

ChatGPT im Lehr-Lern-Labor

- Potenziale eines KI-basierten Assistenten bei der Entwicklung von
Experimentierumgebungen -

Patrick Herz*, Jens Damköhler*, Wolfgang Lutz*, Thomas Trefzger*

*Lehrstuhl für Physik und ihre Didaktik, Universität Würzburg
patrick.herz@stud-mail.uni-wuerzburg.de

Kurzfassung

Mit der Veröffentlichung von ChatGPT im November 2022 hat Künstliche Intelligenz (KI) unzählige neue Möglichkeiten zur Unterstützung unseres Alltags geschaffen. Auch im Bildungsbereich eröffnen sich durch KI neue Möglichkeiten zur Erweiterung bestehender Lehrformate und -methoden, beispielsweise durch den Einsatz von Chatbots als unterstützende Lehrassistenten.

Vor diesem Hintergrund wurde ein KI-gestützter Chatbot für ein Lehr-Lern-Labor auf Basis didaktischer Prinzipien in ChatGPT entwickelt, um Studierende bei der Entwicklung von Experimentierumgebungen für Schülerinnen und Schüler zu unterstützen. Nach einer mehrwöchigen Nutzung durch die Studierenden wurde die Akzeptanz des Chatbots mittels leitfadengestützter Interviews evaluiert. Die Transkripte wurden auf Basis des Technologieakzeptanzmodells (TAM) qualitativ analysiert, um zu erfassen, wie die Studierenden den Chatbot nutzen und bewerten. Die Ergebnisse zeigen, dass die Studierenden den Chatbot intensiv zur Ideenfindung und anfänglichen Planung ihrer Stationen nutzten, während die Nutzung im weiteren Verlauf der Stationsentwicklung abnahm. Insgesamt wurde der Chatbot hinsichtlich seiner Nützlichkeit und Benutzerfreundlichkeit überwiegend positiv bewertet. Zugleich verdeutlichen die Ergebnisse die Grenzen solcher generativer KI-Systeme und die Relevanz individueller sowie kontextueller Einflussfaktoren.

1. Einleitung

Im Jahr 1966 veröffentlichte Joseph Weizenbaum das Computerprogramm ELIZA, das durch einfache Mustererkennung und Textsubstitution eine psychotherapeutische Gesprächssituation simulierte. Obwohl dieser frühe Chatbot keinerlei inhaltliches Verständnis besaß, zeigten Nutzende oft unerwartet starke emotionale Reaktionen bei der Interaktion mit ihm. Weizenbaum selbst war von diesen Reaktionen überrascht und setzte sich kritisch mit der Bewertung der Fähigkeiten solcher Systeme auseinander (Weizenbaum, 1976). ELIZA legte den Grundstein für die jahrzehntelange Weiterentwicklung computergestützter Kommunikation und stellte zugleich eine frühe Vorstufe dessen dar, was heute als generative Künstliche Intelligenz (KI) bezeichnet wird. Unter generativer KI verstehen wir Systeme, die auf Basis umfangreicher Trainingsdaten eigenständig neue Inhalte erzeugen können (Marineau, 2023). Heute, fast sechs Jahrzehnte später, wirkt ELIZA auf uns primitiv und wie ein nettes Spielzeug. Die Systeme, die wir heute verwenden können, sind dies jedoch bei Weitem nicht mehr. Die Fähigkeiten generativer KI sind immens gewachsen. So können große Sprachmodelle (Large Language Models, kurz LLMs) wie etwa GPT-4o von OpenAI, kohärente und kontextbezogene Texte generieren.

Durch diese erweiterten technischen Möglichkeiten wächst allerdings auch die Verantwortung für einen reflektierten und sinnvollen Einsatz, insbesondere im Bildungsbereich. Ein Großteil der Studierenden nutzt

bereits jetzt Chatbots zu verschiedensten Zwecken (Lutz et al. (im Druck), 2024; von Garrel & Mayer, 2025), was auch neue Fragen zur Integration dieser Technologien in der Lehrkräftebildung aufwirft.

Die vorliegende Studie untersucht, wie angehende Physiklehrkräfte einen Chatbot im Rahmen eines fachdidaktischen Seminars nutzen und wahrnehmen. Dabei stehen insbesondere die Aspekte der Nützlichkeit, Benutzerfreundlichkeit sowie die individuellen Einflussfaktoren auf die Technologieakzeptanz im Fokus.

2. Theoretischer Hintergrund

2.1. KI-basierte Unterstützung in der Lehrkräftebildung

Bereits seit drei Jahrzehnten ist die Forschung zu KI in der Bildung präsent (Zawacki-Richter et al., 2019). Seit dem Durchbruch der generativen KI (wie ChatGPT) rückte der Einsatz einer solcher KI jedoch verstärkt in den Fokus der Bildungsforschung (Wang et al., 2024).

Von Garrel & Mayer (2025) kommen in einer im Jahr 2025 bundesweit durchgeführten Studie auf einen Anteil von rund 91,6% der Studierenden, die KI-Tools für ihr Studium nutzen, am häufigsten ChatGPT. Unter den Lehramtsstudierenden liegt die Nutzungsrate in dieser Studie mit 89,2% nur geringfügig darunter. Diese Zahlen heben den breiten Zugang zu KI-Tools in der Hochschulbildung auch unter angehenden Lehrkräften hervor.

Mit dem rasanten Aufstieg generativer KI-Systeme in den letzten Jahren stellt sich zunehmend die Frage, wie diese Entwicklung in der Lehrkräftebildung berücksichtigt werden kann. In der Debatte steht oft die Förderung der KI-Kompetenz von Lehrkräften als wichtiger Forschungsbereich im Vordergrund (vgl. Huwer et al., 2024; KMK, 2024; Vo & Pancratz, 2023). Die Frage nach der Akzeptanz von KI-Systemen bei (angehenden) Lehrkräften findet hingegen weniger Beachtung. Gerade Lehramtsstudierende nehmen dabei eine Doppelrolle ein: Sie sind nicht nur Lernende an der Universität, sondern auch zukünftige Lehrkräfte an Schulen, deren Einstellungen und Erfahrungen mit KI den späteren Unterricht mitprägen können.

Dabei zeigen empirische Befunde, dass die Technologieakzeptanz einen entscheidenden Einfluss auf die Entwicklung der digitalen Kompetenz hat. In einer Studie mit 667 Lehramtsstudierenden konnten Knoth & Haider (2023) nachweisen, dass eine höhere Akzeptanz digitaler Technologien signifikant mit einer höheren Einschätzung der eigenen Kompetenz einhergeht. Zusätzlich zu diesem Einfluss gilt die Technologieakzeptanz seit Langem als verlässlicher Prädiktor für die Absicht, neue Technologien zu nutzen (vgl. King & He, 2006).

Für die Lehrkräftebildung bedeutet dies: Nur wenn die Zielgruppe KI-gestützte Systeme akzeptiert, können geeignete Maßnahmen zur Förderung der KI-Kompetenz greifen. Eine Analyse der Akzeptanz kann dabei aufzeigen, welche Rahmenbedingungen, Nutzungserwartungen oder Herausforderungen für angehende Lehrkräfte eine Rolle spielen.

Die vorliegende Studie setzt genau an diesem Punkt an: Sie untersucht die Technologieakzeptanz generativer KI in der Physiklehrkräftebildung – also dort, wo auch eine erste Förderung von KI-Kompetenz ansetzen kann. Im Zentrum dieser Studie steht daher, wie Lehramtsstudierende der Physik einen Chatbot als Unterstützung bei der Planung und Ausarbeitung von Experimentierstationen erleben und bewerten.

2.2. Das Technologieakzeptanzmodell

Mit dem Technologieakzeptanzmodell (TAM) versucht Davis et. al (1989) die Akzeptanz einer Technologie durch Nutzende systematisch zu erklären, vorherzusagen und letztlich auch gezielt zu verbessern. Das TAM gehört zu den am häufigsten verwendeten theoretischen Modellen zur Untersuchung der Einführung neuer Technologien (Davis & Granić, 2024) - auch im Bildungsbereich (Granić & Marangunic, 2019).

Im Zentrum des TAM (siehe Abb. 2) stehen dabei zwei zentrale Konstrukte: die wahrgenommene Nützlichkeit (WN) und die Wahrgenommene Benutzerfreundlichkeit (WB). Unter der WN versteht Davis et al. (1989) die subjektive Einschätzung einer Person, ob der Einsatz einer Technologie dabei hilft, Aufgaben effizienter oder erfolgreicher zu bewältigen. Die

WB bezieht sich hingegen, auf die Erwartung wie unkompliziert und mühelos sich die Technologie bedienen lässt. Beide Konstrukte wirken sich auf die Nutzungsintention aus, die wiederum als unmittelbarer Prädiktor für die tatsächliche Nutzung gilt (vgl. King & He, 2006). Ergänzt wird das Modell durch sogenannte externe Variablen (EV), die die WN und WB beeinflussen können. Dabei kann es sich beispielsweise um vorangehende Nutzungserfahrungen (Jackson et al., 1997), Persönlichkeitsmerkmale (Svendson et al., 2013) oder Vertrauen (Beldad & Hegner, 2018) handeln.

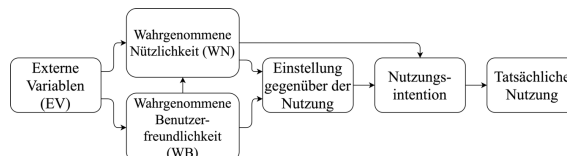


Abb. 1: Technologieakzeptanzmodell (nach Davis et al., 1989). Eigene Darstellung.

Das ursprüngliche TAM wurde in den letzten Jahren vielfach weiterentwickelt. So differenziert das TAM2 (Venkatesh & Davis, 2000) beispielsweise externe Einflussgrößen weiter aus, während die Unified Theory of Acceptance and Use of Technology (UTAUT) (Venkatesh et al., 2003) zudem Elemente weiterer Akzeptanzmodelle integriert. Dennoch zeigen verschiedene Studien, dass das TAM in seiner Klarheit und Validität überzeugen kann (z.B. King & He, 2006; Ma & Liu, 2004). Auch im Bildungskontext belegten Scherer et al. (2019) in einer Metaanalyse, dass das Modell sowohl bei Lehrkräften als auch bei Lehramtsstudierenden zuverlässige Ergebnisse liefert.

Zudem ist das TAM im Hinblick auf die Akzeptanz von KI-Technologien relevant. Eine aktuelle Metaanalyse von Ali et al. (2024) vergleicht das TAM mit dem UTAUT und kommt zu dem Ergebnis, dass das TAM am besten geeignet ist, um die Akzeptanz von KI-basierten Anwendungen im Bildungsbereich zu erfassen. Die Flexibilität des TAMs erlaubt zudem, mögliche neue Einflussfaktoren, die im Kontext von Chatbots auftreten, als externe Variablen zu integrieren.

2.3. Lehrveranstaltungsformat „Lehr-Lern-Labor“ in Würzburg

Das Lehr-Lern-Labor (LLL) an der Universität Würzburg ist ein verpflichtendes fachdidaktisches Veranstaltungsformat für angehende Physiklehrkräfte an Gymnasien und Realschulen. In diesem können die Studierenden zusätzlich zur didaktischen Ausbildung Praxiserfahrungen sammeln. Die Studierenden befinden sich meist im 7. Semester, bringen dabei unterschiedlich viele Erfahrungen aus der Unterrichtspraxis mit. Ziel des LLLs ist es, den Studierenden eine praxisnahe Auseinandersetzung mit der Planung, Durchführung und Reflexion von Lehr-Lernsituationen im Fach Physik zu ermöglichen.

Das Seminar ist in zwei Phasen gegliedert, die sich über insgesamt 14 Semesterwochen à 3 Stunden

erstrecken, und einem iterativen Ablauf folgen (Elscholz & Trefzger, 2017). In der ersten Phase, die bis zur neunten Semesterwoche geht, liegt der Schwerpunkt auf der eigenständigen Entwicklung und Ausgestaltung einer 30-minütigen Experimentierstation für Schulklassen. In der ersten Semesterwoche erhalten die Studierenden einen theoretischen Input zur Planung und Betreuung von Experimentierstationen. Unter anderem werden den Studierenden die Facetten experimenteller Kompetenz nach Nawrath et al. (2011), Grundlagen des forschend entdeckenden Lernens mit dem 5-E Modell (Bybee et al., 2006), der Mehrwert digitaler Medien nach dem SAMR-Modell (Puentedura, 2014) und das Vier-Phasen-Modell der Interessenentwicklung (Hidi & Renninger, 2006) nähergebracht. Auf diesen Konzepten aufbauend entwickeln die Studierenden ab der zweiten Semesterwoche eigenständig eine Station zu einem von den Dozierenden festgelegten Thema der Schulphysik und einer ausgewählten Jahrgangsstufe.

Eine zentrale didaktische Herausforderung für die Studierenden ist dabei, ihr selbst gewähltes Thema an das Vorwissen der SuS der jeweiligen Jahrgangsstufe anzupassen, ohne Inhalte vorwegzunehmen, die noch nicht behandelt wurden. Zusätzlich sollen die Experimente nicht aus dem Unterricht bekannt sein, sondern motivierend und neuartig sein. Dabei sollen die Materialien und Geräte der Physiksammlung verwendet werden. Unter Umständen können auch zusätzliche Materialien angeschafft werden. Die Studierenden sollen ausgehend von eigenen Ideen die Station entwickeln. Für Fragen und Hilfestellungen stehen die Dozierenden zur Verfügung.

In der zweiten Phase des LLLs werden die entwickelten Stationen praktisch mit drei Schulklassen am M!ND-Center der Universität Würzburg erprobt. Die SuS bearbeiten innerhalb eines Vormittags in Kleingruppen von sechs bis acht Personen die Stationen der Studierenden. Die drei Durchführungssitzungen sind jeweils durch eine Reflexionssitzung voneinander getrennt. Darin setzen sich die Studierenden vor dem Hintergrund theoretischer Unterrichtsmodelle mit ihren Erfahrungen aus den vergangenen Durchführungen auseinander (Damköhler et al., 2023) und nehmen gegebenenfalls Anpassungen an ihrer Station vor.

2.4. Der Experiment-Chatbot

Für das LLL wurde auf Basis des Sprachmodells GPT-4o ein Chatbot entwickelt, der die Studierenden während der Konzeption physikalischer Experimentierstationen begleitet. Zunächst besteht seine Aufgabe darin, gemeinsam mit den Studierenden ein geeignetes Thema zu finden, das den Kriterien des LLLs (siehe Kapitel 2.3) genügt. Aufbauend darauf oder anknüpfend an ein eigenes Thema hilft der Chatbot den Studierenden, eine Station mit Experimenten für SuS zu erstellen. Der Chatbot ist im Seminar stets zugänglich und kann die Studierenden somit gezielt individuell unterstützen, insbesondere, da die beiden

Dozierenden nicht alle Studierenden gleichzeitig betreuen können.

Die Idee, einen Chatbot im LLL einzusetzen, basiert auf Erfahrungen aus vorangegangenen Studien (Damköhler et al., 2024a; Damköhler et al., 2024b; Damköhler et al., 2025), in denen im vorangegangenen Semester ein Chatbot als Reflexionscoach in der zweiten Phase des LLLs eingesetzt wurde. Aufbauend auf diesen Studien soll der Chatbot in der vorliegenden Studie bereits in der ersten Phase des Seminars eingesetzt werden.

Dafür wurde ein Prompt verwendet, der durch Rollenzuweisung, strukturierte Ablaufpläne und umfassenden Materialien als Datengrundlage eine adaptive Kommunikation ermöglicht. Die schrittweise Gliederung des Prompts folgt dem Prinzip des Chain-of-Thought-Prompting, welches sich laut Wei et al. (2022) positiv auf die Qualität der Antworten auswirkt. Durch diese sequenzielle Struktur werden die Studierenden in einem iterativen Prozess durch die Ideenfindung und Ausarbeitung der Station geleitet. Zusätzlich kommt das Prinzip des Few-Shot-Promptings zum Einsatz: Dem Chatbot wurden dazu Beispiele von Stationen aus vorangegangenen Semestern gegeben. Laut Brown et al. (2020) wird dadurch die Qualität der Antwort des Chatbots gesteigert.

Um die Qualität des Outputs weiter zu erhöhen, kann der Chatbot auf eine Reihe zusätzlicher Materialien zurückgreifen. Dazu gehören ein digitalisierter Auszug aus den Inhalten des LehrplanPLUS (Staatsinstitut für Schulqualität und Bildungsforschung München, 2025), die Facetten der experimentellen Kompetenz (Nawrath et al., 2011), die Phasen der Motivation (Kircher, 2015), sowie eine Sammlung erprobter Experimente (Sächsischer Bildungsserver, o.D.; PhySX - Physikalische Schulexperimente Wiki, o.D.). Mithilfe dieser Materialien kann der Chatbot Vorschläge curricular einordnen, didaktisch besser fundieren und die Anforderungen an die Stationen besser umsetzen.

3. Forschungsfragen

In der vorliegenden Studie wird das TAM verwendet, um die Nutzungserfahrung von Lehramtsstudierenden der Physik mit einem KI-basierten Chatbot zu analysieren. Im Rahmen des behandelten theoretischen Hintergrunds ergeben sich folgende Forschungsfragen (FF):

FF1. Welches Nutzungsverhalten zeigen die Studierenden bei der Nutzung des Chatbots?

FF2. Wie erleben die Studierenden die Nutzung des Chatbots...

- a. hinsichtlich der wahrgenommenen Nützlichkeit?
- b. hinsichtlich der wahrgenommenen Benutzerfreundlichkeit?

FF3. Welche externen Variablen beeinflussen die Nutzungserfahrung der Studierenden?

4. Methodik

4.1. Erhebungsdesign und Durchführung der Interviewstudie

Im Rahmen des LLL-Seminars im Wintersemester 2024/25 nahmen sechs Lehramtsstudierende der Physik im höheren Fachsemester an der Studie teil. Die Studie bezieht sich ausschließlich auf die erste Phase des Seminars, in der die Studierenden eine 30-minütige Station zum Thema „Elektromagnetismus“ für die 10. Jahrgangsstufe konzipieren sollten. Diese erste Phase ging von Oktober bis Dezember 2024.

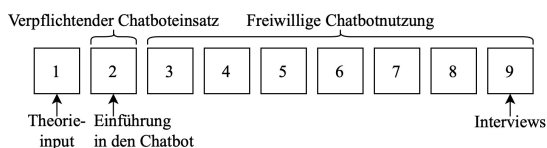


Abb. 2: Ablauf des Chatbot-Einsatzes und der Interviewstudie von Semesterwoche 1 bis 9. Eigene Darstellung.

Nachdem die Studierenden in der ersten Semesterwoche theoretische Inputs erhalten hatten (siehe Kapitel 2.3), wurden sie in der zweiten Semesterwoche in die Nutzung des Experiment-Chatbots eingeführt. Dabei wurden das Ziel des Chatbots, seine Funktionsweise und Tipps zur Kommunikation kurz erläutert. Der genaue Ablauf der Studie kann Abb. 2 entnommen werden.

Nach der Einführung erhielten die Studierenden jeweils einen Laptop mit einem Chatfenster, über das sie mit dem Chatbot, der auf dem Sprachmodell GPT-4o basiert, chatten konnten. Zunächst sollten sie mithilfe des Chatbots erste Ideen für eine Station entwickeln und diese weiter ausarbeiten. Dieser Einsatz war für alle Teilnehmenden verpflichtend, die Dauer der Nutzung blieb jedoch individuell offen. In den nachfolgenden Semesterwochen bis einschließlich der neunten wurde der Chatbot mit demselben Chat erneut bereitgestellt, die Nutzung blieb den Teilnehmenden jedoch überlassen. Parallel dazu konnten die Studierenden auch mit einem Dozierenden über ihre Station sprechen.

Zur Erhebung der Nutzungserfahrung wurde ein leitfadengestütztes Interview entwickelt, das sich an den zentralen Konstrukten des TAMs orientiert. Die Interviews wurden in der letzten Semesterwoche der Planungsphase durchgeführt. Die Teilnahme war freiwillig und unabhängig von der Bewertung.

4.2. Datenaufnahme und -auswertung

Die Gespräche wurden digital aufgezeichnet, mit MaxQDA automatisch transkribiert und anschließend händisch überarbeitet. Dabei kam das Transkriptionssystem mit sprachlicher Glättung nach Dresing & Pehl (2018) zum Einsatz. Alle Transkripte wurden

anonymisiert, so dass die Studierenden nur mit dem Kürzel B1 bis B6 bezeichnet werden.

Für die Auswertung der Interviews wurde eine qualitative Inhaltsanalyse nach Kuckartz & Rädiker (2022) herangezogen. Dabei kamen zwei Analyseformen zum Einsatz: die inhaltlich strukturierende Analyse (für FF1 und FF3) und die evaluative qualitative Analyse (FF2a und FF2b).

Die Hauptkategorien wurden in allen Fällen deduktiv aus dem TAM abgeleitet. Für FF1 wurden die Subkategorien deduktiv auf Basis des Interviewleitfadens entwickelt. Eine dieser Subkategorien wurde im weiteren Verlauf induktiv durch feinere Sub-Subkategorien ergänzt. Für FF2 und FF3 erfolgte die Bildung der Subkategorien vollständig induktiv aus dem Datmaterial.

Das Erstellen der Transkripte und die erste Entwicklung des Kategoriensystems wurden zunächst eigenständig durchgeführt. Anschließend wurden die Kategorien innerhalb der Forschungsgruppe diskutiert und gemeinsam reflektiert.

4.3. Kategoriensystem

Die Hauptkategorien des entwickelten Kategoriensystems lauten in Anlehnung an die Hauptelemente des TAM: Nutzungsverhalten (FF1), Wahrgenommene Nützlichkeit (FF2a), Wahrgenommene Benutzerfreundlichkeit (FF2b) und Externe Variablen (FF3). Jede dieser Hauptkategorien wurde durch Subkategorien weiter differenziert:

4.3.1. Nutzungsverhalten (FF1)

Zur Beschreibung des Umgangs der Studierenden mit dem Chatbot wurden drei Subkategorien gebildet: Nutzungszweck, Nutzungsart und Nutzungshäufigkeit.

Die Subkategorie Nutzungszweck wurde weiter in folgende Sub-Subkategorien untergliedert: Unterstützung bei der Ideen- und Themenfindung, Unterstützung bei der Ausarbeitung, Fachliche Hilfe, Didaktische Unterstützung, zukünftige Nutzung und Sonstiges.

4.3.2. Wahrgenommene Nützlichkeit (FF2a)

Die zweite Hauptkategorie umfasst die Einschätzung der Studierenden, wie sehr und in welcher Hinsicht der Chatbot bei der Arbeit im Seminar geholfen hat. Folgende Subkategorien wurden identifiziert: Unterstützung in der Ideenfindung*, Unterstützung in der Ideenausgestaltung*, Zeitersparnis durch den Chatbot*, Didaktische Unterstützung*, Lehrplanspezifische Unterstützung*, Fachliche Unterstützung*, Vergleich des Chatbots mit Dozierenden, Zukünftige Unterrichtsplanung*, Sonstiges.

Die mit Stern (*) gekennzeichneten Subkategorien wurden zusätzlich mit einer evaluativen Analyse untersucht. Zur Bewertung der Wahrnehmung wurde ein abgestuftes Ausprägungsschema verwendet: sehr negative Wahrnehmung, überwiegend negative

Wahrnehmung, gemischte Wahrnehmung, überwiegend positive Wahrnehmung, sehr positive Wahrnehmung sowie Wahrnehmung nicht klassifizierbar.

4.3.3. Wahrgenommene Benutzerfreundlichkeit (FF2b)

Diese Hauptkategorie bezieht sich auf die Einschätzung, ob und in welcher Hinsicht die Arbeit mit dem Chatbot einfach oder schwierig war. Die Subkategorien lauten: Formulierung von Eingaben*, Benutzeroberfläche*, Darstellung und Umfang der Antworten*, Anpassungsfähigkeit und Präzision*, Sonstiges.

Auch in diesem Fall wurde für die mit Stern (*) versehenen Subkategorien eine evaluative Einschätzung anhand des oben beschriebenen Ausprägungsschemas vorgenommen.

4.3.4. Externe Variablen (FF3)

Zur Erfassung kontextueller oder individueller Einflussfaktoren auf die Nutzungserfahrung wurden folgende Subkategorien identifiziert:

Personenbezogene Merkmale, eigene Ideen zur Station, Vorwissen und Fachkompetenz, unterrichtsbezogene Erfahrung, Erfahrung mit der Nutzung von Chatbots, allgemeine Eigenschaften von Chatbots, organisatorische Rahmenbedingungen des Seminars, Verfügbarkeit von Ressourcen, didaktische und inhaltliche Vorgaben, externe Personen, individuelle Präferenzen und Interessen, individuelle Herangehensweise an die Stationsgestaltung, Vertrauen sowie Sonstiges.

5. Ergebnisse der Analyse

5.1. Nutzungsverhalten (FF1)

5.1.1. Nutzungszweck

Die Auswertung zeigt, dass der Chatbot von allen Studierenden für verschiedene Zwecke genutzt wurde. Ein zentrales Einsatzfeld war die Ideen- und Themenfindung „zur Inspiration [...] in welche Richtung es ungefähr gehen soll“ (B1). Auch bei der Ausarbeitung der Station, teils auf Grundlage eigener Ideen, teils auf Basis mit dem Chatbot entwickelter Ansätze, kam der Chatbot bei allen Studierenden zum Einsatz. Fünf der sechs Studierenden ließen sich bei fachlichen Fragen vom Chatbot helfen. Alle sechs berichteten, dass sie didaktische Unterstützung erhalten oder gesucht haben.

„Ich weiß auch nicht, ob der [Chatbot] dazu da ist, aber manchmal dachte ich mir so, ist das jetzt überhaupt interessant? Passt das zeitlich jetzt so? Sind es jetzt 30 Minuten? Ist das jetzt zu viel in diesen 30 Minuten? Ist das jetzt interessant gestaltet? Ist da ein roter Faden drin? Ja, solche Sachen habe ich ihn auch sehr viel gefragt.“ (B5)

Einzelne Studierende nannten auch andere Verwendungszwecke des Chatbots, etwa zur Erstellung von Grafiken oder als Programmierhilfe. Fünf Studierende äußerten zudem Überlegungen, wie sie einen

solchen Chatbot im Lehrberuf einsetzen könnten, hauptsächlich zur Inspiration bei der Unterrichtsplanung:

„Also ich denke auch später im Unterricht und so ist man eigentlich sehr dankbar, wenn man sich das Ganze ein bisschen verkürzen kann und Stunden-vorbereitungen einfach sich auf jeden Fall die Inspiration und schon mal so ein Grundkonstrukt durch die AI holt.“ (B1)

5.1.2. Nutzungsart

Auch die Art der Interaktion mit dem Chatbot unterschied sich deutlich. Einige Studierende formulierten präzise Anforderungen und forderten wiederholte Anpassungen der Vorschläge, während andere eher offen vorgingen und ihre Anfragen bewusst vage hielten.

„Ich habe einfach genau das geschrieben, wie ich den Aufbau von meinem Versuch haben will und da musste ich halt einen längeren Text schreiben, aber einfach, damit ich halt am Ende eine schöne Zusammenfassung habe.“ (B6)

Zwar wurden konkrete Prompttechniken nicht benannt, doch manche Studierende sprachen von solchen indirekt. Eine strategische Nutzung des Promptings war somit unterschiedlich stark ausgeprägt.

5.1.3. Nutzungshäufigkeit

Die Nutzung des Chatbots konzentrierte sich vor allem auf die ersten Semesterwochen des Seminars (siehe Abb. 3). In der zweiten Semesterwoche war der Einsatz noch verpflichtend, in der darauffolgenden Semesterwochen verwendeten alle Studierenden den Chatbot. Tendenziell nahm die Nutzung mit fortschreitender Konkretisierung deutlich ab. Nur ein Studierender gab an, den Chatbot „im Schnitt einmal pro Sitzung“ (B4) verwendet zu haben. B6 griff in späteren Semesterwochen noch einmal auf den Chatbot zurück. B5 konnte den Chatbot krankheitsbedingt seltener benutzen.

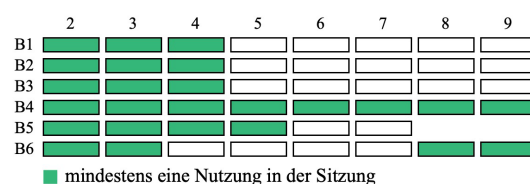


Abb. 3: Visualisierung der Nutzungshäufigkeit von der zweiten bis zur neunten Semesterwoche (Sitzung). Eine Nutzung wurde dann gezählt, wenn ein Teilnehmender erwähnte, dass der Chatbot in dieser Semesterwoche verwendet wurde, unabhängig von der Nutzungsdauer. Eigene Darstellung.

Die Studierenden wurden zudem gebeten, den Anteil der vom Chatbot beigetragenen Ideen im Vergleich zur Eigenarbeit bei ihrer finalen Station zu schätzen.

Tab. 1: Durch die Befragten wahrgenommene Gewichtung zwischen Chatbot-Unterstützung und Eigenleistung in Prozent.

Fall	B1	B2	B3	B4	B5	B6
Chatbot:Eigenleistung	50:50	25:75	65:35	30:70	50:50	20:80

Drei der sechs Studierenden schätzten die Eigenarbeit höher ein (siehe Tab. 1), zwei beschrieben eine gleichmäßige Aufteilung und eine Person gab an, dass der Chatbot den größeren Beitrag geleistet habe.

5.2. Wahrgenommene Nützlichkeit (FF2a)

Wie die detaillierte Übersicht in Tabelle 2 zeigt, nahm die Mehrheit der Studierenden den Chatbot bei der Konzeption der Station als überwiegend hilfreiches Werkzeug wahr.

Tab. 2: Evaluative Analyse der Wahrgenommenen Nützlichkeit. Legende: sehr negative Wahrnehmung (--), überwiegend negative Wahrnehmung (-), gemischte Wahrnehmung (+/-), überwiegend positive Wahrnehmung (+), sehr positive Wahrnehmung (++), Wahrnehmung nicht klassifizierbar (n.k.).

Subkategorie	B1	B2	B3	B4	B5	B6
Unterstützung in der Ideenfindung	+	+	+	+/-	+/-	++
Unterstützung in der Ideenausgestaltung	n.k.	+	++	+/-	++	++
Zeitersparnis durch den Chatbot	++	+	++	++	++	++
Didaktische Unterstützung	-	+	+	-	++	+
Lehrplanspezifische Unterstützung	++	+/-	+	+/-	+/-	+
Fachliche Unterstützung	++	+	++	+	++	+
Zukünftige Unterrichtsplanung	+	+	+	-	+/-	+/-

Bei näherer Betrachtung zeigen sich jedoch Diskrepanzen, sowohl zwischen den Teilnehmenden als auch zwischen den Subkategorien. So nimmt beispielsweise B3 alle Facetten der Nützlichkeit überwiegend positiv bis sehr positiv wahr, während andere (z.B. B4 und B5) kritischere Einschätzungen äußern.

Besonders positiv bewertet wurden die Zeitersparnis und die fachliche Unterstützung durch den Chatbot. Ein Studierender äußerte sich zur Zeitersparnis wie folgt:

„In gewisser Weise ja [...] Also, um erst mal in so eine Ideenfindung dann noch mal zu kommen, hätte man ja dann erst mal viel recherchieren müssen. Oder auch, wenn man physikalisch mal stehen geblieben ist. Auch da habe ich ja vorhin schon erwähnt, dass es recht hilfreich war.“ (B4)

In anderen Subkategorien, wie etwa der didaktischen Unterstützung, fällt das Bild negativer aus. Hier wurden sowohl nützliche Impulse durch den Chatbot als auch eine begrenzte Passgenauigkeit der Vorschläge benannt. Auch in Bezug auf Vorstellungen zur zukünftigen Nutzung von Chatbots für die spätere Unterrichtsplanung äußerten sich die Studierenden differenziert. Zwar wurden Potenziale für den Einsatz in der Unterrichtsplanung erkannt, gleichzeitig wurden aber auch Einschränkungen des Chatbots bezüglich

der Anpassungsfähigkeit an individuelle Situationen genannt.

Auch in Hinblick auf die Frage zu Vor- und Nachteilen des Chatbots gegenüber der Betreuung durch einen Dozierenden, ergibt sich ein differenziertes Bild. Ein zentraler Vorteil des Chatbots liegt in seiner ständigen Verfügbarkeit. Gleichzeitig heben viele Studierende die größere Tiefe und Präzision der Rückmeldung von Dozierenden hervor, insbesondere aufgrund deren schulpraktischer Erfahrung. Ein Studierender berichtete dies folgendermaßen:

„Also ich finde es gut, dass sie [die Dozierenden] eher so für das Physikalische da sind und wenn man da wirklich Nachfragen hat zu den Versuchen. Aber ich kann halt dem Chatbot halt unendlich viele Fragen stellen und auch egal wie nervig die sind und egal wie oft und ja, genau. [...] Also ich finde die ergänzen sich sehr gut.“ (B5)

5.3. Wahrgenommene Benutzerfreundlichkeit (FF2b)

Auch in Bezug auf die WB zeigt sich insgesamt ein überwiegend positives Bild. Tabelle 3 bietet eine detaillierte Übersicht der einzelnen Subkategorien. Dabei lassen sich zwei Tendenzen erkennen: Die eine Seite bewertet die Benutzerfreundlichkeit als durchweg positiv (B2, B3), während die andere ihr deutlich kritischer gegenübersteht (B1, B5, B6).

Tab. 3: Evaluative Analyse der Wahrgenommenen Benutzerfreundlichkeit. Legende: sehr negative Wahrnehmung (--), überwiegend negative Wahrnehmung (-), gemischte Wahrnehmung (+/-), überwiegend positive Wahrnehmung (+), sehr positive Wahrnehmung (++), Wahrnehmung nicht klassifizierbar (n.k.).

Subkategorie	B1	B2	B3	B4	B5	B6
Formulierung von Eingaben	+	++	+	+/-	-	++
Benutzeroberfläche	++	+	+	++	++	-
Darstellung und Umfang der Antworten	-	++	++	n.k.	-	n.k.
Anpassungsfähigkeit und Präzision	+	+	++	+/-	+/-	++

Die Formulierung von Eingaben wurde von vielen Teilnehmenden als intuitiv und problemlos beschrieben. Das gilt ebenso für die Benutzeroberfläche von ChatGPT, über die sie mit dem Chatbot kommunizierten. Gleichzeitig wurde angemerkt, dass längere Chatverläufe unübersichtlich werden können, sodass ältere Nachrichten nur schwer auffindbar sind.

Die Anpassungsfähigkeit und Präzision der Antworten wurden meist positiv wahrgenommen. Falls die Antworten nicht den Vorstellungen der Nutzenden entsprachen, wussten sich die Studierenden durch gezielte Rückfragen oder Umformulierungen zu helfen.

Am unterschiedlichsten fiel die Wahrnehmung der Darstellung und des Umfangs der Antworten aus. Während B2 die Gliederung der Antworten lobte und B3 den knappen Umfang der Antworten als hilfreich empfand, kritisierten andere genau diese Punkte. B1 erschienen die Antworten zu knapp, zudem wurde der

strukturelle Aufbau der Nachrichten negativ wahrgenommen. B5 empfand die Nachrichten als zu umfangreich.

5.4. Externe Variablen (FF3)

Im Rahmen des Interviews wurden zahlreiche externe Einflussfaktoren benannt, die den definierten Subkategorien zugeordnet wurden. Die genannten Variablen erstrecken sich dabei von personenbezogenen Merkmalen und individuellen Vorerfahrungen bis hin zu organisatorischen Rahmenbedingungen.

Alle Studierenden thematisierten ihr eigenes fachliches Vorwissen. So schilderte eine Person:

„Aber wenn man mal so physikalische Fragen hatte und sich jetzt nicht ganz klar war, auf was für ein Konzept/ also wieso sowas funktioniert, [...] dann habe ich auch manchmal den Chatbot gefragt und die Antworten fand ich eigentlich ganz gut, um drüber nachzudenken in der Richtung.“ (B4)

Die Erfahrungen, die mit der Nutzung von Chatbots außerhalb des Seminars gemacht wurden, variierten stark. Während B6 beispielsweise in einem früheren Seminar eine Einführung in Prompttechniken erhalten hat, hat B4 vor dem Seminar noch nie mit einem Chatbot gearbeitet.

Ähnlich unterschiedlich waren die unterrichtlichen Vorerfahrungen. B1 hatte bereits Teilzeit als Physiklehrkraft gearbeitet, B2 und B6 hatten hingegen noch keine Lehrerfahrung im Physikunterricht gesammelt.

Auch die individuellen Herangehensweisen an die Stationsgestaltung prägten die Aussagen der Studierenden. So war es für B3 beispielsweise wichtig, „neue Themen zu finden [...] was halt mal so out of the box ist“ (B3). B4 wollte sich dagegen lieber enger am Lehrplan orientieren.

Darüber hinaus spielte die Verfügbarkeit von Ressourcen für die Studierenden eine Rolle. Einige merkten dabei die ständige Verfügbarkeit des Chatbots an. Gleichzeitig bemängelte B6, dass für eine vergleichbare Nutzung außerhalb des Seminars die kostenpflichtige ChatGPT-Plus-Version erforderlich sei.

6. Diskussion

Die Ergebnisse der Studie zeigen, dass die Studierenden den Chatbot vielseitig genutzt haben. Sowohl aus den Aussagen zur zukünftigen Nutzung als auch aus der beobachteten zeitlichen Verteilung wird deutlich, dass die Studierenden das größte Potenzial eines solchen Chatbots als Inspirationsquelle und zur Ausarbeitung erster Ansätze sehen. Mit zunehmender Praxisarbeit bei der Stationsausgestaltung wurde der Chatbot weniger verwendet und es wurden andere Vorgehensweisen gewählt (z.B. Unterstützung durch Dozierende oder Eigenarbeit).

Die WN, die laut dem TAM das Nutzungsverhalten maßgeblich beeinflusst, wurde überwiegend positiv bewertet. Dabei berichten die Studierenden besonders häufig von einer Zeitersparnis durch den Chatbot.

Inwiefern diese empfundene Zeitersparnis objektiv messbar ist, bleibt offen und könnte Gegenstand zukünftiger Studien sein. Dass die Studierenden die didaktische Unterstützung deutlich negativer empfinden, mag am Systemprompt des Chatbots liegen. Zwar basiert der Chatbot auf didaktischen Prinzipien, er war jedoch nicht explizit darauf ausgelegt, vertiefte didaktische Rückfragen zu beantworten.

Auch die Benutzerfreundlichkeit des Chatbots fiel den Studierenden überwiegend positiv auf, wies dabei aber ähnliche Auffälligkeiten wie die WN auf. Teilnehmende, die den Chatbot als weniger nützlich empfanden, bewerteten seine Benutzerfreundlichkeit ähnlich. Dieser Befund ist im Sinne des TAM plausibel, da die WB als direkter Prädiktor für die WN gilt.

Allgemein lassen sich die beobachteten Diskrepanzen in den Bewertungen zwischen den Studierenden sowie zwischen den einzelnen Subkategorien durch die identifizierten EV erklären. Personenspezifische Variablen wie Vorerfahrungen mit Chatbots können für intraindividuelle Unterschiede in der Bewertung einzelner Subkategorien verantwortlich sein. Dagegen können für die unterschiedliche Bewertung von Subkategorien EV verantwortlich sein, welche Rahmenbedingungen außerhalb des Einflusses der Studierenden umfassen. Dazu zählt auch die bereits angesprochene Priorisierung von Funktionen im Systemprompt des Chatbots.

Beim Vergleich mit der Betreuung durch die Dozierenden wurde der Chatbot von vielen als hilfreiche und jederzeit verfügbare Ergänzung wahrgenommen, jedoch nicht als Ersatz für die Dozierenden. Die persönliche Expertise der Lehrperson schien für die Studierenden unersetzbar zu sein. Auch Damköhler et al. (2025) stellten fest, dass ein Großteil der Studierenden einen Chatbot in der Reflexionsphase nicht als Person wahrnimmt. Eine hybride Lernumgebung, in der ein KI-Werkzeug mit menschlicher Unterstützung kombiniert wird, scheint demnach vielversprechend zu sein.

Die Vielzahl identifizierter EV zeigt, wie vielschichtig und komplex die Einflüsse auf die WN und WB bei der Nutzung eines Chatbots sein können. Durch die Identifikation der EV kann der Chatbot besser an die Bedürfnisse der Studierenden angepasst werden. Insofern liefert diese Studie wichtige Impulse für die vielfach geforderte Förderung von KI-Kompetenz. Eine gezielte Anpassung zukünftiger Chatbots kann zu einer frühzeitig positiven Nutzungserfahrung mit KI bei angehenden Lehrkräften führen.

Limitationen

Bei der Interpretation der Ergebnisse ist zu berücksichtigen, dass der Chatbot ausschließlich während der Seminarsitzungen zugänglich war. Dadurch können die WB, WN und das Nutzungsverhalten potenziell negativ beeinflusst werden. Darüber hinaus handelt es sich um ein spezifisches fachdidaktisches Setting im Physik-Lehramtsstudium, dessen Ergebnisse nicht ohne Weiteres auf andere Einsatzbereiche der

Lehrkräftebildung übertragbar sind. Ein weiterer limitierender Faktor ist, dass kein Vergleich zu einer Kontrollgruppe, die ohne Chatbot-Unterstützung gearbeitet hat, vorgenommen wurde. Dadurch bleibt beispielsweise offen, wie groß der tatsächliche Effizienzzuwachs durch den Chatbot ist. Schließlich ist die Stichprobe mit sechs Teilnehmenden klein, sodass keine Sättigung bei den identifizierten Sub- und Subkategorien zu erwarten ist.

7. Ausblick

Zukünftige Studien könnten den Chatbot auf Basis der Ergebnisse weiterentwickeln und gezielt zur Förderung von KI-Kompetenz in der Lehrkräftebildung einsetzen. Eine API-basierte Lösung würde eine flexiblere Nutzung auch außerhalb des Seminars ermöglichen. Zudem wäre die Messung des Effizienzgewinns durch eine Kontrollgruppe förderlich. Darüber hinaus können weiterführende quantitative Erhebungen dazu beitragen, die Wirkung der identifizierten EV auf die WN und WB systematisch zu untersuchen.

8. Literatur

- Ali, I., Warraich, N. F., & Butt, K. (2024). Acceptance and use of artificial intelligence and AI-based applications in education: A meta-analysis and future direction. *Information Development*, 02666669241257206. <https://doi.org/10.1177/02666669241257206>
- Beldad, A. D., & Hegner, S. M. (2018). Expanding the Technology Acceptance Model with the Inclusion of Trust, Social Influence, and Health Valuation to Determine the Predictors of German Users' Willingness to Continue using a Fitness App: A Structural Equation Modeling Approach. *International Journal of Human-Computer Interaction*, 34(9), 882–893. <https://doi.org/10.1080/10447318.2017.1403220>
- Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., ... Amodei, D. (2020). Language Models are Few-Shot Learners. *Curran Associates Inc.*, 1877–1901.
- Bybee, R. W., Taylor, J. A., Gardner, A., Van Scooter, P., Powell, J. C., Westbrook, A., & Landes, N. (2006). The BSCS 5E Instructional Model: Origins, Effectiveness, and Applications. BSCS.
- Damköhler, J., Elsholz, M., & Trefzger, T. (2023). Förderung der Reflexionskompetenz im Lehr-Lern-Labor: Unterstützung der Entwicklung der Reflexionskompetenz Studierender in einem Lehr-Lern-Labor-Seminar. *PhyDid B - Didaktik der Physik - Beiträge zur DPG-Frühjahrstagung 2023*, 195–198.
- Damköhler, J., Lutz, W., & Trefzger, T. (2024a). Das Würzburger KI-Projekt: ChatGPT als Reflexionscoach im Lehr-Lern-Labor-Seminar Physik (J. Huwer, S. Becker-Genschow, C. Thyssen, A. Finger, L. von Kotzebue, E. Krenser, M. Meier, & T. Bruckermann, Hrsg.; 1. Aufl., S. 70–73). Waxmann.
- Damköhler, J., Lutz, W., & Trefzger, T. (2024b). Der digitale Dozent: ChatGPT als Co-Pilot in der Lehrpersonenbildung. *PhyDid B - Didaktik Der Physik - Beiträge Zur DPG-Frühjahrstagung 2024*, 1(1). <https://ojs.dpg-physik.de/index.php/phydid-b/article/view/1412>
- Damköhler, J., Lutz, W., & Trefzger, T. (2025). ChatGPT als Reflexionscoach. Einblicke in das Würzburger Lehr-Lern-Labor. In H. van Horst (Hrsg.), *Entdecken, lehren und forschen im Schülerlabor. Jahrestagung der Gesellschaft für Didaktik der Chemie und Physik in Bochum 2024*. Im Druck.
- Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. *Management Science*, 35(8), 982–1003.
- Davis, F. D., & Granić, A. (2024). *The Technology Acceptance Model: 30 Years of TAM*. Springer International Publishing. <https://doi.org/10.1007/978-3-030-45274-2>
- Dresing, T., & Pehl, T. (2018). *Praxisbuch Interview, Transkription & Analyse: Anleitungen und Regelsysteme für qualitativ Forschende* (8. Auflage). Eigenverlag.
- Elsholz, M., & Trefzger, T. (2017). Professionalisierung durch Praxisbezug – Begleitforschung zu den Würzburger Lehr-Lern-Laboren. *Gesellschaft für Didaktik der Chemie und Physik Jahrestagung in Zürich 2016*. In C. Maurer (Hrsg.), *Implementation fachdidaktischer Innovation im Spiegel von Forschung und Praxis* (S. 488–491). Universität Regensburg.
- Granić, A., & Marangunić, N. (2019). Technology acceptance model in educational context: A systematic literature review. *British Journal of Educational Technology*, 50(5), 2572–2593. <https://doi.org/10.1111/bjet.12864>
- Hidi, S., & Renninger, K. A. (2006). The Four-Phase Model of Interest Development. *Educational Psychologist*, 41(2), 111–127. https://doi.org/10.1207/s15326985ep4102_4
- Huwer, J., Becker-Genschow, S., Thyssen, C., & Thoms, L.-J. (2024). Kompetenzen für den Unterricht mit und über Künstliche Intelligenz.
- Jackson, C. M., Chow, S., & Leitch, R. A. (1997). Toward an Understanding of the Behavioral Intention to Use an Information System. *Decision Sciences*, 28(2), 357–389. <https://doi.org/10.1111/j.1540-5915.1997.tb01315.x>
- King, W. R., & He, J. (2006). A meta-analysis of the technology acceptance model. *Information &*

- Management, 43(6), 740–755.
<https://doi.org/10.1016/j.im.2006.05.003>
- Kircher, E. (2015). Methoden im Physikunterricht. In E. Kircher, R. Girwidz, & P. Häußler (Hrsg.), *Physikdidaktik: Theorie und Praxis* (3. Aufl., S. 141–192). Springer.
https://doi.org/10.1007/978-3-642-41745-0_5
- KMK (Kultusministerkonferenz). (2021). Handlungsempfehlung für die Bildungsverwaltung zum Umgang mit Künstlicher Intelligenz in schulischen Bildungsprozessen.
https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2024/2024_10_10-Handlungsempfehlung-KI.pdf
- Knoth, S., & Haider, M. (2023). Digitale Kompetenzen und Technologieakzeptanz bei angehenden Grundschullehrkräften. *Merz | Medien + Erziehung*, 67(3), 70–77.
<https://doi.org/10.21240/merz/2023.3.16>
- Kuckartz, U., & Rädiker, S. (2022). *Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung* (5. Aufl.). Beltz Juventa.
- Lutz, W., Damköhler, J., & Trefzger, T. (2025). Künstliche Intelligenz. Akzeptanz und AI-Literacy unter Lehramtsstudierenden. In H. van Horst (Hrsg.), *Entdecken, lehren und forschen im Schülerlabor. Jahrestagung der Gesellschaft für Didaktik der Chemie und Physik in Bochum 2024*. Im Druck.
- Ma, Q., & Liu, L. (2004). The Technology Acceptance Model: A Meta-Analysis of Empirical Findings. *Journal of Organizational and End User Computing*, 16(1), 59–72.
<http://dx.doi.org/10.4018/978-1-59140-474-3.ch006>
- Marineau, K. (2023, April 20). What is generative AI? IBM. <https://research.ibm.com/blog/what-is-generative-ai>
- Nawrath, D., Maisyenka, V., & Schecker, H. (2011). Experimentelle Kompetenz: Ein Modell für die Unterrichtspraxis. *Praxis der Naturwissenschaften - Physik in der Schule*, 60, 42–48.
- PhySX - Physikalische Schulexperimente Wiki. (o. J.). Abgerufen 27. März 2025, von <https://www.physikalische-schulexperimente.de/physo/Hauptseite>
- Puentedura, R. R. (2014, Dezember 10). SAMR, learning and assessment. *rrpweblog*.
<http://www.hippasus.com/rrpweblog/archives/2014/11/28/SAMRLearningAssessment.pdf>
- Sächsischer Bildungsserver. (o. J.). Experimente für den Physikunterricht. Abgerufen 27. März 2025, von <https://www.sachsen.schule/~physik/experi.php>
- Scherer, R., Siddiq, F., & Tondeur, J. (2019). The technology acceptance model (TAM): A meta-analytic structural equation modeling approach to explaining teachers' adoption of digital technology in education. *Computers & Education*, 128, 13–35. <https://doi.org/10.1016/j.compedu.2018.09.009>
- Staatsinstitut für Schulqualität und Bildungsforschung München. (2025). *LehrplanPLUS - Gymnasium—Fachlehrpläne*. https://www.lehrplanplus.bayern.de/schulart/gymnasium/inhalt/fachlehrplaene?w_schulart=gymnasium&wt_1=schulart&w_fach=physik&wt_2=fach
- Svendsen, G. B., Johnsen, J.-A. K., Almås-Sørensen, A.-S., & Vittersø, J. (2013). Personality and technology acceptance: The influence of personality factors on the core constructs of the Technology Acceptance Model. *Behaviour & Information Technology*, 32(4), 323–334.
<https://doi.org/10.1080/0144929X.2011.553740>
- Venkatesh, V., & Davis, F. D. (2000). A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. *Management Science*, 46(2), 186–204.
- Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*, 27(3), 425–478.
<https://doi.org/10.2307/30036540>
- Vo, G. M., & Pancratz, N. (2023, September 20). Vorstellungen von Lehramtsstudierenden zu künstlicher Intelligenz. *INFOS 2023 - Informatikunterricht zwischen Aktualität und Zeitlosigkeit*. <https://doi.org/10.18420/infos2023-005>
- von Garrel, J., & Mayer, J. (2025). Künstliche Intelligenz im Studium—Eine quantitative Längsschnittstudie zur Nutzung KI-basierter Tools durch Studierende [Application/pdf]. Hochschule Darmstadt.
https://doi.org/10.48444/H_DOCS-PUB-533
- Wang, S., Wang, F., Zhu, Z., Wang, J., Tran, T., & Du, Z. (2024). Artificial intelligence in education: A systematic literature review. *Expert Systems with Applications*, 252, 124167.
<https://doi.org/10.1016/j.eswa.2024.124167>
- Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E., Le, Q., & Zhou, D. (2022). Chain-of-Thought Prompting Elicits Reasoning in Large Language Models. *Curran Associates Inc.*, 24824–24837.
- Weizenbaum, J. (1976). *Computer Power and Human Reason: From Judgment to Calculation*. W H Freeman.
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (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), 39.
<https://doi.org/10.1186/s41239-019-0171-0>