt – ein Befund, der die SDT-basierte Designentscheidung empirisch stützt. Autonomieerleben setzt dabei voraus, dass die Aufgabe lösbar erscheint – was die Positionierung im Schwierigkeitsgefüge durch den Elo-Algorithmus zur motivationalen Grundbedingung macht.

Kompetenzerleben bezeichnet das Bedürfnis, sich als wirksam und fähig zu erleben. Der Elo-basierte Algorithmus ist direkt auf dieses Bedürfnis ausgerichtet. Gleichzeitig ist die Art der Rückmeldung entscheidend: Feedback, das den Lernprozess würdigt statt nur das Ergebnis, stärkt das Kompetenzerleben nachhaltiger als bloße Richtig-Falsch-Rückmeldungen (Hattie & Timperley, 2007). Der Chatbot ist entsprechend darauf ausgelegt, Bemühungen anzuerkennen und Denkschritte sichtbar zu machen – auch dann, wenn das Endergebnis noch nicht korrekt ist.

Soziale Eingebundenheit meint das Bedürfnis, in Beziehung zu anderen zu stehen und von bedeutsamen Personen wertgeschätzt zu werden. Es wäre verfehlt, einem KI-Chatbot echte Sozialität zuzuschreiben. Der Chatbot ist kein Ersatz für die Beziehung zwischen Lehrkraft und Lernenden und beansprucht das auch nicht. Seine Funktion ist eine andere: Er soll dort einspringen, wo zeitliche Ressourcen im Unterricht fehlen, um jedem*r Einzelnen individuell zu helfen, und damit Freiraum schaffen – Freiraum, den Lehrkräfte für tiefgehende Gespräche, konzeptuelle Schwierigkeiten und jene Momente nutzen können, die echte zwischenmensch-

liche Interaktion erfordern. Die soziale Eingebundenheit bleibt dabei strukturell Aufgabe der Lehrkraft und des Klassenverbands; der Chatbot ergänzt, entlastet und überbrückt – er substituiert nicht. Cheng et al. (2022) zeigen in diesem Zusammenhang, dass anthropomorphe Gestaltung von Chatbots zwar das Vertrauen und die Akzeptanz bei Nutzenden erhöhen kann, gleichzeitig aber andere Erwartungshaltungen erzeugt als bei klar als Maschine wahrgenommenen Systemen. Für das Design des Chatbots bedeutet dies: Eine zugewandte, responsiv gestaltete Kommunikation ist didaktisch sinnvoll, muss aber transparent machen, dass es sich um ein KI-System handelt – auch um unrealistischen Erwartungen an soziale Reziprozität vorzubeugen. Dass responsives Kommunikationsverhalten digitaler Systeme dennoch als angenehm erlebt werden kann, belegen bereits ältere Studien zur „Computers Are Social Actors-Hypothese“ (Nass & Reeves, 1996) – ein Effekt, der für die Akzeptanz des Chatbots relevant ist, ohne seine grundlegende Funktion zu verschieben.

2.2.3 Zone der proximalen Entwicklung (Vygotsky)

Vygotskys (1978) Konzept der Zone der proximalen Entwicklung unterscheidet zwischen dem aktuellen Entwicklungsstand – dem, was eine Person eigenständig leisten kann – und dem potenziellen Entwicklungsstand – dem, was sie mit Unterstützung zu erreichen vermag. Entscheidend ist das daraus abgeleitete Konzept des „Scaffolding“: temporäre, gezielte Unterstützung, die schrittweise zurückgenommen wird, sobald Lernende eigenständiges Handeln zeigen und so den Cognitive Load sukzessive trainieren.

Im Chatbot-Design konkretisiert sich dieses Prinzip als gestufte Hinweisstruktur: Der Chatbot beginnt mit einer offenen Frage, die zum eigenständigen Nachdenken einlädt. Nur wenn keine produktive Reaktion erfolgt, wird ein erster Hinweis gegeben – dann ein zweiter, spezifischer. Fertige Lösungswege werden grundsätzlich nicht geliefert. Dieses Vorgehen entspricht dem sokratischen Tutoring, wie es in der deutschsprachigen Didaktik etwa von Leisen (2007) als Gesprächsführungsform beschrieben wird, die durch gezieltes Fragen kognitive Aktivität provoziert, anstatt Wissen direkt zu übertragen – und wie es in der englischsprachigen Forschungstradition bei Collins und Stevens (1982) sowie Graesser et al. (1995) systematisch ausgearbeitet wurde. Dabei ist anzumerken, dass sokratisches Tutoring in seiner ursprünglichen Form eine teleologische Gesprächsführung voraussetzt – der Tutor kennt das Ziel und richtet seine Fragen strategisch darauf aus. Ein generatives Sprachmodell operiert anders: Es generiert Folgefragen auf Basis statistischer Muster, nicht auf Basis eines expliziten Gesprächsziels. Diese strukturelle Differenz begrenzt den Anspruch des Chatbots: Er kann sokratisches Tutoring *approximieren*, nicht vollständig realisieren. Lin et al. (2023) weisen in diesem Zusammenhang darauf hin, dass LLM-basierte Systeme im Bildungskontext besonders dann an Grenzen stoßen, wenn eine präzise inhaltliche Steuerung des Dialogs erforderlich ist – ein Befund, der die Notwendigkeit der pädagogischen Konfiguration und thematischen Einschränkung von Lern-Chatbots unterstreicht. Im Quiz-internen KI-Chat der *MatheArena*-Apps wird diese Erkenntnis bereits berücksichtigt (siehe Kap. 4).

Dialogbasierte Tutoringsysteme, die Lernende zur aktiven Auseinandersetzung mit dem Lösungsweg anregen, erzielen gegenüber reinen Feedbacksystemen signifikant höhere Lernzuwächse (VanLehn, 2011). Jeon (2023) ergänzt diesen Befund aus dem Bereich des sprachgestützten Lernens: Chatbot-gestützte dynamische Assessments, die Lernende durch adaptive Folgefragen zur Vertiefung anregen, können diagnostisch wertvolle Einblicke in den Lernstand liefern – ein Potenzial, das auch für den mathematischen Kontext relevant ist. Die Verbindung zur adaptiven Aufgabensteuerung ist dabei funktional: Scaffolding funktioniert nur dann, wenn die Aufgabe tatsächlich in der ZPD liegt. Der Elo-Algorithmus übernimmt diese Positionierung auf algorithmischer Ebene, während der Chatbot die qualitative Begleitung innerhalb dieser Zone gestaltet.

2.2.4 Cognitive Load Theory

Die Cognitive Load Theory (Sweller, 1988; Paas et al., 2010) geht davon aus, dass das menschliche Arbeitsgedächtnis in seiner Kapazität begrenzt ist. Sweller unterscheidet drei Formen kognitiver Belastung: die *intrinsische* Belastung durch die Komplexität des Lernstoffs, die *extrinsische* Belastung durch unzureichend gestaltete Aufgaben oder Erklärungen, und die *lernrelevante* Belastung durch tiefes Verarbeiten und Schemabildung – letztere soll aktiv gefördert werden. Paas et al. (2010) präzisieren in ihrer Weiterentwicklung der Theorie, dass diese drei Formen nicht unabhängig wirken, sondern in ihrer Summe die verfügbare Arbeitsgedächtniskapazität belasten – mit direkten Konsequenzen für das Instruktionsdesign.

Für den Chatbot ergeben sich daraus direkte Gestaltungskonsequenzen, die mit den beiden anderen Theorien verknüpft sind. Intrinsische Belastung hängt mit der Aufgabenschwierigkeit zusammen – hier setzt der Elo-Algorithmus an, indem er Überforderung systematisch reduziert. Extrinsische Belastung entsteht durch unklare oder das Individuum unverständliche Formulierungen oder zu viele Informationen auf einmal – der Chatbot minimiert sie durch knappe, schrittweise Hinweise. Lernrelevante Belastung wird gefördert, indem der Chatbot Lernende zur Verbalisierung von Lösungsschritten, zur Reflexion von Fehlern und zur Herstellung von Bezügen zu bekannten Konzepten auffordert – Strategien, die Hattie und Timperley (2007) als besonders wirksam auf der Prozess- und Selbstregulationsebene des Feedbacks ausweisen. Labadze et al. (2023) bestätigen in ihrer Literaturübersicht zu KI-Chatbots im Hochschulbereich, dass unmittelbares, adaptiv getimtes Feedback zu den zentralen Stärken chatbot-gestützter Lernumgebungen zählt und zugleich, dass eine zu hohe Informationsdichte in Chatbot-Antworten die kognitive Belastung kontraproduktiv steigern kann.

Besondere Relevanz hat in diesem Zusammenhang der bereits erwähnte „Expertise Reversal Effect“: Was für Lernende mit geringem Vorwissen kognitive Entlastung schafft, kann für fortgeschrittene Lernende zur unnötigen Zusatzbelastung werden. Der Chatbot muss daher nicht nur auf das aktuelle Aufgabenniveau, sondern auch auf den Grad der Vorwissensstrukturierung der Lernenden reagieren – ein Anspruch, der in der pädagogischen Konfiguration beider App-Varianten theoretisch berücksichtigt wird.

2.3 KI-Chatbots im Bildungsbereich

Mit der Weiterentwicklung von Bildungstechnologien fließen die von Gallagher et al. (2022) beschriebenen Feedback-Loops des adaptiven Lernens (Anreize für Lernende, Reflexion und Metakognition der Lehrkraft und Maßnahmen der Lehrkraft) zunehmend in die Gestaltung digitaler Lernwerkzeuge ein. KI-gestützte Systeme beginnen, Aspekte dieses Zyklus widerzuspiegeln, indem sie dynamisch auf Benutzereingaben reagieren, was sie im Kontext des Lehrens und Lernens besonders relevant macht. Awang et al. (2025) identifizierten zwanzig kommerzialisierte KI-basierte Mathematik-Lerntools („AI-based mathematical education tools“ – AIME), die KI integrieren, darunter „Khan Academy“, „MathE“, „Zapper“, „GeoGebra“, „Photomath“ oder „Squirrel AI“.

Gokcarslan et al. (2024) heben in ihrer systematischen Literaturrecherche die Vorteile des Einsatzes von Chatbots in der Bildung hervor. Sie stellten fest, dass die häufigsten Vorteile für Lernende „erhöhte Lernmotivation“, „Entwicklung von Sprachkenntnissen“, „verbesserte Lernleistung“ und „personalisierte adaptive Lernumgebungen“ sowie „Zeitersparnis“, „Reduzierung der Arbeitsbelastung“ und „Bereitstellung von Ressourcen“ für Lehrende sind. Labadze et al. (2023) betonen zudem, dass KI-Chatbots Pädagog*innen vor allem bei Routineaufgaben und personalisierter Unterstützung helfen können. Wenn der Chatbot jedoch falsche Informationen oder Anweisungen liefert, könnte er die Schüler*innen irreführen und ihren Lernfortschritt somit sogar behindern (Labadze et al., 2023).

Schon seit der Definition des „Turing-Tests“ im Jahr 1950 beschäftigen sich Forscher*innen damit, wie genau Chatbots menschliche Dialoge nachahmen können und ob sie uns wirklich davon überzeugen können, dass wir mit einer realen Person sprechen (Jeon, 2023). Dies erklärt, warum die rasante Entwicklung von Chatbots Bedenken hinsichtlich der Bildung, der Untergrabung der akademischen Integrität oder der Gefährdung der Lernergebnisse hervorgerufen hat (Martínez-Télez & Camacho-Zuñiga, 2023; Anonymous et al., 2024). Daher hat die Europäische Kommission Leitlinien herausgegeben, um Pädagog*innen dabei zu unterstützen, KI verantwortungsbewusst einzusetzen, wobei sowohl die potenziellen Vorteile als auch die damit verbundenen Risiken hervorgehoben werden (European Commission, Directorate-General for Education, Youth, Sport and Culture, 2022). Auch die UNESCO veröffentlichte sowohl für Schüler*innen (UNESCO, 2024a) als auch für Lehrer*innen (UNESCO, 2024b) KI-Kompetenzrahmen, um sicherzustellen, dass alle am Bildungswesen Beteiligten darauf vorbereitet sind, mit neuen Technologien auf informierte Weise umzugehen. Gokcarslan et al. (2024) stellten trotzdem zahlreiche Kehrseiten von KI-Chatbots im Bildungsbereich für Schüler*innen fest, darunter „Einschränkung der Handlungsmöglichkeiten“, „irreführende Antworten“, „Einschränkung des personalisierten Feedbacks“ und „Unfähigkeit, komplexe Ausdrücke zu verstehen“. Zu den Nachteilen für die Lehrkraft zählen „Originalität und Plagiate“, „Unfähigkeit, den Grad der Bereitschaft zu bestimmen“ und „Schwierigkeiten bei der Entwicklung KI-gestützter Anwendungen“.

3 Didaktisches Design des KI-Chatbots

Der KI-basierte Chatbot der *MatheArena* soll eine bedarfsgerechte, pädagogisch fundierte Lernunterstützung in einer mobilen Lernumgebung bieten, die die oben genannten Theorien miteinbezieht und bestehende Herausforderungen vorwegnimmt. Der Chatbot wurde entwickelt, um Herausforderungen zu bewältigen, die häufig beim Distance Learning oder beim selbstbestimmten Lernen in Mathematik auftreten, darunter mangelndes zeitnahes Feedback, nachlassende Motivation und fehlende Unterstützung durch Lehrpersonen. Durch die zusätzliche Einbettung eines Chatbots in jede einzelne Aufgabe fördert das System die Selbstständigkeit und Ausdauer der Lernenden und gewährleistet gleichzeitig eine hohe didaktische Qualität.

Auf Grundlage der Selbstbestimmungstheorie (Deci & Ryan, 1980) soll der Chatbot die grundlegenden psychologischen Bedürfnisse der Lernenden unterstützen: Autonomie (durch die Möglichkeit der Interaktion im eigenen Tempo), Kompetenz (durch die Bereitstellung von Unterstützung und Feedback) und Verbundenheit (durch die Simulation eines geführten Dialogs). Dabei orientiert sich der Chatbot außerdem an der „Cognitive Load Theory“, indem er prägnante und strukturierte Interaktionen bietet, sowie an Vygotskys „Zone of Proximal Development“, indem er die Unterstützung an den Schwierigkeitsgrad der Lernenden anpasst (Vygotsky & Cole, 1978).

Der Chatbot basiert auf dem sokratischen Unterrichtsmodell und folgt dabei strukturierten Dialogprinzipien, die aktives Lernen fördern sollen (Leisen, 2007). Um die Lernenden zu Lösungen zu führen, werden gezielte Fragen eingesetzt, anstatt direkt Antworten zu geben. Wenn Lernende Schwierigkeiten beim Verständnis haben, bietet das System schrittweise Hinweise, um zu verdeutlichen, wie das zugrunde liegende mathematische Konzept oder Verfahren funktioniert. Sollte auch das nicht zu einer Lösung führen, so werden Teillösungen angeboten, um den Problemlösungsprozess weiter zu unterstützen. Außerdem gibt der Chatbot durchgehend prägnante Antworten (begrenzt auf zwei oder drei Sätze), um die kognitive Belastung zu reduzieren und den Lernenden zu helfen, sich auf die Aufgabe zu konzentrieren. Die Sprache und der Tonfall des Chatbots sind bewusst unterstützend, forschungsbasiert und entwicklungsgerecht gestaltet, wobei mathematische Denkstrategien wie Deduktion, Induktion und logische Schlussfolgerungen angewandt werden.

4 Technische Umsetzung und Integration in *MatheArena*

Der *MatheArena* Chatbot basiert auf Gemini Flash, einem Large Language Model (LLM) von Google, das aufgrund seiner Kosteneffizienz, seiner Mehrsprachigkeit und seiner Eignung für den Einsatz im Bildungsbereich ausgewählt wurde. Um Halluzinationen zu reduzieren und die pädagogische Genauigkeit zu gewährleisten, arbeitet das LLM in einer eingeschränkten Generierungsumgebung, wodurch freie Antworten minimiert und inhaltlich abgestimmte Anleitun-

gen priorisiert werden. Darüber hinaus wird der Interaktionsverlauf in jeden neuen Austausch übernommen, um die Kohärenz und den Kontext für die Lernenden aufrechtzuerhalten.

Um die Sicherheit der Lernenden zu gewährleisten, wurden mehrere Schutzmechanismen implementiert. Erstens verwendet der Chatbot Algorithmen zur Themenüberwachung, die von Mathematik abweichende Eingaben erkennen können. Werden unerwünschte Themen eingegeben, reagiert das System mit einer sanften Umleitung und ermutigt die Lernenden, sich auf die jeweilige mathematische Aufgabe zu konzentrieren. Sollten die Lernenden trotz dieser Weisungen wiederholt versuchen, sich auf abweichende Themen einzulassen, wird die Chat-Sitzung automatisch beendet, um den Fokus auf den Lernprozess zu erhalten. Weiters sind umfassende Datenschutzmaßnahmen vorhanden: Die Daten der Lernenden werden über die API anonymisiert und alle Datenverarbeitungsverfahren entsprechen den DSGVO-Standards.

Der Chatbot ist direkt in die *MatheArena*-Lernoberfläche eingebettet und erscheint als bewegliche Sprechblase, die bei jeder Aufgabe eingeblendet wird. Die Lernenden können mit dem Chatbot entweder per Text oder Sprache interagieren, wobei speech-to-text und text-to-speech zur Wahrung der Barrierefreiheit zur Verfügung stehen. Zusätzlich zur aufgabenspezifischen Integration ist ein universeller Chatbot auch außerhalb der Lernobjekte verfügbar, um metakognitive Fragen, motivierende Unterstützung und zusätzliche Übungsaufgaben zu ermöglichen. So ist der Chatbot vollständig in den Lernprozess von *MatheArena* integriert. Aktuell wurde der Chatbot erfolgreich auf Deutsch, Englisch, Türkisch und Indonesisch getestet. Dabei passt er sich automatisch an die Sprache der Benutzer*innen an und stellt dabei regionale mathematische Notationen sowie die Ausrichtung auf den jeweiligen Lehrplan zur Verfügung.

5 Anwendungsszenarien

Aus der Analyse der Interaktionsprotokolle der *MatheArena*-App werden die Auswirkungen des KI-Chatbots in den verschiedenen Lernsituationen ersichtlich. Nachfolgend werden exemplarische Nutzungsmuster sowie die Chatbot-Funktion aus didaktischer Sicht erörtert.

5.1 Beispielhafte Lerninteraktion

In einer erfolgreichen, sokratischen Situation des Tutorings erkundigen sich Lernende nach mathematischen Konzepten oder Lösungen. Anstelle von direkten Antworten gibt der Chatbot Anleitungen, diese selbst zu finden, indem an Definitionen von Begriffen erinnert und lösbare Teilaufgaben gestellt werden. Im Lauf mehrerer Rückfragen baut der Bot die Lösungsfindung auf und knüpft daran die weitere Fragestellung an, bis das richtige Ergebnis gefunden wurde.

Dieser Ablauf entspricht dem sokratischen Tutoring, weil durch konkrete Fragen des Bots das selbständige Finden des Lösungswegs strukturiert wird. Damit wird die proximale Entwick-

lung in der Praxis realisiert: der Chatbot wirkt als temporäres Scaffolding und nimmt sich sukzessive bei erkennbarem eigenständigem Handeln zurück.

Bei falschen Vorstellungen verdeutlicht sich die Didaktik des Chatbots, indem Lösungen nicht direkt benannt werden, sondern durch spezifische Hinweise auf die Lösungsfindung ausgerichtet werden. Es gibt keine Form einer Bestrafung bei falschen und/oder fehlenden Antworten, stattdessen kehrt der Bot zum Ausgangspunkt zurück. Es wird so Druck vor dem ersten Hindernis genommen und der Schritt zu einer klar benannten Teilaufgabe gelenkt, die eigenständig von Lernenden bearbeitet werden kann. Dieser Ansatz entspricht der Cognitive Load Theory in Form einer Reduktion von extrinsischen Belastungen mittels klarer und schrittweiser Informationen, während zugleich intrinsische Belastungen aufgrund von Teilaufgaben aufrecht gehalten werden.

Einige Interaktionen können ohne eine Lösung enden, wenn die Lernenden wiederholt keine zielführende Antwort liefern. In diesen Fällen bleibt der Chatbot auf der sachlichen Ebene und hält den Dialog mit Wiederholungen bzw. einfachen Rückfragen aufrecht. Sollten darauf aber keine produktiven Reaktionen erfolgen (z. B. wenn Hinweise nicht aufgegriffen oder das Angebot einer helfenden Struktur nicht angenommen wird), erfolgt der Abbruch der Konversation. Hier endet der automatisierte Support, denn der Chatbot ersetzt keine Lehrkraft und deren diagnostischen Einfühlungsfähigkeiten, um in derartigen Situationen abwägen zu können, ob und welche alternativen Vorgehensweisen bzw. Methoden für den individuellen Fall zielführender wären. Das System hat hier nicht die Möglichkeit, emotionale Zustände von Lernenden zu erkennen – die soziale Eingebundenheit und Selbstbestimmungstheorie bleibt ihm vorbehalten.

5.2 Reaktion auf verschiedene Lernstände

Bei der Auswertung von Interaktionsdaten wird deutlich, wie der Chatbot seine Vorgehensweise an vorliegende Schwierigkeitsgrade aufbaut:

- einfache Aufgaben: Erklärungen, schrittweise Hilfe (wenig abstrakt, engmaschiges Scaffolding)
- mittlerer Schwierigkeitsgrad: Fehleranalyse durch den Bot, direktes Abfragen von Lösungen nach selbstständiger Vorarbeit
- fortgeschrittene Aufgaben: hier gibt es kaum Dokumentation über Interaktionen, aber es zeigt sich die Tendenz einer Anregung der Lernenden für metakognitives Nachdenken (mittels offener Fragen oder Reflexion von Strategien)

Dieser Aufbau je nach Aufgabenniveau mit der jeweiligen Form von Unterstützung des Chatbots zeigt dessen adaptive Grundausrichtung.

5.3 Verbindung mit Elo-Schwierigkeitsgrad

Die App verwendet einen Elo-basierten Algorithmus für eine dynamische Auswahl von Aufgaben. Die Schwierigkeitsgrade werden dabei laufend an den jeweiligen Kompetenzwert von Lernenden ausgerichtet. Diese laufende Kalibrierung zeigt sich in den Schwierigkeitsgraden (Beginner, Intermediate, Advanced): wenn Lernende eine Aufgabe mit Erfolg lösen können, ist die nächste Aufgabe tendenziell schwieriger und umgekehrt. Durch diesen algorithmischen Regelkreis werden Lernende immer mit dem für sie optimalen Niveau gefördert und herausgefordert.

Die automatische Anpassung wird durch den Chatbot mit einer qualitativen Ebene ergänzt, denn während durch den Elo-Algorithmus die Steuerung der Aufgaben erfolgt (quantitative Dimension), trägt der Chatbot zur prozessbegleitenden Unterstützung bei (qualitative Dimension), je nach vorliegendem Schwierigkeitsgrad. Bei einfacheren Aufgaben für ‚Beginner‘ tendiert der Bot zu stärkeren Hilfestellungen in Form von schrittweisen Tipps, Wiederholungen von Definitionen oder der engmaschigen Strukturierung des Lösungswegs, um die kognitive Belastung extrinsisch zu reduzieren. Bei mittleren und fortgeschrittenen Aufgaben wird diese Unterstützung verringert und dafür öfter zur selbstständigen Analyse aufgefordert, etwa anhand von offenen Fragen zum Anregen der metakognitiven Denkweise. Das System stellt so eine Verbindung der algorithmischen Adaptivität mit der didaktisch Gesprächsführung her, die sich abgestuft an dem individuellen Bedarf einer Hilfestellung von Lernenden ausrichtet.

6 Diskussion

Die beispielhaften Interaktionsdaten liefern die Basis zur Reflexion der Potenziale sowie Grenzen von KI-Chatbots im Mathematikunterricht.

6.1 Chancen für mathematisches Denken und Problemlösekompetenz

Die Analyse von Interaktionsprotokolle zeigt die Vielfalt an kognitiven Aktivitäten, die durch den Chatbot ausgelöst werden. Häufig gelingt es den Lernenden, die Lösung durch die gezielten Rückfragen und/oder abgestuften Hinweise eigenständig zu finden. Der größte Teil der Interaktionen trägt zur Klärung von mathematischen Grundbegriffen bei, indem der Chatbot wie ein Werkzeug für Begriffe genutzt werden kann, wodurch mathematische Grundvorstellungen solide weiterentwickelt werden können.

Die Problemlösekompetenz wird insbesondere in Dialogen gefördert, wenn Lernende durch den Chatbot angeregt werden, die Rechenschritte anzugeben und ihre Fehler zu reflektieren. Anstelle von fertigen Lösungen werden Fragen gestellt, um sich in kleinen Schritten mit der Aufgabe auseinanderzusetzen und die eigene Vorgehensweise kritisch zu reflektieren.

Dies ist nach der Lehr-Lernforschung eine besonders wirksame Strategie, weil der Fokus auf dem Prozess der Erkenntnisgewinnung und einem nachhaltigen Verständnis liegt.

Eine Studie von Li (2025) deutet darauf hin, dass Schüler*innen, die die Interaktion mit Chatbots als einfach und nützlich empfinden, eher Interesse am Lernprozess entwickeln, auch in einem Fach wie Mathematik, das häufig Unsicherheit und Ängste hervorruft. Lernende mit einer höheren Selbstwirksamkeit sind sicherer im Umgang mit KI-gestützten Tools, was wiederum ihre positive Bewertung der Nützlichkeit und Benutzerfreundlichkeit dieser Tools verstärkt (Li, 2025).

6.2 Rolle von KI als Tutor*in vs. Lehrer*in

Das Auswerten der Interaktionsdaten verdeutlicht, dass der Chatbot als Unterstützung (anstatt Ersatz) für Lehrkräfte dient, der praktisch als Werkzeug genutzt werden kann. Sie/Er kann durch Erklärungen von einfachen Definitionen, Wiederholungen von grundlegenden Verfahren und dem Bereitstellen von Übungen Lehrkräfte von diesen zeitintensiven Unterstützungstätigkeiten entlasten und dadurch Freiraum schaffen für pädagogische Kernaufgaben, die nur ein Mensch erfüllen kann.

Das soziale Einbinden von Lernenden ist weiterhin die Domäne von zwischenmenschlichen Interaktionen. In Chat-Protokollen sind kaum emotionale Aspekte vorhanden, da sie vor allem auf sachliche Inhalte und Aufgaben ausgerichtet sind. Das entspricht der systemimmanenten Funktion dieses Chatbots als Lerntool.

Gleichzeitig zeigt sich bei vereinzelt unangemessenen oder nicht auf eine Aufgabe bezogenen Äußerungen, dass persönliche Beziehungen durch Chatbots nicht ersetzt werden können/sollen. In diesen Fällen werden die Grenzen von automatisierten Systemen ersichtlich, die beim Umgang mit sozial-emotionalen Lernbereichen eine Nachbereitung durch Menschen (in Form eines pädagogischen Einwirkens und/oder begleitende Gespräche) unersetzlich machen. Der Chatbot ist zwar fähig, störende bzw. irritierende Interaktionen zu erkennen und unterbricht diese, aber bei den Lehrkräften verbleibt die pädagogische Verantwortung zur Einordnung und Bearbeitung dieser Verhaltensweisen.

6.3 Grenzen: Fehlinterpretationen, Abhängigkeit, Monitoring

Obwohl das System unter pädagogischen Gesichtspunkten konfiguriert wurde, hat eine automatisierte Unterstützung entsprechende Grenzen. So kann der Chatbot bei ungenauen bzw. mehrdeutigen Eingaben mit Sicherheitsmeldungen reagieren, doch ist er nicht fähig, die konkrete kognitive Hürde zu erkennen, aufgrund der eine vermeintlich zusammenhanglose Antwort gegeben wird. An diesem Punkt zeigt sich der Unterschied zu einem herkömmlichen Intelligent Tutoring System mit konkretem Konzept zu Schüler*innen. Der Chatbot besitzt keinen derartigen diagnostischen Ansatz einer Wissensstruktur, sondern kann lediglich auf Sprachausdruck reagieren.

Weiters ist es möglich, dass die Nutzung abgebrochen wird, wenn sich durch die Interaktion nicht unmittelbar eine Lösung ergibt. Dieses Verhalten deutet auf die instrumentelle Nutzung des Systems hin, bei dem nur ein schnelles Abrufen der Ergebnisse im Interesse liegt. Dies steht im Widerspruch zum eigentlichen Zweck eines selbstständigen Denkens und verletzt das Prinzip eines sokratischen Tutorings. Es ist eine Herausforderung, bei den Lernenden die Bereitschaft zu einem dialogischen Prozess zugewinnen. Außerdem hat die Analyse gezeigt, dass Chats teilweise inhaltlich nicht zur Aufgabe passen oder sogar nicht angemessene Äußerungen enthalten. Obwohl der Chatbot diese Vorfälle durch die integrierten Sicherheitsmechanismen erkennt/unterbricht, bleiben menschliche Nachbearbeitungen unverzichtbar, um Missbrauch zu vermeiden und Lernende angemessen bei vorliegenden Verständnisproblemen zu unterstützen.

6.4 Technologieakzeptanz (TAM)

Die ersten Hinweise bezüglich der Akzeptanz des Chatbots können aus Nutzungsdaten abgeleitet werden. Die meisten mit Erfolg abgeschlossenen Interaktionen lassen vermuten, dass das System hilfreich und nützlich ist, denn die Lernenden kommen wieder und führen eigenständig Aufgaben mit dem Bot aus. Die hohe Zahl von gelösten Aufgaben kann als positiv für den subjektiv empfundenen Nutzen interpretiert werden, bei dem Nutzungsbereitschaft gefördert wird.

Die mitunter plötzlichen Abbrüche der Interaktionen liefern dagegen Hinweise auf mögliche Schwierigkeiten bei der Bedienung. Es kann etwa sein, dass der Dialog nicht nach den Erwartungen von Lernenden verläuft, z. B. wenn Rückfragen umständlich sind, Hilfestellungen nicht rasch genug erfolgen. Diese Abbrüche deuten an, dass manche die Nutzerfreundlichkeit als gering wahrnehmen.

Diese Interpretationen basieren derzeit auf indirekten Ableitungen aus den Protokollen der Interaktionen. Für eine valide Erfassung der realen Akzeptanz des Chatbots müssen die Einflussfaktoren für die Nutzung und Zufriedenheit identifiziert werden und es sind konkrete Befragungen der Nutzer*innen notwendig. So kann die Wirkung, der Nutzen und die Benutzerfreundlichkeit des Chatbots aus der Sicht von Lernenden untersucht werden.

7 Fazit & Ausblick

Die Analyse der Interaktionsprotokolle liefert die Bestätigung der konzeptionellen Idee, wie ein pädagogisch gerahmter KI-Chatbot niedrigschwellig den Mathematikunterricht unterstützen kann. Die Daten zeigen den vielfach erfolgreichen Beitrag des Chatbots, wenn Lernende selbstständig Lösungen erarbeiten sollen. Das Tool reagiert auf die verschiedenen Lernvoraussetzungen und passt die Unterstützung auf diese an. Durch das Kombinieren mit den adaptiven Aufgabenalgorithmus und dessen laufende Adaptierung des Schwierigkeitsgrads auf den

jeweiligen Leistungsstand wird ein Gesamtsystem geboten, mit dem eine differenzierte und individuell ausgerichtete Förderung möglich wird.

Die Datenauswertung hat weitere spezifische Aspekte zur Optimierung aufgezeigt, etwa in Bezug auf das Erkennen von falschen Vorstellungen. Das System könnte eine bessere Fähigkeit aufweisen, um neben den falschen Aussagen auch die Wissenslücken dahinter genauer zu erfassen, damit konkret darauf eingegangen werden kann. Ein anderer Punkt wäre das meta-kognitive Gestalten von Dialogen. Bislang wurde vor allem auf die schrittweise Lösungsfindung abgezielt, doch könnten durch zusätzliche Reflexionsfragen nach einer abgeschlossenen Aufgabe Lernende dazu angeregt werden, das individuelle Vorgehen zu reflektieren und Lösungsstrategien zu bewerten. Des Weiteren hat sich gezeigt, dass bei einem sehr niedrigen Vorwissen vermehrt strukturierte Hilfestellungen notwendig wären. Dies könnte den Zugang durch den Ausbau eines gestuften Scaffoldings (z. B. mittels noch kleinerer Teilaufgaben oder ergänzender Veranschaulichungen) erleichtern.

Die soweit mögliche Analyse der Interaktionsdaten bietet einen wertvollen Einblick in die Nutzung des Chatbots in der Praxis und dessen charakteristische Merkmale. Notwendig wäre für zusätzliche Erkenntnisse eine kontrollierte Wirksamkeitsstudie; für den belastbaren Nachweis der Lernwirksamkeit des Systems ist eine randomisierte Feldstudie in Planung, bei der ein Vergleich von Lernenden mit/ohne Chatbot-Unterstützung unternommen wird. Ergänzend dazu erfolgen standardisierte Tests, die vor und auch nach der Interventionsphase durchgeführt werden. Zusätzlich werden qualitative Interviews mit Lehrkräften und Lernenden geführt, damit die subjektive Wahrnehmung der Unterstützung und die Akzeptanz der Technologie tiefergehend erfasst wird. Darüber hinaus werden die Lernpfade einer kombinierten Analyse unterzogen, bei der Veränderungen von algorithmisch geschätzten Kompetenzwerten inhaltlich mit den Interaktionen der Chats in Beziehung gesetzt werden. Diese Daten sollen die Weiterentwicklung des Chatbots ermöglichen, um die hier dargelegten Grenzen systematisch zu öffnen. So wird ein Beitrag geleistet, dass individualisiertes und selbstbestimmtes Mathematiklernen langfristig gesichert wird.

Anmerkungen

Dieses Projekt wurde durch Fördermittel der Österreichischen Forschungsförderungsgesellschaft FFG unterstützt.

Der vorliegende Artikel wurde mithilfe von KI-Tools verbessert: ChatGPT für die Ideenentwicklung und Verfeinerung, DeepL für Übersetzungen und Grammarly für Grammatik und Stil. Diese Tools trugen zur Klarheit und Kohärenz bei und stellten gleichzeitig die Einhaltung ethischer Standards und die Integrität der Autor*innen sicher.

Literatur

- Awang, L., Yusop, F., & Danaee, M. (2025). Current practices and future direction of Artificial Intelligence in mathematics education: A systematic review. *International Electronic Journal of Mathematics Education*. <https://doi.org/10.29333/iejme/16006>
- Benesch, T., Plank, V., & Infanger, E.-M. (2023). Einsatz von MatheArena Junior in der Sekundarstufe 1: Fallbeispiel 1. Klasse. *Wissenschaft, Wissenschaftlichkeit und Öffentlichkeit*, (11), 52–55.
- Cheng, X., Zhang, X., Cohen, J., & Mou, J. (2022). Human vs. AI: Understanding the impact of anthropomorphism on consumer response to chatbots from the perspective of trust and relationship norms. *Information Processing & Management*, 59(3):102940. <https://doi.org/10.1016/j.ipm.2022.102940>
- Corbett, A. T., & Anderson, J. R. (1994). Knowledge tracing: Modeling the acquisition of procedural knowledge. *User Modeling and User-Adapted Interaction*, 4(4), 253–278. <https://doi.org/10.1007/BF01099821>
- Deci, E. L., & Ryan, R. M. (1980). Self-determination theory: When mind mediates behavior. *The Journal of Mind and Behavior*, 1(1), 33–43.
- European Commission, Directorate-General for Education, Youth, Sport and Culture (2022). *Ethical guidelines on the use of Artificial Intelligence (AI) and data in teaching and learning for educators*. <https://data.europa.eu/doi/10.2766/153756>
- Gallagher, M. A., Parsons, S. A., & Vaughn, M. (2022). Adaptive teaching in mathematics: A review of the literature. *Educational Review*, 74(2):298–320. <https://doi.org/10.1080/00131911.2020.1722065>
- Gokcearslan, S., Tosun, C., & Erdemir, Z. G. (2024). Benefits, challenges, and methods of artificial intelligence (AI) chatbots in education: A systematic literature review. *International Journal of Technology in Education*, 7. <https://doi.org/10.46328/ijte.600>
- Graesser, A. C., Person, N. K., & Magliano, J. P. (1995). Collaborative dialogue patterns in naturalistic one-to-one tutoring. *Applied Cognitive Psychology*, 9(6), 495–522.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112.
- Infanger, E.-M., Infanger, G., Lavicza, Z., & Sobieczky, F. (2022). Applying Time-Inhomogeneous Markov Chains to Math Performance Rating. In G. Kotsis, A. M. Tjoa, I. Khalil, B. Moser, A. Taudes, A. Mashkoor, J. Sametinger, J. Martinez-Gil, F. Sobieczky, L. Fischer, R. Ramler, M. Khan, & G. Czech (Hrsg.), *Database and Expert Systems Applications—DEXA 2022 Workshops* (Bd. 1633, S. 11–21). Springer International Publishing. https://doi.org/10.1007/978-3-031-14343-4_2
- Jeon, J. (2023). Chatbot-assisted dynamic assessment (CA-DA) for L2 vocabulary learning and diagnosis. *Computer Assisted Language Learning*, 36(7), 1338–1364. <https://doi.org/10.1080/09588221.2021.1987272>
- Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist*, 38(1), 23–31.
- Labadze, L., Grigolia, M., & Machaidze, L. (2023). Role of AI chatbots in education: Systematic literature review. *International Journal of Educational Technology in Higher Education*, 20, 56. <https://doi.org/10.1186/s41239-023-00426-1>
- Leisen, J. (2007). Unterrichtsgespräch: Fragend-entwickelnder Unterricht, sokratischer Dialog und Schülergespräche. In *Physik Methodik für die Sekundarstufen* (pp. 115–132). Cornelsen Verlag Scriptor.
- Li, K. (2025). Factors influencing learners' learning interest when using AI chatbots-assisted math learning in higher education. *Frontiers in Psychology*, 16, 1716913. <https://doi.org/10.3389/fpsyg.2025.1716913>

- Lin, C. C., Huang, A. Y., & Lu, O. H. (2023). Artificial intelligence in intelligent tutoring systems toward sustainable education: A systematic review. *Smart Learning Environments*, 10(1), 41.
- Ma, W., Adesope, O. O., Nesbit, J. C., & Liu, Q. (2014). Intelligent tutoring systems and learning outcomes: A meta-analysis. *Journal of Educational Psychology*, 106(4), 901–918.
<https://doi.org/10.1037/a0037123>
- Martínez-Téllez, R., & Camacho-Zuñiga, C. (2023). *Enhancing mathematics education through AI Chatbots in a Flipped Learning Environment*. 1–8. <https://doi.org/10.1109/weef-gedc59520.2023.10343838>
- Moral-Sánchez, S. N., Ruiz Rey, F. J., & Cebrián-de-la-Serna, M. (2023). Analysis of artificial intelligence chatbots and satisfaction for learning in mathematics education. *IJERI: International Journal of Educational Research and Innovation*, (20), 1–14.
<https://doi.org/10.46661/ijeri.8196>
- OECD. (2026). *Student performance (PISA)*. <https://www.oecd.org/en/topics/student-performance-pisa.html>
- Paas, F., van Gog, T., & Sweller, J. (2010). Cognitive load theory: New conceptualizations, specifications, and integrated research perspectives. *Educational Psychology Review*, 22(2), 115–121. <https://doi.org/10.1007/s10648-010-9133-8>
- Pelánek, R. (2016). Applications of the Elo rating system in adaptive educational systems. *Computers & Education*, 98, 169–179. <https://doi.org/10.1016/j.compedu.2016.03.017>
- Saferinternet.at. (2024). *Studie: Werkzeug, Ratgeber, Bezugsperson – 94 Prozent der Jugendlichen nutzen KI-Chatbots*. <https://www.saferinternet.at/presse-detail/studie-werkzeug-ratgeber-bezugsperson-94-prozent-der-jugendlichen-nutzen-ki-chatbots>
- Schneider, W., & Artelt, C. (2010). Metacognition and mathematics education. *ZDM Mathematics Education*, 42(2), 149–161.
- Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review*, 31(2), 261–292.
- UNESCO. (2024a). *AI competency framework for students*. <https://doi.org/10.54675/JKJB9835>
- UNESCO. (2024b). *AI competency framework for teachers*. <https://www.unesco.org/en/articles/ai-competency-framework-teachers>
- VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, 46(4), 197–221.
- Vygotsky, L. S., & Cole, M. (1978). *Mind in society: Development of higher psychological processes*. Harvard University Press.