Diyalog 2025/ 2: 711-729 (Research Article)

ChatGPT im Deutschunterricht: Eine experimentelle Studie zur Wirksamkeit KI-gestützter Lehrmethoden

Aygül Şahin Toptaş, Ankara

https://doi.org/10.37583/diyalog.1827931

Abstract (Deutsch)

Im Zeitalter des rasanten digitalen Wandels wird Künstliche Intelligenz (KI) zunehmend in Bildungskontexte integriert und eröffnet neue Wege für personalisierten, motivierenden Sprachunterricht. Die vorliegende experimentelle Studie prüft die Wirksamkeit von ChatGPT-gestützten Lehraktivitäten im Deutschunterricht der Sekundarstufe II. Im Frühjahrssemester 2024/25 nahmen an einer öffentlichen Schule in der Türkei 50 Schüler:innen der elften Klasse teil, die mithilfe einer Randomisierung gleichmäßig (n = 25 / 25) auf Experimental- und Kontrollgruppe verteilt wurden. Beide Gruppen absolvierten vier Wochen lang je zwei Unterrichtsstunden zur A1.2-Einheit „Im Einkaufszentrum“. Während die Kontrollgruppe konventionellen, lehrplanorientierten Unterricht erhielt, arbeitete die Experimentalgruppe mit wöchentlich von ChatGPT generierten Dialogen, Wortschatzaufgaben und Schreibaktivitäten. Ein validierter 20-Item-Leistungstest (Cronbach-α = .84) wurde als Prä- und Posttest eingesetzt. Abhängige t-Tests zeigten in der Experimentalgruppe einen hoch signifikanten Leistungszuwachs (M = 12,40 → 16,85; p < .001), der den moderaten Anstieg der Kontrollgruppe (M = 12,10 → 13,45; p = .022) deutlich übertraf. Ein unabhängiger t-Test der Posttests bestätigte den Vorteil der KI-gestützten Instruktion (p < .001). Die Ergebnisse belegen, dass ChatGPT das Lernengagement steigert, individualisierte Rückmeldungen ermöglicht und Wortschatz-, Grammatik- sowie Produktionsleistungen signifikant verbessert. Zugleich unterstreichen sie die zentrale Rolle der Lehrkraft bei der didaktischen Anpassung KI-generierter Materialien an konkrete Unterrichtsbedürfnisse.

Schlüsselwörter: *ChatGPT, künstliche Intelligenz, Deutschunterricht, experimentelle Forschung, digitales Lernen.*

Abstract (English)

ChatGPT in German Language Teaching: An Experimental Study on the Effectiveness of AI-Supported Instructional Methods

In the era of rapid digital transformation, artificial intelligence (AI) is being increasingly integrated into educational contexts, opening up new paths for personalized and motivating language teaching. The present *experimental* study examines the effectiveness of ChatGPT-supported teaching activities in upper-secondary German classes. During the spring semester of 2024/ 25, 50 eleventh-grade students at a public school in Türkiye participated and were randomly assigned in equal numbers (n = 25 / 25) to an experimental group and a control group. Over four weeks both groups completed two lessons per week on the A1.2 unit *“Im Einkaufszentrum.”* While the control group received conventional, curriculum-based instruction, the experimental group worked with weekly dialogues, vocabulary tasks and writing activities

Einsendedatum: 22.08.2025 Freigabe zur Veröffentlichung: 15.12.2025

generated by ChatGPT. A validated 20-item achievement test (Cronbach's $\alpha = .84$) was administered as a pre- and post-test. Paired-samples *t*-tests revealed a highly significant gain in the experimental group ($M = 12.40 \rightarrow 16.85$; $p < .001$) that clearly exceeded the moderate increase in the control group ($M = 12.10 \rightarrow 13.45$; $p = .022$). An independent-samples *t*-test of the post-test scores confirmed the advantage of AI-supported instruction ($p < .001$). These results demonstrate that ChatGPT enhances learner engagement, enables individualized feedback and significantly improves vocabulary, grammar and productive language performance. At the same time, they underscore the pivotal role of teachers in adapting AI-generated material to concrete classroom needs.

Keywords: *ChatGPT, Artificial Intelligence, German Language Teaching, Experimental Research, Digital Learning.*

EXTENDED ABSTRACT

The rapid digital transformation of education has created new opportunities for innovation in foreign language teaching. Among the emerging technologies, artificial intelligence (AI) has become particularly prominent, offering tools that can reshape teaching and learning processes in fundamental ways. One of the most notable innovations in this field is ChatGPT: a large language model that capable of generating natural language, producing interactive dialogues, and creating learning materials tailored to specific teaching contexts. Despite the growing focus on AI in education, there has been a lack of experimental studies examining its impact on secondary school German language teaching. This study addresses this gap by investigating the effectiveness of ChatGPT-based activities in enhancing student performance in German classes.

The main objective of the study was to evaluate whether the integration of ChatGPT-supported materials into German language lessons at the secondary school level leads to higher learning gains compared to conventional methods. The study focused on practical classroom applications, such as generating dialogues, writing activities, and interactive tasks, which could be implemented directly in teaching. The central research question was whether using ChatGPT for lesson planning and material preparation would significantly improve students' achievement in German as a foreign language.

The research followed a quasi-experimental design. A total of 50 eleventh-grade students from a public secondary school participated in the study. They were randomly assigned to one of two groups: an experimental group of 25 students and a control group of 25 students. The intervention was implemented over a four-week period during the spring semester of the 2024/2025 academic year. Both groups were taught by the same teacher to ensure consistency. While the control group received lessons based on conventional methods, the experimental group used materials generated with ChatGPT. These materials included short dialogues for role-playing, sentence exercises to practice vocabulary and grammar, buyer–seller conversations to develop communicative competence, and short writing prompts to strengthen productive skills. The materials were adapted to the curriculum unit "Im Einkaufszentrum" and aligned with the learning objectives set for each week.

Student performance was measured through pre- and post-tests administered to both groups. The pre-test took place before the beginning of the intervention, and the post-test was conducted after the four-week programme. Statistical analyses were performed using parametric tests, as the data met the assumptions of normal distribution. Paired-sample t-tests were used to examine within-group differences between pre- and post-test scores, and independent-sample t-tests were applied to compare post-test results between the two groups. In addition, effect sizes were calculated to provide further evidence of the magnitude of the observed differences.

The results revealed clear improvements in both groups, but the learning gains were much stronger in the experimental group. Students who participated in ChatGPT-supported lessons demonstrated a significant improvement in their test scores compared to their performance before the intervention. In contrast, the control group also improved, but to a smaller degree. The comparison of the post-test scores confirmed a statistically significant difference in favour of the experimental group, indicating that the use of ChatGPT-based materials had a stronger impact on learning outcomes. These results suggest that integrating AI-supported activities can enhance student performance in German as a foreign language beyond that is achieved with traditional methods alone.

The study also highlighted the importance of the teacher's role in adapting and implementing AI-generated materials. While ChatGPT was able to produce diverse and engaging examples, the responsibility for selecting, modifying, and integrating these materials into classroom instruction remained with the teacher. This ensured that the activities were pedagogically appropriate and aligned with the curriculum goals. Therefore, the findings suggest that the effective use of AI in education depends not only on the technology itself but also on how it is embedded within teaching practice.

However, several limitations must be considered when interpreting the results. The study was conducted with a relatively small sample of 50 students from a single school, which restricts the

generalizability of the findings. Additionally, the four-week intervention period was too short to draw conclusions about long-term effects. Finally, the study relied solely on quantitative data, which provided valuable insights into performance outcomes, but did not capture students' subjective experiences or perceptions. Future research should include larger and more diverse samples, longer interventions, and qualitative data to better understand how students perceive and engage with AI-based learning tools.

In conclusion, this study provides evidence that ChatGPT-supported materials can significantly improve learning outcomes in German as a foreign language at the secondary school level. By enhancing test performance, stimulating student motivation, and supporting active classroom participation, AI can serve as a valuable tool to complement traditional teaching. The findings underline the potential of integrating innovative technologies into regular school curricula and highlight the importance of developing pedagogical strategies that make effective use of AI in education. This research contributes to the growing discussion on the role of AI in language teaching, offering practical guidance to teachers and policymakers who are looking to implement digital transformation in education.

Einleitung

Der Fremdsprachenunterricht ist ein vielschichtiger Prozess, der nicht nur auf den Erwerb sprachlicher Kompetenzen abzielt, sondern auch darauf, diese Kompetenzen im soziokulturellen Kontext zu verankern. Sprache fungiert dabei als Träger kultureller Werte und Normen und stellt ein zentrales Mittel zur Etablierung effektiver interkultureller Kommunikation dar. In diesem Zusammenhang bieten digitale Technologien und Anwendungen Künstlicher Intelligenz (KI) nicht nur Werkzeuge zur Unterstützung von Lehrprozessen, sondern auch Chancen zur Förderung kultureller Sensibilität (Topalovic u. a. 2024). In den letzten Jahren haben rasante Fortschritte im Bereich der KI-Technologien tiefgreifende Veränderungen in Bildungssystemen bewirkt. Künstliche Intelligenz wird im Allgemeinen als Computersystem definiert, das in der Lage ist, menschliches Denken, Lernen, Problemlösen und Entscheidungsverhalten zu simulieren (Russell / Norvig 2016). Der zunehmende Einfluss digitaler Technologien im Bildungsbereich verändert die Qualität der Lehrprozesse grundlegend; KI-Anwendungen führen zu einem multidimensionalen Wandel – von den Lehrstrategien der Lehrkräfte bis hin zu den individuellen Lernerfahrungen der Schüler:innen (Kozma 2003; Luckin u. a. 2016; Holmes u. a. 2019). Dabei darf nicht außer Acht gelassen werden, dass die digitale Transformation des Fremdsprachenunterrichts bereits zuvor durch Innovationen wie Web 2.0- und Web 3.0-Werkzeuge geprägt wurde, die neue Formen der Interaktion und Kollaboration eröffnet haben (Güngenci 2024). KI-Anwendungen wie ChatGPT knüpfen an diese Entwicklungen an und erweitern sie durch adaptive und generative Funktionen. Insbesondere in kommunikationsorientierten Fächern wie dem Fremdsprachenunterricht tragen KI-basierte Werkzeuge dazu bei, die sprachliche Exposition der Lernenden in der Zielsprache zu erhöhen, sofortiges Feedback zu liefern, Lernprozesse zu individualisieren und die Interaktion dynamischer zu gestalten (Chen u. a. 2020). In diesem Sinne ist KI nicht nur ein Werkzeug zur Erleichterung des Lernens, sondern fungiert zunehmend als aktiver Lernpartner, der es den Lernenden ermöglicht, sich stärker in ihre eigenen Lernprozesse einzubringen (Luckin u. a. 2016; Holmes u. a. 2019).

Im Kontext des Fremdsprachenunterrichts ermöglichen KI-Anwendungen den Lernenden den Zugang zur Zielsprache in schnellerer, vielfältigerer und interaktiverer Weise, indem sie authentische, alltagsnahe Sprachverwendungen bereitstellen. Intelligente Sprachlern-Apps, Chatbots und automatische Feedbacksysteme spielen eine wirksame Rolle bei der Entwicklung grundlegender sprachlicher Kompetenzen wie Aussprache, Grammatik, Schreiben und Wortschatz (Sun 2023; Chen u. a. 2023; Du / Daniel 2024; Mohammed / Khalid 2025). In diesem Zusammenhang betonen Şahin Toptaş (2024) das Potenzial der anpassbaren Struktur von Chatbots, sich an die individuellen Bedürfnisse und Sprachlevels der Lernenden anzupassen, und heben hervor, dass solche Technologien im Deutschunterricht als effektive Hilfsmittel eingesetzt werden können. Ebenso bieten KI-basierte Lernplattformen wie Duolingo und Babbel personalisierte Inhalte, die auf die individuellen Bedürfnisse der Lernenden zugeschnitten sind, und tragen so zu einem differenzierten Lernprozess bei (Godwin-Jones 2021). Durch Technologien wie automatische Spracherkennung, maschinelle Übersetzung, Chatbots und Content-Generatoren wird die Exposition gegenüber der

Zielsprache erhöht, wodurch sich produktive Sprachfähigkeiten gezielter weiterentwickeln lassen (Godwin-Jones 2022; Zawacki-Richter u. a. 2019). Ein besonders vielversprechendes Beispiel für KI-gestützte Sprachmodelle stellt ChatGPT dar, das durch Funktionen wie natürliche Sprachproduktion, Textvervollständigung, Szenarienentwicklung und die Generierung interaktiver Inhalte ein bemerkenswertes Potenzial für Unterrichtsumgebungen bietet (Kasneci u. a. 2023). Solche Anwendungen, die insbesondere Unterstützung bei der Gestaltung aufgabenorientierter Lernaktivitäten, Wortschatzarbeit, Leseverständnis und kreativem Schreiben bieten, fördern die aktive Teilnahme der Lernenden am Lernprozess und ermöglichen eine stärker personalisierte, lernerzentrierte Lernerfahrung, die über konventionelle Methoden hinausgeht (Smutny / Schreiberova 2020; Zawacki-Richter u. a. 2019).

Trotz der pädagogischen Potenziale, die KI-gestützte Anwendungen bieten, ist die Anzahl datenbasierter Studien zum Einsatz dieser Technologien im Fremdsprachenunterricht nach wie vor begrenzt. Besonders im Kontext des Deutschunterrichts nimmt die Zahl der Studien zur Integration großer Sprachmodelle wie ChatGPT erst allmählich zu. So untersuchte Kumari (2024) den Einfluss von ChatGPT auf akademische Schreibkompetenzen im Fach Deutsch und stellte fest, dass dieses Werkzeug die Motivation der Lernenden steigern könne, der Schreibprozess jedoch durch Faktoren wie „begrenzte Erfahrung“ und „individuellen Schreibstil“ erschwert werden könne. Michael u. a. (2025) analysierten die Wirkung generativer KI auf kollaborative Schreibprozesse und zeigten, dass ChatGPT signifikante Veränderungen in der Textqualität, den Überarbeitungsverhalten und der Interaktion zwischen Lernenden bewirkt habe. Die Studie hebt hervor, dass solche Werkzeuge bestehende Modelle des kollaborativen Schreibens um neue Dimensionen erweitern. Studien, die sich speziell mit dem Einsatz von KI im Deutschunterricht beschäftigen, betonen insbesondere die positiven Beiträge solcher Werkzeuge im Bereich des Leseverstehens, des Schreibens und der Materialentwicklung. Akol Göktaş (2024) unterstreicht, dass ChatGPT in der Lage ist, Texte gemäß dem Gemeinsamen Europäischen Referenzrahmen (GER) zu erstellen, vielfältige Übungsformate zu entwickeln, unmittelbares Feedback zu geben und Inhalte zu personalisieren. Gleichzeitig betont er die zentrale Rolle der Lehrkraft bei der effektiven Nutzung dieser Technologie im Unterricht. In ähnlicher Weise hebt Öztürk (2024) hervor, dass ChatGPT ein effizientes Instrument zur Erstellung von Lehrmaterialien darstellt, jedoch die generierten Inhalte vor der Verwendung didaktisch gefiltert und angepasst werden müssten. Şahin Toptaş (2023) sowie Tuğlu (2024) haben die bestehenden und potenziellen Einsatzmöglichkeiten von KI-Anwendungen im Deutschunterricht sowohl aus theoretischer als auch aus praxisorientierter Perspektive untersucht. Diese Studien, die sich auf das Potenzial von ChatGPT zur Förderung der Schreib- und Lesekompetenz im Fach Deutsch konzentrieren, zeigen, dass solche Werkzeuge sowohl lernendenzentrierte als auch aufgabenbasierte Lehrstrategien wirksam unterstützen können. Im Kontext der Türkei ist insgesamt ein zunehmender Trend zum KI-gestützten Fremdsprachenunterricht erkennbar. In diesem Zusammenhang zeigt die systematische Übersicht von Köksal und Yılmaz (2024), dass die Mehrheit der in der Türkei durchgeführten Studien zur KI im Sprachunterricht qualitativ angelegt ist

und sich insbesondere auf Werkzeuge der natürlichen Sprachverarbeitung wie ChatGPT konzentriert. Şentürk (2024) hebt hervor, dass KI-Tools den Lehrkräften Zeitersparnis ermöglichen und dadurch eine stärkere Fokussierung auf schülerorientierte Lehr-Lern-Prozesse zulassen, betont jedoch zugleich die Notwendigkeit, die KI-Kompetenz von Lehrkräften weiter zu stärken. Ähnliche Ergebnisse zeigen weitere Studien aus der Türkei (Öztürk 2024; Şahin Toptaş 2023; Tuğlu 2024; Akol Göktaş 2024), die hervorheben, dass KI-generierte Inhalte im Unterricht nur dann effektiv eingesetzt werden können, wenn sie durch die pädagogische Anleitung der Lehrkraft ergänzt und gesteuert werden. Vor diesem Hintergrund besteht weiterhin ein deutlicher Bedarf an empirischen, insbesondere experimentellen Untersuchungen zum Einsatz von KI-Anwendungen im Deutschunterricht.

Diese Übersicht an Beiträgen verdeutlicht insbesondere im Hinblick auf den Deutschunterricht, dass es an experimentellen Studien mangelt, die sich auf den Einsatz von KI-Anwendungen im Klassenzimmer konzentrieren. Zudem wird die zentrale Rolle der Lehrkraft bei der pädagogischen Steuerung und Umsetzung solcher Technologien hervorgehoben. Gleichzeitig hat die Integration KI-basierter Technologien in das Bildungssystem in der Türkei zunehmend strategische Bedeutung erlangt. Künstliche Intelligenz (KI) gilt weltweit als transformative Technologie mit Einflussbereichen, die von Bildung über Gesundheit und Finanzen bis hin zur Industrie reichen (vgl. OECD 2019). Im Bildungswesen tragen KI-gestützte Lernsysteme durch Datenanalytik, automatisierte Feedback-Mechanismen und personalisierte Lernangebote dazu bei, konventionelle Unterrichtsansätze grundlegend zu verändern (Pedro u. a. 2019). Auch in der Türkei nimmt die Bedeutung der KI-Integration im Bildungssystem stetig zu, im Einklang mit globalen Entwicklungen. Das türkische Bildungsministerium (BM 2024) führt verschiedene Projekte und berufliche Weiterbildungsprogramme durch, um die KI-Kompetenz von Lehrkräften zu stärken und die technologiebezogenen Lernfähigkeiten der Schüler:innen zu fördern. In diesem Rahmen wurden unter anderem praxisnahe Materialien wie das Handbuch „Eğitimde Yapay Zekâ Araçları: Öğretmen El Kitabı“ (*KI-Werkzeuge in der Bildung: Handbuch für Lehrkräfte*, BM 2024) veröffentlicht, das Lehrkräften die effektive Nutzung von KI-Tools im Unterricht erleichtern soll.

Parallel zu diesem Transformationsprozess in der Türkei zielt die vorliegende Studie darauf ab, die Wirksamkeit KI-gestützten Unterrichts im Fremdsprachenunterricht anhand eines experimentellen Ansatzes zu untersuchen. Vor diesem Hintergrund wurde die Studie sowohl im Einklang mit den Digitalisierungsstrategien der Türkei als auch zur Schließung einer identifizierten Forschungslücke im Bereich experimenteller Untersuchungen konzipiert. Im Mittelpunkt steht dabei das Potenzial des Sprachmodells ChatGPT im Fremdsprachenunterricht. Ziel der Studie ist es, den Einsatz von ChatGPT, einem KI-gestützten Sprachmodell, im Fremdsprachenunterricht – insbesondere im Deutschunterricht auf Sekundarstufenniveau – zu analysieren und dessen Einfluss auf die schulischen Leistungen der Lernenden zu bewerten. Die Relevanz dieser Arbeit ergibt sich aus der zunehmend zentralen Rolle von KI im Bildungsbereich sowie der Notwendigkeit, das pädagogische Potenzial solcher Technologien systematisch zu untersuchen. Obwohl ChatGPT als innovatives Werkzeug im Bildungsbereich große

Aufmerksamkeit erregt hat, gibt es bisher nur wenige experimentelle und datengestützte Studien zu seinem Einsatz im regulären Unterricht. Dies gilt insbesondere für Zielsprachen wie Deutsch, in denen die Zahl praxisorientierter Studien noch gering ist – ein Umstand, der dieser Untersuchung zusätzlichen wissenschaftlichen Wert verleiht. Ziel dieser Studie ist es daher, die Effekte KI-gestützter Lehrmethoden auf den Fremdsprachenunterricht mittels eines experimentellen Designs zu erfassen und ein evidenzbasiertes Rahmenmodell für den Einsatz innovativer Technologien im Sprachunterricht zu entwickeln. Zu diesem Zweck wurde ein experimentelles Design gewählt, in dessen Rahmen ChatGPT-basierte Lehraktivitäten entwickelt und über einen Zeitraum von vier Wochen im regulären Deutschunterricht auf Sekundarstufenniveau umgesetzt wurden. Während die Experimentalgruppe wöchentliche Aktivitäten erhielt, die mithilfe von ChatGPT erstellt wurden, wurde die Kontrollgruppe konventionell unterrichtet. Zur Bewertung der Wirksamkeit der Intervention wurden in beiden Gruppen ein Prätest und ein Posttest durchgeführt. Die zentrale Forschungsfrage lautet dementsprechend:

- Führen KI-gestützte Lehraktivitäten mit ChatGPT im Deutschunterricht der Sekundarstufe zu einem signifikant höheren Lernerfolg im Vergleich zum konventionellen Unterricht?

Methode

Forschungsdesign

Diese Studie wurde im Rahmen eines quantitativen Forschungsansatzes mit einem experimentellen Design durchgeführt, das aus Prä- und Posttests mit Kontrollgruppe besteht. Das experimentelle Design ermöglicht eine randomisierte Zuweisung von Schüler:innen zu den Gruppen, wodurch potenzielle Verzerrungen minimiert werden (Fraenkel u. a. 2012). Im Rahmen der Untersuchung wurden zwei Gruppen gebildet: eine Experimentalgruppe und eine Kontrollgruppe. Der Experimentalgruppe wurde ein KI-gestütztes Lehrprogramm angeboten, das aus ChatGPT-basierten Inhalten und Aktivitäten bestand, während der Unterricht in der Kontrollgruppe nach konventionellen Lehrmethoden durchgeführt wurde. Zu Beginn und am Ende der Intervention wurde in beiden Gruppen ein Deutsch-Leistungstest durchgeführt, um die Veränderungen in der akademischen Leistung der Schüler:innen zu messen. Ziel dieser Untersuchung ist es, die Wirkung eines durch Künstliche Intelligenz – konkret ChatGPT – unterstützten Lehransatzes auf den Lernerfolg empirisch nachzuweisen. Solche experimentellen Studien sind im Bildungsbereich weit verbreitet, da sie Rückschlüsse auf Kausalzusammenhänge ermöglichen (Creswell / Creswell 2018). Darüber hinaus erlaubt dieses Design die Erprobung von Unterrichtsinterventionen unter realen schulischen Bedingungen, was die Anwendungsvalidität der Ergebnisse erhöht. Die Studie wurde von der Ethikkommission der [Institution] genehmigt (Protokollnummer: 2025/676, Datum: 15.04.2025); die Teilnahme der Schüler:innen erfolgte freiwillig und basierte auf einem schriftlichen Einwilligungsformular, das vor Beginn der Untersuchung eingeholt wurde.

Teilnehmer:innen

Die Studie wurde im Frühjahrssemester des Schuljahres 2024/2025 mit Schüler:innen der 11. Klasse durchgeführt, die am Ibni Sina Anadolu Lisesi in Sincan (Ankara) am Deutschunterricht teilnahmen. Insgesamt gab es elf 11. Klassen, aus denen zufällig zwei Klassen für die Teilnahme an der Studie ausgewählt wurden. Die Zuweisung zu den Gruppen erfolgte zufällig: Jede Gruppe (Experimental- und Kontrollgruppe) bestand aus 25 Schüler:innen. Diese Zahl wurde gewählt, um eine gleichmäßige Verteilung der Teilnehmer:innen auf beide Gruppen zu gewährleisten. Beide Gruppen wurden von derselben Lehrkraft auf Grundlage derselben Unterrichtseinheit unterrichtet, jedoch in verschiedenen Zeiträumen, um äußere Variablen zu minimieren und die Vergleichbarkeit der Gruppen zu gewährleisten. Die Experimentalgruppe wurde mit einem ChatGPT-gestützten Unterrichtsprogramm unterrichtet, während die Kontrollgruppe im Rahmen des regulären Curriculums mit konventionellen Methoden unterrichtet wurde. Auf diese Weise sollte sichergestellt werden, dass externe Variablen, die den Lernprozess beeinflussen könnten, minimiert werden.

Datenerhebungsinstrumente

Zur Erhebung quantitativer Daten wurde in dieser Studie ein von *der Forscher*in entwickelter Deutsch-Leistungstest eingesetzt. Der Test basierte auf der dritten Einheit mit dem Titel *„Im Einkaufszentrum"* des Lehrwerks *Deutsch macht Spaß – A1.2*, das vom BM (Bildungsministerium der Republik Türkiye) offiziell an Schulen verwendet wird. Ziel des Tests war es, den Wortschatz und die Grammatikkenntnisse der Schüler:innen zu dieser spezifischen Unterrichtseinheit zu überprüfen. Der entwickelte Test bestand aus insgesamt 20 Aufgaben: 10 Fragen zum Wortschatz, 7 Fragen zur Grammatik und 3 Aufgaben zur Bildung korrekter Dialoge. Um die Inhaltsvalidität sicherzustellen, wurde der Test von zwei Fachdidaktiker:innen für Deutsch als Fremdsprache begutachtet und anhand ihrer Rückmeldungen entsprechend überarbeitet. Zur Überprüfung der Durchführbarkeit und der Trennschärfe der Items wurde eine Pilotstudie mit einer Vergleichsgruppe durchgeführt, die nicht Teil der eigentlichen Untersuchung war; dabei ergab sich eine interne Konsistenz von Cronbach-$\alpha = 0.84$, was auf eine hohe Reliabilität hindeutet. Diese Pilotierung diente der Bewertung der Verständlichkeit der Aufgabenstellungen sowie der zeitlichen Angemessenheit des Tests. Der Leistungstest wurde mit maximal 20 Punkten bewertet. Jede richtige Antwort wurde mit einem Punkt gewertet; falsche oder nicht beantwortete Aufgaben erhielten keine Punkte. Die Auswertung erfolgte vollständig objektiv auf Grundlage des richtig–falsch-Schemas. Der entwickelte Test wurde sowohl zu Beginn (Prätest) als auch am Ende (Posttest) der Intervention in beiden randomisiert zugewiesenen Gruppen eingesetzt, um die Wirkung des KI-gestützten Unterrichts auf die Leistungen der Schüler:innen vergleichend analysieren zu können. Vor der Teilnahme an der Studie wurden alle Schüler:innen und ihre Erziehungsberechtigten mittels eines Einwilligungsformulars über die Studie aufgeklärt und um ihre Zustimmung gebeten.

Durchführungsprozess

Die Durchführung der Studie erstreckte sich über einen Zeitraum von vier Wochen im Frühjahrssemester des Schuljahres 2024/2025. Da gemäß dem schulischen Curriculum jede Einheit in einem Monat abgeschlossen wird, wurde die Umsetzung entsprechend geplant. In beiden Gruppen fand der Deutschunterricht zweimal wöchentlich im regulären Klassenraum statt. Die Unterrichtsinhalte basierten auf der dritten Einheit *„Im Einkaufszentrum"* des Lehrwerks *Deutsch macht Spaß – A1.2* (OGM Materyal 2025). Für die Experimentalgruppe wurden ChatGPT-gestützte Materialien und Aktivitäten entsprechend den wöchentlichen Themen entwickelt. In jeder Woche wurde ChatGPT zunächst mit der Erstellung eines Unterrichtsplans beauftragt und anschließend mit der Generierung ergänzender Übungen, kurzer Dialoge und spielerischer Aufgaben betraut. Die generierten Materialien orientierten sich am GER-Niveau A1 und an den curricularen Vorgaben der Einheit „Im Einkaufszentrum". Dazu gehörten kurze, repetitive Dialoge für Rollenspiele („Wie findest du …?", „Kann ich Ihnen helfen?"), einfache Vokabel- und Zuordnungsaufgaben sowie Schreibaktivitäten in Form kurzer narrativ-deskriptiver Texte („Ich habe einen roten Rock gekauft."). Die Aufgaben waren bewusst niedrigschwellig gestaltet, um die kognitive Belastung der Lernenden gering zu halten und die Übereinstimmung mit den Kann-Beschreibungen des GER sicherzustellen. Diese Inhalte wurden an die jeweiligen Lernziele der Woche angepasst und individualisiert auf die sprachlichen Anforderungen der Einheit *„Im Einkaufszentrum"* zugeschnitten. So wurde ChatGPT beispielsweise in der ersten Woche von der Lehrkraft gebeten, Dialoge zur Kleidungsvorstellung unter Verwendung der Struktur *„Wie findest du …?"* zu erstellen, um die Meinungsäußerung über Kleidung zu fördern. Diese Dialoge wurden im Unterricht als Partnergespräche und Rollenspiele eingesetzt. In der zweiten Woche lag der Fokus auf der Beschreibung von Farben und Größen. Dementsprechend generierte ChatGPT Beispielsätze wie *„Der Pullover ist blau und groß."*, welche die Schüler:innen in Zuordnungs- und Satzergänzungsübungen verwendeten. In der dritten Woche wurden Einkaufsgespräche thematisiert. Die Lehrkraft gab ChatGPT den Auftrag, Verkäufer-Kunde-Dialoge mit Wendungen wie „*Kann ich Ihnen helfen?*" und „*Ich möchte ... kaufen*." zu generieren. Diese Dialoge wurden anschließend von den Lernenden in Rollenspielen zur Förderung ihrer kommunikativen Kompetenz eingesetzt. Ein von ChatGPT erstelltes Beispiel:

> *Verkäufer: Guten Tag! Kann ich Ihnen helfen?*
> *Kunde: Ja, ich suche eine blaue Jacke.*
> *Verkäufer: Hier, bitte. Möchten Sie sie anprobieren?*
> *Kunde: Ja, gerne. Sie passt gut. Ich nehme sie.*

In der vierten und letzten Woche wurden produktive Sprachfähigkeiten durch Schreibaufgaben gezielt gestärkt. ChatGPT generierte kurze Texte mit Sätzen wie *„Ich habe einen roten Rock gekauft."*, in denen Einkaufserlebnisse beschrieben wurden. Diese dienten den Schüler:innen als Vorlage für eigene Einkaufsgeschichten, die anschließend im Plenum vorgetragen wurden. Alle diese Aktivitäten wurden so konzipiert, dass sie die

sprachlichen und thematischen Lernziele jeder Woche unterstützen und die aktive Beteiligung der Schüler:innen fördern. Die ChatGPT-basierten Inhalte wurden aufgabenorientiert gestaltet, um sowohl sprachliche als auch kommunikative Fähigkeiten systematisch weiterzuentwickeln. In der Kontrollgruppe wurden die gleichen Themen mit konventionellen Unterrichtsmethoden behandelt, die von der Lehrkraft vorbereitet wurden. Beide Gruppen wurden von derselben Lehrkraft unterrichtet; der Unterricht erfolgte jedoch zu versetzten Zeiten. In der ersten Woche, vor Beginn des eigentlichen Unterrichts, wurde in beiden Gruppen ein Prätest durchgeführt. Nach Abschluss der vier Unterrichtswochen wurde ein Posttest außerhalb der regulären Unterrichtszeit eingesetzt, um die Lernfortschritte zu messen.

Datenanalyse

Für die Analyse der quantitativen Daten wurde das Statistikprogramm SPSS 25.0 verwendet. Die Datenauswertung basierte auf den Leistungspunkten, die die Schüler:innen in den Prä- und Posttests erzielt hatten. Vor Beginn der Analyse wurde die Verteilung der Testergebnisse auf Normalverteilung überprüft; hierzu wurde der Shapiro-Wilk-Test herangezogen, zusätzlich wurde die Varianzhomogenität mit dem Levene-Test kontrolliert. Dabei wurde festgestellt, dass die Verteilung die Voraussetzungen für die Anwendung parametrischer Tests erfüllt. Dementsprechend wurden bei den statistischen Analysen parametrische Verfahren eingesetzt. Um Veränderungen im Leistungsniveau der Schüler:innen zu ermitteln, wurden innerhalb jeder Gruppe die Prä- und Posttest-Ergebnisse mit dem t-Test für abhängige Stichproben analysiert. Zum Vergleich der Posttestergebnisse zwischen der Experimental- und der Kontrollgruppe wurde der t-Test für unabhängige Stichproben herangezogen. Neben den Signifikanzwerten wurden auch Effektstärken (Cohen's d) berechnet, um die praktische Bedeutsamkeit der Ergebnisse zu verdeutlichen. Für alle statistischen Analysen wurde ein Signifikanzniveau von $\alpha = 0{,}05$ festgelegt. Die gewonnenen Ergebnisse wurden verwendet, um den Einfluss der KI-gestützten Unterrichtsintervention auf die schulische Leistung im Fach Deutsch vergleichend darzustellen und zu interpretieren.

Ergebnisse

Die deskriptiven Statistiken zu den Ergebnissen der Prä- und Posttests im Deutschleistungstest wurden gruppenweise separat berechnet. Die Mittelwerte und Standardabweichungen der Prä- und Posttestergebnisse der Experimentalgruppe wurden einander gegenübergestellt; dieselbe Analyse wurde auch für die Kontrollgruppe durchgeführt. Für die Experimentalgruppe ergab sich ein Mittelwert von $M = 12.40$ mit einer Standardabweichung von $SD = 2.31$ im Prätest, während im Posttest ein Mittelwert von $M = 16.85$ und eine Standardabweichung von $SD = 1.90$ erzielt wurde. Diese Ergebnisse deuten auf eine deutliche Leistungssteigerung der Experimentalgruppe nach der Intervention hin. In der Kontrollgruppe lag der Mittelwert im Prätest bei $M = 12.10$ mit $SD = 2.15$; im Posttest wurde ein Mittelwert von $M = 13.45$ mit $SD = 2.80$ erreicht.

Obwohl auch hier ein Leistungsanstieg beobachtet wurde, fiel dieser im Vergleich zur Experimentalgruppe deutlich geringer aus. Eine Übersicht über die Testergebnisse beider Gruppen ist in Tabelle 1 dargestellt.

Gruppe	**Testart**	**n**	**Mittelwert (*M*)**	**Standardabweichung (*SD*)**
Experimentalgruppe	Prätest	25	12.40	2.31
Experimentalgruppe	Posttest	25	16.85	1.90
Kontrollgruppe	Prätest	25	12.10	2.15
Kontrollgruppe	Posttest	25	13.45	2.80

Tab. 1: Deskriptive Statistiken der Prä- und Posttestergebnisse der Experimental- und Kontrollgruppe im Deutschleistungstest

In der Studie wurde ein t-Test für abhängige Stichproben durchgeführt, um zu ermitteln, ob es innerhalb der Experimental- und der Kontrollgruppe signifikante Unterschiede zwischen den Prä- und Posttestergebnissen gab. Darüber hinaus wurden die Posttestergebnisse beider Gruppen mittels eines t-Tests für unabhängige Stichproben verglichen, um festzustellen, ob ein signifikanter Leistungsunterschied zwischen der Experimental- und der Kontrollgruppe besteht.

In der Experimentalgruppe wurde ein signifikanter Unterschied zwischen den Ergebnissen des *Prätests* ($M = 12.40$, $SD = 2.31$) und des *Posttests* ($M = 16.85$, $SD = 1.90$) festgestellt ($t(24) = -5.87$, $p < .001$, d = 1.17). Dieses Ergebnis deutet darauf hin, dass die ChatGPT-gestützten Unterrichtsaktivitäten wirksam zur Verbesserung der Deutschleistungen der Schüler:innen beigetragen haben. Auch in der Kontrollgruppe zeigte sich ein signifikanter Anstieg zwischen dem *Prätest* ($M = 12.10$, $SD = 2.15$) und dem *Posttest* ($M = 13.45$, $SD = 2.80$) ($t(24) = -2.45$, $p = .022$, d = .49); dieser Zuwachs war jedoch im Vergleich zur Experimentalgruppe deutlich geringer. Der Gruppenvergleich der *Posttest*-Ergebnisse ergab ebenfalls einen signifikanten Unterschied zwischen der Experimental- und der Kontrollgruppe ($t(48) = 5.12$, $p < .001$, d = 1.48), wobei der Unterschied zugunsten der Experimentalgruppe interpretiert wurde. Die Ergebnisse der Analysen sind in Tabelle 2 im Detail dargestellt.

Testart	**Gruppe**	***t (df)***	***p*-Wert**	**Signifikanz**	**Effektstärke (d)**
Prätest – Posttest	Experimentalgruppe	5.87 (24)	$p < 0.001$	Signifikant	1.17 (groß)
Prätest – Posttest	Kontrollgruppe	2.45 (24)	$p=0.022$	Signifikant	.49 (mittel)
Posttest (Experimental vs. Kontroll)	Zwischen den Gruppen	5.12 (48)	$p<0.001$	Signifikant	1.48 (sehr groß)

Tab. 2: Ergebnisse der t-Tests für Prä- und Posttest in Experimental- und Kontrollgruppe

Die erhobenen Befunde belegen den Einfluss des KI-gestützten Unterrichts auf die Schülerleistungen. In diesem Zusammenhang ist hervorzuheben, dass die Lehrkraft ChatGPT in der Unterrichtsvorbereitung einsetzte, indem sie Dialoge, Übungen und Schreibaufgaben generieren ließ. Die Schüler:innen arbeiteten nicht direkt mit der KI, sondern nutzten die von ChatGPT erstellten Materialien im Unterricht. Diese Ergebnisse werden im folgenden Abschnitt im Lichte der aktuellen Fachliteratur diskutiert.

Diskussion

Die in dieser Studie gewonnenen Ergebnisse zeigen, dass KI-gestützte Unterrichtsansätze – insbesondere auf ChatGPT basierende Inhalte – die schulischen Leistungen von Schüler:innen im Fach Deutsch deutlich verbessern können. In der Experimentalgruppe wurde am Ende der vierwöchigen Intervention ein statistisch hoch bedeutsamer Anstieg zwischen den Ergebnissen des *Prätests* und des *Posttests* festgestellt ($t(24) = 5.87$, $p < .001$). Obwohl auch in der Kontrollgruppe ein signifikanter Anstieg beobachtet wurde ($t(24) = 2.45$, $p = .022$), fiel dieser im Vergleich zur Experimentalgruppe geringer aus. Der klar erkennbare Unterschied zwischen den *Posttest*-Ergebnissen beider Gruppen zugunsten der Experimentalgruppe ($t(48) = 5.12$, $p < .001$) belegt deutlich den positiven Einfluss des ChatGPT-gestützten Unterrichts auf die schulische Leistung der Lernenden. Diese Ergebnisse stimmen mit bisherigen Forschungsergebnissen überein. Besonders Chen et al. (2020) und Sun (2023) betonen, dass KI-gestützte Lehrmittel die Lernleistung steigern und durch unmittelbares Feedback, personalisierte Inhalte sowie aufgabenbasierte Strukturen positiv zum Lernprozess beitragen. In der vorliegenden Studie zeigte sich insbesondere, dass ChatGPT aufgabenbasierte Materialien (z. B. Dialoge, Rollenspiele, Schreibaufgaben) generierte, die im Unterricht eingesetzt wurden. Eine direkte Personalisierung der Inhalte oder ein unmittelbares Feedback durch die KI fanden hingegen nicht statt; die Rückmeldungen wurden von der Lehrkraft gegeben. ChatGPT trug jedoch durch die Vielfalt der erzeugten Beispiele dazu bei, die Lernmotivation der Schüler:innen zu steigern. Zudem orientierten sich die generierten Inhalte am GER-Niveau A1 und waren in Hinblick auf curriculare Lernziele pädagogisch anschlussfähig gestaltet. Die Ergebnisse unterstützen zudem das von Kasneci et al. (2023) sowie Smutny / Schreiberova (2020) hervorgehobene Potenzial großer Sprachmodelle (LLMs) im Fremdsprachenunterricht. In diesen Studien wird betont, dass Werkzeuge wie ChatGPT als effektive Lernpartner zur Förderung produktiver Sprachfähigkeiten (z. B. Dialogführung, kreatives Schreiben) fungieren. In ähnlicher Weise zeigen auch die Ergebnisse von Şahin Toptaş (2024), dass KI-gestützte Chatbots an individuelle Lernbedürfnisse angepasst werden können und so das Maß an sprachlicher Exposition im Zielsprachenunterricht erhöhen. Ähnliche Ergebnisse berichtet auch eine qualitative Untersuchung im Hochschulkontext, die zeigt, dass Lehrende ChatGPT als kognitives Werkzeug im Materialentwicklungsprozess nutzen, um Aufgabenformate und Übungsoptionen zu generieren (Öksüz Zerey 2025). Die vorliegende Studie ergänzt diese Perspektive, indem sie empirisch darlegt, wie ChatGPT-basierte Materialien im schulischen Fremdsprachenunterricht zur Leistungssteigerung beitragen. Im Vergleich zu

bisherigen Untersuchungen liefert die vorliegende Studie eine differenzierte Perspektive: Sie zeigt, dass auch ohne direkte Personalisierung durch die KI und ohne unmittelbares Feedback ein deutlicher Leistungszuwachs erzielt werden kann. Ein weiterer Beitrag liegt in der experimentellen Anwendung im Fach Deutsch als Fremdsprache in der Türkei – einem Bereich, der bislang in der internationalen Forschung kaum empirisch untersucht wurde. Damit ergänzt die Studie die bestehende Literatur und trägt zur internationalen Diskussion über die Implementierung von KI in seltener untersuchten Fremdsprachenfächern bei.

Der in der Kontrollgruppe beobachtete begrenzte Leistungszuwachs deutet darauf hin, dass konventionelle Lehrmethoden nach wie vor eine Wirkung auf das Lernen haben, jedoch ohne technologische Unterstützung nur begrenzt wirksam sind. Darüber hinaus zeigt die Tatsache, dass in beiden Gruppen der Unterricht von derselben Lehrkraft mit denselben Inhalten durchgeführt wurde und dennoch allein die methodische Differenz zu signifikanten Leistungsunterschieden führte, die pädagogische Wirksamkeit der gewählten Unterrichtsstrategie. Diese Ergebnisse unterstützen die von Pedro u. a. (2019) und Holmes u. a. (2019) betonte Notwendigkeit, KI-gestützte Systeme systematisch in den Unterrichtsprozess zu integrieren. Es ist jedoch zu berücksichtigen, dass die Teilnehmerzahl mit insgesamt 50 Schüler:innen relativ gering war. Diese begrenzte Stichprobengröße schränkt die Generalisierbarkeit der Ergebnisse ein und erlaubt lediglich vorsichtige Schlussfolgerungen. Für zukünftige Studien wäre daher eine größere und vielfältigere Stichprobe wünschenswert, um die Aussagekraft der Befunde zu erhöhen.

Insgesamt zeigen die Ergebnisse dieser Studie, dass künstliche Intelligenz im Fremdsprachenunterricht nicht nur als unterstützendes Element, sondern als ein Instrument betrachtet werden kann, das das Lernen vertieft und differenziert. Damit solche Anwendungen jedoch erfolgreich umgesetzt werden können, sind Lehrkraftanleitung, Inhaltsanpassung und pädagogische Angemessenheit unerlässlich (Şahin Toptaş 2023; Öztürk 2024; Akol Göktaş 2024). In diesem Zusammenhang ist es unabdingbar, KI-Werkzeuge nicht lediglich als Technologie zu betrachten, sondern sie bewusst und zielgerichtet im Sinne pädagogischer Zielsetzungen einzusetzen.

Im Kontext der Türkei wird die Integration von KI-gestützten Unterrichtsanwendungen in das Bildungssystem zunehmend zu einer bildungspolitischen Priorität. Das türkische Bildungsministerium (BM) entwickelt Richtlinien und Materialien zur Förderung der KI-Kompetenz von Lehrkräften und zur Verbreitung des technologiegestützten Unterrichts. In diesem Rahmen werden unter anderem Lehrerhandbücher, Fortbildungsprogramme und didaktische Materialempfehlungen bereitgestellt (BM 2024). Vor diesem nationalen Hintergrund kommt der vorliegenden Studie besondere Bedeutung zu, da sie an einer staatlichen Schule durchgeführt wurde und eine experimentelle Anwendung von KI-gestützten Inhalten im regulären Deutschunterricht umfasste. In einer Phase, in der sich die digitalen Transformationsstrategien und KI-gestützten Unterrichtskonzepte in der Türkei weiterentwickeln, kann diese Untersuchung sowohl für die Unterrichtspraxis als auch für bildungspolitische Strategien richtungsweisend sein. Dass ein experimenteller Einsatz im

Bereich Deutsch als Fremdsprache – einer im türkischen Bildungssystem eher marginalisierten Zielsprache – zu positiven Ergebnissen führt, verdeutlicht die Notwendigkeit weiterer ähnlicher Studien. Angesichts der zunehmenden Bedeutung von KI in der türkischen Bildungspolitik und des bestehenden Bedarfs an datenbasierten, praxisorientierten Forschungsarbeiten im Deutschunterricht stellt diese Studie ein bedeutendes Beispiel dar, das sowohl auf lokaler als auch auf internationaler Ebene zum wissenschaftlichen Diskurs beitragen kann.

Fazit und Empfehlungen

Diese Studie hat gezeigt, dass KI-gestützte Unterrichtsmethoden – insbesondere auf ChatGPT basierende Anwendungen, die Lernleistungen von Deutschlernenden der Sekundarstufe steigern können. Die Ergebnisse der Prä- und Posttests belegen, dass die Experimentalgruppe, die mit ChatGPT-unterstützten Inhalten und Aktivitäten unterrichtet wurde, im Vergleich zur Kontrollgruppe, die konventionell unterrichtet wurde, deutliche Lernfortschritte erzielt hat. Dieses Ergebnis deutet darauf hin, dass KI-Werkzeuge insbesondere in interaktiven, spielerischen und aufgabenbasierten Lernsettings die Aufmerksamkeit der Lernenden erhöhen, ihre Beteiligung am Unterricht stärken und die Lernmotivation steigern mit positiven Auswirkungen auf den Lernerfolg. Die dynamischen und individualisierbaren Lernumgebungen, die durch KI-gestützte Tools ermöglicht werden, erlauben eine differenziertere Unterrichtsgestaltung, die stärker auf die individuellen Bedürfnisse der Lernenden eingehen kann. Große Sprachmodelle wie ChatGPT verfügen über das Potenzial, reichhaltige Inhalte zur Förderung grundlegender sprachlicher Kompetenzen wie Wortschatz, Satzbildung und produktives Schreiben zu generieren. Damit diese Technologien jedoch pädagogisch wirksam sein können, ist nicht nur der technologische Aspekt entscheidend, sondern auch die didaktische Anpassung, die Entwicklung zielgerichteter Unterrichtsszenarien und die Gestaltung von Aktivitäten, die den Lernbedürfnissen entsprechen. In diesem Zusammenhang lässt sich festhalten, dass der bewusste und zielgerichtete Einsatz von KI-Technologien im Fremdsprachenunterricht wertvolle pädagogische Möglichkeiten bietet. Um dieses Potenzial jedoch effektiv zu nutzen, müssen Lehrkräfte sowohl technologisch als auch didaktisch geschult werden. Nur dann kann der Einsatz von KI im Unterricht nachhaltig wirksam sein. Ein professioneller, reflektierter und sicherer Umgang mit KI-Anwendungen setzt voraus, dass Lehrpersonen Vertrauen in die Technologie entwickeln, sich mit Einsatzszenarien vertraut machen und eine kritische Perspektive einnehmen. Vor diesem Hintergrund sind Fortbildungen zur KI-Kompetenz, didaktische Leitfäden, Best-Practice-Beispiele und der Aufbau professioneller Austauschplattformen für Lehrkräfte von zentraler Bedeutung. Darüber hinaus zeigen die Ergebnisse dieser Untersuchung – die im Kontext einer im türkischen Bildungssystem eher marginal vertretenen Zielsprache wie Deutsch durchgeführt wurde –, dass technologiegestützte Ansätze nicht auf weit verbreitete Sprachen beschränkt bleiben sollten. Die Anpassungsfähigkeit von ChatGPT an unterschiedliche Sprachen und seine Fähigkeit zur Erstellung effektiver, lernendenzentrierter Inhalte legen nahe, dass solche

Technologien auch in mehrsprachigen und multikulturellen Bildungskontexten effektiv eingesetzt werden können. Gerade in Sprachen wie Deutsch, die oft nur mit begrenzter Wochenstundenzahl unterrichtet werden, kann der technologische Support den Lernprozess wesentlich fördern und bereichern.

Abschließend lässt sich festhalten, dass KI-gestützte Unterrichtsansätze mit ihrem Potenzial für personalisiertes Lernen, unmittelbares Feedback und die Erstellung lernendenzentrierter Inhalte einen bedeutenden Stellenwert im modernen Fremdsprachenunterricht einnehmen. Damit dieses Potenzial jedoch nachhaltig in das Bildungssystem integriert werden kann, bedarf es nicht nur technischer Infrastrukturinvestitionen, sondern auch der gezielten Stärkung der Lehrer:innenkompetenzen, der Entwicklung unterstützender bildungspolitischer Strategien sowie praxisorientierter Forschungsarbeiten, die evidenzbasierte Entscheidungsprozesse ermöglichen.

Einschränkungen und Ausblick auf zukünftige Forschung

Diese Studie weist einige Einschränkungen auf. Erstens wurde die Untersuchung ausschließlich an einer staatlichen Schule mit einer begrenzten Anzahl von Schüler:innen durchgeführt. Die Stichprobe umfasst lediglich zwei zufällig ausgewählte 11. Klassen derselben Schule, was trotz Randomisierung die Vielfalt der Teilnehmenden einschränkt und die Übertragbarkeit der Ergebnisse limitiert. Darüber hinaus erstreckte sich der Anwendungszeitraum lediglich über vier Wochen und konzentrierte sich auf eine einzige Unterrichtseinheit. Aus diesem Grund können keine fundierten Aussagen über die langfristigen Auswirkungen des KI-gestützten Unterrichts getroffen werden. Eine weitere Einschränkung besteht darin, dass die Studie ausschließlich auf quantitativen Daten basiert. Um die Lernerfahrungen der Schüler:innen sowie ihre emotionalen Reaktionen und Wahrnehmungen besser zu verstehen, könnten auch qualitative Daten (z. B. Schüler:innengespräche, Beobachtungsprotokolle) erhoben und ein Mixed-Methods-Design gewählt werden. Zukünftige Studien sollten ähnliche experimentelle Anwendungen in verschiedenen Schultypen, mit unterschiedlichen Altersgruppen und in weiteren Fremdsprachen durchführen. Darüber hinaus wäre es sinnvoll, KI-gestützten Unterricht nicht nur im Hinblick auf kognitive Lernziele, sondern auch in Bezug auf seine Auswirkungen auf Motivation, digitale Kompetenzen und kritisches Denken in multidimensionalen Forschungsmodellen zu untersuchen. Solche Ansätze könnten der Fachliteratur wertvolle Beiträge liefern.

Literaturverzeichnis

Akol Göktaş, Selma (2024): Zur Vermittlung des Leseverstehens im DaF-Unterricht mit KI-Tools am Beispiel von ChatGPT. Asutay, Hikmet / Akol Göktaş, Selman / Arabacı Candan, Binnur / Şentürk, Rahim (Eds.): *Dil ve eğitimde yapay zeka. Künstliche Intelligenz in der Sprache und Bildung*. Trakya Üniversitesi Yayınevi, 15-35.

Bildungsministerium der Republik Türkei [BM] (2024): *Eğitimde kullanılan yapay zekâ araçları: Öğretmen el kitabı* [KI-Werkzeuge im Bildungswesen: Handbuch für Lehrkräfte]. https://yegitek.meb.gov.tr/meb_iys_dosyalar/2024_07/12172521_meb_nek_yapay_zek_araclari_12072024_web.pdf.

Bildungsministerium der Republik Türkei [BM] (2024): *Eğitimde yapay zekâ uygulamaları uluslararası forumu raporu* [Bericht des internationalen Forums über KI-Anwendungen im Bildungsbereich]. https://yegitek.meb.gov.tr/meb_iys_dosyalar/2024_09/11104346_meb_egitimde_uyz_formu_raporu_web_28082024_tr.pdf.

Chen, Howard/ Hao-Jan, Yang/ Christine, Ting-Yu/ Lai, Kyle Kuo-Wei (2023): Investigating college EFL learners' perceptions toward the use of Google Assistant for foreign language learning. *Interactive Learning Environments, 31*(3), 1335-1350. https://doi.org/10.1080/10494820.2020.1833043.

Chen, Lijia/ Chen, Pingping/ Lin, Zhijian (2020): Application of artificial intelligence in education: A review. *IEEE Access, 8*, 75264–75278. https://doi.org/10.1109/ACCESS.2020.2988510.

Creswell, John W./ Creswell, J. David (2018): *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.

Du, Jinming/ Daniel, Ben Kei (2024): Transforming language education: A systematic review of AI-powered chatbots for English as a foreign language speaking practice. *Computers and Education: Artificial Intelligence, 6*, 100230. https://doi.org/10.1016/j.caeai.2024.100230.

Fraenkel, Jack R./ Wallen, Norman E./ Hyun, Helen H. (2012): *How to design and evaluate research in education* (8th ed.): McGraw-Hill Higher Education.

Godwin-Jones, Robert (2021): Big data and language learning: Opportunities and challenges. *Language Learning & Technology, 25* (1), 4-19.

Godwin-Jones, Robert (2022): Chatbots in language learning: AI systems on the rise. *Intelligent CALL, granular systems, and learner data: Short papers from EUROCALL*, *2022*, 124-128.

Güngenci, Meryem Melike (2024): Web 2.0 ve web 3.0 araçlarının yabancı dil eğitiminde rolü. In A. Şahin Toptaş (Ed.): *Dijital çağda yabancı dil eğitimi*. Pegem, 63-81.

Holmes, Wayne/ Bialik, Maya/ Fadel, Charles (2019): *Artificial intelligence in education promises and implications for teaching and learning*. Center for Curriculum Redesign.

Kasneci, Enkelejda/ Sessler, Kathrin/ Küchemann, Stefan/ Bannert, Maria/ Dementieva, Daryna,/ Fischer, Frank/ Gasser, Urs/ Groh, Georg/ Günnemann, Stephan/ Hüllermeier, Eyke/ Krusche, Stephan/ Kutyniok, Gitta/ Michaeli, Tilman/ Nerdel, Claudia/ Pfeffer, Jürgen/ Poquet, Oleksandra/ Sailer, Michael/, Schmidt, Albrecht/ Seidel, Tina/ Stadler, Matthias/ Kasneci, Gjergji (2023): ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and individual differences*, *103*, 102274. https://doi.org/10.1016/j.lindif.2023.102274.

Köksal, Handan/ Yılmaz, Deren (2024): Yabancı dil eğitiminde yapay zekâ uygulamalarına yönelik bir doküman incelemesi. Asutay, Hikmet / Akol Göktaş, Selma / Arabacı Candan, Binnur / Şentürk, Rahim (Ed.): *Dil ve eğitimde yapay zeka. Künstliche Intelligenz in der Sprache und Bildung*. Trakya Üniversitesi Yayınevi, 1-14.

Kozma, Robert B. (2003): Technology and classroom practices: An international study. *Journal of research on technology in education*, *36*(1), 1-14. https://doi.org/10.1080/15391523.2003.10782399.

Kumari, Payal (2024): The Impact of AI Models on Students' Writing Competence and Writing Motivation in the Context of Learning German as a Foreign Language in India. *International Journal for Multidisciplinary Research*, *6* (2), 1-11. https://doi.org/10.36948/ijfmr.2024.v06i02.18516.

Luckin, Rosemary/ Holmes, Wayne (2016): *Intelligence unleashed: An argument for AI in education*. Pearson.

Michel, Marije/ Bazhutkina, Iryna/ Abel, Niklas/ Strobl, Carola (2025): Collaborative writing based on generative AI models: Revision and deliberation processes in German as a foreign language. *Journal of Second Language Writing*, *67*, 101185. https://doi.org/10.1016/j.jslw.2025.101185.

Mohammed, Shireen Jamal/ Khalid, Maryam Waleed (2025): Under the world of AI-generated feedback on writing: mirroring motivation, foreign language peace of mind, trait emotional intelligence, and writing development. *Language Testing in Asia, 15*(1), 7. https://doi.org/10.1186/s40468-025-00343-2.

OECD (2019): *Artificial intelligence in society*. OECD Publishing. https://doi.org/10.1787/eedfee77-en

OGMMateryal (2025): *Deutsch macht spaß.* https://ogmmateryal.eba.gov.tr/panel/panel/EKitapUniteOnizle.aspx?Id=6174.

Öksüz Zerey, Merve (2025): Developing AI-supported classroom materials: Lecturer insights and practical strategies. *The Journal of Language Teaching and Learning, 15*(2), 140-156.

Öztürk, Ali Osman (2024): Öğretiminde ChatGPT kullanımına dair bazı öneriler. H. Asutay, S. Akol Göktaş, B. Arabacı Candan, / R. Şentürk (Eds.): *Dil ve eğitimde yapay zeka. Künstliche Intelligenz in der Sprache und Bildung* (pp. 105-129). Trakya Üniversitesi Yayınevi.

Pedro, Francesc/ Subosa, Miguel/ Rivas, Axel/ Valverde, Paula (2019): *Artificial Intelligence in education: Challenges and opportunities for sustainable development* (Working Papers on Education Policy No. 7). UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000366994.

Russell, Stuart/ Norvig, Peter (2016): *Artificial intelligence: a modern approach*. Pearson.

Şahin Toptaş, Aygül (2023): Yapay Zekânın Yabancı Dil Öğretiminde Kullanımı: Almanca Örneği. A. Deregözü (Ed.): *Contemporary Research in Language Education*. Konya: Eğitim Yayınevi, 77-96.

Şahin Toptaş, Aygül (2024): Dijital çağda yabancı dil eğitiminin dönüşümü. Şahin Toptaş, Aygül (Ed.): *Dijital çağda yabancı dil eğitimi*. Pegem, 1-14.

Şentürk, Rahim (2024): Fremdsprachenunterricht im Wandel: Die Potenziale und Chancen von künstlicher Intelligenz für Lehrer. Asutay, Hikmet / Akol Göktaş, Selman / Arabacı Candan, Binnur / Şentürk, Rahim (Eds.): *Dil ve eğitimde yapay zeka. Künstliche Intelligenz in der Sprache und Bildung*. Trakya Üniversitesi Yayınevi, 92-104.

Smutny, Pavel/ Schreiberova, Petra (2020): Chatbots for learning: A review of educational chatbots for the Facebook Messenger. *Computers & Education*, *151*, 103862.

Sun, Weina (2023): The impact of automatic speech recognition technology on second language pronunciation and speaking skills of EFL learners: a mixed methods investigation. *Frontiers in Psychology, 14*, Article 1210187. https://doi.org/10.3389/fpsyg.2023.1210187.

Topalovic, Marijana Drinovac/ Valcic, Marija/ Herceg, Jadranka (2024): Digital transformation and artificial intelligence in teaching foreign language for specific purposes: A proposal of a culture teaching strategy model. In Rassasmeethes, B. / Meeampol, S. / Suwannarat, P. (Eds.): *Economic and social development (Book of Proceedings), 112th International Scientific Conference on Economic and Social Development*, 355-365.

Tuğlu, Yıldırım (2024): Türkiye'de Almanca eğitimi: Yapay zekâ ve gelecek perspektifi. Asutay, Hikmet / Akol Göktaş, Selman / Arabacı Candan, Binnur / Şentürk, Rahim (Eds.): *Dil ve eğitimde yapay zeka. Künstliche Intelligenz in der Sprache und Bildung*. Trakya Üniversitesi Yayınevi, 136-152.

Zawacki-Richter, Olaf / Marín, Victoria I. / Bond, Melissa / Gouverneur, Franziska (2019): Systematic review of research on artificial intelligence applications in higher education–where are the educators? *International journal of educational technology in higher education*, *16*(1), 1-27.