

Fortbildung zur Förderung der Planungsfähigkeit mit Hilfe von KI

Ausgangslage

Mit der Veröffentlichung von ChatGPT und der einhergehenden Omnipräsenz (vgl. Maslej et al., 2025), hat sich die Diskussion von Künstlicher Intelligenz (KI) in Schule intensiviert (z. B. SWK, 2023; KMK, 2024). Beispielsweise wird Potential in der Unterstützung von Lehrkräften mit Large Language Models (LLM) in der Unterrichtsplanung gesehen (z. B. Neumann, 2024; SWK, 2023), welches in ersten Studien empirisch bestärkt wird (z. B. Hashem et al., 2024; Brott & Egerer, 2024). Gleichzeitig wird betont, dass Anwendende für den souveränen Einsatz von KI, spezifische Kompetenzen benötigen (Long & Magerko, 2020), weshalb der Ausbau von Fortbildungsangeboten für Lehrkräfte gefordert wird (SWK, 2023).

Zielsetzung und Rahmenbedingungen

Ziel dieses Projektes ist es daher, eine Lehrkräftefortbildung zum Einsatz generativer KI bei der Physikunterrichtsplanung zu konzipieren, welche i) bedürfnisgerecht ist, ii) inhaltlich fundierte Strategien mit hohem Nutzwert vermittelt, iii) allgemeinen Kriterien wirksamer Fortbildungen entspricht und iv) einem Framework der AI-Literacy unterliegt. Dazu wurde zunächst i) eine Bestands- und Bedürfniserhebung bei praktizierenden Lehrkräften durchgeführt. Im Weiteren wurde aufgrund der fehlenden empirischen Evidenz zu den Auswirkungen von Promptstrategien bei der Unterrichtsplanung ii) eine Exploration durchgeführt. Da die Qualität der Ausgaben von generativer KI durch die Eingabeaufforderungen (sog. Prompts) steuerbar ist (White et al., 2023), sollen Beispiele und Promptstrategien identifiziert werden, mit denen Unterricht empirisch begründet erfolgreich geplant werden kann. Im Hinblick auf iii) werden Kriterien wirksamer Fortbildungen bei der Fortbildungskonzeption berücksichtigt (Lipowsky & Rzejak, 2021). Insbesondere wird Wert auf einen hohen Praxisbezug gelegt, weshalb die Lehrkräfte in der Fortbildung selbst Unterrichtsplanungen mit generativer KI erstellen sollen. Schließlich bedarf ein angemessener Umgang mit KI einer eigenen Kompetenz, der AI-Literacy (Long & Magerko, 2020; Ng et al. 2021), welche als iv) Framework in der Fortbildung eingebunden werden soll. Nachfolgend werden die Bestands- und Bedürfniserhebung sowie die Exploration näher erläutert.

Bestands- und Bedürfniserhebung

Zur bedürfnisgerechten Konzeption wurde zuerst eine Bestands- und Bedürfniserhebung mit Lehrkräften durchgeführt, wobei Einsatzbereiche, Erfahrungen, Einstellungen und Bedürfnisse analysiert wurden (Cramer, 2025). In einem ersten Schritt wurden $N = 5$ Lehrkräfte bzw. Fachleitungen mit Erfahrung in der Nutzung von KI bei der Unterrichtsplanung in einem leitfadengestützten Interview befragt, um grundlegende Einsatzgebiete zu identifizieren, bei denen generative KI Lehrkräfte bei der Physikunterrichtsplanung unterstützen kann. Mit einer fokussierten Interviewanalyse nach Kuckartz und Rädiker (2024) wurden acht Einsatzbereiche identifiziert: Planen von Unterrichtsreihen, -stunden und -phasen, Erstellen von Arbeitsblättern, Arbeitsaufträgen,

Prüfungsaufgaben und differenzierenden Arbeitsaufträgen sowie das Erstellen von Abbildungen für den Unterricht.

In einem zweiten Schritt wurden praktizierende Lehrkräfte in einer quantitativen, digitalen Fragebogenerhebung gebeten, ihre Vorerfahrungen zu den identifizierten Einsatzbereichen anzugeben, das Potential in der Qualitäts- und Effizienzsteigerung einzuschätzen und anzugeben, inwiefern sie Bedürfnisse für eine Lehrkräftefortbildung zu dem Thema haben. Dazu wurden zum einen allgemeine Aussagen und zum anderen die konkreten Einsatzbereiche beschrieben und auf einer 6-stufigen Likert-Skala beurteilt. Obgleich die Interviewten kein Potential im Einsatz von generativer KI zur Planung von Experimenten gesehen haben, wurden Experimente in der Umfrage aufgenommen, da diese als zentral im Physikunterricht gelten. Die Umfrage wurde auch für Chemielehrkräfte geöffnet, da die identifizierten Einsatzbereiche wenig physikspezifisch sind und davon auszugehen ist, dass sich die Potentiale auch auf die Chemieunterrichtsplanung übertragen lassen, was in der Auswertung über ein Gruppenvergleich bestätigt wurde. Die Umfrage konnte von $N = 73$ Lehrkräften ($N = 38$ Physik, $N = 31$ Chemie $N = 4$ unbekannt) ausgewertet werden, die zum Großteil (60 %; 21 % Gesamtschule; 11 % Realschule; 8 % andere) an Gymnasien in NRW unterrichten. Die Umfrage hat gezeigt, dass 81 % der Lehrkräfte überwiegend nur wenig Erfahrung mit generativer KI zur Unterrichtsplanung haben. Dennoch sehen 81 % Potential in der Unterstützung, wobei Lehrkräfte, die mehr Erfahrung mit der Technologie haben, das Potential signifikant höher einschätzen als solche mit weniger Erfahrung. Das Potential wird mehr in der Effizienzsteigerung als in der Qualitätssteigerung gesehen. In Bezug auf die konkreten Einsatzbereiche, sehen sie vor allem Potential in der Unterstützung bei der Erstellung von differenzierenden Arbeitsaufträgen, Arbeitsblättern und der Planung von Unterrichtsstunden und -phasen. Wie auch bei den Interviews wird kein Potential in der Planung von Experimenten gesehen. Für eine Fortbildung interessiert die Lehrkräfte insbesondere, der richtige Umgang mit KI-Tools und weniger, wie der technische Hintergrund ist. Insgesamt zeigen die meisten Lehrkräfte (61%) Interesse an einer Fortbildung unabhängig von Einsatzbereichen.

Exploration: Umgang mit generativer KI zur Unterrichtsplanung

Da es wenig Evidenz für den richtigen Umgang mit generativer KI zur Unterrichtsplanung gibt und auch die Interviews keine konkreten Empfehlungen hervorgebracht haben, wurde eine Explorationsstudie durchgeführt, bei der Strategien im Umgang mit generativer KI literaturbasiert entwickelt und die KI-Ausgaben systematisch bewertet wurden (Bothur, 2025). Dabei wurden die realitätsnahe Planungsaufgabe und das Kodiermanual des Planungstests von Schröder (2022) verwendet, anhand dessen eine Unterrichtsstunde zum 3. Newtonschen Axiom mit einem Experiment zu planen ist und das Planungsprodukt anschließend mittels eines validierten Kodiermanuals bewertet wird (vgl. ebd.). Für die Exploration wurde das Experiment der Planungsaufgabe mit ChatGPT und acht verschiedenen Promptstrategien geplant. Strategie P1 beginnt mit einem intuitiven Prompt ohne Kontext, wobei der Kontext über P2, P3 und P4 immer ausführlicher wird. Als 5. Strategie (P5) wurde das PREP-Verfahren (Prompt, Role, Explicit, Parameter) von Fitzpatrick (2023) getestet. In der 6. Strategie (P6) wurden konkrete Ziele für die Planung im Prompt vorgegeben. Schließlich wurde eine Schachtelung in Teilprompts (P7) erprobt. Um die umfangreichen Prompts zu kürzen, wurden in Strategie (P8) Gestaltungsprinzipien wie Nummerierungen und Trennung zwischen Kontextinformationen angewendet, um den Prompt zu verdichten.

Jede dieser acht Strategien wurde in drei verschiedenen Chat-Fenstern durchgeführt und jeweils ein Mittelwert der Scores gebildet, die sich aus der Kodierung mit dem Manual von Schröder (2022) ergeben, womit ein Vergleich zwischen den Strategien ermöglicht wird.

Um die Vermutung zu überprüfen, dass Reasoning-Modelle (wie ChatGPT o4-mini) besser für komplexe Aufgaben geeignet sind als GPT-Modelle (wie ChatGPT 4o) (OpenAI, o. D.), wurden zunächst alle acht Strategien mit beiden Modellen angewendet und miteinander verglichen. Dabei zeigt sich, dass das o4-mini Modell bei jeder Strategie einen höheren Score als das GPT-Modell erreicht (bis zu 4,3 Punkte Unterschied von möglichen 18), was darauf hindeutet, dass sich das Reasoning-Modell für den vorliegenden Fall besser zur Unterrichtsplanung eignet, was eine erste Handlungsempfehlung für die Fortbildung darstellt.

Erste Ergebnisse der Analyse der Promptstrategien zeigen, dass die Art der Prompts Einfluss auf die Planungsqualität der Ausgaben hat. Prompts, bei denen mehr Kontext angegeben wurde, führten meistens zu besseren Ausgaben. Auch das Beschreiben konkreter Zielvorgaben hat zu besseren Ergebnissen geführt. Die Strategie P7 (Schachtelung) konnte bei beiden Modellen nach dem letzten Teilprompt die meisten Punkte erreichen (o4-mini: 18; GPT: 16), wobei auch die Strategie P8 (Verdichtung) in beiden Fällen mit am besten abgeschlossen hat (o4-mini: 17; GPT: 12). Die Strategie nach dem PREP-Verfahren (P6) hat in beiden Fällen fast die niedrigste Punktzahl (o4-mini: 9; GPT: 8) erreicht und kann für diesen Einsatz damit nicht empirisch gestärkt werden. Als Handlungsempfehlungen, können damit die Gestaltungsprinzipien aus Strategie P8, konkrete Kontext- und Zielangaben sowie eine Schachtelung in Teilprompts angegeben werden. Hier ist zu limitieren, dass noch weitere Auswertungen ausstehen und der Fokus nur auf der Planung des Experiments liegt.

Aufbau der Fortbildung

Die Ergebnisse der Bedürfniserhebung der Lehrkräfte nach einer Fortbildung mit Fokus auf der Anwendung generativer KI ohne der Erarbeitung technischer Hintergründe steht im Spannungsverhältnis zu Modellen der AI-Literacy. Dabei wird für einen sinnvollen Umgang mit generativer KI zunächst eine allgemeine Wissensbasis vorausgesetzt (z. B. Annapureddy et al., 2025; Ng et al., 2021). Außerdem bedarf es noch einer Kompetenz, KI-Produkte zu bewerten (ebd). Daher wird die Fortbildung in zwei Zyklen aus Input, Erproben und Reflektieren aufgebaut. Im ersten Zyklus wird ein allgemeiner Input über Grundlagen von KI gegeben. Darunter Grundbegriffe wie KI, AGI (Artificial General Intelligence), LLM und generativer KI sowie einer vereinfachten Erklärung der Inhaltsgenerierung und der Limitationen der Modelle insbesondere bezüglich ihrer Erklärbarkeit und Interpretierbarkeit (Bauckhage et al., 2021). Anschließend erproben die Lehrkräfte ChatGPT, indem sie einen Teil einer Unterrichtsplanung mit KI durchführen und ihre Erfahrungen reflektieren. Im zweiten Zyklus werden die fundierten Handlungsempfehlungen aus der Exploration (s.o.) beschrieben, auf Basis dieser die Lehrkräfte erneut eigenen Unterricht planen. Im Anschluss werden diese Erfahrungen reflektiert und insgesamt vor der Fragestellung nach Potential und Limitationen der Nutzung von generativer KI zur Unterrichtsplanung diskutiert.

Ausblick

Die ersten Fortbildungen werden im Herbst 2025 angeboten und richten sich an alle Lehrkräfte, die Physik unterrichten. Die Fortbildung wird halbtags in Präsenz angeboten. Um den Kompetenzaufbau bezüglich der generativen AI-Literacy zu überprüfen, wird die Fortbildung mit einer adaptierten und übersetzten Form des *generative AI literacy assessment test (GLAT)* von Jin et al. (2025) evaluiert und fortlaufend aktualisiert.

Literatur

- Annapureddy, R., Fornaroli, A. & Gatica-Perez, D. (2025). Generative AI literacy: Twelve defining competencies. *Digital Government: Research and Practice*, 6(1), 1–21. <https://doi.org/10.1145/3685680>
- Bauckhage, C., Fürnkranz, J. & Paaß, G. (2021). Vertrauenswürdigen, transparentes und robustes maschinelles Lernen. In G. Görz, U. Schmid, & T. Braun (Hrsg.), *Handbuch der Künstlichen Intelligenz* (6. Aufl., S. 429–508). Walter de Gruyter. <https://doi.org/10.1515/9783110659948-013>
- Bothur, M. (2025). *Einfluss unterschiedlicher Promptstrategien auf die Qualität KI-generierter Unterrichtsplanungen für den Physikunterricht*. [Unveröffentlichte Masterarbeit]. Universität Paderborn. <https://doi.org/10.13140/RG.2.2.21709.29925>
- Brott, M. & Egerer, C. (2024). CUKI – Chemie Unterricht geplant durch künstliche Intelligenz. In H. van Vorst (Hrsg.), *Frühe naturwissenschaftliche Bildung. Gesellschaft für Didaktik der Chemie und Physik, Jahrestagung in Hamburg 2023* (S. 126–129).
- Cramer, K. (2025). *Nutzen generativer künstlicher Intelligenz in der Physikunterrichtsplanung aus Sicht von Personen mit schulpraktischem Hintergrund*. [Unveröffentlichte Masterarbeit]. Universität Paderborn. <https://doi.org/10.13140/RG.2.2.32686.83523>
- Fitzpatrick, D. (2023). *The AI classroom: The ultimate guide to artificial intelligence in education*. TeacherGoals Publishing.
- Hashem, R., Ali, N., El Zein, F., Fidalgo, P. & Abu Khurma, O. (2023). AI to the rescue: Exploring the potential of ChatGPT as a teacher ally for workload relief and burnout prevention. *Research and Practice in Technology Enhanced Learning*, 19, 023. <https://doi.org/10.58459/rptel.2024.19023>
- Jin, Y., Martinez-Maldonado, R., Gašević, D. & Yan, L. (2025). GLAT: The generative AI literacy assessment test. *Computers and Education: Artificial Intelligence*, 9, 100436.
- Kuckartz, U. & Rädiker (2024). *Qualitative Inhaltsanalyse. Methoden, Praxis, Umsetzung mit Software und künstlicher Intelligenz* (6. Auflage). Beltz Juventa.
- KMK (2024). *Handlungsempfehlung für die Bildungsverwaltung zum Umgang mit Künstlicher Intelligenz in schulischen Bildungsprozessen*.
- Lipowsky, R. & Rzejak, D. (2021). Welche Art von Fortbildung wirkt? In B. Jungkamp & M. Pfaffertott (Hrsg.), *Was Lehrkräfte lernen müssen—Bedarfe der Lehrkräftebildung in Deutschland* (1. Aufl., S. 19–38). Friedrich-Ebert-Stiftung.
- Long, D. & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–16. <https://doi.org/10.1145/3313831.3376727>
- Maslej, N., Fattorini, L., Perrault, R., Gil, Y., Parli, V., Kariuki, N., Capstick, E., Reuel, A., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Niebles, J. C., Shoham, Y., Wald, R., Walsh, T., Hamrah, A., Santarlasci, L., ... & Oak, S. (2025). *Artificial intelligence index report 2025*.
- Neumann, K., Kuhn, J. & Drachsler, H. (2024). Generative Künstliche Intelligenz in Unterricht und Unterrichtsforschung – Chancen und Herausforderungen. *Unterrichtswissenschaft*, 52(2), 227–237.
- Ng, D. T. K., Leung, J. K. L., Chu, S. K. W. & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, 2, 100041.
- OpenAI. (o. D.). *Reasoning best practices—OpenAI API*. <https://platform.openai.com>
- Schröder, J. (2022). *Entwicklung eines Performanztests zur Messung der Fähigkeit zur Unterrichtsplanung bei Lehramtsstudierenden im Fach Physik*. Logos Verlag Berlin GmbH.
- SWK (2023). *Large Language Models und ihre Potenziale im Bildungssystem. Impulspapier der Ständigen Wissenschaftlichen Kommission der Kultusministerkonferenz* (S. 31) [Impulspapier]. SWK : Bonn.
- White, J., Fu, Q., Hays, S., Sandborn, M., Olea, C., Gilbert, H., Elnashar, A., Spencer-Smith, J. & Schmidt, D. C. (2023). *A prompt pattern catalog to enhance prompt engineering with ChatGPT* (No. arXiv:2302.11382). arXiv. <https://doi.org/10.48550/arXiv.2302.11382>