

DIDAKTISCHE PROMPTS FÜR STUDIUM UND LEHRE ENTWICKELN – VERTIEFUNG UND PRAXISHILFE

INHALT

Einleitung	2
Materialüberblick	3
Wege zum eigenen Prompt	4
Das Entwurfsraster als Fundament	5
Modellwahl	6
Prinzipien für die Entwicklung didaktischer Prompts	8
A – Grundsätzliche Ausrichtung des Chatbots festlegen	8
B – Aufbau des Dialogs gestalten	11
C – Qualität der Ausgaben des Chatbots erhöhen	18
Baugerüst für das eigenständige Schreiben von Prompts	23
Prompts erproben und überarbeiten	24
Fehlersuche: Vom beobachteten Problem zum passenden Prinzip	26
Schluss	28
Literatur	29

AUTORINNEN: Gunda Mohr

DATUM: September 2026

URL: <https://uhh.de/hul-selma125>

LIZENZ:  Creative-Commons Namensnennung –
Weitergabe unter gleichen Bedingungen 4.0 International Lizenz

EMPFOHLENE ZITIERWEISE

Mohr, G. (2026). Didaktische Prompts für Studium und Lehre entwickeln – Vertiefung und Praxishilfe.
HUL-Selbstlernmaterial. <https://uhh.de/hul-selma125>

EINLEITUNG

Dieses Material soll dabei helfen, Prompts so zu gestalten, dass sie das Erreichen von Lehr-Lern-Zielen sinnvoll unterstützen, ohne eigenständiges Denken, Urteilen und Lernen zu verdrängen. Didaktische Prompts machen das Gespräch mit dem Chatbot zu einem Gestaltungselement für Studium und Lehre.

Wenn ein Chatbot nicht nur ein einzelnes Ergebnis liefern, sondern einen mehrschrittigen Lern- oder Arbeitsprozess begleiten soll, ist meist ein ausführlicherer, didaktisch strukturierter Prompt sinnvoll. Er kann den Chatbot anweisen, die Nutzenden schrittweise durch eine Aufgabe zu führen, passende Rückfragen zu stellen, auf typische Schwierigkeiten zu reagieren oder Reflexion anzuregen. So können beispielsweise Studierende passgenau und im richtigen Moment beim Selbststudium unterstützt werden. Solche didaktischen Prompts können einen Entwicklungsaufwand von mehreren Stunden erfordern. Ob dieser Aufwand gerechtfertigt ist, hängt vom Lernziel, vom Zusatznutzen und von der Wiederverwendbarkeit ab.

Beispiel für einen komplexen Prompt: [KI im Studium – deine Hilfe für fundierte Entscheidungen](#)

Ein Leitgedanke zieht sich durch dieses Material: Bei der Entwicklung didaktischer Prompts gibt es eine sinnvolle Arbeitsteilung zwischen Mensch und Chatbot. Die didaktisch anspruchsvolle Denkarbeit – Ziele klären, den Unterstützungsprozess entwerfen, Schwierigkeiten antizipieren, Erfolgskriterien festlegen – gehört in die Hand der Lehrenden. Sie ist im [Entwurfsraster](#) (siehe „[Materialüberblick](#)“) systematisch angelegt. Die Übersetzung dieser Überlegungen in einen funktionierenden Prompttext ist dagegen weitgehend Handwerk, bei dem aktuelle Sprachmodelle gut unterstützen können. Für komplexe didaktische Dialogprompts ist diese Arbeitsteilung bislang nicht direkt empirisch untersucht; sie stützt sich vor allem auf praktische Erprobung.

Einen indirekten Hinweis liefert Y. Zhou et al. (2023): Dort erreichten automatisch erzeugte Instruktionen bei 24 von 24 bzw. 17 von 21 untersuchten Aufgaben gleichwertige oder bessere Ergebnisse als die von Menschen formulierten. Zwei Einschränkungen begrenzen die Übertragung: Untersucht wurden kurze Instruktionen für eng umrissene Aufgaben, keine mehrschrittigen Dialogprompts. Außerdem erzeugte das Verfahren viele Alternativen und wählte anschließend anhand gelabelter Daten den messbar besten aus. Die Leistung beruhte damit wesentlich auf dieser Erzeugungs- und Auswahl-Schleife und nicht allein auf der Formulierung eines einzelnen Prompts.

Umso wichtiger ist die anschließende Testphase. Die Delegation der Formulierungsarbeit kürzt also nicht die didaktische Reflexion ab, sondern nur deren Verschriftlichung in Promptform – geprüft werden muss das Ergebnis in jedem Fall.

Dieses Dokument ist entsprechend **nicht als Text gedacht, der von vorne bis hinten durchgearbeitet werden muss**. Es dient zwei Zwecken: Es erklärt, warum bestimmte Promptgestaltungen funktionieren, und es hilft bei der gezielten Fehlersuche, wenn ein Prompt in der Erprobung nicht das tut, was er soll. Für den schnellen Weg zum funktionierenden Prompt gibt es ein separates, kurzes [Schnellpfad-Dokument](#) (siehe „[Materialüberblick](#)“).

MATERIALÜBERBLICK

Dieses Dokument ist Teil eines Materialpakets. Damit klar ist, welche Datei wofür gedacht ist, hier eine Übersicht:

Datei	Inhalt	Wann nutzen?
<u>Schnellpfad-Dokument</u>	Kurzer Ablaufleitfaden: Entwurfsraster erstellen → Prompt erzeugen lassen → erproben und nachschärfen	Standardweg für die meisten Anwendungsfälle
<u>Entwurfsraster</u>	Vorlage für die didaktische Planung: Block A für den ersten Entwurf, Block B nur bei Relevanz; Grundlage für Promptgenerierung und Test	Schritt 1 – selbst ausfüllen oder mit dem Planungsassistenten-Prompt erstellen
<u>Entwurfsraster – ausgefülltes Beispiel</u>	Beispielhaft ausgefülltes Entwurfsraster; zeigt auch, wie knapp die Antworten sein können	Zur Orientierung – nicht als Arbeitsvorlage verwenden
<u>Planungsassistenten-Prompt</u>	Prompt, der die Kernentscheidungen nacheinander abfragt und am Ende das ausgefüllte Entwurfsraster als reinen Text ausgibt	Schritt 1 – Alternative zum selbstständigen Ausfüllen des Entwurfsrasters
<u>Meta-Prompt</u>	Standard-Prompt, der aus dem ausgefüllten Entwurfsraster einen ersten Promptentwurf erstellt und Annahmen sowie Prüfstellen ausgibt	Schritt 2 – Weg 1 (Standardvorgehen nach dem Entwurfsraster)
<u>Kurzversion des Meta-Prompts</u>	Kürzere Alternative für einfache, überschaubare Anwendungsfälle oder wenn für die Prompterstellung nur ein leistungsschwächeres Modell zur Verfügung steht	Schritt 2 – Weg 1 bei einfachen Fällen
<u>Test-Checkliste</u>	Beobachtungshilfe und Protokoll für mehrere Testrunden; dokumentiert vor allem relevante Auffälligkeiten	Schritt 3 – ab dem ersten Promptentwurf
<u>Test-Checkliste – ausgefülltes Beispiel</u>	Beispiel für die Dokumentation mehrerer Testrunden und die Auswertung beobachteter Auffälligkeiten	Zur Orientierung bei der ersten Erprobung

Datei	Inhalt	Wann nutzen?
Dieses Dokument (Didaktische Prompts für Studium und Lehre entwickeln – Vertiefung und Praxis-hilfe)	Prinzipien mit Begründungen, Beispielen und übertragbaren Vorlagen; Diagnose-Tabelle zur Fehlersuche; Baugerüst für das eigenständige Schreiben	Schritt 2 – Weg 4; außerdem Vertiefung und Fehlersuche
<u>Prompt-Workbook</u>	Sammlung erprobter Prompts	Schritt 2 – Wege 2 und 3; außerdem als Fundus und Beispielquelle

WEGE ZUM EIGENEN PROMPT

Vor der Promptentwicklung sollten Datenschutz, Prüfungsregeln, zugelassene Werkzeuge, Chancengleichheit beim Zugang, Barrierefreiheit, Verantwortung und Folgen möglicher Fehler geklärt werden.

Erst danach wird die bewusste Entscheidung über den Weg bei der Promptentwicklung getroffen. Die didaktische Planung im [Entwurfsraster](#) ist für alle Wege der Ausgangspunkt. Sie kann selbst ausgefüllt oder mithilfe des [Planungsassistentz-Prompts](#) im Dialog erstellt werden; in beiden Fällen liegt anschließend dasselbe ausgefüllte Raster vor. Für die eigentliche Prompterstellung stehen vier Wege zur Verfügung; sie lassen sich auch kombinieren. Am Ende steht immer die Erprobung (siehe Abschnitt [„Prompts erproben und überarbeiten“](#)).

1. **Prompt vom Chatbot erzeugen lassen (Standardweg).** Das ausgefüllte [Entwurfsraster](#) und der [Meta-Prompt](#) werden genutzt, um mit Hilfe eines leistungsfähigen Chatbots daraus einen ersten Promptentwurf zu erstellen. Für einfache, überschaubare Anwendungsfälle kann stattdessen die [Kurzversion des Meta-Prompts](#) genutzt werden. Dieser Weg ist im [Schnellpfad-Dokument](#) Schritt für Schritt beschrieben und für die meisten Anwendungsfälle der effizienteste. Die didaktische Qualität hängt dabei an der Qualität der Inhalte des Entwurfsrasters, nicht an eigenen Formulierungskünsten im Prompting.
2. **Einen vorhandenen Prompt anpassen.** Ein ähnlicher Prompt aus dem [Prompt-Workbook](#) wird als Ausgangsbasis genutzt und an die eigenen Bedürfnisse angepasst. Das lohnt sich vor allem, wenn ein Workbook-Prompt dem eigenen Vorhaben bereits nahekommt.
3. **Einen lernbegleitenden Prompt direkt im Dialog entwickeln.** Im Prompt-Workbook befindet sich ein Prompt, mit dem ein erster Entwurf schrittweise im Gespräch entwickelt werden kann. Dieser Weg ist vom [Planungsassistentz-Prompt](#) zu unterscheiden: Der Planungsassistentz-Prompt endet mit einem ausgefüllten [Entwurfsraster](#); hier entsteht bereits der lernbegleitende Prompt. Das ausgefüllte Raster dient dabei als inhaltliche Grundlage für die Antworten der Nutzenden.

4. **Vollständig selbst schreiben.** Das ist der aufwendigste Weg, führt aber zum tiefsten Verständnis dafür, wie Prompts wirken – und ist die richtige Wahl, wenn maximale Kontrolle über jede Formulierung gewünscht ist. Dafür stehen in diesem Dokument die Prinzipien mit ihren übertragbaren Vorlagen sowie das Baugerüst zur Verfügung.

Unabhängig vom Weg gilt: Auch von Chatbots erzeugte oder aus dem Workbook übernommene Prompts müssen getestet werden, da sie bei der Nutzung zu falschen Gesprächsverläufen, unerwünschten Abbrüchen etc. führen können. Wenn erprobte Prompts im Kollegium oder in gemeinsamen Sammlungen geteilt werden, können andere darauf aufbauen, statt bei null zu beginnen. Orte zum Teilen eigener Prompts sind zum Beispiel das [Prompt-Workbook](#) oder der [Offene Prompt-Katalog des KI-Campus](#).

DAS ENTWURFSRASTER ALS FUNDAMENT

Bevor es an das Formulieren oder Generieren des Prompts geht, werden im Block A des [Entwurfsrasters](#) die Kernentscheidungen für einen ersten Entwurf geklärt:

- Einsatzkontext: Wer soll den Prompt nutzen, in welcher Lern- oder Arbeitssituation und wobei entsteht Unterstützungsbedarf?
- Ziel: Was sollen die Nutzenden am Ende besser können oder erreicht haben?
- Arbeitsteilung und Unterstützung: Wobei und wie soll der Chatbot unterstützen, was sollen die Nutzenden selbst leisten und was soll der Chatbot nicht übernehmen?
- Fachliche Grundlage: Welche Inhalte, Kriterien, Definitionen, Theorien, Modelle oder Quellen soll der Chatbot verwenden?
- Informationsbedarf: Welche Informationen benötigt der Chatbot von den Nutzenden und wann soll er sie jeweils erfragen?
- Gesprächsablauf und Ergebnis: Wie soll das Gespräch grob ablaufen und welches Ergebnis soll am Ende vorliegen?
- Technische Rahmenbedingungen: In welchen Zielmodellen und Sprachen soll der Prompt funktionieren und wie umfangreich soll der Dialog ungefähr sein?
- Block B – nur wenn relevant: Gibt es besondere Eingaben, Situationen oder Verzweigungen, die berücksichtigt werden müssen? Welches Fehlverhalten des Chatbots wäre besonders kritisch? Woran lässt sich erkennen, dass der Prompt in diesem Anwendungsfall besonders gut funktioniert?

Genau diese Fragen strukturieren das [Entwurfsraster](#). Das ausgefüllte Raster erfüllt drei Funktionen zugleich: Es dokumentiert die didaktische Planung, dient als Grundlage für die Promptgenerierung bzw. -formulierung und liefert mit den besonderen Erfolgskriterien in Block B – sofern solche für den Anwendungsfall relevant sind – zusätzliche Prüfpunkte für den späteren Test. Allgemeine Beobachtungskriterien stehen bereits in der [Test-Checkliste](#). Das Raster ist damit die Klammer des Entwicklungsprozesses. Die Antworten fließen in den Promptentwurf ein und bestimmen auch dessen Aufbau; es können z. B. bestimmte Schrittabfolgen, Informationsabfragen oder Verzweigungen notwendig werden.

Kleiner Tipp: Es lohnt sich, das ausgefüllte [Entwurfsraster](#) auch nach der erfolgreichen Promptentwicklung weiter aufzubewahren, da es ein hilfreicher Ausgangspunkt bei der Entwicklung ähnlicher Prompts in der Zukunft sein kann.

MODELLWAHL

Zunächst werden die wichtigsten praktischen Hinweise zur Modellwahl zusammengefasst. Die anschließenden hervorgehobenen Passagen erläutern die zugrunde liegenden Forschungsbefunde und ihre Grenzen.

Praktische Hinweise

- Verfügbarkeit für die Nutzenden: Entscheidend ist nicht das beste existierende Modell, sondern das Modell, das den späteren Nutzenden tatsächlich zur Verfügung steht. Diese Angabe wird im [Entwurfsraster](#) in Block A unter „Technische Rahmenbedingungen“ erfasst.
- Leistungsfähigkeit bei komplexen Anweisungen: Komplexere didaktische Prompts sollten in einem leistungsfähigen, möglichst aktuellen Modell erprobt werden, das mehrteilige Anweisungen und längere Dialoge im konkreten Zielsetting zuverlässig befolgt. Eine große nominelle Kontextlänge allein ist kein ausreichendes Auswahlkriterium.
- Mehrschrittige Dialoge: Benötigte Informationen sollten möglichst früh erhoben werden, bevor der Chatbot sich auf endgültige Lösungen oder Bewertungen festlegt. Ein einzelner gelungener Testdurchlauf sagt wenig über die Verlässlichkeit eines komplexen Prompts aus. Hat sich ein längerer Dialog erkennbar auf einer frühen Fehlannahme festgefahren, ist ein Neustart mit zusammengefassten Angaben oft sinnvoller als weiteres Nachsteuern.
- Rahmenbedingungen: Welches Modell konkret am besten passt, hängt außerdem vom Einsatzkontext, von der gewünschten Sprache, von Datenschutz- und Lizenzfragen sowie der hochschulinternen Verfügbarkeit ab.
- Gegenprüfung bei Problemen: Wenn ein Prompt unerwartet versagt, kann ein Test in einem zweiten, leistungsfähigeren Modell helfen zu unterscheiden, ob eher der Prompt oder das verwendete Modell die Ursache ist.
- Veränderlichkeit der Modelle: Modelle werden fortlaufend aktualisiert, teils unter gleichbleibendem Namen. Ein erfolgreich getesteter Prompt kann sich nach einem solchen Wechsel anders verhalten; ein erneuter Kurztest ist dann sinnvoll (siehe „[Test-Checkliste](#)“).

Komplexe Anweisungen stellen selbst leistungsfähige Modelle vor messbare Schwierigkeiten. Wen et al. (2024) untersuchten in einem Benchmark (einem standardisierten Aufgabensatz zum Vergleich von Modellen), wie gut Modelle mehrere miteinander verknüpfte Vorgaben befolgen. Besonders schwierig waren Aufgaben, bei denen Bedingungen voneinander abhingen oder ineinander verschachtelt waren; mit zunehmender Verschachtelung nahm die Leistung ab.

Die berichteten Befunde beruhen auf den chinesischsprachigen Daten; eine englische Fassung wurde anschließend ebenfalls bereitgestellt. Auch auf CELLO, einem überwiegend chinesischsprachigen Benchmark mit komplexen Aufgabenbeschreibungen sowie langen, teils verrauschten und heterogenen Eingaben, blieben selbst leistungsfähige damalige Modelle deutlich unter vollständiger Anweisungsbefolgung (He et al., 2024). He et al. verglichen jedoch nicht systematisch eine kurze und eine lange Version derselben Aufgabe. Die Studie zeigt deshalb nicht, dass allein die Eingabelänge die schlechteren Ergebnisse verursacht hat.

Mehrschrittige Gespräche können zusätzlich dadurch problematisch werden, dass wichtige Informationen erst nach und nach vorliegen. In einer großen Simulationsstudie mit über 200.000 Dialogen, 15 Modellen und sechs analytischen Generierungsaufgaben sank die Leistung im Mittel um 39 %, wenn relevante Aufgabeninformationen nicht vollständig zu Beginn vorlagen, sondern über mehrere Gesprächsschritte nachgereicht wurden (Laban et al., 2026). In einem Zusatzexperiment mit 31 Instruktionen und zwei Modellen trat der Einbruch bereits auf, sobald die Information auf zwei Gesprächsschritte verteilt wurde; eine feinere Aufteilung verstärkte ihn nicht wesentlich. Im Hauptexperiment sank die Best-Case-Leistung deutlich weniger (–16 %) als die Zuverlässigkeit der Ergebnisse: Die Unzuverlässigkeit nahm um 112 % zu, und zwischen bestem und schlechtestem Durchlauf derselben Aufgabe lagen im Mittel 50 Prozentpunkte. Die Autoren führen dies unter anderem darauf zurück, dass Modelle früh Annahmen treffen, vorschnell Lösungen vorschlagen und später daran festhalten. Eine niedrigere Temperatur – eine Einstellung, die beeinflusst, wie stark Antworten variieren – behob das Problem in einem weiteren Zusatzexperiment nicht nennenswert; auch die beiden getesteten Reasoning-Modelle (Modelle mit besonderem Fokus auf komplexe Denk- und Schlussfolgerungsaufgaben) waren in vergleichbarem Ausmaß betroffen. Untersucht wurden englischsprachige analytische Aufgaben mit simulierten Nutzenden, keine didaktischen Hochschuldialoge.

Auch die Position von Informationen in längeren Eingaben kann eine Rolle spielen. Liu et al. (2024) fanden in kontrollierten Experimenten eine U-förmige Kurve: Informationen am Anfang und Ende einer Eingabe wurden besser genutzt als Informationen in der Mitte. Varianten mit erweitertem Kontextfenster – also einer größeren maximal verarbeitbaren Eingabelänge – nutzten Informationen innerhalb des von beiden Varianten verarbeitbaren Bereichs nicht besser als die jeweiligen Standardvarianten. Untersucht wurde allerdings das Auffinden von Informationen in Eingabedokumenten, nicht die Befolgung von Promptregeln. Die in diesem Material empfohlene Wiederholung weniger besonders wichtiger Kernregeln ist deshalb eine Robustheitsheuristik, keine direkt durch diese Studie geprüfte Gegenmaßnahme. Auch ausdrücklich angefordertes schrittweises Denken verbessert die Anweisungsbefolgung nicht automa-

tisch: Bei Li et al. (2025) verschlechterten Chain-of-Thought-Aufforderungen – also Aufforderungen, einen Lösungsweg ausdrücklich schrittweise zu entwickeln – die Einhaltung von Vorgaben auf IFEval bei 13 von 14 und auf Complex-Bench bei allen getesteten Modellen; bei bestimmten Format- und Wortwahlvorgaben konnten sie zugleich helfen. Mehrere solche Reasoning- bzw. Thinking-Varianten schnitten bei der Anweisungsbefolgung schlechter ab als ihre jeweiligen Basisvarianten. Untersucht wurden jedoch einzelne Aufgaben mit klar prüfbaren Vorgaben, keine mehrschrittigen didaktischen Dialoge.

PRINZIPIEN FÜR DIE ENTWICKLUNG DIDAKTISCHER PROMPTS

Die folgenden Prinzipien können dazu beitragen, komplexe Prompts so zu formulieren, dass sie in der Praxis gut funktionieren. Dabei geht es zunächst um die grundsätzliche Ausrichtung des Chatbots, dann um den Aufbau des Dialogs und danach um die Erhöhung der Qualität der Ausgaben des Chatbots.

Wenn im [Prompt-Workbook](#) geeignete Beispiele zu den Prinzipien existieren, sind diese jeweils als wörtliches Zitat eingefügt. Der Link in der Klammer dahinter führt zu dem vollständigen Prompt im Workbook.

Die hier formulierten Prinzipien beruhen primär auf praktischer Erprobung und werden punktuell durch Befunde aus der Lernforschung, der Mensch-Computer-Interaktion und der Forschung zu maschinellem Lernen gestützt. Eine spezifische empirische Evidenzbasis für didaktische Prompts in der Hochschullehre fehlt bisher noch weitgehend; entsprechend sind die Prinzipien als reflektierte Heuristiken zu lesen, nicht als evidenzgesicherte Standards.

Es gibt zwei Zugänge zu den Prinzipien: **systematisch** über die folgenden Beschreibungen – für alle, die sich das Repertoire erschließen wollen – und **problemorientiert** über die Diagnose-Tabelle im [Abschnitt „Fehlersuche“](#) – für alle, die aus der Erprobung eines Prompts kommen und ein konkretes Problem beheben wollen.

A – GRUNDSÄTZLICHE AUSRICHTUNG DES CHATBOTS FESTLEGEN

Die Prinzipien in diesem Abschnitt steuern das spätere Verhalten des Chatbots im Gespräch und beschreiben damit auch das didaktische Fundament.

Prinzip	Funktion in einem Satz
Chatbot eine Rolle und einen Arbeitsmodus zuweisen	Legt fest, aus welcher Perspektive und in welcher Art der Chatbot unterstützt.
Ziel- statt bloße Themenangabe machen	Benennt den Zweck des Gesprächs, nicht nur das Thema.

Prinzip	Funktion in einem Satz
Didaktisches Ziel konsequent berücksichtigen	Richtet die Unterstützung am angestrebten Kompetenzaufbau aus.
Eigenständigkeit bewusst schützen	Verhindert, dass der Chatbot zentrale Denk- und Lernarbeit abnimmt.

CHATBOT EINE ROLLE UND EINEN ARBEITSMODUS ZUWEISEN

Eine Rollenbezeichnung kann Perspektive, Ton und Kommunikationsstil des Chatbots bündeln. Für die didaktische Steuerung ist eine Rollenbezeichnung allein jedoch meist zu unbestimmt. Deshalb ist es sinnvoll, die gewünschte Rolle zusätzlich in konkrete, beobachtbare Verhaltensregeln zu übersetzen – zum Beispiel: zunächst Fragen stellen, gestufte Hinweise geben, nur anhand vorgegebener Kriterien bewerten oder keine eigenen Lösungsideen einbringen. Auch für die objektive Wissensleistung zeigte sich kein verlässlicher Vorteil einer Rollenangabe allein: In einer systematischen Untersuchung mit 162 Rollen, neun Modellen mit offenen Gewichten und 2.410 Multiple-Choice-Wissensfragen führte keine der untersuchten Rollenangaben zu einer statistisch signifikanten Verbesserung der Trefferquote gegenüber der Kontrollbedingung ohne Rolle; einzelne Rollen verschlechterten die Leistung sogar (Zheng et al., 2024). Zwei Einschränkungen sind wichtig: Gemessen wurde ausschließlich die Richtigkeit bei objektiven Wissensfragen – nicht Tonfall, Gesprächsführung oder didaktische Qualität; und geprüft wurden nur reine Rollenetiketten, keine Rollen mit ausformulierten Verhaltensregeln. Dass die hier empfohlene Übersetzung in beobachtbare Regeln besser wirkt, ist damit nicht belegt, sondern beruht auf praktischer Erprobung – ein guter Kandidat für einen eigenen Vergleichstest.

Eine Variante des Rollen-Prinzips wäre, dem Chatbot – ähnlich wie bei Kreativitätstechniken, die gezielt unterschiedliche Perspektiven vorgeben – verschiedene Rollen und Arbeitsmodi mitzugeben, die der Chatbot an verschiedenen definierten Stellen nutzen soll.

Beispiel:¹ „Deine Rolle ist die eines Coaches. Deine Aufgabe ist, meine Selbstreflexion zu unterstützen. Dabei bringst du keine eigenen inhaltlichen Ideen ein, sondern stellst ausschließlich Fragen.“ ([Fragender Coach](#))

Übertragbare Vorlage: „Arbeite in folgendem Modus: [Rolle oder Funktion]. Das bedeutet konkret: [beobachtbare Verhaltensregeln]. Vermeide dabei: [unerwünschtes Verhalten].“

ZIEL- STATT BLOßE THEMENANGABE MACHEN

Ein guter Prompt benennt nicht nur das Thema, sondern auch den Zweck des Gesprächs. Die klare Zielbestimmung erhöht die Chance, dass der Prompt genau die Art von Unterstützung liefert, die tatsächlich gebraucht wird.

¹ Wird in diesen Beispielen und den übertragbaren Vorlagen von „ich“ geschrieben, ist immer die Person gemeint, die den Prompt nutzt. Dies können je nach Prompt Lehrende oder Studierende sein.

Beispiel: „Ich möchte herausfinden, wie ich wichtige Aspekte, aber ggf. auch neue Aspekte, die ich bisher eher weniger bedacht habe, noch stärker in meiner Lehre berücksichtigen kann.“

([Lehr-Ideen-Café](#))

Übertragbare Vorlage: „Ziel dieses Gesprächs ist, dass ich am Ende [konkretes Ergebnis / konkrete Einsicht / konkrete Entscheidungsgrundlage] habe.“

DIDAKTISCHES ZIEL KONSEQUENT BERÜCKSICHTIGEN

Didaktische Prompts werden besonders stark, wenn sie vom gewünschten Lehr-Lern-Ziel ausgehend gedacht werden. Es geht also nicht nur um das Ergebnis am Ende des Gesprächs, sondern vor allem darum, welche Kompetenzen die Nutzenden mit Unterstützung des Chatbots auf dem Weg dorthin aufbauen sollen. Durch diesen Fokus wird klarer, was im Gesprächsverlauf passieren soll und was vermieden werden sollte.

Beispiel: „Sehr wichtig ist mir, dass die Studierenden bei der Nutzung des Prompts in einem Chatbot selbst die Kompetenz aufbauen, die gestellte Aufgabe zu lösen und auch erkennen, was der Vorteil für sie ist, dies selbst zu tun.“ ([Prompt zur Entwicklung lernbegleitender Prompts](#))

Übertragbare Vorlage: „Richte deine Unterstützung konsequent darauf aus, dass [Zielgruppe] im Verlauf des Gesprächs [gewünschtes Wissen / gewünschte Fertigkeiten / gewünschte Einstellungen] entwickelt.“

EIGENSTÄNDIGKEIT BEWUSST SCHÜTZEN

Im Kontext Studium und Lehre ist es wichtig, dass der Chatbot angewiesen wird, zentrale Denk- und Lernarbeit nicht abzunehmen. Eine Befragung von Wissensarbeitenden zeigte, dass höheres Vertrauen in generative KI mit weniger selbstberichtetem kritischem Denken zusammenhing. Zugleich hing höheres aufgabenspezifisches Vertrauen in die eigene Fähigkeit mit mehr selbstberichtetem kritischem Denken zusammen (Lee et al., 2025); aus dem Studiendesign folgt allerdings keine Kausalaussage. Ein verwandter Befund zur Bedeutung der konkreten Gestaltung der KI-Unterstützung stammt aus einem großen Feldexperiment mit knapp 1.000 Schülerinnen und Schülern im Mathematikunterricht: Wer beim Üben uneingeschränkter Zugang zu einem GPT-4-basierten Chatbot hatte, schnitt zwar in den Übungsphasen besser, in der anschließenden Prüfung ohne KI-Zugang jedoch schlechter ab als die Vergleichsgruppe ohne KI. Eine zweite Variante desselben Chatbots nutzte aufgabenspezifische didaktische Leitplanken und sollte auf Grundlage der von Lehrkräften vorbereiteten Materialien gestufte Hinweise geben, statt Lösungen offenzulegen. Sie vermied den negativen Prüfungseffekt weitgehend (Bastani et al., 2025). Zur Einordnung gehört, wie diese Leitplanken konkret aussahen: Der Prompt der Tutor-Variante enthielt für jede einzelne Übungsaufgabe die korrekte Lösung (ausdrücklich zur Vermeidung von Halluzinationen), eine Sammlung typischer Schülerfehler und die dazu passenden Hinweise. Die Autoren beschreiben diesen Aufbau als arbeitsintensiv. Die Wirkung beruhte also nicht allein auf der Anweisung ‚gib Hinweise statt Lösungen‘, sondern auf einer inhaltlichen Verankerung des Prompts im Aufgabenmaterial. Für die eigene Promptentwicklung heißt das: Wo Lösungswege, Bewertungskriterien oder typische Fehlvorstellungen vorliegen und ohne Rechte- oder Datenschutzprobleme eingefügt werden können,

erhöht das die Chance, dass die Hilfen wirklich tragen (siehe Prinzip [„Didaktische Hilfen gezielt einbringen lassen“](#)). Allerdings war die Leitplanken-Variante in der abschließenden Prüfung statistisch nicht von der Gruppe ohne KI-Zugang zu unterscheiden – ein positiver Lerneffekt zeigte sich also nicht. Didaktische Leitplanken sind nach diesem Befund zunächst ein Mittel, um Schaden zu vermeiden; dass sie darüber hinaus das Lernen verbessern, ist bislang nicht gezeigt. Auch wenn diese Befunde aus einem spezifischen Schulmathematik-Kontext stammen, passen sie zum Kern dieses Prinzips: In der untersuchten Situation war ein Chatbot mit aufgabenspezifischen Leitplanken, der gestufte Unterstützung statt direkter Lösungen geben sollte, für die anschließende eigenständige Prüfungsleistung weniger problematisch als ein weitgehend unbeschränkter Chatbot. Ob sich dies auf andere Fächer, Altersgruppen oder didaktische Promptformen überträgt, muss gesondert geprüft werden.

Zur ehrlichen Einordnung gehört allerdings: In der Studie von Bastani et al. (2025) konnten die Lernenden in der Leitplanken-Bedingung ausschließlich diese eine Chatbot-Variante nutzen; der Zugriff auf andere Websites und Anwendungen war untersagt und die Einhaltung wurde während der Sitzungen aktiv überwacht. Im Studienalltag können Studierende einen lernbegleitenden Prompt dagegen jederzeit umgehen – etwa indem sie ihn abwandeln oder schlicht einen neuen Chat ohne den Prompt beginnen. Didaktische Prompts sind deshalb ein Unterstützungsangebot, kein Kontrollinstrument. Ihre Wirkung hängt wesentlich davon ab, dass Studierende den Nutzen der geführten Bearbeitung für ihr eigenes Lernen erkennen – genau dies unterstützen die Prinzipien „Thema, Vorgehen, Nutzen und Ziel des Dialogs kommunizieren lassen“ und „Nutzende durch den Prozess führen lassen“.

Beispiel: „Es ist sehr wichtig, dass du mir keine Arbeit abnimmst und mich alles selbst machen lässt, da ich sonst nichts dabei lerne.“ ([Forschungsfrage formulieren](#))

Übertragbare Vorlage: „Unterstütze mich so, dass ich die Aufgabe selbst bewältigen lerne. Nimm mir das Denken und Entscheiden nicht ab, sondern fördere meine Eigenständigkeit.“

B – AUFBAU DES DIALOGS GESTALTEN

Neben der Grundausrichtung ist es auch wichtig zu beschreiben, wie sich der Chatbot in einzelnen Situationen verhalten soll. Hierdurch lässt sich die didaktische Gestaltung verfeinern und typisches problematisches Gesprächsverhalten des Chatbots reduzieren.

Den folgenden Prinzipien liegen zentrale Merkmale guter Lernunterstützung zugrunde: adaptive Passung, schrittweise Zurücknahme der Hilfen und Verantwortungsübergabe an die Lernenden. Diese Merkmale werden in einem Review zur Unterrichtsinteraktion als drei Kernmerkmale von Scaffolding herausgearbeitet (van de Pol et al., 2010). Die Wirksamkeitsevidenz war zu diesem Zeitpunkt noch begrenzt. Eine Meta-Analyse von 144 experimentellen Studien zu computerbasiertem Scaffolding berichtet im Mittel positive Effekte auf kognitive Lernergebnisse (Belland et al., 2017). Eine spätere Unterrichtsstudie zeigt zugleich, dass stärker adaptive Unterstützung nicht automatisch besser wirkt; ihre Wirkung hing unter anderem davon ab, wie viel Zeit Lernende zuvor eigenständig gearbeitet hatten (van de Pol et al., 2015). Diese Befunde beziehen sich nicht auf LLM-gestützte Dialoge und belegen daher nicht unmittelbar die Wirksamkeit der hier vorgeschlagenen Promptgestaltung.

Prinzip	Funktion in einem Satz
<u>Thema, Vorgehen, Nutzen und Ziel des Dialogs kommunizieren lassen</u>	Sorgt für Orientierung und Motivation zu Gesprächsbeginn.
<u>Nutzende durch den Prozess führen lassen</u>	Erklärt an jeder Stelle, was passiert und warum.
<u>Kontext zuerst erheben, Antwort später geben lassen</u>	Macht Vorschläge passgenauer statt allgemein.
<u>Relevanz von Ehrlichkeit betonen lassen</u>	Verbessert die Eingabequalität der Nutzenden.
<u>Schrittfolge statt Gesamtauftrag beschreiben</u>	Zerlegt komplexe Aufgaben in nacheinander bearbeitete Schritte.
<u>Interaktion drosseln</u>	Begrenzt Fragen und Impulse pro Antwort.
<u>Umgang mit unklaren oder unerwarteten Antworten regeln</u>	Verhindert, dass der Chatbot auf dünner Grundlage weiterarbeitet.
<u>Wahlpunkte und Verzweigungen verwenden</u>	Gibt Nutzenden Einfluss auf den Gesprächsverlauf.
<u>Konsistenz im Gesprächsverlauf sichern</u>	Hält frühere Entscheidungen im weiteren Verlauf stabil.
<u>Iterationen zur Weiterentwicklung einbauen</u>	Behandelt erste Lösungen als Entwurf mit Überarbeitungsschleifen.
<u>Mehrere Optionen statt nur einer Antwort fordern</u>	Fördert Vergleichen, Prüfen und Auswählen.
<u>Didaktische Hilfen gezielt einbringen lassen</u>	Steuert Hinweise, gestufte Hilfen und Reflexionsimpulse.
<u>Warnhinweise ausgeben lassen</u>	Macht auf Folgen ungünstiger Entscheidungen und geltende Regeln aufmerksam.
<u>Reflexion am Ende einer Lernsituation anregen lassen</u>	Sichert das Gelernte durch abschließende Reflexion.
<u>Umsetzung und nächste Schritte absichern lassen</u>	Unterstützt den Übergang vom Ergebnis ins Handeln.
<u>Gesprächsende klar steuern</u>	Legt fest, wann und wie der Dialog endet.

THEMA, VORGEHEN, NUTZEN UND ZIEL DES DIALOGS KOMMUNIZIEREN LASSEN

Wenn die Nutzenden den Prompt verwenden, ohne ihn vorher zu lesen, kennen sie Ziel, Nutzen und Vorgehen noch nicht. Daher sollte der Chatbot dazu gebracht werden, diese Informationen zu Beginn des Gesprächs zu geben. Gerade wenn Studierende den Prompt nicht selbst gewählt haben, kann das dazu beitragen, das Verständnis für die Aufgabe und die Motivation zu fördern.

Beispiel: „Stelle dich, deine Aufgabe und deine Rolle kurz vor und erkläre mir knapp Ziel und Nutzen dieses Gesprächs.“ ([KI-Antworten kritisch prüfen](#))

Übertragbare Vorlage: „Stelle dich zu Beginn kurz in deiner Rolle vor und erkläre mir knapp, worum es in diesem Gespräch geht, wie du vorgehen wirst, welches Ziel das Gespräch hat und warum dieses Vorgehen für mich hilfreich sein kann. Formuliere dies klar, verständlich und motivierend. Beginne erst danach mit dem ersten Schritt.“

NUTZENDE DURCH DEN PROZESS FÜHREN LASSEN

Die Nutzenden sollten wissen, was sie im Gesprächsverlauf erwartet und warum die einzelnen Schritte sinnvoll sind. Deshalb sollte der Chatbot an den passenden Stellen im Prompt angewiesen werden, dies den Nutzenden zu erläutern. Das fördert bei den Nutzenden das Verständnis für den Ablauf und verringert die Wahrscheinlichkeit, dass sie Informationen zu früh eingeben und sich über scheinbare Wiederholungen wundern.

Beispiel: „Nun betrachten wir nacheinander die drei Prüfbereiche Inhaltliche Qualität, Quellen und Belege und Passung zur eigenen Frage und zum eigenen Verwendungszweck, um eine differenziertere Einschätzung der KI-Antwort entwickeln zu können.“

([KI-Antworten kritisch prüfen](#))

Übertragbare Vorlage: „Gib mir zu Beginn jedes neuen Schritts einen kurzen Überblick darüber, worum es darin geht und warum er sinnvoll ist. Erkläre mir danach jeweils an den passenden Stellen kurz, was wir als Nächstes tun, warum dieser einzelne Schritt sinnvoll ist und worauf es dabei ankommt. Formuliere dies knapp, gut verständlich und nur so ausführlich, wie es für die Orientierung nötig ist.“

KONTEXT ZUERST ERHEBEN, ANTWORT SPÄTER GEBEN LASSEN

Viele gute Prompts funktionieren deshalb gut, weil der Chatbot zunächst den relevanten Kontext erhebt, bevor er Vorschläge macht. Dadurch wird das Ergebnis passgenauer und bleibt weniger allgemein.

Beispiel: „Frage mich nach: meinem Studienfach und Semester, meinen bisherigen Erfahrungen mit KI-Tools (genutzte Tools, jeweils Grad der Erfahrung damit), dem Themenbereich, um den es gleich gehen wird, meinem inhaltlichen Vorwissen darin.“

([KI im Studium – deine Hilfe für fundierte Entscheidungen](#))

Übertragbare Vorlage: „Bevor du Vorschläge machst, erhebe bitte zunächst den nötigen Kontext. Frage mich dazu Schritt für Schritt nach [notwendige Informationen für eine passende Unterstützung].“

RELEVANZ VON EHRlichkeit BETONEN LASSEN

In didaktischen Dialogen mit einem Chatbot ist es oft so, dass die Qualität späterer Empfehlungen oder Einschätzungen stark davon abhängt, wie ehrlich, differenziert und realistisch die Nutzenden kommunizieren. Der Chatbot sollte daher im Prompt aufgefordert werden, die Nut-

zenden darauf hinzuweisen. Dies gilt besonders bei Reflexions-, Beratungs- oder Entscheidungsprozessen, in denen der Chatbot nicht nur Informationen verarbeitet, sondern auf Basis der Selbstbeschreibung der nutzenden Person Unterstützung strukturiert. Gerade dort, wo eine realistische Selbsteinschätzung schwerfällt oder Wissenslücken ungenannt benannt werden, kann dieser Hinweis dazu beitragen, dass die gemachte Eingabe zutreffender und das Unterstützungsangebot des Chatbots dadurch passgenauer wird.

Beispiel: „Weise mich darauf hin, dass deine späteren Einschätzungen nur so gut sein können, wie meine Antworten ehrlich sind.“ ([KI im Studium – deine Hilfe für fundierte Entscheidungen](#))

Übertragbare Vorlage: „Erkläre mir knapp, welche Angaben für eine passende Unterstützung benötigt werden. Bitte mich um möglichst zutreffende und ausreichende Antworten, ohne mich zu sensiblen oder identifizierenden Details zu drängen. Weise mich darauf hin, dass Fragen übersprungen und personenbezogene Angaben verallgemeinert werden können. Erkläre mir sachlich, welche Einschränkungen sich daraus für deine Unterstützung ergeben.“

SCHRITTFOLGE STATT GESAMTAUFTRAG BESCHREIBEN

Komplexe Aufgaben können für den Chatbot besser bearbeitbar und im Dialog leichter steuerbar werden, wenn der Prompt sie in klar getrennte, nacheinander zu bearbeitende Teilschritte zerlegt, statt nur einen Gesamtauftrag zu formulieren. Für bestimmte formale Denk- und Schlussfolgerungsaufgaben zeigte sich, dass eine Zerlegung in einfachere, sequenziell gelöste Teilprobleme die Bearbeitung durch das Modell verbessern kann (D. Zhou et al., 2023). Untersucht wurden dabei keine didaktischen Dialoge.

In didaktischen Dialogen kann eine sichtbar gemachte Schrittfolge den Nutzenden außerdem als Verfahrensmodell dienen.

Beispiel: „Bitte gehe immer erst zum nächsten Schritt weiter, wenn ich dir die Antwort zum vorherigen Schritt gegeben habe und dieser Schritt abgeschlossen ist.“
([Forschungsfrage formulieren](#))

Übertragbare Vorlage: „Arbeite die Aufgabe in klar getrennten Schritten nacheinander ab. Gehe immer erst dann weiter, wenn der vorherige Schritt abgeschlossen ist.“

INTERAKTION DROSSELN

Chatbots neigen dazu, viele Fragen, Anregungen oder Informationen in eine einzige Antwort zu packen. Daher werden viele Dialoge besser, wenn der Chatbot angewiesen wird, den Nutzenden nur eine Frage oder nur wenige Impulse auf einmal auszugeben. Der Vorteil ist, dass der Dialog übersichtlicher bleibt und die Nutzenden nicht mit zu vielen Anforderungen auf einmal überfordert werden.

Beispiel: „Bitte stelle immer nur eine Frage auf einmal.“
([Prompt zur Entwicklung lernbegleitender Prompts](#))

Übertragbare Vorlage: „Stelle immer nur eine Frage auf einmal, außer wenn für den aktuellen Schritt unbedingt mehrere Angaben gleichzeitig nötig sind.“

UMGANG MIT UNKLAREN ODER UNERWARTETEN ANTWORTEN REGELN

Ein Prompt kann festlegen, wie der Chatbot reagieren soll, wenn Antworten unklar, unvollständig oder nicht passend sind. So bleibt der Prozess auch dann brauchbar, wenn die Nutzenden nicht so antworten wie erwartet. Ohne solche Regeln gehen Chatbots bei unklaren Eingaben oft trotzdem weiter, füllen Lücken eigenmächtig oder produzieren Ergebnisse auf zu dünner Grundlage.

Beispiel: „Erst wenn ich deine Frage aus Schritt 1 hinreichend beantwortet habe, mache mit Schritt 2 weiter.“ ([Schrittweise Bewertungsraster für eine effiziente Beurteilung von studentischen Arbeitsergebnissen erstellen](#))

Übertragbare Vorlage: „Wenn meine Antwort unklar, unvollständig oder nicht passend ist, frage gezielt nach, bevor du weitermachst. Gehe erst dann zum nächsten Schritt über, wenn die notwendige Information vorliegt.“

WAHLPUNKTE UND VERZWEIGUNGEN VERWENDEN

Ein guter Prompt muss nicht starr linear sein. Es kann sehr sinnvoll sein, an bestimmten Stellen Auswahlmöglichkeiten zu geben und den weiteren Verlauf davon abhängig zu machen. Auf diese Weise bestimmen die Nutzenden stärker den Verlauf des Dialogs, was den Dialog passender und authentischer machen kann. Das kann dazu beitragen, dass Nutzende Prozess und Ergebnis stärker als eigene Entscheidung erleben und eher mittragen.

Beispiel: „Dabei frage mich, welche der folgenden Leitfragen ich bei der Weiterentwicklung zuerst fokussieren möchte.“ ([Ideen für einladende Aufgabenstellungen bekommen](#))

Übertragbare Vorlage: „Gib mir mehrere klar unterscheidbare Optionen. Lass mich zunächst wählen, bevor du weitermachst. Richte die nächsten Schritte an meiner Auswahl aus.“

KONSISTENZ IM GESPRÄCHSVERLAUF SICHERN

Bei mehrschrittigen Prompts kann es sinnvoll sein, den Chatbot ausdrücklich dazu anzuhalten, frühere Entscheidungen und bereits festgelegte Richtungen im weiteren Gesprächsverlauf konsequent zu berücksichtigen. So wird eher vermieden, dass der Chatbot den Verlauf eigenmächtig verändert oder frühere Festlegungen aus dem Blick verliert.

Beispiel: „Wenn du mir nach der Auswahl einer neuen Frage neue Antwortoptionen gibst, sollten diese immer zu meinen bisher getroffenen Entscheidungen passen.“ ([Ideen für einladende Aufgabenstellungen bekommen](#))

Übertragbare Vorlage: „Bleibe konsistent zu meinen bisherigen Entscheidungen und verändere eingeschlagene Richtungen nur, wenn ich das ausdrücklich möchte.“

ITERATIONEN ZUR WEITERENTWICKLUNG EINBAUEN

Viele Ergebnisse profitieren davon, dass eine erste Lösung nicht sofort als Endergebnis behandelt, sondern gezielt als Entwurf weiterentwickelt wird. Dazu kann der Prompt mehrere Überarbeitungsrunden oder Verfeinerungsschleifen vorsehen. Solche Iterationen behandeln die erste Lösung bewusst als Entwurf und machen Überarbeitungsschritte und Entscheidungen sichtbar. Ob sie im jeweiligen Anwendungsfall auch zu besseren Ergebnissen oder Lernprozessen führen, sollte in der Erprobung geprüft werden.

Beispiel: „Nutze meine Antwort, um 2 optimierte, detailliertere Vorgehensvarianten zu skizzieren.“ ([KI im Studium – deine Hilfe für fundierte Entscheidungen](#))

Übertragbare Vorlage: „Schritt 1: Erstelle einen ersten Vorschlag. Bitte mich anschließend um eine Rückmeldung dazu. Schritt 2: Nutze meine Rückmeldung, um einen überarbeiteten Vorschlag zu erstellen. Frage mich anschließend, ob ich noch Änderungswünsche habe und wenn ja, welche konkreten Änderungswünsche ich habe, oder ob ich im Gesamtprozess weitermachen möchte. Wenn ich noch Änderungswünsche benenne, wiederhole Schritt 2. Wenn ich keine Änderungswünsche mehr habe, fahre im Gesamtprozess fort.“

MEHRERE OPTIONEN STATT NUR EINER ANTWORT FORDERN

Gerade in didaktischen Kontexten ist es oft hilfreich, bewusst mehrere Vorschläge erzeugen zu lassen. Der Vorteil ist, dass Nutzende so eher vergleichen, prüfen und auswählen, statt einen ersten Vorschlag einfach unüberlegt zu übernehmen.

Beispiel: „Wenn ich dir diese Informationen gegeben habe, verfasse bitte 3 Versionen der Nachricht. Dies ermöglicht mir, die passendste auszusuchen.“ ([Willkommensmail](#))

Übertragbare Vorlage: „Erstelle bitte [Anzahl] unterschiedliche Versionen / Vorschläge, damit ich sie vergleichen und gezielt auswählen und weiterentwickeln kann.“

DIDAKTISCHE HILFEN GEZIELT EINBRINGEN LASSEN

Lernförderliche Prompts geben nicht nur Fragen und Schritte vor, sondern bauen auch didaktische Unterstützung direkt ein und steuern diese. Der Chatbot wird angewiesen, an bestimmten Stellen Hinweise, Zusatzinformationen, gestufte Hilfen oder Reflexionsimpulse zu geben. Dabei kann vorgegeben werden, auf welche Quellen der Chatbot sich beziehen soll, und es können Beispiele für passende Hilfestellungen formuliert werden. Damit wird die gewünschte Art der Hilfestellung konkreter beschrieben und kann im anschließenden Test gezielt daraufhin geprüft werden, ob sie zum Lehr-Lern-Ziel passt.

Wenn der Anwendungsfall mehrere vergleichbare Bearbeitungsschritte, Übungsfälle oder Durchgänge vorsieht, kann die Unterstützung im Verlauf bewusst zurückgenommen werden: zunächst ausführlichere Hilfe, später zuerst ein eigener Versuch und nur noch ein knapper Hinweis. Wie stark und über welchen Zeitraum diese Unterstützung zurückgenommen wird, hängt vom Lernziel und vom Aufbau des Einsatzes ab.

Beispiel: „Wenn ich eine Unsicherheit auswähle, gib mir dazu ein oder zwei möglichst konkrete und realistische Ideen, wie ich diese Unsicherheit weiter klären könnte. Wenn es sinnvoll ist und die Ideen sich deutlich unterscheiden, kannst du auch noch eine dritte Möglichkeit nennen. Achte darauf, dass die Vorschläge: genau zu der ausgewählten Unsicherheit passen, mit möglichst überschaubarem Aufwand umsetzbar sind, mir helfen, gezielt weiterzuarbeiten, sich deutlich voneinander unterscheiden, falls du mehr als einen Vorschlag nennst, so beschrieben sind, dass auch eine Person, die so etwas noch nie gemacht hat, nachvollziehen kann, wie sie dabei praktisch vorgehen kann; dazu sollen die nötigen einzelnen Schritte so beschrieben werden, dass eine direkte Umsetzung möglich ist. Solche Anregungen können zum Beispiel sein: eine zu umfangreiche Unsicherheit in kleinere Teilfragen zerlegen, damit nicht alles auf einmal geklärt werden muss, ...“ ([KI-Antworten kritisch prüfen](#))

Übertragbare Vorlage: „Wenn [bestimmte Schwierigkeit] auftritt, gib [Art der Hilfe, z. B. Hinweise, Zusatzinformationen, gestufte Hilfen, Reflexionsimpulse oder konkrete nächste Schritte]. Nutze dafür [Quelle / Material / vorgegebene Kriterien / Beispielhilfen]. Formuliere die Unterstützung so, dass sie [z. B. verständlich, konkret, realistisch, direkt umsetzbar und passend zum Lehr-Lern-Ziel] ist.“

WARNHINWEISE AUSGEBEN LASSEN

Der Prompt sollte den Chatbot anweisen, bei ungünstigen Entscheidungen der Nutzenden auf mögliche negative Folgen hinzuweisen sowie ggf. vorhandene relevante institutionelle, fachliche oder lehrbezogene Richtlinien ausdrücklich in Erinnerung zu rufen. Dadurch werden die Nutzenden angeregt, ein nicht empfehlenswertes Vorgehen vor der Umsetzung noch einmal zu überdenken oder im Idealfall so anzupassen, dass ein empfehlenswertes Vorgehen daraus wird.

Ein Sonderfall solcher Warnhinweise betrifft den Datenschutz: Gerade bei Reflexions- und Beratungsprompts kann der Chatbot angewiesen werden, die Nutzenden zu Beginn darauf hinzuweisen, keine sensiblen persönlichen Angaben und keine personenbezogenen Daten Dritter in das Gespräch einzugeben.

Beispiel: „Bitte weise sicherheitshalber außerdem darauf hin, dass der Einsatz von KI mit der Lehrperson besprochen werden sollte und es wichtig ist, sich an die Vorgaben der Hochschule, des Fachbereichs und der Lehrperson zu halten.“

([KI im Studium – deine Hilfe für fundierte Entscheidungen](#))

Übertragbare Vorlage: „Weise sicherheitshalber darauf hin, dass [getroffene Entscheidung] zu [mögliche negative Auswirkungen] führen kann und deshalb [Präventionsmaßnahmen] durchgeführt werden sollten.“

REFLEXION AM ENDE EINER LERNSITZUNG ANREGEN LASSEN

Gerade bei didaktischen Prompts kann es sinnvoll sein, den Dialog mit einer kurzen Reflexion zu beenden. So wird bewusst gemacht und festgehalten, was gelernt wurde und was eventuell noch nicht geschafft ist. Die Förderung von Selbstbewertung und Reflexion ist eines der sieben

von Nicol und Macfarlane-Dick (2006) formulierten Prinzipien guten Feedbacks und wird dort in ein Modell selbstregulierten Lernens eingeordnet. Die Empfehlung, einen Chatbot-Dialog mit einer Reflexion abzuschließen, ist eine didaktische Übertragung.

Beispiel: „Ask students to reflect on their learning after completing the simulation.“

[\(Diskussionen üben und Feedback erhalten\)](#)

Übertragbare Vorlage: „Bitte schließe den Prozess mit einer kurzen Reflexion ab, indem du mich fragst: Was hast du gelernt? Was ist dir noch unklar? Was nimmst du für die nächste Arbeitsphase mit?“

UMSETZUNG UND NÄCHSTE SCHRITTE ABSICHERN LASSEN

Ein Prompt kann am Ende den Übergang in konkretes Handeln unterstützen, wenn dies aus didaktischer Sicht angebracht ist. Dadurch wird aus guten Überlegungen eher gelebte Praxis.

Beispiel: „Erstelle aus der gewählten Variante einen konkreten Umsetzungsplan ohne Zeitangaben, aber mit der Überschrift ‚dein Umsetzungsplan‘ und allen einzelnen notwendigen Schritten.“

[\(KI im Studium – deine Hilfe für fundierte Entscheidungen\)](#)

Übertragbare Vorlage: „Leite aus dem Ergebnis zum Schluss einen konkreten nächsten Schritt oder einen kurzen Umsetzungsplan ab.“

GESPRÄCHSENDE KLAR STEUERN

Ein Prompt wird oft robuster, wenn er nicht nur den Einstieg und Verlauf, sondern auch das Ende des Dialogs festlegt. So weiß der Chatbot, wann der Prozess abgeschlossen ist und wie er den Abschluss gestalten soll. Dies vermeidet, dass der Prozess durch weitere, wenig hilfreiche Anschlussangebote unnötig verlängert oder verwässert wird.

Beispiel: „Schließe am Ende mit Erfolgswünschen für das Gelingen der Aufgabe ab und mache keine weiteren Angebote. Außerdem weise mich noch freundlich hervorgehoben darauf hin, dass es hilfreich sein könnte, wenn ich mir die gefundene Lernstrategie an einer passenden Stelle abspeichere, sodass sie mir auch noch später zur Verfügung steht.“

[\(KI im Studium – deine Hilfe für fundierte Entscheidungen\)](#)

Übertragbare Vorlage: „Schließe den Prozess am Ende in folgender Form ab: [gewünschte Abschlussform]. Mache danach keine weiteren Angebote und öffne keinen neuen Gesprächsschritt.“

C – QUALITÄT DER AUSGABEN DES CHATBOTS ERHÖHEN

Die folgenden Prinzipien können dazu beitragen, dass die Ergebnisse passender, verlässlicher, differenzierter und somit besser nutzbar werden.

Nicht für jedes der folgenden Prinzipien liegt derzeit ein passendes, erprobtes Beispiel aus dem Prompt-Workbook vor. In diesen Fällen wird bewusst nur die übertragbare Vorlage gezeigt.

Prinzip	Funktion in einem Satz
<u>Ausgabeformat eindeutig definieren</u>	Macht Ergebnisse direkt weiterverwendbar.
<u>Didaktische und fachliche Qualitätskriterien direkt in den Prompt integrieren</u>	Gibt dem Chatbot Maßstäbe für gute Antworten mit.
<u>Unangemessene Verkürzungen und Vereinfachungen vermeiden</u>	Verhindert künstliche Eindeutigkeit bei unsicheren Sachverhalten.
<u>Transparenz von Unsicherheiten stärken</u>	Lässt Annahmen, Unsicherheiten und Prüfbedarf offenlegen.
<u>Kritische Selbstprüfung vor der Ausgabe fordern</u>	Lässt Antworten vor der Ausgabe an Kriterien prüfen.
<u>Unterrepräsentierte Perspektiven gezielt mit einbeziehen</u>	Verringert blinde Flecken durch dominante Sichtweisen.
<u>Strukturelle Antworttendenzen des Chatbots gezielt ausgleichen</u>	Wirkt systematischen Verzerrungen in Antwortmustern entgegen.
<u>Bestätigungstendenzen des Chatbots gezielt begrenzen (Sycophancy)</u>	Reduziert bloß gefällige statt ehrlicher Rückmeldungen.
<u>Weitere Optimierungen nur bei echtem Mehrwert anbieten</u>	Verhindert unnötige Überarbeitungsrunden und Verschlimmbesserungen.

AUSGABEFORMAT EINDEUTIG DEFINIEREN

Wenn ein Ergebnis in einer bestimmten Form weiterverwendet werden soll, lohnt es sich, das gewünschte Format im Prompt sehr präzise festzulegen. Dazu können Formatvorgaben, Längenbegrenzungen und bei Bedarf auch konkrete Beispiele für die Zielstruktur gehören. Mit solchen Angaben wird das Ergebnis gleich so ausgegeben, dass es direkt weitergenutzt werden kann und nicht erst noch mühselig nachbearbeitet werden muss.

Beispiel: „WICHTIG: Halte dabei unbedingt bei jeder Frage folgende Hinweise zum Format ein: In der ersten Zeile steht die Frage ohne Nummer und ohne das Wort Frage davor.“

([Selbsttests anfertigen \(H5P Single Choice Test\)](#))

Übertragbare Vorlage: „Gib das Ergebnis in folgendem Format aus: [Formatbeschreibung]. Halte dich genau daran. Die Ausgabe soll maximal [Länge / Umfang] haben. Hier ist ein kurzes Beispiel für die gewünschte Struktur: [Beispiel].“

DIDAKTISCHE UND FACHLICHE QUALITÄTSKRITERIEN DIREKT IN DEN PROMPT INTEGRIEREN

Ein Prompt wird oft besser, wenn er dem Chatbot nicht nur eine Arbeitsanweisung gibt, sondern zusätzlich im Prompt deutlich macht, woran der Chatbot erkennt, dass seine Antwort gut ist. So erhält der Chatbot fachliche und didaktische Kriterien für die Qualität seiner Antwort. Wenn er sich in seiner Antwort ausdrücklich auf diese Kriterien bezieht, werden diese für die Nutzenden sichtbar. Dadurch wird auch für die Nutzenden deutlicher, worauf es bei guten Antworten ankommt.

Beispiel: „Bitte unterstütze mich dabei, eine einladende Aufgabenstellung zu entwickeln, die Neugier weckt, Verknüpfungen zu Vorwissen und ggf. auch der Praxis oder Erfahrungswelt der Studierenden fördert, Herausforderungen beinhaltet, die Kreativität anregt und die Zusammenarbeit fördert.“ ([Ideen für einladende Aufgabenstellungen bekommen](#))

Übertragbare Vorlage: „Richte deine Antwort an den folgenden Qualitätskriterien aus: [Kriterien]. Beziehe dich in deiner Antwort ausdrücklich auf diese Kriterien und mache sichtbar, inwiefern sie erfüllt sind.“

UNANGEMESSENE VERKÜRZUNGEN UND VEREINFACHUNGEN VERMEIDEN

Chatbots erzeugen oft Antworten, die fachlich plausibel und präzise wirken, ohne dass sie inhaltlich wirklich tragfähig sind. Aussagen können verkürzt oder zu stark vereinfacht sein – oder genauer formuliert, als es die Grundlage erlaubt. Es kann daher sinnvoll sein, den Chatbot ausdrücklich anzuweisen, kontroverse oder unsichere Sachverhalte nicht unnötig eindeutig, präzise oder fachlich glatt darzustellen. Damit wird der Chatbot ausdrücklich angewiesen, problematische Verkürzungen und künstliche Genauigkeit zu vermeiden.

Übertragbare Vorlage: „Formuliere fachliche Einschätzungen nur so eindeutig und präzise, wie es die Grundlage tatsächlich erlaubt. Vereinfache komplexe oder umstrittene Sachverhalte nicht unangemessen und vermeide es, unsichere Einschätzungen künstlich exakt oder fachlich glatter darzustellen, als sie sind.“

TRANSPARENZ VON UNSICHERHEITEN STÄRKEN

Eine umfassende Übersichtsarbeit dokumentiert für verschiedene Aufgaben der natürlichen Sprachgenerierung, dass Modelle flüssig formulierte Inhalte erzeugen können, die einer zugrunde liegenden Quelle widersprechen, von ihr nicht gestützt werden oder anderweitig unzutreffend sind (Ji et al., 2023). Zur begrifflichen Einordnung dieses Verhaltens sei ergänzend auf eine technikphilosophische Analyse verwiesen – ausdrücklich keine empirische Studie: Hicks et al. (2024) argumentieren dort, das Modell folge durchgehend demselben wahrscheinlichkeitsbasierten Generierungsprinzip und sei gleichgültig gegenüber der Frage, ob das Erzeugte zutrifft. Daher ist es gerade bei anspruchsvolleren Themen sinnvoll, Unsicherheiten, Annahmen und Prüfbedarf vom Chatbot explizit benennen zu lassen. Ebenso wichtig kann es sein, ihm ausdrücklich zu untersagen, bei fehlender breiter Datenbasis einfach irgendwelche Quellen, Fakten oder andere Angaben auszugeben. Solche Regeln adressieren das Risiko falscher oder unbelegter Angaben und können für die Nutzenden sichtbar machen, was besonders kritisch geprüft werden sollte. Wie zuverlässig sie Falschangaben im jeweiligen Zielmodell tatsächlich reduzieren, sollte im Test geprüft werden.

Übertragbare Vorlage: „Nutze für überprüfbare Aussagen vorrangig die im Prompt freigegebenen Quellen oder Materialien. Trenne klar zwischen belegten Aussagen, plausiblen Ableitungen und ungesicherten Vermutungen. Erfinde keine Quellenangaben. Kennzeichne nicht verifizierte Angaben als prüfbedürftig. Weise darauf hin, dass auch eine vom Chatbot geäußerte Sicherheit keine externe Prüfung ersetzt.“

KRITISCHE SELBSTPRÜFUNG VOR DER AUSGABE FORDERN

Der Chatbot kann dazu angehalten werden, Ergebnisse vor der Ausgabe anhand konkreter Kriterien zu prüfen. Qualitätssteigerungen sind vor allem für iterative Verfahren gezeigt worden, bei denen ein Modell zunächst Feedback zu seiner Ausgabe erzeugt und diese anschließend in weiteren Schritten überarbeitet (Madaan et al., 2023). Daraus folgt jedoch nicht, dass eine einmalige allgemeine Aufforderung zur Selbstprüfung zuverlässig hilft. Wenn Modelle ihre Antworten allein aus sich heraus korrigieren sollten, also ohne Rückmeldung von außen, verbesserten sie ihre Leistung bei Denk- und Schlussfolgerungsaufgaben in einer anderen Untersuchung nicht verlässlich; teilweise verschlechterte sie sich sogar (Huang et al., 2024). Für die praktische Promptgestaltung ist es daher plausibler, eine Selbstprüfung an konkrete Prüfkriterien zu binden als nur allgemein zur Selbstkorrektur aufzufordern. Auch dann sollte das Ergebnis bei relevanten Entscheidungen weiterhin extern kontrolliert werden.

Übertragbare Vorlage: „Prüfe deine Antwort vor der Ausgabe noch einmal kritisch anhand [genannter Ziele / Kriterien] und verbessere sie, wenn nötig.“

UNTERREPRÄSENTIERTE PERSPEKTIVEN GEZIELT MIT EINBEZIEHEN

Chatbots orientieren sich an Mustern in ihren Trainingsdaten. Dadurch können dominante Sichtweisen in Antworten leichter sichtbar werden als andere. Es kann daher sinnvoll sein, den Chatbot ausdrücklich anzuweisen, auch relevante unterrepräsentierte oder marginalisierte Perspektiven zu berücksichtigen. Das kann z. B. bei gesellschaftlichen, ethischen, politischen oder pädagogischen Fragen helfen, ein Thema differenzierter zu betrachten und blinde Flecken zu verringern.

Übertragbare Vorlage: „Beziehe in deine Antwort nicht nur dominante, sondern auch relevante unterrepräsentierte oder marginalisierte Perspektiven ein. Weise darauf hin, wenn bestimmte Sichtweisen in gängigen Darstellungen oft zu wenig berücksichtigt werden.“

STRUKTURELLE ANWORTTENDENZEN DES CHATBOTS GEZIELT AUSGLEICHEN

Chatbots können in ihren Antworten bestimmte Optionen, Deutungen oder Vorgehensweisen systematisch plausibler, attraktiver oder naheliegender erscheinen lassen als andere. Das liegt nicht nur an der konkreten Frage, sondern auch an typischen Antwortmustern des Chatbots. Es kann daher sinnvoll sein, den Chatbot ausdrücklich anzuweisen, solche wahrscheinlichen Verzerrungen aktiv auszugleichen, statt sie unbemerkt in die Antwort einfließen zu lassen. Das ist besonders dann hilfreich, wenn der Chatbot bei Abwägungen, Entscheidungen oder Reflexionen unterstützen soll.

Beispiel: „Du bist dir bewusst, dass du selbst eine KI bist und daher strukturell dazu neigst, KI-Nutzung attraktiver darzustellen als sie ist. Diesem Bias arbeitest du aktiv entgegen, indem du KI-Optionen und Nicht-KI-Optionen gleich konkret, gleich attraktiv und gleich kritisch formulierst.“
([KI im Studium – deine Hilfe für fundierte Entscheidungen](#))

Übertragbare Vorlage: „Berücksichtige, dass du bestimmte Optionen oder Sichtweisen aufgrund typischer Antwortmuster möglicherweise leichter plausibel oder attraktiv erscheinen lässt als andere. Arbeite diesem Bias aktiv entgegen, indem du relevante Alternativen gleich konkret, gleich kritisch und gleich differenziert darstellst.“

BESTÄTIGUNGSTENDENZEN DES CHATBOTS GEZIELT BEGRENZEN (SYCOPHANCY)

Chatbots können dazu neigen, ihre Antworten an den geäußerten Ansichten oder Präferenzen der nutzenden Person auszurichten, ihre Sichtweise zu bestätigen oder Rückmeldungen entsprechend positiver beziehungsweise negativer zu formulieren. Dieses Verhalten wird als Sycophancy bezeichnet. Sharma et al. (2024) wiesen unterschiedliche Formen davon bei fünf damals aktuellen KI-Assistenten in mehreren offenen Textgenerierungsaufgaben nach.

Die Befunde deuten darauf hin, dass menschliche Präferenzurteile zur Entstehung dieses Verhaltens beitragen können: Antworten, die den Ansichten der Nutzenden entsprachen, wurden eher bevorzugt. In einem weiteren Experiment entschieden sich sowohl Menschen als auch Präferenzmodelle in einem nicht unerheblichen Anteil der Fälle für überzeugend formulierte, sycophantische Antworten statt für wahrheitsgemäße Korrekturen. Die Studie legt allerdings keine alleinige Ursache fest; auch andere Bestandteile des Trainings können zu dem Verhalten beitragen (Sharma et al., 2024).

Für Lern-, Reflexions- oder Entscheidungsprozesse kann das problematisch sein, weil kritische Rückmeldungen dadurch abgeschwächt oder mögliche Probleme zu wenig deutlich benannt werden. Daher kann es sinnvoll sein, den Chatbot ausdrücklich aufzufordern, die nutzende Person nicht vor allem zu bestätigen oder ihr nach dem Mund zu reden.

Übertragbare Vorlage: „Gib mir keine bloß bestätigenden Rückmeldungen. Sei ehrlich, kritisch und konkret, auch wenn deine Einschätzung für mich unbequem sein könnte.“

WEITERE OPTIMIERUNGEN NUR BEI ECHTEM MEHRWERT ANBIETEN

Chatbots neigen bei Überarbeitungsprozessen dazu, auch dann neue Versionen anzubieten, wenn diese gegenüber dem bisherigen Entwurf keinen echten Mehrwert bringen oder sogar schlechter sind. Es kann daher sinnvoll sein, den Chatbot ausdrücklich anzuweisen, nur dann eine neue optimierte Version anzubieten, wenn sie gegenüber dem bisherigen Stand tatsächlich eine erkennbare Verbesserung darstellt. So wird eher vermieden, dass unnötig zusätzliche Zeit der Nutzenden verbraucht wird und Verschlimmbesserungen vorgenommen werden.

Übertragbare Vorlage: „Biete nur dann eine überarbeitete Version an, wenn sie gegenüber meinem bisherigen Entwurf inhaltlich, sprachlich oder funktional tatsächlich besser ist. Wenn keine echte Verbesserung möglich ist, sage das offen.“

BAUGERÜST FÜR DAS EIGENSTÄNDIGE SCHREIBEN VON PROMPTS

Das folgende Baugerüst richtet sich an alle, die den Weg „Selbst schreiben“ gewählt haben (siehe [„Wege zum eigenen Prompt“](#)). Es hilft dabei, einen eigenen didaktischen Prompt Schritt für Schritt zu entwickeln. Wer den Standardweg über die Promptgenerierung nutzt, braucht das Baugerüst nicht – dort übernimmt der Chatbot die Strukturierung.

Wie Entwurfsraster und Baugerüst zusammenhängen: Das [Entwurfsraster](#) klärt die didaktische Planung – also was im Gespräch passieren soll und warum. Das Baugerüst hilft anschließend dabei, diese Planung in einen Prompttext zu übersetzen. Die Inhalte des ausgefüllten Rasters fließen deshalb direkt in die Bausteine des Gerüsts ein:

Entwurfsraster (didaktische Planung)	Baugerüst (Promptarchitektur)
Einsatzkontext	Ausgangslage und Ziel; Kontext
Ziel	Ausgangslage und Ziel
Arbeitsteilung und Unterstützung	Rolle und Arbeitsweise des Chatbots
Fachliche Grundlage	Kontext; ggf. Qualitätsanforderungen
Informationsbedarf	Kontext
Gesprächsablauf und Ergebnis	Ablauf; Ausgabeformat; Abschluss
Technische Rahmenbedingungen	beeinflusst die Gesamtformulierung (Länge, Redundanz, Komplexität)
Besondere Eingaben, Situationen und Verzweigungen (Block B)	Ablauf (Hilfen, Verzweigungen, Regeln)
Besonders kritisches Fehlverhalten / besondere Erfolgskriterien (Block B)	Qualitätsanforderungen

Nicht jeder Prompt braucht alle zuvor beschriebenen Prinzipien. Für viele Zwecke reichen wenige bewusst ausgewählte Elemente aus. Welche davon sinnvoll sind, hängt vom jeweiligen Ziel und von den Ergebnissen des [Entwurfsrasters](#) ab. Das Baugerüst unterstützt dabei, die zentralen Elemente eines Prompts zunächst knapp zu skizzieren. Als Ergänzung können die übertragbaren Vorlagen genutzt werden.

Ausgangslage und Ziel

Ich bin ... [Lehrperson / Studierende:r / ...]

Unterstütze mich dabei, ... [konkretes Ziel / gewünschte Unterstützung]

Wichtig ist mir, dass am Ende ... [gewünschtes Ergebnis / gewünschter Lern- oder Arbeitsgewinn]

Rolle und Arbeitsweise des Chatbots

Arbeite als ... [Rolle / Arbeitsmodus]

Wichtig ist mir, dass du ... [gewünschte Arbeitsweise]

Kontext

Frage mich / die Studierenden nach folgenden Informationen: ...

Erfrage jede Information erst dann, wenn sie für den nächsten Schritt benötigt wird.

oder: Berücksichtige dabei insbesondere: ...

Ablauf

Gehe in folgenden Schritten vor: ...

Gehe immer erst dann weiter, wenn der vorherige Schritt abgeschlossen ist.

Qualitätsanforderungen

Achte darauf, dass dein Ergebnis ...

Ausgabeformat

Gib das Ergebnis als ... aus. [gewünschtes Format]

Abschluss

Fasse am Ende ... zusammen / leite daraus ... ab. [gewünschter Abschluss]

Wenn das Gerüst ausgefüllt ist, lässt sich aus den Elementen des Baugerüsts, den ausgewählten übertragbaren Vorlagen und ggf. weiteren eigenen Ergänzungen ein erster Promptentwurf erstellen. Oft genügt es, dabei die Reihenfolge logisch anzupassen und einzelne Übergänge genauer zu beschreiben.

Wenn nicht sicher ist, wie zuverlässig das verwendete Modell komplexe Anweisungen befolgt, gilt besonders: Prompt eher kurz halten, einfache Sprache verwenden, nur wenige Verzweigungen einbauen und die wichtigsten Regeln (z. B. ‚nur eine Frage auf einmal‘, ‚keine Lösungen vorgeben‘) an mehreren Stellen im Prompt wiederholen.

Gerade bei komplexeren Prompts ist nicht entscheidend, dass sofort eine perfekte Endfassung entsteht. Wichtiger ist, dass die zentralen Elemente enthalten sind.

PROMPTS ERPROBEN UND ÜBERARBEITEN

Gute Prompts entstehen selten in einem großen Wurf, sondern in der Regel durch mehrfache Erprobung und Überarbeitung. Es ist daher sinnvoll, Prompt-Entwicklung von Anfang an als iterativen Gestaltungsprozess zu verstehen. Das gilt unabhängig vom gewählten Weg – auch und gerade für Prompts, die ein Chatbot erzeugt hat.

Die [Test-Checkliste](#) enthält eine Beobachtungshilfe für den Gesprächsverlauf und ein Protokoll für mehrere Testrunden. Während des Tests dient die Kriterienliste zur Orientierung; nach dem Durchlauf werden nur relevante Auffälligkeiten dokumentiert. Besondere Erfolgskriterien aus Block B des [Entwurfsrasters](#) sowie gegebenenfalls Annahmen und Prüfstellen aus der Prompterzeugung werden zusätzlich berücksichtigt. Die folgenden Hinweise begründen das Vorgehen:

- Warum einen engagierten und einen schwierigen Durchlauf testen? Im engagierten Durchlauf wird so geantwortet, wie es bei einer Person zu erwarten wäre, die sich ernsthaft auf den Lernprozess einlässt. Im schwierigen Durchlauf werden knappe, unklare, ausweichende oder lösungsfordernde Antworten erprobt. So lässt sich sowohl prüfen, ob der normale Ablauf trägt, als auch, wie zuverlässig der Prompt das gewünschte Verhalten des Chatbots hervorruft. Für beide Durchläufe sollte ein konkreter, realistischer Testfall aus dem vorgesehenen Einsatzbereich verwendet werden.
- Warum jeweils einen neuen Chat nutzen und zusätzlich die Einstellungen beachten? Für jeden Test sollte ein neuer Chat verwendet werden. So wird verhindert, dass der bisherige Gesprächsverlauf das Ergebnis beeinflusst. Ein neuer Chat allein reicht jedoch nicht immer aus: Manche Chatbots nutzen zusätzlich Informationen aus früheren Gesprächen oder aus persönlichen Einstellungen und passen ihr Verhalten daran an. Auch solche persönlichen Einflüsse sollten für den Test ausgeschaltet werden. So lässt sich besser prüfen, ob der Prompt für sich genommen wie geplant funktioniert.
- Warum den Dialogverlauf prüfen, nicht nur das Ergebnis? Gerade bei schrittweisen Prompts lohnt es sich, nicht nur das Endergebnis, sondern auch den Dialogverlauf zu prüfen: Benimmt sich der Chatbot im Gespräch wie geplant? Ist der Dialog unnötig lang? Sind Anweisungen an die Nutzenden unklar oder doppelt? Benötigen die Nutzenden mehr Hilfen oder andere Hilfen im Gesprächsverlauf? Macht der Chatbot Dinge, die er eigentlich vermeiden sollte? Passt das Ausgabeformat? Werden im Prompt benannte Qualitätskriterien vom Chatbot eingehalten?
- Warum mehrfach testen? Sprachmodelle erzeugen auch bei identischen Eingaben unterschiedliche Antworten. Einzelne Testdurchläufe können daher ebenso zufällig gelingen wie zufällig misslingen. Aussagekräftiger wird die Erprobung über mehrere Durchläufe hinweg; Überarbeitungen sollten sich vorrangig an Problemen orientieren, die wiederholt auftreten – nicht an einmaligen Ausreißern. Soll der Prompt in mehreren Zielmodellen funktionieren, müssen die Tests außerdem in diesen Modellen wiederholt werden.
- Warum nachschärfen statt neu schreiben? Oft ist es einfacher, einen bestehenden Prompt gezielt nachzuschärfen, statt ihn neu zu schreiben. Die Diagnose-Tabelle im Abschnitt „[Fehlersuche](#)“ führt vom beobachteten Problem zum passenden Prinzip. Mögliche Ansatzpunkte sind z. B.: eine Frage präziser formulieren, einen Wahlpunkt ergänzen, eine Endregel klarer setzen, eine Hilfe einbauen oder typisches Fehlverhalten des Chatbots begrenzen. Genau dadurch wächst nach und nach ein Prompt, der nicht nur gut klingt, sondern auch in der Praxis tragfähig ist. Und wenn ein Prompt trotz Überarbeitung unerwartet versagt: erst in einem leistungsfähigeren Modell gegenprüfen (siehe „[Modellwahl](#)“), bevor der Prompt verworfen wird.
- Warum auch andere testen lassen? Hilfreich kann darüber hinaus sein, Prompts nicht nur selbst zu testen. Wenn Studierende mit einem Prompt arbeiten sollen, lohnt sich eine kleine Erprobung durch Personen, die nicht so vertiefte fachliche Kenntnisse haben wie die entwickelnde Person selbst. Schon wenige Rückmeldungen zeigen oft, ob Formulierungen missverständlich sind, ob ein Schritt zu viel abverlangt oder ob an einer Stelle eine wichtige Unterstützung fehlt. Gerade bei Prompts, die Lernprozesse begleiten sollen, ist diese Außensicht sehr wertvoll.

Die Checkliste prüft in erster Linie, ob der Prompt im vorgesehenen System und Anwendungsfall funktional arbeitet. Sie beurteilt nicht, ob der geplante Einsatz zulässig oder fachlich bzw. didaktisch vertretbar ist; das muss unabhängig davon vorab anhand der geltenden Rahmenbedingungen geklärt werden. Ob die Nutzung tatsächlich zu Kompetenzaufbau oder Transfer beiträgt, lässt sich ebenfalls nicht allein aus dem Chatverlauf ableiten. Dafür sind ergänzende Beobachtungen oder Arbeitsproben aus einem begrenzten Einsatz mit der Zielgruppe nötig.

FEHLERSUCHE: VOM BEOBACHTETEN PROBLEM ZUM PASSENDEN PRINZIP

Diese Tabelle ist der problemorientierte Zugang zu den Prinzipien. Sie ist vor allem für die Situation gedacht, dass ein Prompt in der Erprobung nicht wie gewünscht funktioniert: In der [Test-Checkliste](#) werden Auffälligkeiten nach dem Durchlauf dokumentiert; hier findet sich der Ansatzpunkt zur Behebung. Die genaue Begründung, ein Beispiel und eine übertragbare Vorlage stehen jeweils im zugehörigen [Prinzipienabschnitt](#). Wenn mehrere Prinzipien genannt sind, empfiehlt es sich, mit dem erstgenannten zu beginnen und nur so viele Anpassungen wie nötig vorzunehmen – jede zusätzliche Anweisung verlängert den Prompt und kann die Befolgung anderer Anweisungen schwächen (siehe Abschnitt [„Modellwahl“](#)).

Beobachtetes Problem	Passende Prinzipien
Der Chatbot gibt zu schnell eine fertige Lösung.	▪ Eigenständigkeit bewusst schützen ▪ Schrittfolge statt Gesamtauftrag ▪ Kontext zuerst erheben
Der Chatbot fragt zu viel auf einmal.	▪ Interaktion drosseln
Die Ergebnisse sind zu allgemein oder passen nicht zur Situation.	▪ Kontext zuerst erheben ▪ Ziel- statt Themenangabe ▪ Qualitätskriterien integrieren
Tonfall, Anspruchsniveau und Beispiele passen nicht zur im Entwurfsraster beschriebenen Zielgruppe.	▪ Ziel- statt Themenangabe ▪ Didaktisches Ziel konsequent berücksichtigen ▪ Qualitätskriterien integrieren
Wahlpunkte werden übergangen, oder der Verlauf passt nicht zur getroffenen Auswahl.	▪ Wahlpunkte und Verzweigungen verwenden ▪ Konsistenz im Gesprächsverlauf sichern ▪ ggf. Modellwahl prüfen

Beobachtetes Problem	Passende Prinzipien
Der Chatbot hält die Schrittfolge nicht ein oder überspringt Schritte.	▪ Schrittfolge statt Gesamtauftrag ▪ Umgang mit unklaren Antworten regeln ▪ ggf. Modellwahl prüfen
Der Chatbot arbeitet bei knappen oder unklaren Eingaben einfach weiter.	▪ Umgang mit unklaren Antworten regeln ▪ Relevanz von Ehrlichkeit betonen
Nutzende verstehen nicht, worum es geht, oder sind wenig motiviert.	▪ Thema, Vorgehen, Nutzen und Ziel kommunizieren ▪ Nutzende durch den Prozess führen
Nutzende verlieren im längeren Dialog die Orientierung.	▪ Nutzende durch den Prozess führen ▪ Interaktion drosseln
Nutzende sind an einzelnen Stellen überfordert.	▪ Didaktische Hilfen gezielt einbringen ▪ Interaktion drosseln
Der Chatbot gibt Quellen, Namen oder Daten an, die nicht existieren oder falsch sind.	▪ Transparenz von Unsicherheiten stärken
Der Chatbot gibt fachlich nicht tragfähige Antworten.	▪ Didaktische und fachliche Qualitätskriterien direkt in den Prompt integrieren
Antworten wirken eindeutiger, glatter oder sicherer, als es der Sache angemessen ist.	▪ Unangemessene Verkürzungen vermeiden ▪ Transparenz von Unsicherheiten stärken
Der Chatbot bestätigt nur und übt keine Kritik.	▪ Bestätigungstendenzen begrenzen (Sycophancy)
Antworten sind einseitig oder blenden relevante Sichtweisen aus.	▪ Unterrepräsentierte Perspektiven einbeziehen ▪ Strukturelle Antworttendenzen ausgleichen
Der Chatbot vergisst frühere Entscheidungen oder wechselt die Richtung.	▪ Konsistenz im Gesprächsverlauf sichern ▪ ggf. Modellwahl prüfen (Kontextlänge)
Erste Vorschläge werden von Nutzenden unkritisch übernommen.	▪ Mehrere Optionen fordern ▪ Iterationen einbauen

Beobachtetes Problem	Passende Prinzipien
Nutzende treffen ungünstige Entscheidungen ohne Hinweis auf Folgen oder Regeln.	▪ Warnhinweise ausgeben lassen
Die Ausgabe muss aufwendig nachbearbeitet werden, um weiterverwendbar zu sein.	▪ Ausgabeformat eindeutig definieren
Der Dialog endet nicht bzw. der Chatbot macht immer neue Angebote.	▪ Gesprächsende klar steuern ▪ Weitere Optimierungen nur bei echtem Mehrwert
Am Ende bleibt unklar, was gelernt wurde oder wie es weitergeht.	▪ Reflexion am Ende anregen ▪ Umsetzung und nächste Schritte absichern

SCHLUSS

Nicht jeder Prompt muss von Anfang an komplex und ausgefeilt sein. Oft ist es sinnvoller, mit einem kleinen, überschaubaren Anwendungsfall zu beginnen und zunächst nur einige wenige Prinzipien bewusst zu nutzen. Schon die klare Benennung des Ziels, eine gezielte Kontextabfrage oder ein sauber formulierter Abschluss können die Qualität eines Prompts deutlich verbessern. Erst wenn deutlich wird, dass ein Prompt einen Prozess genauer steuern, Nutzende gezielter begleiten oder bestimmte Schwierigkeiten besser abfangen soll, lohnt sich ein aufwendigerer Aufbau.

Die hier beschriebenen Prinzipien sind daher nicht als starre Vorgaben zu verstehen, die immer vollständig berücksichtigt werden müssen. Sie bilden vielmehr ein Repertoire von Gestaltungsmöglichkeiten, aus dem je nach Ziel und Kontext das ausgewählt werden kann, was für den jeweiligen Fall sinnvoll ist.

Mit einer bewussteren Entwicklung eigener Prompts und deren Weitergabe an Studierende und andere Lehrende wird die KI-Nutzung in der Lehre nicht automatisch einfacher. Häufig wird sie aber sinnvoller. Denn dann geht es nicht mehr nur darum, ob ein Chatbot etwas erzeugen kann, sondern wie die Nutzung so gestaltet werden kann, dass sie Lehren und Lernen tatsächlich unterstützt. In manchen Situationen kann die beste Unterstützung für das Lernen allerdings auch ein bewusster Verzicht auf den KI-Einsatz sein.

LITERATUR

Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakci, Ö., & Mariman, R. (2025). Generative AI without guardrails can harm learning: Evidence from high school mathematics. *Proceedings of the National Academy of Sciences*, 122(26), Article e2422633122.

<https://doi.org/10.1073/pnas.2422633122>

Belland, B. R., Walker, A. E., Kim, N. J., & Lefler, M. (2017). Synthesizing results from empirical research on computer-based scaffolding in STEM education: A meta-analysis. *Review of Educational Research*, 87(2), 309–344. <https://doi.org/10.3102/0034654316670999>

He, Q., Zeng, J., Huang, W., Chen, L., Xiao, J., He, Q., Zhou, X., Liang, J., & Xiao, Y. (2024). Can large language models understand real-world complex instructions? *Proceedings of the AAAI Conference on Artificial Intelligence*, 38(16), 18188–18196.

<https://doi.org/10.1609/aaai.v38i16.29777>

Hicks, M. T., Humphries, J., & Slater, J. (2024). ChatGPT is bullshit. *Ethics and Information Technology*, 26, Article 38. <https://doi.org/10.1007/s10676-024-09775-5>

Huang, J., Chen, X., Mishra, S., Zheng, H. S., Yu, A. W., Song, X., & Zhou, D. (2024). Large language models cannot self-correct reasoning yet. In *The Twelfth International Conference on Learning Representations (ICLR 2024)*. <https://openreview.net/forum?id=lkmD3fKBPO>

Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y. J., Madotto, A., & Fung, P. (2023). Survey of hallucination in natural language generation. *ACM Computing Surveys*, 55(12), Article 248. <https://doi.org/10.1145/3571730>

Laban, P., Hayashi, H., Zhou, Y., & Neville, J. (2026). LLMs get lost in multi-turn conversation. In *The Fourteenth International Conference on Learning Representations (ICLR 2026)*.

<https://openreview.net/forum?id=VKGTGGcwl6>

Lee, H.-P., Sarkar, A., Tankelevitch, L., Drosos, I., Rintel, S., Banks, R., & Wilson, N. (2025). The impact of generative AI on critical thinking: Self-reported reductions in cognitive effort and confidence effects from a survey of knowledge workers. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (Article 1121, pp. 1–22). Association for Computing Machinery. <https://doi.org/10.1145/3706598.3713778>

Li, X., Yu, Z., Zhang, Z., Chen, X., Zhang, Z., Zhuang, Y., Sadagopan, N., & Beniwal, A. (2025). When thinking fails: The pitfalls of reasoning for instruction-following in LLMs. In *The Thirty-Ninth Annual Conference on Neural Information Processing Systems (NeurIPS 2025)*.

<https://openreview.net/forum?id=w5uUvxp81b>

Liu, N. F., Lin, K., Hewitt, J., Paranjape, A., Bevilacqua, M., Petroni, F., & Liang, P. (2024). Lost in the middle: How language models use long contexts. *Transactions of the Association for Computational Linguistics*, 12, 157–173. https://doi.org/10.1162/tacl_a_00638

Madaan, A., Tandon, N., Gupta, P., Hallinan, S., Gao, L., Wiegrefe, S., Alon, U., Dziri, N., Prabhunoye, S., Yang, Y., Gupta, S., Majumder, B. P., Hermann, K., Welleck, S., Yazdanbakhsh,

A., & Clark, P. (2023). Self-Refine: Iterative refinement with self-feedback. In *Advances in Neural Information Processing Systems 36 (NeurIPS 2023)* (pp. 46534–46594). Curran Associates.

<https://doi.org/10.52202/075280-2019>

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), 199–

218. <https://doi.org/10.1080/03075070600572090>

Sharma, M., Tong, M., Korbak, T., Duvenaud, D., Askell, A., Bowman, S. R., Cheng, N., Durmus, E., Hatfield-Dodds, Z., Johnston, S. R., Kravec, S., Maxwell, T., McCandlish, S., Ndousse, K., Rausch, O., Schiefer, N., Yan, D., Zhang, M., & Perez, E. (2024). Towards understanding sycophancy in language models. In *The Twelfth International Conference on Learning Representations (ICLR 2024)*. <https://openreview.net/forum?id=tvhaxkMKAn>

van de Pol, J., Volman, M., & Beishuizen, J. (2010). Scaffolding in teacher–student interaction: A decade of research. *Educational Psychology Review*, 22(3), 271–296.

<https://doi.org/10.1007/s10648-010-9127-6>

van de Pol, J., Volman, M., Oort, F., & Beishuizen, J. (2015). The effects of scaffolding in the classroom: Support contingency and student independent working time in relation to student achievement, task effort and appreciation of support. *Instructional Science*, 43(5), 615–641.

<https://doi.org/10.1007/s11251-015-9351-z>

Wen, B., Ke, P., Gu, X., Wu, L., Huang, H., Zhou, J., Li, W., Hu, B., Gao, W., Xu, J., Liu, Y., Tang, J., Wang, H., & Huang, M. (2024). Benchmarking complex instruction-following with multiple constraints composition. *Advances in Neural Information Processing Systems*, 37, 137610–137645.

<https://doi.org/10.52202/079017-4371>

Zheng, M., Pei, J., Logeswaran, L., Lee, M., & Jurgens, D. (2024). When "a helpful assistant" is not really helpful: Personas in system prompts do not improve performances of large language models. In *Findings of the Association for Computational Linguistics: EMNLP 2024* (pp. 15126–15154). Association for Computational Linguistics.

<https://doi.org/10.18653/v1/2024.findings-emnlp.888>

Zhou, D., Schärli, N., Hou, L., Wei, J., Scales, N., Wang, X., Schuurmans, D., Cui, C., Bousquet, O., Le, Q., & Chi, E. (2023). Least-to-most prompting enables complex reasoning in large language models. In *The Eleventh International Conference on Learning Representations (ICLR 2023)*.

<https://openreview.net/forum?id=WZH7099tgfM>

Zhou, Y., Muresanu, A. I., Han, Z., Paster, K., Pitis, S., Chan, H., & Ba, J. (2023). Large language models are human-level prompt engineers. In *The Eleventh International Conference on Learning Representations (ICLR 2023)*.

<https://openreview.net/forum?id=92gvk82DE->