



Universität Regensburg

Einsatz von GPT-4 als Lerncoach: Messung der Auswirkungen auf Motivation und Selbstwirksam- keit bei Studierenden

Masterarbeit im Fach Digital Humanities am Lehrstuhl für
Informationswissenschaft

Vorgelegt von: Daniel Alexander Vikete
Adresse: Oberländerstraße 16, Regensburg 93051
E-Mail (Universität): Daniel-Alexander.Vikete@stud.uni-regensburg.de
E-Mail (privat): alexander.vikete@gmx.de
Matrikelnummer: 2025386
Erstgutachter: PD Dr. David Elsweiler
Zweitgutachter: Prof Dr. Udo Kruschwitz
Betreuer*innen: PD Dr. David Elsweiler/ Selina Meyer M.A.
Laufendes Semester: WS 23/24
Abgegeben am: 15.01.2024
GitHub: <https://github.com/AlexVike/Masterarbeit>

Inhaltsverzeichnis

1	Einleitung	8
2	Theoretischer Rahmen	10
2.1	Coaching	10
2.1.1	Coaching-Tools	11
2.1.2	Online-Coaching	13
2.1.3	Effekte des Coachings	16
2.2	Lerncoaching	17
2.2.1	Kommunikation im Lerncoaching	20
2.2.2	Phasen des Lerncoachings	23
2.2.3	Limitationen des Lerncoachings	28
2.3	Motivation, Selbstwirksamkeit	28
2.4	Chatbots	31
2.4.1	Generative Pre-trained Transformer (GPT)	32
2.4.2	Few-Shot-Learning	35
3	Aktueller Forschungsstand	37
3.1	Gesundheitswesen	37
3.2	Industrie	38
3.3	Bildungswesen	40
3.4	Coaching	41
4	Präzisierung der Forschungsfrage	43
5	Methodik	44
5.1	Lerncoaching-Chatbot	44
5.2	Forschungsdesign	45
5.2.1	Vorstudie	46
5.2.2	Hauptstudie	48
5.3	Stichprobenziehung	49
5.3.1	Zielgruppe und Auswahlkriterien	50
5.3.2	Rekrutierung der Teilnehmenden	50
5.3.3	Größe und Struktur der Stichprobe	51
5.4	Datenerhebung	52
5.5	Software der Chatbots	53

5.5.1	Front-End: Telegram.....	54
5.5.2	Back-End: GPT-4, SQLite-Datenbank.....	54
5.6	Datenauswertung.....	54
6	Ergebnisse.....	56
7	Diskussion.....	59
7.1	Ergebnisse zwischen den beiden Gruppen	59
7.2	Ergebnisse innerhalb der Gruppen.....	60
7.3	Analyse Chatlogs.....	62
7.4	Praktische Implikationen	63
8	Abschluss.....	65
8.1	Ausblick	65
8.2	Limitationen.....	66
8.3	Fazit	67
9	References.....	68
Anhang A: Aufbau der Vorstudie		80
A1 Beispielchat einer Vorstudie		80
Anhang B: Prompt- und Chatbeispiele.....		83
B1 Prompts an GPT-4.....		83
B2 Chatbeispiele		87

Abbildungsverzeichnis

Abbildung 1: Aufbau Erwartungswert-Modell nach McKeachie et al. [1987]	29
Abbildung 2: Ablauf der Studie	49
Abbildung 3: Flussdiagramm der Chatbots	53
Abbildung 4: Die Differenz und statistische Werte für jede Strategie.....	59
Abbildung 5: Wordlcoud LC-Gruppe links, NC-Gruppe rechts.....	63

Tabellenverzeichnis

Tabelle 1 Fragearten im Lerncoaching	22
Tabelle 2 Auswertung der Vorstudie nach den Phasen des Lerncoachings.....	47
Tabelle 3: Alter, Geschlecht und Studienfach der Teilnehmenden der Vorstudie	51
Tabelle 4: Demografische Daten der Teilnehmenden der Hauptstudie	51
Tabelle 5: Mittelwerte und Standardabweichungen der beiden Gruppen	56
Tabelle 6: Unterschiede der Motivation zwischen den beiden Gruppen.....	57
Tabelle 7: Unterschiede der Motivation innerhalb der beiden Gruppen.....	58
Tabelle 8: Menschliche/Chatbot Durchschnittliche Dauer & Nachrichtenlänge.....	62
Tabelle 9: Beispielchat einer Vorstudie	82
Tabelle 10: Beispiel Lerncoaching-Chatbot.....	89
Tabelle 11: Beispiel Neutraler-Chatbot	91
Tabelle 12: Übersetzung Motivated Strategies for Learning Questionnaire.....	94

Zusammenfassung

Angesichts der sich wandelnden Bildungslandschaft, verstärkt durch die Herausforderungen der Covid-19-Pandemie und die Integration digitaler Methoden, erscheint die Entwicklung eines Coaching-Chatbots besonders vielversprechend. Besonders im Kontext neuer Technologien wie ChatGPT gewinnt diese Idee an Bedeutung. In dieser Arbeit wird das State-of-the-Art-Modell GPT-4 verwendet, wobei Few-Shot-Learning verwendet wird. Ziel ist es zu untersuchen, ob GPT-4 mittels Few-Shot-Learning effektiv als Coaching-Tool fungieren kann, um Studierenden der Universität Regensburg bei der Steigerung ihrer Motivation und Selbstwirksamkeit zu unterstützen. Für die Durchführung dieser Studie wird zunächst eine Vorstudie mit einem Wizard-of-Oz-Design realisiert, um Daten für das Few-Shot-Learning zu sammeln. Anschließend wird der entwickelte Coaching-Chatbot in einem Between-Group Design überprüft. Hierbei interagiert die Kontrollgruppe mit einer Basisversion von GPT-4, die ohne spezifisches Coaching-Framework und Few-Shot-Learning arbeitet. Sowohl die Kontroll- als auch die Experimentalgruppe füllen vor und nach der Intervention Fragebögen aus, um Veränderungen in Motivation und Selbstwirksamkeit zu erfassen. Die Ergebnisse zeigen, dass es keinen signifikanten Unterschied zwischen den beiden Gruppen gibt. Interessanterweise wurde jedoch in beiden Gruppen eine Steigerung der Selbstwirksamkeit beobachtet. Dies legt nahe, dass GPT-4 auch mit minimalen Anweisungen und ohne spezifisches Few-Shot-Learning effektiv darin sein kann, Studierende bei der Erreichung ihrer Lernziele zu unterstützen. Diese Erkenntnis unterstreicht das Potenzial, dass bereits einfache Implementierungen von Coaching-Chatbots mit geringem Aufwand in das Curriculum einer Universität integriert werden könnten.

Abstract

In light of the evolving educational landscape, further intensified by the challenges of the Covid-19 pandemic and the integration of digital methods, the development of a coaching chatbot seems particularly promising. Especially in the context of new technologies like ChatGPT, this idea gains significance. In this work, the state-of-the-art model GPT-4 is used, employing Few-Shot-Learning. The goal is to investigate whether GPT-4 can effectively function as a coaching tool using Few-Shot-Learning to assist students of the University of Regensburg in enhancing their motivation and self-efficacy. For conducting this study, a preliminary study using a Wizard-of-Oz design will be realized to gather data for Few-Shot-Learning. Subsequently, the developed coaching chatbot will be evaluated in a between-group design. Here, the control group interacts with a basic version of GPT-4, operating without a specific coaching framework and Few-Shot-Learning. Both the control and experimental groups will fill out questionnaires before and after the intervention to measure changes in motivation and self-efficacy. The results show that there is no significant difference between the two groups. However, an increase in self-efficacy was observed in both groups. This suggests that GPT-4, even with minimal instructions and without specific Few-Shot-Learning, can be effective in helping students achieve their learning goals. This finding highlights the potential that even simple implementations of coaching chatbots could be integrated into a university's curriculum with minimal effort.

1 Einleitung

Die Bildungslandschaft wandelt sich zunehmend. Dabei entwickeln sich Dozierende von reinen Wissensvermittler*innen zu Lernbegleiter*innen [112]. Dies führt zu einer veränderten Kommunikation zwischen Lernenden und Lehrenden. Ein weiterer bedeutender Wandel ist die fortschreitende Digitalisierung in der Bildung [35]. Die Covid-19-Pandemie hat diesen Digitalisierungstrend verstärkt, insbesondere durch den längeren Zeitraum des Distanzunterrichts an verschiedenen Schulen [43]. Allerdings wurde deutlich, dass viele Schüler*innen Defizite in ihren mathematischen und leserischen Fähigkeiten aufwiesen. Dies wird durch die aktuelle PISA-Studie der OECD [2023] bestätigt.

Lerncoaching ermöglicht es, auf diese Veränderungen gezielt einzugehen. Dabei verbindet es die Aspekte von Lernen und Coaching, wobei der Fokus auf individuellem Lernen liegt [30]. Zusätzlich werden lernbeeinflussende Faktoren wie Motivation und Selbstwirksamkeit im Lerncoaching adressiert. Diese Konstrukte können einen Einfluss auf den Lernerfolg haben [55].

Eine weitere Entwicklung, die einen Einfluss auf das Bildungswesen und die Gesellschaft hatte, ist ChatGPT. Das Programm war eine der am schnellsten wachsenden Anwendungen im Internet. Nach der Einführung am 30. November 2022 hat es schnell an Popularität gewonnen, dies wird durch die Zahl von 100 Millionen Nutzern innerhalb von nur zwei Monaten belegt wird [25]. ChatGPT zeichnet sich durch die Fähigkeit aus, eine Vielzahl von Aufgaben im Bereich der natürlichen Sprachverarbeitung (NLP) zu bewältigen [29]. Dazu gehören die Übersetzung von natürlicher Sprache in Code, das Beantworten von Fragen zu fast jedem Thema, das Erstellen von Geschichten sowie das Lösen von Standardproblemen im Bereich NLP wie Textklassifikation und Übersetzung [29]. Aktuell gilt GPT-4 als State-of-the-Art in diesem Bereich [66]. Die zunehmende Verbreitung von Chatbots in der Allgemeinbevölkerung, wie z.B. Siri oder Google Assistant, die in den meisten Mobiltelefonen integriert sind, zeigt ebenfalls den wachsenden Einfluss solcher Technologien. Chatbots imitieren menschliche Gespräche und helfen Nutzer*innen, Ziele zu erreichen, indem sie Text- oder Sprachdaten verarbeiten und generieren [59].

Angeichts der Entwicklungen im Bildungsbereich und den neuen Technologien scheint die Erforschung eines Chatbots, der einen Lerncoaching-Ansatz verfolgt,

sinnvoll zu sein. Um den Lerncoaching Chatbot zu untersuchen, erfolgt eine Vorstudie mit einem Wizard-of-Oz-Design zur Datensammlung und Validierung des Studiendesigns. Die Hauptstudie nutzt ein Between-Group Design, um die Effekte des Lerncoaching-Chatbots (LC) im Vergleich zu einem Neutralen-Chatbot (NC) zu bewerten. Der LC basierte auf den Ergebnissen der Vorstudie und durchläuft die Phasen des Lerncoachings, während der NC minimale, neutrale Anweisungen erhält. Die Studie nutzt einen Fragebogen, um die Motivation der Teilnehmer*innen vor und nach der Studie zu messen, wobei die Hauptstudie die Effekte des LC im Vergleich zu einem NC untersucht.

Um die Ziele dieser Arbeit zu erreichen, wird mit einer Erläuterung der Schlüsselbegriffe im theoretischen Rahmen begonnen. Anschließend wird der aktuelle Forschungsstand beleuchtet, um ein Verständnis der derzeitigen Situation zu gewährleisten. Diese beiden Abschnitte führen zur Präzisierung der Forschungsfrage. Darauf folgt eine Beschreibung der angewandten Methodik, bevor die Ergebnisse präsentiert werden. Die anschließende Diskussion mündet in einen Ausblick und eine Reflexion über die Limitationen der Arbeit.

2 Theoretischer Rahmen

In diesem Abschnitt werden die theoretischen Grundlagen der Konzepte Coaching, Lerncoaching, Motivation, Selbstwirksamkeit und Chatbots dargelegt. Diese Informationen erleichtern das Verständnis der Forschungsarbeit.

2.1 Coaching

Coaching besitzt einen multidisziplinären Charakter [5]. Dies bedeutet, dass der Begriff aus unterschiedlichen wissenschaftlichen Disziplinen beeinflusst wird. Zusätzlich existieren unterschiedliche Formen des Coachings, die jeweils einen anderen Fokus besitzen [28]. Zu den verschiedenen Coachingmethoden gehören zielorientiertes, systemisches, kognitiv-behaviorales, lösungsorientiertes, positiv-psychologisches, ergebnisorientiertes und narratives Coaching [28]. Eine besonders bekannte und traditionelle Form ist das Führungskräfte-Coaching [39, 42, 87], während das Lerncoaching als weniger bekannte Methode gilt [30, 62, 71]. Das Lerncoaching, eine Reaktion auf sich wandelnde Lernumgebungen [71], stellt eine relativ neue und noch wenig erforschte Coachingform dar. Es erweist sich jedoch im Kontext dieser Arbeit als sehr geeignet und wird in einem weiteren Kapitel detaillierter beschrieben. Neben den unterschiedlichen Coachingmethoden variiert auch die Art ihrer Durchführung. Es existieren verschiedene Formate wie Einzel-Coaching, Gruppen-Coaching, Coaching mit mehreren Coaches und Online-Coaching [17]. Somit sind viele Definitionen in der Literatur zu finden, die sich leicht voneinander unterscheiden können [6]. Einige sehen Coaching als eine Form der personenzentrierten Beratung. Andere sehen es nicht als Beratung im herkömmlichen Sinne, sondern eher als Unterstützung zur Förderung der Selbsthilfefähigkeiten [28]. Es ist oft schwierig, Coaching von anderen Beratungsformen zu unterscheiden. Coaching differenziert sich von Disziplinen wie Psychotherapie, Supervision, Mentoring, Beratung und Training. Diese fokussieren sich mehr auf Beratung, Behandlung oder Wissensvermittlung [17]. Trotzdem kann Coaching Elemente aus diesen Bereichen integrieren.

Nach Bachkirova and Kauffman [2009] macht es Sinn Coaching mit folgenden Fragen zu definieren: „Für was wird es verwendet?“, „Was beinhaltet es?“, „In welchen Kontext wird es verwendet?“, „Für wen wird es verwendet?“. Mit diesen Fragen ist es möglich die Diversität des Coachings einzufangen und im Kontext einer Arbeit zu

definieren. In der Studie von Rauen [2014] finden sich Antworten auf diese Fragen, die im Verlauf dieser Forschungsarbeit als Arbeitsdefinition herangezogen werden.

Coaching ist ein interaktiver Prozess, der auf individueller Unterstützung basiert. Es ist lösungs- und zielfokussiert mit dem Ziel der Selbstentwicklung. Der Coach bietet keine direkten Lösungen, sondern fördert Selbstreflexion und Eigeninitiative des Klienten. Die Beziehung zwischen Coach und Klient, auch Coachee genannt, ist von Vertrauen und Freiwilligkeit geprägt, ohne manipulative Techniken. Coaching erfolgt in mehreren, zeitlich begrenzten Sitzungen.[86]

Zu den weiteren Schlüsselkonzepten des Coachings zählen die Förderung von Hilfe zur Selbsthilfe und die Entwicklung von Zielsetzungen [86, 87, 111]. Dies spiegelt sich in der Ausrichtung des Coachings wider, das auf die Förderung der Eigeninitiative und einen lösungszentrierten Ansatz abzielt. Diese Konzepte finden sich in verschiedenen Coachingformen wieder.

Obwohl der Ablauf eines Coachings je nach Form variiert, gliedert er sich typischerweise in drei Hauptphasen. Die Vorphase umfasst die Kontaktaufnahme. Die Hauptphase beinhaltet die Klärung der Ausgangssituation sowie die Festlegung von Zielen oder die Entwicklung von Lösungen. Die Abschlussphase ist durch einen Abschluss und eine Evaluation gekennzeichnet. [15]

2.1.1 Coaching-Tools

Ein Coaching-Tool ist eine Kombination aus verschiedenen Methoden, Techniken und Interventionen. Es wird speziell für bestimmte Anwendungsbereiche entwickelt und verfolgt klare Ziele. Diese Tools dienen dazu, effektive Kommunikation zu fördern, indem sie auf einer grundlegenden Ebene mit Interventionen beginnen, diese zu Techniken erweitern und schließlich zu umfassenderen Methoden zusammenführen. Die Vielfalt und Komplexität von Coaching-Tools ergibt sich aus dieser vielschichtigen Zusammensetzung und der Anpassung an unterschiedliche Bedürfnisse und Situationen. [16]

Die Tools im Coaching besitzen vier Kategorien, die in den unterschiedlichen Phasen eingesetzt werden können [16]. Dies beinhaltet diagnostische Tools, reflexive Tools, Tools für die Zielsetzung und Tools zur Veränderung.

Ein wichtiger Bestandteil all dieser Tools ist die Gesprächsführung im Coaching. Dabei werden Methoden wie das Paraphrasieren, der naive Dialog oder das Einstreuen

von Ideen verwendet. Mithilfe der Gesprächsführung kann auch der Beziehungsaufbau gefördert werden. Dabei werden Methoden wie das Pacing und das Leading verwendet. Pacing im Coaching bedeutet, dass der Coach sich auf die Klienten einstellt, indem er oder sie dessen Körpersprache und Sprache aufmerksam beobachtet und nachempfiehlt. [16]

Ein weiteres Element des Coachings, welches in allen Tools verwendet werden kann, ist der Gebrauch spezieller Fragetypen. Selbstaufmerksamkeitsverstärkende Fragen zielen darauf ab, den Klienten über ihre eigenen Gefühle und Gedanken nachdenken zu lassen. Dem gegenüber stehen selbstaufmerksamkeitsunterbrechende Fragen, die den Fokus eher auf den Coach lenken, aber seltener hilfreich sind, da sie die Aufmerksamkeit vom Klienten wegverlagern. Direkte Fragen können nützlich sein, bergen aber das Risiko, dass sich die Klienten bedrängt fühlen. Daher sollten sie nur bei einer stabilen Beziehung eingesetzt werden. Indirekte oder offene Fragen sind weniger konfrontativ und für den Aufbau einer Beziehung besser geeignet. Zirkuläre Fragen ermutigen die Klienten, die Perspektiven anderer Personen in Betracht zu ziehen. Suggestive Fragen, die meist vorgegebene Antworten enthalten, sollten vorsichtig verwendet werden, um Manipulation zu vermeiden. Rhetorische oder hypothetische Fragen regen die Klienten dazu an, ihr Anliegen aus verschiedenen Blickwinkeln zu betrachten. Während naive Fragen den Klienten helfen, ihre Probleme klarer zu formulieren, indem der Coach zugibt, nicht alles zu wissen. Schließlich helfen Skalierungsfragen den Klienten, Veränderungen präzise zu benennen. [16]

Die diagnostischen Tools umfassen Techniken der Rekonstruktion. Dabei versetzen sich Klienten mit Unterstützung der Coaches in vergangene Situationen zurück, um diese erneut zu durchleben. Dies ermöglicht es den Klienten auch, Systeme, in denen sie sich bewegen, zu visualisieren. Durch diese visuelle Annäherung können die Formen von Beziehungen besser erkannt werden. Zudem eröffnet das Hineinversetzen in andere Personen die Möglichkeit, neue Perspektiven zu gewinnen. Diese Techniken sind nützlich, um Klienten im Coachingprozess zu helfen und Schlussfolgerungen zu ziehen. Sie sind vor allem in der Hauptphase des Coachings von Vorteil. Weitere diagnostische Methoden beinhalten das Analysieren von Zusammenhängen, Testverfahren, Karriere- und Rollenanalysen. [16]

Im Coaching bieten reflexive Werkzeuge einen Rahmen für Feedback von Wahrnehmungsverzerrungen.

Die Aufgabe des Coaches ist es, Klienten zu helfen, ihre Wahrnehmungs-, Beurteilungs- und Verhaltensmuster zu reflektieren. Indem sie Generalisierungen in den Aussagen der Klienten aufzeigen, fördern Coaches eine tiefere Überprüfung und Einsicht in deren Perspektiven. Metaphern können verwendet werden, um komplexe Themen anzusprechen. Gleichzeitig wird die Selbstreflexion gefördert, indem eine Umgebung geschaffen wird, die eine offene Auseinandersetzung mit Emotionen und Gedanken ermöglicht. Durch diesen Prozess werden Muster der Wahrnehmung und Bewertung untersucht und eventuell angepasst. [16]

Bei der Zielsetzung können Ziele über das ganze Coaching gesetzt werden, oder auch für einzelne Sitzungen. Wichtig dabei, um die Ziele umzusetzen, können Transferaufgaben sein, die die Klient*innen zwischen den Sitzungen absolvieren können. [16]

Im Coaching werden Tools zur Veränderung genutzt, um neue Verhaltensweisen zu fördern. Rollenspiele helfen, Konflikte zu verstehen und Verhalten zu ändern. Die Flexibilisierung von Verhaltensmustern erfolgt durch unerwartete Coach-Interventionen oder Rollenspiele. Reframing setzt Verhaltensmuster in neue Kontexte, um Perspektiven zu öffnen. Konfrontation und Implosion fördern schnelle Erkenntnisse. Humor und Provokation erweitern Problemperspektiven, erfordern jedoch eine starke Coach-Klient-Beziehung und gutes Timing. [16]

Zusammengefasst kann gesagt werden, dass alle Coaching-Tools in den unterschiedlichen Phasen des Coachings verwendet werden können. Viele dieser Tools basieren auf der Kommunikation der Klienten und Coaches. Einige Techniken, vor allem bei den diagnostischen Tools erfordern jedoch neben der Kommunikation ein weiteres Medium, um die Maßnahmen aufzuzeichnen.

2.1.2 Online-Coaching

Das Online-Coaching besitzt wie das Coaching unterschiedliche Definitionsmöglichkeiten [22]. Häufig verbirgt sich hinter dem Online-Coaching ein Coaching-Ansatz über Telefon, Chat, E-Mail oder Video [7, 23]. Es werden auch kollaborative Plattformen genutzt, die beispielsweise das Arbeiten mit Avataren in 3-D-Räumen oder die Verwendung verschiedener Coaching-Tools ermöglichen [7]. Dies macht sowohl synchrone als

auch asynchrone, textbasierte Arbeit oder Mischformen möglich. Mischformen treten auch auf, wenn Online-Coaching mit persönlichem Coaching kombiniert wird, welches als Blended-Coaching bezeichnet wird [7]. Das Online-Coaching ist in der Literatur seit Mitte der 2010er Jahre präsent [7, 97]. Diese Entwicklung zeigt die schnellen Veränderungen auf, die durch Technologie ermöglicht wurden. Siegrist [2006] beschreibt, wie auch die heutige Literatur, dass es sowohl synchrone als auch asynchrone Kommunikationsmedien im Online-Coaching gibt. Jedoch merkt der Autor an, dass die am häufigsten genutzte Form des Online-Coachings der E-Mail-Verkehr ist, also eine Form des asynchronen Coachings

Laut Berninger-Schäfer [2018] ist das Online-Coaching in vier Stufen gegliedert. In der ersten Stufe wird Online-Coaching als die Nutzung von Telekommunikation, E-Mails, Chats oder Videofunktionen definiert. Der Einsatz von Plattformen, die eine Vielzahl von Kommunikationskanälen und zusätzlichen Kooperationsmöglichkeiten bieten, ist die zweite Stufe. Jedoch sind sie nicht speziell auf das Coaching ausgerichtet.

In der dritten Phase des Coachings kommen spezielle Tools zum Einsatz. Diese können einzeln oder kombiniert genutzt werden. Sie sind entweder isoliert oder in einen festgelegten Coaching-Prozess integriert. Die vierte und letzte Stufe kombiniert verschiedene Elemente einer Plattform. Dazu gehören coaching-spezifische Werkzeuge, Prozesse und das Management der Coaching-Aktivität.

Ribbers and Waringa [2015] entwickelten einen neuen Prozess für Online-Coaching. Dieser ist in drei Phasen gegliedert. Ein wesentlicher Unterschied zum traditionellen Coaching ist der stärkere Fokus auf Zieldefinition. Zudem wird angestrebt, diese Ziele in kleinen, überschaubaren Schritten zu erreichen. Dieser Fokus auf einen schrittweisen Erfolg steht im Mittelpunkt des Online-Coaching-Prozesses. Für das Erstellen der Ziele können Methoden wie SMART oder Skalierungsfragen verwendet werden. Dazu in einem späteren Kapitel mehr.

Bei schriftlichen Formen des Coachings, wie E-Mail-Kommunikation oder Chat-Coaching, sind verschiedene Vorteile erkennbar. Wenn Klienten Nachrichten verfassen, können sie diese sorgfältig überarbeiten und sich Zeit zum Nachdenken nehmen. Das Verfassen von Nachrichten kann ein Gefühl der Anonymität [100] auslösen. Dies kann dazu führen, dass Coachees mehr Informationen teilen. Das Schreiben steigert die Motivation zur Problembewältigung. Es fördert die Selbstreflexion und das

Selbstbewusstsein, indem es das Verständnis für eigene Emotionen und Gedanken verbessert [7, 88]. Dies erhöht die Selbstwahrnehmung und Selbstregulation und unterstreicht den therapeutischen Effekt des Schreibens. [7, 88]

Allerdings gibt es bei den schriftlichen Formen des Coachings wichtige Faktoren zu berücksichtigen. Coaches müssen sich aktiv bemühen, die Texte der Klienten zu verstehen. Dabei ist es wichtig zu erkennen, welche Fakten beschrieben werden, welche Emotionen im Text sichtbar sind und ob Metaphern verwendet werden [7]. Dies hilft dabei zu verstehen, welche Interventionen am besten geeignet sind [7]. Bei den Antworten der Coaches ist auf korrekte Rechtschreibung [7] zu achten. Es ist wichtig, klar formulierte Fragen zu stellen. Zudem sollten die Antworten der Klienten gespiegelt werden, um Verständnis und Empathie zu zeigen [7].

Neben den schriftlichen Formen besteht auch die Möglichkeit, über Online-Plattformen an Videoanrufen teilzunehmen. Diese finden im Rahmen von sogenannten Blended-Coaching-Formaten statt [7, 23, 100]. Dabei kann auch der Einsatz von VR-Brillen eine Rolle spielen [100]. Diese Technologie ermöglicht es, Situationen, in denen sich Klienten befinden, zu simulieren. Zusätzlich können VR-Technologien eingesetzt werden, um spezifische Coaching-Tools zu nutzen [100]. Insgesamt lassen sich die meisten in einem früheren Abschnitt beschriebenen Coaching-Methoden nicht in der ersten Stufe des Online-Coachings umsetzen [7]. Hierfür ist meist ein multimodaler Ansatz erforderlich, um visuelle Aufgaben durchführen zu können. Somit ist in Stufe eins die Kommunikation zwischen Coach und Coachee im Vordergrund. Fragetechniken müssen verwendet werden, um ein erfolgreiches Coaching durchzuführen.

Aus den gesammelten Informationen lassen sich Vor- und Nachteile des Online-Coachings ableiten. Online-Coaching bietet Zeitersparnis [7, 97, 100] und ermöglicht es Klienten, flexibel zu entscheiden [7, 88, 97, 100], wann und wo das Coaching stattfindet. Dies kann zu einer gesteigerten Selbstwirksamkeit [7] führen. Insbesondere bei textbasierten Coaching-Formen wird eine intensivere Selbstreflexion [7] gefördert. Die Anonymität und der geringe Aufwand unterstützen zudem den niedrigherschweligen Zugang zum Online-Coaching [7, 100]. Allerdings gibt es im Online-Coaching auch Herausforderungen. Es kann leichter zu Unterbrechungen kommen [7], und Missverständnisse entstehen möglicherweise schneller [7]. Besonders in textbasierten Formen des Online-

Coachings fehlen Mimik und Gestik, die in vielen Coaching-Ansätzen eine wichtige Rolle spielen [7, 86, 97].

2.1.3 Effekte des Coachings

Professionelles Coaching ist mittlerweile ein anerkanntes Konzept, unterstützt durch umfassende Forschung. Viele Studien untersuchen die Wirkungen von Coaching [12, 13, 39, 56, 95, 106, 110].

Theeboom et al. [2014] fanden heraus, dass Coaching-Interventionen alle untersuchten Bereiche positiv beeinflussen. Dazu gehören Leistung, Fähigkeiten, Wohlbefinden, Bewältigungsstrategien, Arbeitseinstellungen und zielorientierte Selbstregulation. Ein interessanter Aspekt ist die Verbesserung der Selbstwirksamkeit. Diese Ergebnisse unterstreichen, dass Coaching auch die Selbstwirksamkeit steigern kann.

Auch Shirley C. Sonesh et al. [2015] fanden positive Effekte von Coaching. Ihre Studie zeigt Verbesserungen in verschiedenen Bereichen. Dazu gehören die Beziehungen zwischen Coach und Coachee sowie die Arbeitsallianz. Die Coachees selbst erlebten Fortschritte bei der Zielerreichung. Coaching bewirkt auch Einstellungsänderungen. Auf der Arbeitsebene führt es zu mehr Motivation und Selbstwirksamkeit. Persönlich kommt es zu einer Reduktion von Stress und einem gesteigerten Glücksempfinden.

In ihrer Studie stellten Jones et al. [2016] fest, dass Coaching positive Auswirkungen auf organisatorische Ergebnisse hat. Einschließlich spezifischer Bereiche wie Fertigkeiten, emotionale Aspekte und individuelle Leistungen. Die affektiven Ergebnisse betreffen Einstellungen und Motivation einschließlich Selbstwirksamkeit und Zufriedenheit. Die auf Fertigkeiten basierenden Ergebnisse beziehen sich auf die Entwicklung neuer Fähigkeiten, beobachtet am Arbeitsplatz.

Laut Burt and Talati [2017] hat Coaching positive Auswirkungen in Bereichen wie Leistung, Wohlbefinden, Bewältigungsstrategien, Arbeitseinstellungen und zielgerichteter Selbstregulation. Die Studie fand keine signifikanten Unterschiede in den Auswirkungen des Coachings aufgrund des Alters der Teilnehmenden oder der Messmethoden. Insgesamt wurde festgestellt, dass Coaching eine moderat signifikante positive Wirkung auf die Coachees hat.

Wang et al. [2022] stellten in ihrer Analyse fest, dass Coaching signifikante positive Auswirkungen auf kognitive Ergebnisse, insbesondere in Bezug auf die allgemeine

wahrgenommene Selbstwirksamkeit und Zielerreichung, sowie auf die von anderen bewertete Leistung und das psychische Wohlbefinden am Arbeitsplatz hatte. Im Gegensatz dazu zeigte Coaching positive, aber nicht signifikante Effekte auf affektive Ergebnisse und die selbstbewertete Leistung.

Cannon-Bowers et al. [2023] untersuchten verschiedene Aspekte, um die Wirksamkeit von Coaching zu bewerten. Sie stellten fest, dass der theoretische Ansatz des Coachings, ob prozess- oder ergebnisorientiert, keinen signifikanten Unterschied in der Wirksamkeit zeigte. Ebenso ergaben die Analysen, dass die Art der Ergebnismessung keinen signifikanten Unterschied in der Effektivität des Coachings aufwies. Bezüglich der Coaching-Typen zeigten sich bei affektiven Maßnahmen sowohl für prozess- als auch für ergebnisorientiertes Coaching ähnliche positive Auswirkungen. Die Analyse ergab auch, dass die Quelle der Bewertung, ob selbstberichtet, von Vorgesetzten oder Arbeiter*innen, keinen signifikanten Unterschied in den positiven Ergebnissen zeigte. Face-to-Face-Coaching und virtuelles Coaching waren beide mit positiven Ergebnissen verbunden. Es wurde kein signifikanter Unterschied zwischen den beiden festgestellt. Schließlich fanden sie heraus, dass weder die Anzahl der Coaching-Sitzungen noch die Gesamtstunden des Coachings signifikante Prädiktoren für die Gesamtwirksamkeit des Coachings waren.

Zuletzt zu erwähnen ist der Fokus auf die Zielformulierung im Coaching. Gemäß Locke and Latham [1990] sind spezifische, messbare und erreichbare Ziele entscheidend für Leistungssteigerung und Motivation. Ein klarer Zeitrahmen und die persönliche Relevanz der Ziele verstärken Engagement [50] des Coachees.

Schließlich hat sich professionelles Coaching als effektive Methode herausgestellt. Leistung, Wohlbefinden, Bewältigungsstrategien, Arbeitseinstellungen und vor allem Selbstwirksamkeit können alle verbessert werden.

2.2 Lerncoaching

Lerncoaching kombiniert die Konzepte von „Lernen“ und „Coaching“. Dabei steht das individuelle Lernen und die optimale Abstimmung zwischen lernender Person und Lerninhalt im Vordergrund [62, 71]. Jedoch gibt es noch lernbeeinflussende Faktoren, (z.B. Motivation) die im Lerncoaching bearbeitet werden können [30]. Laut Hardeland [2023] ist das Ziel des Lerncoachings, den lernenden Personen dazu zu befähigen den

Lernprozess selbstregulierend zu steuern. Dafür müssen Lernende aktiv teilnehmen und Verantwortung für den Lernprozess übernehmen. „Coaching“ stellt die Form der Beratung dar. Ein Coach versucht nach dem Prinzip der Hilfe zur Selbsthilfe zu arbeiten. Die Lernenden werden als Experten gesehen, somit findet eine horizontale Beratung statt [30]. Der Coach geht davon aus, dass die Lernenden sowohl die Lösung als auch die erforderlichen Ressourcen haben, obwohl sie möglicherweise nicht bewusst verfügbar sind. Dies bedeutet, dass der Coach keine Lösungen diktiert, sondern sich auf die Denkweise der Klient*innen konzentriert. Dadurch wird sichergestellt, dass jede Coaching-Maßnahme speziell auf die Lernaktivitäten der Lernenden abgestimmt ist [62].

Um ein Lerncoaching durchzuführen, müssen bestimmte Bedingungen erfüllt sein [30]:

- Die Lernenden müssen freiwillig am Lerncoaching teilnehmen
- Der Lerncoach muss eine Neutralität bewahren
- Die Inhalte einer Lerncoachingsitzung müssen vertraulich behandelt werden
- Ein Lerncoach versucht transparent zu arbeiten

Die Ziele des Lerncoachings können vielseitig sein. Wie schon am Anfang beschrieben, bezieht sich das Coaching nicht nur auf das Lernen. Hardeland [2023] formuliert unterschiedliche Ziele die während dem Lerncoaching verfolgt werden können. Die Lernkompetenz und die lernspezifischen Handlungen können verbessert werden. Wie im klassischen Coaching können laut Hardeland [2023] Fähigkeiten, Ressourcen und Potenziale der Coachees weiterentwickelt werden. Die Autorin teilt diese Ziele in Kompetenzen auf, die die Teilnehmer*innen erlangen können:

- Sozialkommunikative Kompetenz → Kommunikationsfähigkeit
- Personalkompetenz → Frustrationstoleranz, Reflexionsfähigkeit, Selbstwahrnehmung, Selbstbewusstsein, Selbstwirksamkeit, Eigenverantwortlichkeit, Zielstrebigkeit/Motivation
- Fachkompetenz → Problemlösungsfähigkeit
- Methodenkompetenz: → lernspezifisches Fachwissen

Laut Pallasch and Hameyer [2012] steht im Lerncoaching die Optimierung der Passung zwischen den Lernenden und dem Lernstoff als Ziel im Vordergrund. Diese Passung entsteht nicht nur durch die lernende Person und dem Lerngegenstand, sondern

auch durch das Lernarrangement und der Lernsituation. Es ist ein Zusammenspiel aller Faktoren. Lerncoaching legt den Schwerpunkt auf bedeutungsvolles und motiviertes Lernen [71]. Das Hauptziel besteht darin, das Selbstmanagement des Lernenden zu stärken und die Fähigkeit zur Problemlösung durch das Coaching zu verbessern [71].

Nicolaisen unterteilt in seinen unterschiedlichen Arbeiten [61, 62, 112] die Ziele im Lerncoaching auf die Lernenden und die Lehrpersonen.

Im Hinblick auf die Lernenden verfolgt dieses Konzept mehrere wichtige Ziele. Erstens steht die Stärkung der Selbstwirksamkeitserfahrungen im Mittelpunkt [61, 62, 112]. Dies bedeutet, dass die Lernenden ermutigt werden, an ihre Fähigkeiten zu glauben und das Vertrauen in ihre eigenen Fertigkeiten zu stärken. Darüber hinaus geht es darum, selbstregulatorische Kompetenzen zu erwerben, insbesondere wenn es zu emotionalen oder motivationalen Tiefs kommt [61, 112]. Dies soll den Lernenden helfen, ihre Lernprozesse besser zu kontrollieren und zu bewältigen. Die Identifizierung persönlicher Ressourcen ist ein weiteres Ziel [61]. Dies umfasst die Erkennung und Verwendung der individuellen Stärken und Fähigkeiten der Schüler. Darüber hinaus strebt das Konzept das Erlangen lernstrategischer Kompetenz an [61, 112]. Somit können sie ihr Lernen optimieren und ihre Ziele effizienter erreichen. Zusätzlich wird das präzise Erfassen von Lernschwierigkeiten, das Optimieren von Lernprozessen, das Entwickeln effektiver Lernstrategien und das Lösen von Lernblockaden im Lerncoaching behandelt [62]. Das Hauptziel des Lerncoachings ist es, Selbststeuerung im Lernprozess zu fördern. [112]. Lernende sollen dadurch ihre Lernziele und -aktivitäten selbst planen. Dies unterstützt Studierende in ihrer persönlichen und akademischen Entwicklung und stärkt ihre Lernfähigkeiten.

Weiterhin soll Lerncoaching die Kompetenzen der Coaches in verschiedenen Bereichen verbessern [62, 112]. Dazu zählen verbesserte Kommunikationsfähigkeiten [112], das Eingehen auf das individuelle Erleben der Lernenden [112], das präzise Erfassen von Lernschwierigkeiten [112] und die Förderung einer ressourcenorientierten Lernbegleitung [112]. Gleichzeitig sollten Lerncoachs grundlegende Fähigkeiten erwerben. Dazu gehören die Reflexion ihrer persönlichen Haltung [62], die Klärung ihrer Rolle [62], die Fähigkeit zur professionellen Gestaltung von Kommunikation [62], die Anwendung ressourcen- und lösungsorientierter Arbeitsmethoden [62] sowie die Fähigkeit zur angemessenen Abgrenzung [62] und die Beherrschung verschiedener Coaching-Methoden

[62]. Dies hilft, Schülerbedürfnisse zu verstehen und Lernprozesse besser zu unterstützen.

2.2.1 Kommunikation im Lerncoaching

Ein wichtiger Aspekt im Lerncoaching ist das aktive Zuhören [30, 61, 62]. Zum einen versucht der Lerncoach die Informationen der Klienten bewusst aufzunehmen, zum anderen wird den Klienten durch eine offene Körperhaltung [30], nonverbale Signale [30], oder durch Äußerungen wie „Ja“ oder „Mmh“ Interesse signalisiert [62]. Laut Nicolaisen [2017] können Lernproblematiken vielschichtig sein. Aus diesem Grund ist es wichtig, dass die Klienten vor allem zu Beginn einen größeren Redeanteil besitzen. Sie sollen möglichst viele Informationen zu der bestehenden Problematik beschreiben. So wird sichergestellt, dass das richtige Anliegen erfasst wird. Dies ist vor allem für Phase zwei des Lerncoachings von Bedeutung.

Ein weiterer Teil der Kommunikation im Lerncoaching ist das Beachten der nonverbalen und paraverbalen Äußerungen der Klienten [30, 62]. Ein Lerncoach achtet auf die Körperhaltung, Gestik und Mimik. Der Lerncoach kann mögliche Unterschiede zwischen diesen Ebenen erkennen und gegebenenfalls ansprechen, indem der Lerncoach die verbalen und nonverbalen Signale mit den Informationen des Klient*innen verknüpft. [30, 62]

Das Zusammenfassen und Widerspiegeln von Aussagen der Klienten sind wichtige Techniken im Lerncoaching [30, 61, 62]. Das Zusammenfassen im Lerncoaching verfolgt mehrere Ziele. Zum einen ermöglicht es dem Lerncoach, sich zu vergewissern, ob die Wahrnehmung des Gesagten korrekt ist und ob die Studierenden richtig verstanden werden. Zum anderen fühlen sich die Klient*innen durch diesen Prozess akzeptiert und verstanden. Weiterhin trägt das Zusammenfassen dazu bei, das Geschilderte zu strukturieren und bestimmte Aspekte herauszukristallisieren. Der Lerncoach kann mehrere Themen zusammenfassen, oder unterschiedliche Themen miteinander verbinden. Somit ist das Zusammenfassen vor allem wichtig, um die Komplexität der Thematik zu vermindern [62]. Beim Widerspiegeln sind nun die Themen schon zusammengefasst und die Komplexität wurde gemindert [62]. Somit werden in diesem Schritt konkrete Themen widergespiegelt. Dies erfolgt durch das Paraphrasieren bzw. Verbalisieren [30, 62]. Beim Paraphrasieren werden die sachlichen Informationen vom Lerncoach in

komprimierter Form wiedergegeben, wobei möglichst wenige Worte verwendet werden [62]. Beim Verbalisieren reflektiert der Lerncoach sowohl den unausgesprochenen als auch den ausdrücklich geäußerten emotionalen Gehalt einer Aussage [30].

Mit den bisher beschriebenen Methoden kann der Lerncoach den Klient*innen unterstützen, Zusammenhänge klarer zu erfassen [62]. Deshalb werden die Coachees motiviert, über neue Informationen zu reflektieren. Dies kann zu Pausen führen, in denen nicht gesprochen wird. Es ist wichtig, dass der Lerncoach diese Stille zulässt, da sie ein bedeutender Bestandteil der Gesprächsführung ist [62].

Ein weiteres Kommunikationsinstrument ist das Aufstellen und Diskutieren von Hypothesen [62]. Durch die Anwendung der bereits beschriebenen Gesprächselemente verschafft sich der Lerncoach einen Überblick über die Anliegen des Klienten. Hierbei können Hypothesen formuliert werden, die den Klienten helfen, ihre Sichtweise zu erweitern [62]. Dies schafft neue Anknüpfungspunkte und ermöglicht dem Coach, neue Interventionen zu planen [62]. Es ist jedoch entscheidend, die Hypothesen gemeinsam mit den Klienten zu erörtern, da sie oft auf den Vorerfahrungen des Coaches basieren [62].

Abschließend sollen nun Fragearten erläutert werden. Fragen können verwendet werden, um Ressourcen im Klienten zu aktivieren [62]. Zusätzlich kann es verwendet werden, um Lösungen anzuregen [62]. Somit sind diese Fragearten geeignet, um in Phase vier des Prozesses eine Lösung anzuregen. Die Fragen sind in Tabelle 1 beschrieben. Sie werden definiert, die Ziele der Fragen werden aufgezeigt und für jede Frage wurde ein Beispiel verfasst.

Frageart	Beschreibung	Ziele der Frageart	Beispiel
Offene Fragen [30, 62]	Fragen, die nicht mit Ja oder Nein beantwortet werden können und zu ausführlicheren Antworten anregen.	Informationen sammeln, Denkprozesse anregen, Selbstreflexion fördern.	"Wie fühlst du dich dabei?"
Ausnahme suchen [30]	Fragen nach Momenten oder Situationen, in denen das Problem nicht präsent war oder weniger ausgeprägt war.	Ressourcen identifizieren, Hoffnung auf Veränderung wecken, erfolgreiche Strategien erkennen.	"Wann war das Problem zuletzt nicht so präsent?"
Skalierungsfrage [30]	Fragen, bei denen der Klient eine Situation oder ein Gefühl auf einer Skala von z.B. 1 bis 10 bewertet.	Einschätzung der Intensität oder des Fortschritts erhalten, Entwicklungen sichtbar machen.	"Auf einer Skala von 1 bis 10, wie sicher fühlst du dich?"
Zirkuläre Frage [30, 61]	Fragen, die Unterschiede oder Veränderungen in Beziehungen oder zwischen Zeitpunkten beleuchten.	Systemische Zusammenhänge aufdecken, Muster erkennen, Perspektivenwechsel anregen.	"Wie hat sich dein Verhältnis zu X seitdem verändert?"
Hypothetische Frage [30]	Fragen, die eine hypothetische Situation beschreiben, in der sich etwas verändert hat oder verändern könnte.	Neue Lösungsmöglichkeiten denken, Kreativität anregen, alternative Szenarien in Betracht ziehen.	"Was würdest du tun, wenn das Problem morgen weg wäre?"
Wunderfrage [30]	Fragen nach einer idealen Zukunft, in der ein Wunder geschehen ist und das Problem gelöst wurde.	Ziele und Wünsche klären, Motivation stärken, Lösungsfokus etablieren.	"Wenn morgen ein Wunder geschieht, was wäre anders?"
Reframing [30]	Fragen, die dazu dienen, eine Situation, ein Problem oder ein Verhalten in einem neuen Licht oder Kontext zu sehen.	Bedeutungen neu interpretieren, festgefahrene Sichtweisen aufbrechen, positivere Perspektiven entwickeln.	"Könnte man das auch als eine Stärke sehen?"
Zurückblicken [62]	Fragen, die den Klienten dazu anregen, auf vergangene Ereignisse oder Erfahrungen zurückzublicken.	Lernprozesse und Entwicklungen erkennen, vergangene Strategien identifizieren, Zusammenhänge verstehen.	"Wie bist du in der Vergangenheit damit umgegangen?"
In die Zukunft blicken [62]	Fragen, die den Klienten dazu anregen, über künftige Ereignisse, Ziele oder Entwicklungen nachzudenken.	Zukunftsorientierte Planung, Zielsetzung, Motivation und Hoffnung stärken.	"Wo siehst du dich in einem Jahr in dieser Angelegenheit?"
Extreme erwägen [62]	Fragen, die extreme Szenarien oder Auswirkungen eines Verhaltens oder einer Entscheidung beleuchten.	Bewusstsein für Konsequenzen schärfen, Risiken und Chancen evaluieren, Entscheidungsfindung unterstützen.	"Was wäre das schlimmste/beste Ergebnis in dieser Situation"

Tabelle 1 Fragearten im Lerncoaching

Kommunikation ist im Lerncoaching wesentlich. Sie verbindet Lerncoach und Klient. Ein erfahrener Lerncoach nutzt verschiedene Techniken. Diese ermöglichen eine effiziente Gesprächsführung. Das Hauptziel dieser Kommunikationsfähigkeiten ist die Unterstützung der Klienten, um Einsicht in die Lernschwierigkeiten zu gewinnen und passende Lösungen sowie Strategien zu deren Bewältigung zu finden. Diese Kommunikationswerkzeuge können durchgängig während des Coachings angewandt werden und sind in allen Phasen relevant.

2.2.2 Phasen des Lerncoachings

Obwohl es möglich ist, in jeder Lerncoaching-Sitzung alle Phasen zu durchlaufen, ist dies nicht immer erforderlich. Die Phasen können über mehrere Sitzungen verteilt sein [61, 62]. Lerncoachinggespräche folgen laut Nicolaisen [2017] einem strukturierten Ablauf, der in fünf Phasen unterteilt ist:

Die erste Phase, „Einstieg“, dient dazu, den Kontakt zwischen dem Lerncoach und den Lernenden herzustellen. Hier werden wichtige Rahmenbedingungen festgelegt, wie beispielsweise die Zeit und den Ablauf des Gesprächs. Auch die Rollen beider Parteien werden geklärt, Vertraulichkeitsvereinbarungen getroffen und der genaue Auftrag für das Coaching definiert. [62]

In der zweiten Phase, „Thema erfassen“, steht die präzise Formulierung des zu bearbeitenden Themas im Mittelpunkt. Dabei werden verschiedene Techniken der Gesprächsführung eingesetzt. [62]

Die dritte Phase, „Ziel erfassen“, konzentriert sich auf die Zielformulierung im Dialog zwischen Lerncoach und Lernendem. Hierbei wird angestrebt, klare und handlungswirksame Ziele zu definieren, die im Coachingprozess verfolgt werden sollen.

In der vierten Phase, „Lösungen“, steht die Bearbeitung des erfassten Themas und der begleitenden Aspekte im Vordergrund. Dabei wird ein ressourcen- und lösungsorientierter Ansatz gewählt. Die praktische Umsetzung des zuvor ermittelten Ziels im Alltag steht dabei im Fokus. [62]

Schließlich erfolgt in der fünften Phase, dem „Abschluss“, das Zusammenfassen des erreichten Fortschritts im Bearbeitungsprozess. Es werden konkrete Vereinbarungen für die nächsten Schritte getroffen, und es kann die Planung einer weiteren Sitzung erwogen werden. Die Möglichkeit einer gegenseitigen Rückmeldung rundet die Gesprächsphase

ab. [62] Diese strukturierten Phasen dienen laut Nicolaisen [2017] als Leitfaden für erfolgreiche Lerncoachinggespräche.

Gemäß Hardeland [2023] besteht der Lerncoaching-Zyklus aus zwei bis sechs Sitzungen, wobei jede einzelne Sitzung eine Dauer von 15 bis 90 Minuten haben kann [30]. Die Zeitabstände zwischen den Sitzungen variieren von einer bis zu sechs Wochen, abhängig von den Bedürfnissen der Klient*innen und der spezifischen Thematik [30]. Die Autorin hat den Lerncoaching-Prozess in sechs Phasen unterteilt, im Gegensatz zu Nicolaisen [2017], der fünf Phasen beschrieben hat: "1. Kontakt herstellen", "2. Anliegen erfassen", "3. Lerncoaching-Ziel(e) festlegen", "4. Lösungen/Strategien entwickeln", "5. Ergebnisse zusammenfassen/einzelne Sitzung beenden" und "6. Reflexion und gegenseitiges Feedback". Zu erwähnen ist, dass Hardeland [2023] die letzte Phase von Nicolaisen [2017] in zwei Schritte unterteilt. Dennoch sollte angemerkt werden, dass Nicolaisen [2017] in seiner letzten Phase sowohl die Zusammenfassung als auch die Reflexion behandelt. Dadurch sind fast keine Unterschiede erkennbar.

2.2.2.1 Einstieg/Kontakt herstellen

Im ersten Schritt soll ein Kontakt zwischen Lernenden und Lerncoach hergestellt werden [62]. Dies soll eine bewusste und aktive Kooperation zwischen Klienten und Coach ermöglichen, auch als therapeutische Allianz bekannt [48].

Rapport ist ein Schlüsselbegriff in therapeutischen Beziehungen. Diese Beziehung wird als eine vertrauensvolle und harmonische Verbindung zwischen Therapeuten und Klienten verstanden. Sie basiert auf Zusammenarbeit, Kommunikation und Empathie des Therapeuten sowie auf gegenseitigem Verständnis und Respekt [48]. Ebenso wird der Begriff "Patientenrapport" als eine harmonische Beziehung definiert. Aufgrund der gemeinsamen Grundthemen wie Zusammenarbeit, Gegenseitigkeit, Gleichberechtigung und Wachstum werden diese Begriffe oft als austauschbar betrachtet. [48]

Leach [2005] fasst in seiner Literaturarbeit Informationen aus verschiedenen Forschungsarbeiten zusammen, die aufzeigen, wie der Rapport verbessert werden kann. Fachleute nutzen unterschiedliche Techniken und Verhaltensweisen, um Vertrauen, Kommunikation und die Beziehung zu Klienten zu verbessern. Dazu zählen aktives Zuhören und die Berücksichtigung der Klientenbedürfnisse. Wichtig ist auch, Urteile Fachjargon und eine autoritäre Haltung zu vermeiden. Auf persönlicher Ebene sollten Fachleute verlässlich, offen, flexibel und unterstützend sein, wobei Freundlichkeit, Echtheit

und Empathie eine Rolle spielen. Die Verwendung offener Fragen und die Erklärung von Gründen für Verfahren und Entscheidungen sind ebenfalls hilfreich, um den Rapport zu stärken.

Der Gesprächsbeginn zielt darauf ab, Vertrauen zu schaffen und Rapport aufzubauen. Danach erfolgt eine Rollenklärung. Der Klient oder die Klientin wird darüber informiert, dass der Coach in einer Rolle ähnlich einem/einer Trainer*in im Sport agiert, während der Klient wie ein Sportler handelt. Der Schwerpunkt liegt darauf, gemeinsam eigenes Verhalten oder schwierige Lernsituationen zu analysieren und Verbesserungsmöglichkeiten zu erkunden. [62]

Wichtig ist, dass der Coach betont, dass die Rolle nicht darin besteht, dem Gesprächspartner etwas abzunehmen, sondern dem Klienten dabei zu begleiten, gute Wege und Lösungen für sich selbst zu finden. Die Botschaft ist, dass alles, was der Gesprächspartner benötigt, bereits in ihm oder ihr selbst steckt. Diese Rollentransparenz zielt darauf ab, mögliche Unsicherheiten beim Gesprächspartner zu vermeiden und die Zusammenarbeit in einer kooperativen Atmosphäre zu erleichtern. [62]

Neben dem Rapport und der Rollenklärung werden zu Beginn organisatorische Themen bearbeitet. Die zeitliche Dauer bzw. der Ablauf der Coachingsitzungen werden erläutert [62]. Ist ein Kontakt hergestellt, die Rollen geklärt und die organisatorischen Themen bearbeitet, kann der Prozess in die nächste Phase übergehen.

2.2.2.2 Thema/Anliegen erfassen

In der zweiten Phase des Lerncoachings steht die Herausarbeitung des Themas oder des Anliegens des Coachees im Vordergrund. Die Kommunikation zwischen Lerncoach und Coachee, gestützt auf die Grundbausteine der Kommunikation aus Abschnitt 2.2.1, ermöglicht es dem Coachee, sich seiner oder ihrer Thematik bewusst zu werden. Nicolaisen [2017] betont, dass die klare Erkenntnis des Themas oft durch verschiedene Faktoren erschwert wird. Hier erweisen sich aktives Zuhören, Zusammenfassen und Widerspiegeln als besonders wertvoll. Der Coach unterstützt die Klienten dabei, die Komplexität der Problematik zu entschlüsseln, sodass sich am Ende dieser Phase ein klarer Schwerpunkt herausbildet, an dem der Coachee weiterarbeiten möchte.

2.2.2.3 Ziel erfassen

Laut Nicolaisen [61, 62] konzentriert sich die Zielsetzung primär im Bereich des Lerncoachings auf die Steuerung von Aktivitäten. Sie lenkt und konkretisiert die freigesetzte Energie. Das Bewältigen von Lernschwierigkeiten wird effektiver, wenn die Zwecke und Ziele des Lerncoachings klar definiert sind. Dafür stellt der Autor Kriterien, die bei einer Zielsetzung befolgt werden können auf. Jedoch müssen nicht alle Kriterien in jeder Zielsetzung verarbeitet werden.

Ein Kriterium laut Nicolaisen [2017] ist die Zielerfassung auf der Haltungs- und Verhaltensebene. Verhaltensziele konzentrieren sich auf bestimmte Verhaltensweisen in spezifischen Situationen und bieten äußere Kriterien für messbaren Erfolg, ohne jedoch das innere Erleben zu berücksichtigen. Haltungsziele hingegen beziehen die Identität, Werte und Motive des Lernenden in die Zielformulierung ein. Ein wichtiger Aspekt der Zielsetzung ist die SMART-Formulierung eines Ziels [30, 61, 62, 102]. SMART ist ein Akronym und steht für: Spezifisch, Messbar, Anspruchsvoll/Attraktiv, Realistisch und Terminiert [62]. Die SMART-Formulierung hilft dabei ein sehr konkretes und spezifisches Ziel zu formulieren [30]. Ein Beispiel für eine Zielformulierung könnte wie folgt lauten: „Innerhalb der nächsten sechs Monate strebe ich an, durch regelmäßiges Üben - 30 Minuten dreimal pro Woche (realistisch) und basierend auf meinen Grundkenntnissen (spezifisch) - einen flüssigen, fünfminütigen Vortrag auf Spanisch zu halten (messbar), um im nächsten Urlaub in Spanien selbstbewusst kommunizieren zu können (anspruchsvoll und attraktiv).“. Auf Basis der Arbeit von Storch and Krause [2007] formulieren Nicolaisen [2017] und Hardeland [2023] handlungswirksame Zielformulierungen für das Lerncoaching. Ein Aspekt davon ist es Ziele zu formulieren die ausschließlich unter eigener Kontrolle stehen. Als zweiter Punkt beschreiben die Autoren Annäherungsziele als weiteres Kriterium für Ziele. Lernende sollen ihre Ziele mit einem positiven Zielzustand formulieren. Als letztes Kriterium sollten Zielformulierungen durch positive nonverbale somatische Marker, wie Körperreaktionen und -empfindungen, unterstützt werden, um zu überprüfen, ob das Ziel für den Lernenden stimmig ist, da emotionale Bewertungen und körperliche Reaktionen eng miteinander verknüpft sind [30, 62].

Weiterhin ist es für Nicolaisen [2017] wichtig, dass Ziele im Gegenwartsmodus formuliert werden, um eine größere Identifikation und Motivation zu erzeugen. Zuletzt

sollten die Größe und der Kontext von Zielen berücksichtigt werden, um sicherzustellen, dass sie weder zu überwältigend noch zu trivial sind und in den richtigen Kontext eingebettet sind, um die höchste Wirksamkeit zu erreichen

2.2.2.4 Lösungen

Nachdem ein Anliegen herausgearbeitet wurde, kann ein präzises Ziel in der letzten Phase erfasst werden. In diesem Abschnitt werden die bisherigen Informationen und das formulierte Ziel genutzt, um passende Lösungen für den Klienten zu finden.

Der Coach kann Fragemethoden als Interventionsmöglichkeit benutzen. Mithilfe der Suche nach Ausnahmen kann der/die Klient*in überprüfen, ob es Situationen gibt, in denen das Problem nicht vorhanden ist. Dieser lösungsorientierte Ansatz hilft den Klienten ihr Ziel umzusetzen, mit Ressourcen, die schon vorhanden sind. Zirkuläre Fragen können dabei helfen eine neue Sichtweise zu erlangen. Zusätzlich regt es den Klienten, Beziehungen und Wechselwirkungen in seinem sozialen Umfeld zu reflektieren. Eine weitere Methode ist die Benutzung der Wunderfrage. Diese Fragetechnik eröffnet eine Vision einer idealen Zukunft, wodurch der Klient Schritte zur Verwirklichung dieser Vision entwickeln kann. Durch Reframing können Situationen neu interpretiert und dadurch neue Handlungsoptionen aufgezeigt werden. Das Reflektieren vergangener Erfahrungen hilft dem Klienten, zukünftige Ressourcen und Strategien zu identifizieren. Das Vorausschauen motiviert zur Planung und Realisierung künftiger Ziele.

Um den Klienten zur Ideenfindung zu inspirieren, kann der Coach Hypothesen anbieten und diese mit dem Klienten erörtern. So kann der Coachee neue Blickwinkel entdecken und mit dem Coach Lösungen erarbeiten.

2.2.2.5 Abschluss

In der folgenden Ausführung wird die finale Phase des Lerncoachings gemäß Nicolaisen [2017] detailliert besprochen. Dabei fließen auch Erkenntnisse aus den letzten zwei Phasen des Lerncoachings nach Hardeland [2023] ein, da sie mit Nicolaisens Ansatz in Einklang stehen.

Am Ende jeder Sitzung werden die bisher erreichten Ergebnisse durch den Lerncoach zusammengefasst [30, 62]. Gemeinsam mit den Klienten wird das weitere Vorgehen auf Basis des formulierten Ziels geplant. Für eine Umsetzung im Alltag kann der Coach spezielle Fragen stellen, um die Handlung des Coachees vorzubereiten. Beispiele

nach Nicolaisen [2017], „In welcher konkreten Situation wirst Du das Ergebnis des Coachings einsetzen?“, „Wie genau wirst Du vorgehen?“, „Was wirst Du tun?“. Neben der inhaltlichen Vertiefung der abschließenden Phase des Lerncoachings werden auch organisatorische Elemente berücksichtigt und konstruktives Feedback ausgetauscht. Bei Bedarf wird eine weitere Sitzung koordiniert, während Coach und Klient sich gegenseitig wertvolle Rückmeldungen bieten [30, 61, 62].

2.2.3 Limitationen des Lerncoachings

Es ist zu erkennen, dass das Konstrukt Lerncoaching nur bedingt Literatur aufweisen kann. Die Theorie, Ziele und Wirkungen des Lerncoachings werden genau beschrieben [30, 61, 62, 71, 74], jedoch fehlen wissenschaftliche Überprüfungen zu den genauen Wirkungen des Lerncoachings. Es ist aber zu erwähnen, dass das Lerncoaching auf dem Coaching basiert und es viele Überschneidungen gibt. Zum einen ist der Prozess der Phasen sehr ähnlich zu anderen Coaching-Arten. Zum anderen wird die Kommunikation des Coachings auch im Lerncoaching übernommen. Zusätzlich haben mehrere Literaturanalysen [12, 13, 39, 56, 95, 106, 110] gezeigt, dass Coaching eine positive Wirkung auf unterschiedliche Konstrukte haben kann. Darunter die Selbstwirksamkeit und auch die Motivation. Cannon-Bowers et al. [2023] zeigten außerdem, dass die unterschiedlichen theoretischen Ansätze des Coachings keinen Einfluss auf den Erfolg hatten. Somit kann man davon ausgehen, dass auch das Lerncoaching einen positiven Einfluss auf diese Konstrukte haben kann. Dennoch wäre es ratsam das Konstrukt in Zukunft noch genauer zu untersuchen.

2.3 Motivation, Selbstwirksamkeit

Motivation und Selbstwirksamkeit sind zentrale Konzepte in der Psychologie. [113]. Verschiedene psychologische Schulen wie Behaviorismus, Humanistische Psychologie und Tiefenpsychologie haben diese Konzepte beeinflusst und weiterentwickelt. Jeder Ansatz bietet eine einzigartige Perspektive auf die Ursachen und Mechanismen von Motivation und Selbstwirksamkeit. Sie unterscheiden sich jedoch in ihren methodischen Ansätzen und theoretischen Grundlagen. Pintrich and Groot [1990] entdeckten, dass motivationale Elemente einen signifikanten Zusammenhang zwischen der kognitiven Beteiligung der Schüler*innen und ihren akademischen Leistungen im Klassenzimmer hatten. Es wurde festgestellt, dass kognitive Beteiligung und Schülerleistung positiv mit

Selbstwirksamkeit korrelieren. Diese Erkenntnisse betonen die Bedeutung von Motivation und Selbstwirksamkeit. Sie sind entscheidend, um Schüler*innen zu unterstützen und ihre Lernprozesse zu fördern.

In dieser Arbeit wird das Erwartungswert-Modell von McKeachie et al. [1987] verwendet um die Konstrukte zu beschreiben. Das Modell ist ein Resultat aus einer Literaturübersicht. Erwartungswertmodelle sind kognitive Motivationsmodelle, die sich von denen der Psychodynamik, der Ich-Psychologie, der Lerntheorie, des Triebs und der Humanpsychologie unterscheiden [55]. Sie konzentrieren sich darauf, wie der subjektive Wert von Aufgaben und Erfolgserwartungen die Motivation beeinflussen [55, 75]. Die Literaturübersicht bezieht sich auf Universitätsstudierende, somit stellt sie den passenden Rahmen für diese Arbeit. Der in der Methodik beschriebene und in dieser Arbeit verwendete Fragebogen leitet sich aus dieser Theorie ab.

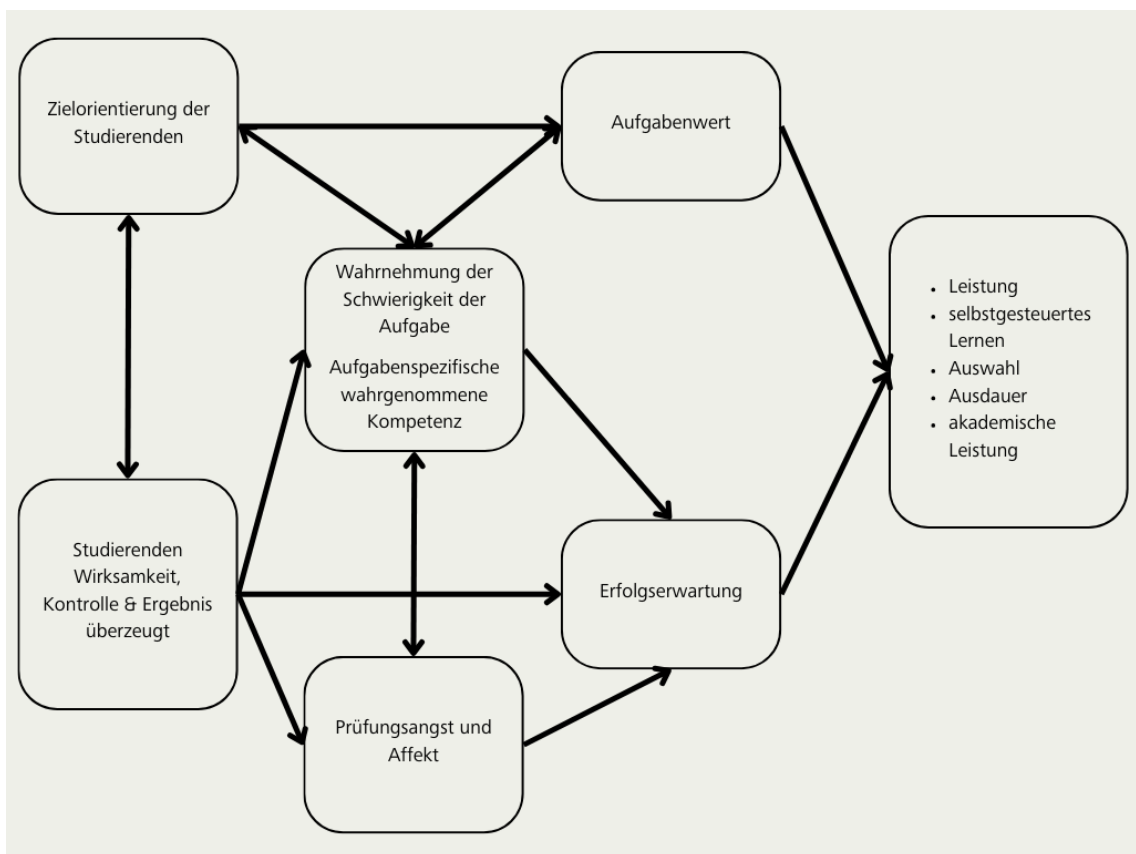


Abbildung 1: Aufbau Erwartungswert-Modell nach McKeachie et al. [1987]

In Abbildung 1 kann der Ablauf des Erwartungswert-Modells beobachtet werden. McKeachie et al. [1987] unterteilen das Modell in zwei Kanälen. Der Aufgabenwert Kanal und der Erfolgserwartungskanal. Diese werden nun auf Basis von McKeachie et al. [1987] erläutert:

Wahrgenommene Selbstkompetenz und Wahrnehmung der Schwierigkeit der Aufgabe laufen direkt in die Erfolgserwartung. Ersteres wird als Einschätzung der Lernenden hinsichtlich ihrer Fähigkeit, eine bestimmte Aufgabe zu bewältigen definiert. Jedoch unterscheidet sich die Selbstkompetenz und die Erfolgserwartung voneinander. Die Interaktion zwischen der Wahrnehmung der Schwierigkeit der Aufgabe durch den Lernenden und der wahrgenommenen Kompetenz ergibt die Erfolgserwartung. Die Selbstkompetenz steht im Zusammenhang mit der Selbstwirksamkeit. Die Autoren weisen daraufhin, dass das Konstrukt mit der Leistung der Lernenden korreliert. Ein positives Selbstkonzept resultiert in höheren Leistungen der Studierenden.

Bei der Wahrnehmung der Aufgabenschwierigkeit wird zwischen der subjektiven und objektiven Aufgabenschwierigkeit unterschieden. Die subjektive Wahrnehmung der Schwierigkeit kann das Verhältnis zwischen der tatsächlichen Schwierigkeit der Aufgabe und der Erfolgserwartung der Studierende und ihrem späteren Leistungsverhalten modulieren.

Prüfungsangst ist eng mit der wahrgenommenen Selbstkompetenz verbunden und steht in einem negativen Zusammenhang mit der Erfolgserwartung. Die Prüfungsangst besitzt eine kognitive und eine emotionale Komponente, wobei im Modell die kognitive Komponente eine übergeordnete Rolle spielt. Eine hohe Prüfungsangst führt zu weniger effektiven Lernfähigkeiten, die zu einer schlechteren Vorbereitung der Studierenden führt. Daraus folgt auch eine Verschlechterung der Leistung. Die kognitive Komponente ist somit ein Vermittler des Einflusses der Angst auf die Leistung. Ein weiterer Aspekt der Prüfungsangst sind aufgabenfremde Gedanken. Bei einer erhöhten Prüfungsangst kann es bei den Lernenden dazu führen, dass Gedanken während einer Prüfung abschweifen, somit wird die Fähigkeit zur Testdurchführung vermindert. Die kognitiven Fähigkeiten der Studierenden steigern sich, wenn effektive Lernstrategien angewandt werden und Ablenkungen durch abschweifende Gedanken minimiert sind, wodurch beide Aspekte miteinander verbunden werden. Ein weiterer Aspekt ist der Affekt. Die emotionale Komponente beeinflusst die Studierende in ihren Überzeugungen gegenüber der Wirksamkeit, der Kontrolle und der Ergebnisse.

Der letzte Aspekt des Erfolgserwartungsmodells bezieht sich auf die Überzeugungen von Studierenden hinsichtlich ihrer Wirksamkeit, Kontrolle und erwarteten Ergebnisse. Es hat sich gezeigt, dass Studierende mit einer starken Kontrollüberzeugung

bessere Leistungen erbringen. Nach der Attributionstheorie beeinflusst die Zuschreibung von Ursachen für Erfolg und Misserfolg zukünftige Erwartungen. Drei Hauptdimensionen - Lokus (Ort der Kontrolle), Stabilität und Kontrollierbarkeit - sowie Selbstwirksamkeitsüberzeugungen sind entscheidend für die Erwartungshaltung und die Leistung von Studierenden.

Insgesamt geht der untere Pfad des Modells davon aus, dass Lernende mit einer hohen Erfolgserwartung für eine bestimmte Aufgabe tendenziell erfolgreicher sind im Abschluss dieser Aufgabe, verglichen mit Studierenden, die eine geringe Erfolgserwartung haben. Dabei wird diese durch die Mediatoren zuvor beeinflusst.

Im oberen Teil der Abbildung ist der Pfad des Aufgabenwertes zu finden. Dabei mündet die Zielorientierung der Studierenden in den Aufgabenwert. Dieser lässt sich noch in die folgenden Komponenten unterteilen: Zielerreichungswert, Interessenwert und Nutzwert. Das Konzept des Zielerreichungswert bezieht sich auf die Überzeugung eines Lernenden, dass eine Aufgabe eine Herausforderung darstellt. Der Interessenwert hingegen hängt mit dem intrinsischen Interesse der Studierenden am Inhalt der Aufgabe zusammen. Im Gegensatz dazu bezieht sich der Nutzwert auf den praktischen oder instrumentellen Aspekt einer Aufgabe, also darauf, wie nützlich sie ist, um ein bestimmtes Ziel zu erreichen. Die Zielorientierung der Studierenden können langfristige Ziele oder kurzfristige Ziele beinhalten. Sie beeinflussen die Wahrnehmung und die Auswahl der Aufgaben der Studierenden.

2.4 Chatbots

Chatbots zielen darauf ab, menschliche Konversationen nachzuahmen. Sie verarbeiten und generieren natürlichsprachliche Daten über Text- oder Sprachschnittstellen. Ihr Ziel ist es, Nutzer*innen zu helfen, ein bestimmtes Ziel zu erreichen oder ein spezifisches Bedürfnis zu befriedigen. [59]

Ein frühes Beispiel ist der Chatbot ELIZA, der 1966 von Joseph Weizenbaum entwickelt wurde [96]. Das Programm antwortete auf Fragen durch Mustererkennung und täuschte viele seiner Nutzer*innen erfolgreich, sodass diese glaubten, mit einem echten Menschen zu kommunizieren [96]. Mittlerweile werden Chatbots verwendet die auf maschinelles Lernen basieren. Maschinelles Lernen besteht aus Algorithmen, die statistische Modelle erstellen, die zur Vorhersage von Antworten verwendet werden. Das

Modell wird meist mit einem großen Datensatz trainiert [14]. Im Unterschied zu Chatbots, die auf Mustererkennung basieren, werden Antworten bei fortgeschrittenen Systemen aus der Eingabe des Benutzers synthetisiert.

2.4.1 Generative Pre-trained Transformer (GPT)

Transformer-Modelle [109], die 2017 erstmals vorgestellt wurden, bestehen aus sogenannten Transformer-Blöcken. Jeder dieser Blöcke umfasst lineare Schichten, feedforward networks und self-attention layers. Bei Methoden, die auf Aufmerksamkeit basieren, ist das Kernstück die Einschätzung der Relevanz eines Elements im aktuellen Kontext, indem es mit einer Reihe anderer Elemente verglichen wird. Insbesondere bei der self-attention erfolgt der Vergleich mit anderen Elementen derselben Sequenz. Die Ergebnisse dieser Vergleiche werden genutzt, um auf der Basis des gegenwärtigen Inputs einen Ausgang zu erzeugen [40]. Für die Realisierung dieser Methode werden im Transformer sogenannte multihead attention layers eingesetzt. Diese Schichten ermöglichen die Verarbeitung von Wörtern in einem Satz, die auf unterschiedliche Weise miteinander in Verbindung stehen. Dank der multihead attention layers ist das Modell fähig, Daten aus verschiedenen Repräsentationssubräumen simultan an unterschiedlichen Positionen zu verarbeiten [109].

Das Sprachmodell GPT-4, welches in dieser Arbeit verwendet wird, basiert auf dem Generative Pre-trained Transformer, kurz GPT, den OpenAI im Jahr 2018 vorstellte [82]. Dieses Modell nutzt die Transformer-Architektur und kombiniert sie mit einem generativen Pre-Training-Ansatz. Ein Sprachmodell wird mit großen Mengen von Textdaten trainiert und anschließend für spezifische Aufgaben der natürlichen Sprachverarbeitung feinjustiert [82].

GPT-2, eine Weiterentwicklung von GPT, hat 1,5 Milliarden Parameter im Vergleich zu den etwa 100 Millionen Parametern von GPT [83]. Der Trainingsdatensatz von GPT-2 wurde erweitert und umfasst über 8 Millionen Dokumente mit einem Gesamttextvolumen von 40 GB [83].

GPT-3, der Nachfolger von GPT-2, besitzt 175 Milliarden Parameter [11]. Auch der Trainingsdatensatz wurde erweitert und umfasst mehr Sprachen als nur Englisch. Texte aus Büchern und Wikipedia wurden ebenfalls verwendet. GPT-3 zeigte starke Leistungen in vielen Bereichen der natürlichen Sprachverarbeitung, wie Übersetzung,

Fragebeantwortung und Lückentext-Aufgaben. Dies gilt für Zero-Shot-, One-Shot- und Few-Shot-Situationen. [11]. Allerdings hat GPT-3, wie fast alle großangelegten Sprachmodelle, Probleme mit Voreingenommenheit, Fairness und Repräsentativität [11].

Laut Ouyang et al. [2022] wurden die Probleme, die bei GPT-3 auftraten, in InstructGPT minimiert. Ein verbreitetes Problem großer Sprachmodelle ist ihre Fehlausrichtung [4]. Das bedeutet, dass sie zwar eine Aufgabe gut ausführen können, aber die Absicht des Benutzers missverstehen, was zu unbeabsichtigten Handlungen des Systems führt. Dies äußert sich darin, dass die Antworten eines Sprachmodells begrenzt nützlich sind. Um dieses Problem zu lösen, wurde das Modell mittels Reinforcement Learning from Human Feedback (RLHF) trainiert [69].

Der erste Schritt dieses Prozesses, das Supervised Fine-Tuning (SFT), wird von Ouyang et al. [2022] beschrieben. Demonstrationsdaten werden gesammelt, um ein überwachtes Modell zu trainieren, das als SFT-Modell bezeichnet wird. Prompts werden ausgewählt und einer Gruppe menschlicher Bewerter*innen präsentiert, die dann das erwartete Ergebnis für das Prompt notieren. Die Prompts wurden entweder direkt von den Bewerter*innen erstellt oder von OpenAI API-Anfragen bezogen. Die Bewerter erhielten drei Anweisungen zum Schreiben der Prompts. Einfache Prompts sollten willkürliche Aufgaben mit einer gewissen Vielfalt enthalten, Few-Shot-Prompts sollten Anweisungen mit mehreren Fragen und Antworten beinhalten, und benutzerbasierte Prompts sollten den Nutzungsfällen entsprechen, die durch die OpenAI API erstellt wurden. Es gibt etwa 13.000 Trainingsanfragen in der SFT-Sammlung. Diese bildeten die Datengrundlage für das Training des SFT-Modells. Da die Datenmenge nicht ausreicht, kann angenommen werden, dass die Ergebnisse nach dem SFT noch nicht zufriedenstellend sind [84]. Um dieses Problem zu lösen, sollten die Bewerter*innen nun die Ausgaben des SFT-Modells bewerten, um ein Belohnungsmodell zu erstellen.

Im zweiten Schritt wird das Belohnungsmodell durchgeführt. Das Ziel ist es, ein Belohnungsmodell direkt aus den Daten zu gewinnen, um ein System zu schaffen, das menschliches Verhalten nachahmt. Dies kann das Problem der Fehlausrichtung des Sprachmodells lösen. Dafür wird ein Datensatz von 33.000 Prompts verwendet, die als Eingabe in das SFT-Modell gegeben werden. Das Modell erzeugt mehrere Ausgaben für jede Eingabe, die von den Bewerter*innen von der besten bis zur schlechtesten bewertet werden. Diese Daten werden verwendet, um das Belohnungsmodell zu trainieren. Da

jeder Prompt zu mehreren Antworten führt, gibt es wesentlich mehr bewertete Paare als Fragen im Datensatz, auf dem der Algorithmus trainiert wird. [70]

Der dritte und letzte Schritt, das Reinforcement Learning, wird ebenfalls von Ouyang et al. [2022] beschrieben. Die SFT-Methode wird durch Optimierung des Belohnungsmodells mittels Reinforcement Learning verbessert. Hierfür wird Proximal Policy Optimization (PPO) verwendet. Ein neuer Prompt wird aus dem Datensatz entnommen, und PPO liefert ein Ergebnis. Das Vorwärtsmodell gibt den Belohnungswert für eine Ausgabe zurück, der zur Aktualisierung des Modells verwendet wird.

InstructGPT wurde mittels menschlicher Eingaben evaluiert. Dabei wurden Hilfsbereitschaft, Wahrhaftigkeit und Unschädlichkeit untersucht [69]. Die Bewerter bevorzugten die Ausgaben von InstructGPT gegenüber GPT-3, obwohl GPT-3 100-mal mehr Parameter besitzt. Das Modell zeigt zudem Verbesserungen in Bezug auf Wahrhaftigkeit und Toxizität im Vergleich zu GPT-3, jedoch nicht bei Voreingenommenheit.

Als Nachfolger von GPT-3 ist GPT-4 konzipiert. Der technische Bericht von OpenAI zu GPT-4 wurde am 16. März 2023 auf arXiv veröffentlicht [66]. OpenAI gibt an, dass das System ähnlich wie InstructGPT mit menschlichem Feedback trainiert wurde.

Allerdings enthält der technische Bericht keine detaillierten Informationen zur Architektur von GPT-4 [66]. Stattdessen werden im GPT-4 Technical Report Paper auf arXiv hauptsächlich die Fähigkeiten, Einschränkungen und Sicherheitsmerkmale von GPT-4 beschrieben [66].

In Bezug auf Prüfungen erzielt GPT-4 Ergebnisse, die in den meisten professionellen und akademischen Tests vergleichbar mit menschlicher Leistung sind [66]. Besonders hervorgehoben wird die Leistung bei der Uniform Bar Examination, in der GPT-4 zu den besten 10% der Teilnehmer zählte, im Gegensatz zu GPT-3.5, der zu den schlechtesten 10% gehörte [66].

Zur Bewertung von GPT-4 wurden auch traditionelle Benchmarks verwendet, die speziell für die Evaluierung von Sprachmodellen entwickelt wurden. Hierbei übertraf das System bestehende Sprachmodelle [66]. Ein neues Merkmal von GPT-4 im Vergleich zu seinem Vorgänger GPT-3.5 ist die Fähigkeit, visuelle Eingaben zu verarbeiten. Das System akzeptiert Prompts, die sowohl Bilder als auch Text enthalten, allerdings werden keine detaillierten Informationen zu den Fähigkeiten dieser Funktion gegeben [66].

Trotz der Fortschritte hat GPT-4 weiterhin Einschränkungen. Es erfindet immer noch Fakten und macht Fehler in Erklärungen [66]. Dennoch übertrifft es andere Modelle in offenen Standards wie TruthfulQA, die die Fähigkeit eines Modells bewerten, zwischen falschen Aussagen und der Wahrheit zu unterscheiden [49]. Solche Tests sind normalerweise für große Sprachmodelle eine Herausforderung [49]. Abschließend ist zu erwähnen, dass GPT-4 82% weniger antwortet als GPT-3, wenn Fragen gestellt werden, die nicht zum erlaubten Inhalt des Systems gehören, und nur 0,73% toxischen Inhalt erzeugt, verglichen mit den 6,48% von GPT-3.5 [66].

Mit der verbesserten Sprachverarbeitungs- und Kontextverständnisfähigkeiten ist GPT-4 ideal für den Chatbot in dieser Arbeit. Die Verbesserung der multihead attention layers und das Training mit menschlichem Feedback haben zu erheblichen Fortschritten bei der Entwicklung von GPT zu GPT-4 geführt, so dass GPT-4 in den im Vergleich zu anderen Sprachmodellen State-of-the-Art ist [66]. Diese umfassen eine bessere Verarbeitung von Sprachkontexten, eine geringere Fehlausrichtung und eine geringere Menge an schädlichen Inhalten. GPT-4 eignet sich für zuverlässige Chatbots, die eine natürliche und sichere Interaktion ermöglichen.

2.4.2 Few-Shot-Learning

Few-Shot-Learning bezeichnet den Ansatz, bei dem einem Modell zunächst eine begrenzte Anzahl von Beispielen für eine spezifische Aufgabe gezeigt wird, bevor die eigentliche Aufgabe gestellt wird [11]. In der Arbeit von Brown et al. [2020] wurde gezeigt, dass große Sprachmodelle wie GPT-3 gute Ergebnisse mit Few-Shot-Learning erzielen. GPT-4 setzt diese Entwicklung fort und zeigt verbesserte Leistungen auf akademischen Benchmarks. Im Vergleich mit den besten State-of-the-Art-Modellen, die sowohl für spezifische Benchmarks trainiert wurden als auch solchen, die im Few-Shot-Kontext evaluiert wurden, übertrifft GPT-4 bestehende Sprachmodelle in allen Benchmarks. Dies gilt für alle Datensätze, mit der Ausnahme von DROP, einem Benchmark, der auf Leseverständnis und Arithmetik fokussiert ist.

Zum Zeitpunkt dieser Arbeit existiert kein spezifischer Datensatz für Lerncoaching. Angesichts der Fähigkeit von GPT-4, mit nur wenigen Beispielen State-of-the-Art-Ergebnisse zu erzielen, erweist sich das Modell als ideal, um einen Chatbot zu entwickeln, der

die Funktionen eines Lerncoachs nachahmt. Jedoch müssen dafür Beispiele für das Few-Shot-Learning erstellt werden.

3 Aktueller Forschungsstand

Im zweiten Abschnitt dieser Arbeit wird der aktuelle Forschungsstand aufgezeigt. Dabei wird darauf eingegangen in welchen Domänen Chatbots verwendet werden. Es soll eine Forschungslücke aufgezeigt werden, in der diese Arbeit einen Beitrag leistet.

3.1 Gesundheitswesen

Chatbots im Gesundheitswesen haben verschiedene Ziele. Sie unterstützen Patienten und Gesundheitspersonal [8, 37, 47, 57, 91, 108]. Sie trainieren medizinisches Fachpersonal, aber auch Patienten [37, 57]. Chatbots helfen älteren Menschen [37]. Sie sind auch bei Diagnosen hilfreich [37, 57]. Zudem unterstützen sie die Ausbildung von Personal [37, 57]. Außerdem tragen sie zur Prävention im Gesundheitswesen bei [37]. Der Fokus bei den Zielen liegt bei den Patient*innen [37]

Die Benutzung von Chatbots im Gesundheitswesen ist laut einer Literaturübersicht aus dem Jahr 2018 ein aufstrebendes Forschungsgebiet [47]. Dabei haben die Autor*innen nur wenige Studien gefunden, bei denen uneingeschränkte Eingabemöglichkeiten in der menschlichen Sprache mit den Chatbots möglich waren. Die Autor*innen sagen aus, dass die meisten Chatbots regelbasiert funktionieren und denen anderen Domänen hinterher liegen. Hingegen stellte die Studie von Tudor Car et al. [2020], lediglich zwei Jahre später, fest, dass die Mehrheit der Chatbots zu diesem Zeitpunkt auf Ansätzen des maschinellen Lernen aufbauten, wodurch die Möglichkeiten der Nutzereingaben erheblich erweitert wurden. Ein ähnlicher Trend kann bei den Studien von Schachner et al. [2020] und Bin Sawad et al. [2022] erkannt werden. Beide Studien untersuchen Chatbots in Bezug auf Patienten, die eine chronische Krankheit besitzen. Die erste Studie konnte im Jahr 2020 noch wenige Ergebnisse berichten. Die zweite Studie hingegen zeigt, dass Chatbots einen positiven Einfluss auf die Unterstützung von Patient*innen haben können. Patient*innen scheinen eine erhöhte Zufriedenheit beim Interagieren mit Chatbots zu erfahren, im Vergleich zu fortlaufenden Nachuntersuchungen aufgrund ihrer chronischen Krankheiten. Die Nutzererfahrung wurde von Bin Sawad et al. [2022] als meist gemessenes Kriterium festgestellt. In den Studien, die Chatbots verwenden, um Patient*innen mit chronischen Krankheiten zu unterstützen, spiegelt die Nutzererfahrung die positiven Auswirkungen und die Verbesserung der Lebensqualität wider. Diese vier

Literaturübersichten zeigen die rasche Weiterentwicklung der Chatbots in der Domäne des Gesundheitswesens.

Die meisten Chatbots im Gesundheitswesen funktionieren Multimodal oder Textbasiert [37]. Nur die wenigsten benutzen Audio als Ein- oder Ausgabe [37]. Einige der Chatbots verfolgen einen Beratungs- oder Coachingansatz [37]. Diese sollen den Patient*innen oder dem Personal in ihren täglichen Aufgaben helfen. Sie können kurze Interventionen ermöglichen, um das Verhalten der Patient*innen zu beeinflussen. Zusätzlich versuchen sie ein Verhältnis mit den Patient*innen aufzubauen, um sie zu Veränderungen zu ermutigen.

Ein Problem welches in nur wenigen Studien behandelt wird, ist die Patientensicherheit [44, 47]. Die Empfehlungen von Chatbots können Fehler beinhalten, die vor allem im Gesundheitswesen ein großes Problem darstellen. Ein weiteres Problem besteht darin, dass die meisten Studien keine einheitlichen Evaluationen der Chatbots durchführen [8, 47, 91]. Zusätzlich werden technische Details der Chatbots kaum berichtet [8].

Große Sprachmodelle haben vor allem seit der Veröffentlichung von ChatGPT einen Einfluss auf die Verwendung von Chatbots im Gesundheitswesen [90]. Sie vereinfachen den klinischen Arbeitsablauf und könnten Kosten sparen [90]. Zusätzlich kann die Effizienz in der Gesundheitsversorgung gesteigert werden [90]. Chatbots, die auf großen Sprachmodellen basieren, helfen bei der Erstellung effizienter Entlassungszusammenfassungen. Dies reduziert die Dokumentationslast im Gesundheitswesen [72, 90]. Sie verbessern Diagnostik, Krankheitsvorhersagen und Arzneimittelforschung [90]. Zusätzlich stellen sie auch allgemein verständliche Gesundheitsinformationen zur Verfügung [36, 90].

3.2 Industrie

Innerhalb der Industrie werden Chatbots in unterschiedlichen Situationen verwendet. Jedoch sind bis zum Zeitpunkt dieser Arbeit noch keine Literaturübersichten zu diesem Sektor verfasst worden. Dies zeigt, dass die Thematik auch in der Industrie ein neues Forschungsthema ist. Einzelne Forschungsarbeiten [54, 58], die das Thema der Chatbots im Ganzen erfassen, teilen die Chatbots in der Industrie in folgende Kategorien ein, ohne ins Detail zu gehen: E-Commerce, Finanzen, Unterstützung bei Verhandlungen,

Unterstützung der Arbeit. Zu diesen Kategorien sind einzelne Forschungsarbeiten zu finden [1, 32, 38, 107, 116].

Chatbots werden im Bereich der Geschäftsunterstützung als softwarebasierte Werkzeuge entwickelt, die Mitarbeiter*innen bei internen Geschäftsprozessen helfen sollen [58]. Das Hauptziel besteht darin, eine effiziente und einfach zu bedienende Schnittstelle für traditionelle Geschäftsprozesse zu entwickeln, um eine teilweise automatisierte Durchführung dieser Prozesse zu ermöglichen [58]. Dazu gehören Finanzen, Verhandlungen und Arbeitsunterstützung [58]. In E-Commerce werden Chatbots verwendet um Kund*innen bei Fragen Antworten zu liefern oder sie bei ihrem Kaufprozess zu unterstützen [58].

Ein Problem bei der Verwendung von Chatbots in Unternehmen besteht darin, dass sensible Daten des Unternehmens und der Kund*innen möglicherweise in der Datenbank des Bots gespeichert sind und somit eine potentielle Sicherheitslücke entsteht [10, 46].

Obwohl nur wenige Forschungsarbeiten zu den einzelnen Sektoren existieren, werden Chatbots von einigen großen Firmen verwendet. Dabei zu nennen sind zum Beispiel Ikea, Zalando oder die World Health Organization [20]. Dies verdeutlicht, dass insbesondere in der Industriebranche die Praxis der Forschung vorangeht. Folglich besteht hier ein Nachholbedarf.

Laut der Studie von Eloundou et al. [2023] könnten fast 80% der amerikanischen Arbeitskräfte von der Einführung von Chatbots mit großen Sprachmodellen (LLM) betroffen sein, wobei mindestens 10% ihrer Arbeitsaufgaben beeinflusst werden könnten. Die Auswirkungen scheinen sich über verschiedene Einkommensklassen zu erstrecken, wobei insbesondere Berufe mit höherem Einkommen anfälliger für die Fähigkeiten von LLMs und LLM-getriebener Software sein könnten [19, 21, 80]. Diese Entwicklungen beschränken sich nicht nur auf Sektoren mit jüngstem Produktivitätswachstum. Die Studie von Eloundou et al. [2023] deutet darauf hin, dass der Einsatz von LLMs wie GPT-4 es ermöglichen könnte, etwa 15% aller Aufgaben amerikanischer Arbeitnehmer*innen schneller und dennoch qualitativ hochwertig zu erledigen. Bei der Betrachtung verschiedener Branchen zeigen sich insbesondere in der Informationsverarbeitung hohe Expositionsgrade, während Fertigung, Landwirtschaft und Bergbau geringere Expositionen aufweisen [21]. Die Forschung legt nahe, dass der Einfluss von LLMs wie GPT-4

wahrscheinlich weitreichend sein wird, und es wird erwartet, dass die wirtschaftlichen Auswirkungen von LLMs weiterhin zunehmen werden [21].

3.3 Bildungswesen

Chatbots haben im Bildungsbereich zunehmend an Bedeutung gewonnen, insbesondere in den letzten Jahren. Dies hat zur Entstehung zahlreicher Literaturübersichten geführt [33, 34, 68, 73, 99, 115].

Chatbots werden laut Pérez et al. [2020] regelmäßig in Bildungseinrichtungen verwendet. Die Studie unterscheidet dabei zwei Hauptkategorien: dienstleistungsorientierte und lehrorientierte Chatbots. Lehrorientierte Chatbots sind besonders vielseitig, werden für verschiedene Altersgruppen entwickelt und zeigen ihre Effektivität vor allem im Sprachunterricht. Sie können auf mehreren Plattformen implementiert werden. Jedoch heben Hobert und Meyer von Wolff [2019] hervor, dass vor allem mobile Anwendungen verwendet werden, um Chatbots im Bildungswesen umzusetzen.

Wollny et al. [2021] haben in ihrer Literaturübersicht die Ziele der Chatbots zusammengefasst. Die meisten Bots haben als Ziel die Fähigkeiten der Lernenden zu verbessern. Dabei agieren sie als Unterstützung bei dem Lernprozess. Ein weiteres Ziel ist die Steigerung der Effizienz der Bildung. Dies geschieht anhand von der Automatisierung wiederkehrender Aufgaben, um so Zeit einzusparen. Die zwei letzten Ziele sind die Steigerung der Motivation der Lernenden und Erhöhung der Verfügbarkeit von Bildung. Letzteres soll zeitliche Flexibilität ermöglichen, ohne auf physische Präsenz angewiesen zu sein. Wollny et al. [2021] arbeiten Rollen heraus, die die Chatbots einnehmen können. In der Lernrolle, werden Chatbots verwendet, um Lehre zu betreiben oder die Lernenden beim Lernen zu helfen. Dies kann realisiert werden, indem die Bots direkt in Lehrpläne integriert oder als ergänzende Hilfsmittel angeboten werden, beispielsweise zur Simulation eines Brieffreundes. Chatbots in der Unterstützungsrolle erleichtern den Studienalltag durch einfachen Zugang zu Informationen und Prozessautomatisierung. In ihrer Mentoren Rolle fokussieren sich Chatbots auf die persönliche Entwicklung der Studierenden. Dabei werden sie ermutigt, über ihren Fortschritt nachzudenken und ihre Lernprozesse zu reflektieren. Bei den Mentoring Methoden von Chatbots unterscheidet man Scaffolding, Empfehlungen und Informationsvermittlung. Diese reichen von der Förderung von Lebenskompetenzen über die Empfehlung von Teamkolleg*innen bis hin zur

Bereitstellung von Informationen. Thematisch decken die Chatbots Bereiche wie selbst-reguliertes Lernen, Lebens- und Lernkompetenzen ab, indem sie beispielsweise das Selbstwertgefühl stärken oder transformative Lernansätze vermitteln. Die Rolle, die in der Forschung bis jetzt am wenigsten untersucht wird, ist die Mentoren Rolle.

Die Lern- und Mentorrollen von Wollny et al. [2021] lassen sich der lehrorientierten Kategorie nach Pérez et al. [2020] zuordnen. Im Gegensatz dazu gehört die Unterstützungsrolle zur dienstleistungsorientierten Kategorie. Folglich lassen sich in der Literatur zu Chatbots im Bildungswesen zwei Hauptkategorien identifizieren, die drei unterschiedliche Rollen mit verschiedenen Zielen umfassen.

LLMs wie ChatGPT haben einen Einfluss auf das Bildungswesen und die Chatbots die darin verwendet werden. Lehrkräfte können das System zur Vorbereitung von Präsentationen, Erstellung von Aufgaben und Tests verwenden [79]. Schüler*innen wiederum können Fragen zu Themen stellen, Texte umschreiben lassen und Ideen sammeln [79]. Allerdings gibt es in der wissenschaftlichen Gemeinschaft auch Bedenken, dass die Bildung und ihre Integrität durch solche Systeme gefährdet sein könnten [41, 104]. Kritiker*innen der Technologie im Bildungsbereich argumentieren, dass ChatGPT das Plagieren von Hausaufgaben und akademischen Arbeiten erleichtert und eine Bedrohung für Online-Prüfungen darstellt [104]. Um dem entgegenzuwirken, gibt es bereits Plagiatsoftware, die Texte von Sprachmodellen erkennen soll.

3.4 Coaching

Der Einsatz von Chatbots im Coaching-Bereich ist bisher nur unzureichend erforscht. Der Fokus der Literatur liegt zum einen auf solche, die Klient*in bei psychologischen Problemen unterstützen [8, 37, 47, 108] und somit dem Gesundheitswesen zuzuordnen sind, und zum anderen auf Mentoring-Chatbots. [3, 60, 94, 115] Letztere sind zwar dem Coaching ähnlich, da die Konstrukte sich Methoden teilen, gehören jedoch primär in den Bereich des Bildungswesens.

Stella Kanatouri [2021] hebt in ihrer Arbeit hervor, dass die in der Coaching-Praxis verwendeten Chatbots nicht in der Lage sind, freie Gespräche zu führen. Außerdem betont sie, dass diese Bots keine Körpersprache deuten oder auf die Emotionen der Klient*innen eingehen können. Jedoch zeigten Ortega-Ochoa et al. [2023] in einer aktuellen Studie, dass es empathische Chatbots gibt, die sowohl die Lernleistung als auch die

Wahrnehmung von Lernenden positiv beeinflussen können. Dies zeigt die rasche Weiterentwicklung der Technologie im Coaching Sektor.

Tess ist einer der bekanntesten Chatbots im Coaching-Bereich [101]. Der Chatbot antwortet Nutzer*innen mit Aussagen, die von psychischen Gesundheitsexperten entwickelt wurden. Diese Aussagen sollen Empathie und Mitgefühl imitieren. Tess passt Interventionen an die Emotionen und Anliegen der Nutzer*innen an. Dies ähnelt der Anpassungsfähigkeit von Therapeuten bei verschiedenen Klient*innen. StudiCoachBot entwickelt von Mai et al. [2022] soll mithilfe eines Coaching Ansatzes Studierende bei ihrer Prüfungsangst unterstützen. Die Forscher*innen stellen fest, dass das Chatbot-Coaching in seiner jetzigen Form den Studierenden bereits hilft, über ihre Prüfungsangst zu reflektieren. Der StudiCoachBot wird von den Studierenden gut angenommen. Es scheint, dass die Studierenden es leicht finden, sich einem Chatbot gegenüber zu öffnen. Das Konzept der Dialoge des Chatbots wurde mithilfe einer Wizard-of-Oz Studie getestet, um den Chatbot auf Basis von Rasa [85] zu realisieren.

Die Verwendung von LLMs ist in der Coaching Literatur bisher nicht zu finden. Einzelne Aspekte des Coachings wie das Reflektieren von Aussagen [98], oder das Reframing [93] werden mithilfe von GPT Modellen untersucht. Jedoch nicht das ganzheitliche Konstrukt Coaching.

4 Präzisierung der Forschungsfrage

Chatbots waren schon vor der Entwicklung von Sprachmodellen wie GPT-4 ein wichtiges Thema. Mit der Einführung von GPT-4 hat sich das Forschungsfeld schnell entwickelt. Es erstreckt sich über die verschiedenen Sektoren. Dazu gehören das Gesundheitswesen, die Industrie und der Bildungsbereich. Große Sprachmodelle werden die Arbeitswelt beeinflussen. Dies ist eine wichtige Erkenntnis für die Industrie. Studierende müssen die Veränderungen durch diese Sprachmodelle verstehen. Es ist wichtig, diese Technologie gründlich zu prüfen. Sie sollte erst nach einer wissenschaftlich fundierten Prüfung in Unternehmen eingesetzt werden. Die Einführung eines Chatbot-basierten Lerncoachings birgt geringere Risiken als deren Einsatz im Gesundheitswesen oder in der Industrie. Im Bildungsbereich haben mehrere Studien die positiven Auswirkungen von Chatbots auf Studierende nachgewiesen. Dabei wurde allerdings dem Mentoring durch Chatbots, das dem Coaching am nächsten steht, bisher weniger Aufmerksamkeit geschenkt. In Bezug auf das Coaching mittels Chatbots, insbesondere unter Einbeziehung neuester Technologien, besteht eine Forschungslücke.

Aus dem theoretischen Rahmen dieser Arbeit geht hervor, dass Coaching, insbesondere in den Bereichen Motivation und Selbstwirksamkeit, positive Effekte haben kann. Diese Faktoren spielen in der modernen Bildung eine entscheidende Rolle. Angesichts des Potenzials von Lerncoaching, wird die Bedeutung eines Lerncoaching-Chatbots, der auf neuester Technologie wie GPT-4 basiert, besonders deutlich. GPT-4 hat in zahlreichen Benchmarks State-of-the-Art Ergebnisse erzielt und erscheint daher ideal geeignet, um einen Lerncoaching-Chatbot mithilfe von Few-Shot-Learning zu entwickeln.

Die Identifikation der Forschungslücke aus der vorhandenen Literatur, zusammen mit den Erkenntnissen aus dem theoretischen Rahmen, betont die Notwendigkeit und das Potenzial für die Entwicklung eines solchen Chatbots. Daraus ergibt sich folgende Forschungsfrage dieser Arbeit:

„Kann GPT-4 mittels Few-Shot-Learning als Lerncoach fungieren, um sowohl die Motivation als auch die Selbstwirksamkeit von Lernenden zu steigern?“

5 Methodik

Dieser Abschnitt beschreibt die Methodik der Studie und den Aufbau des Lerncoaching-Chatbots, einschließlich Forschungsdesign, Stichprobenziehung, Datenerhebung, verwendete Softwarelösungen und Datenauswertung.

5.1 Lerncoaching-Chatbot

Der LC kann als Online-Coaching der Stufe eins betrachtet werden, da nur über einen Chat kommuniziert wird. Somit stehen die Kommunikationstechniken des Lerncoachings im Vordergrund. Das Zusammenfassen und Widerspiegeln von Aussagen spielt eine noch größere Rolle, da Mimik, Gestik, oder auch zustimmende Geräusche bei einem Chatbot wegfallen. Wie im theoretischen Rahmen bisher erläutert wurde, können in jeder Lerncoachingsitzung alle Phasen durchlaufen werden [61, 62]. Zusätzlich wurde gezeigt, dass die Dauer des Coachings keinen Einfluss auf den Erfolg hat [30]. Es werden bei dem Lerncoaching-Chatbot alle Phasen in einer Sitzung durchlaufen. Außerdem wird eine zweite optionale Sitzung angeboten, die auch alle Phasen des Lerncoachings durchläuft.

In der ersten Phase steht die Begrüßung und die Rollenverteilung im Mittelpunkt. Damit jede Person die gleiche Ausgangslage besitzt wird dieser Teil standardisiert. Jedoch wird darauf geachtet nicht zu viele Informationen preiszugeben, da der LC mit einem NC verglichen wird. Somit wurde ein Text gewählt, welcher die Rollen klärt, jedoch bei beiden Chatbots verwendet werden kann: *„Hallo! Ich bin Mila, dein Lernassistent. Mein Ziel ist es, dich auf deiner Lernreise mit Informationen und Unterstützung zu begleiten. Ob du Fragen zu einem Thema hast, Lernressourcen suchst oder Tipps benötigst, lass es mich wissen. Worüber möchtest du heute sprechen?“*.

Bei Phase zwei wird ein Anliegen mit den Klienten geklärt. Dieses Anliegen ist bei jeder Person individuell und kann somit nicht standardisiert werden. Aus diesem Grund wird GPT-4 mithilfe von Few-Shot-Beispielen den Nutzer*innen auf ihre Anfragen antworten. Der Übergang zwischen Phase zwei und drei wird von einer Skalierungsfrage gekennzeichnet: *„Auf einer Skala von 1 bis 10, wobei 1 ‚unzureichend informiert‘ und 10 ‚ausreichend informiert‘ bedeutet, wie gut fühlst du dich informiert, um ein klares Ziel zu diesem Thema zu formulieren?“*. Skalierungsfragen sind eine wichtige Fragetechnik, vor allem im Online Coaching [88]. Zusätzlich wird der bisherige Chatverlauf von GPT-3.5

zusammengefasst und dem Teilnehmer*in präsentiert (GPT-3.5 wird für Zusammenfassungen genutzt, um Ressourcen zu sparen)

In Phase drei wird ein Ziel mit den Klient*innen formuliert. Die genaue Zielformulierung ist ein wichtiger Teil des Coachings und des Lerncoachings. GPT-4 kann trotz Few-Shot-Learning von der zugewiesenen Rolle abweichen. Somit wird dieser Teil bei jeder Person standardisiert. Jede Person kann ein individuelles Ziel besitzen, jedoch ist es möglich auf Basis der SMART Regel den Prozess der Zielformulierung zu vereinheitlichen. Zu jedem Aspekt der SMART Regel wird auf Basis der formulierten Theorie eine Frage gestellt, S: *„Was genau möchtest du in deinen eigenen Worten erreichen?“*, M: *„Woran würdest du merken, dass du dein Ziel erreicht hast?“*, A: *„Warum ist dieses Ziel wichtig für dich?“*, R: *„Ist das Ziel realistisch für dich?“*, T: *„Bis wann möchtest du dein Ziel erreicht haben?“*. Anschließend werden die Beteiligten aufgefordert, auf Basis der gesammelten Informationen und der SMART Regeln ein genaues Ziel zu formulieren.

Die Lösungsfindung auf Basis des Zieles steht in Phase vier im Mittelpunkt. Dabei besitzt jede Person ein anderes Ziel, mit einer andern Ausgangssituation und unterschiedlichen Ressourcen. Somit wird in diesem Teil GPT-4 auf Basis der Ziele den Teilnehmer*innen antworten. Dabei wird der Fokus im Online-Coaching auf einen schrittweisen Erfolg eines Ziels von Ribbers und Waringa [2015] übernommen. Um dies zu gewährleisten, wird jeder Person zu Beginn der Phase folgende Frage gestellt: *„Zum Abschluss besprechen wir gemeinsam, wie du dein definiertes Ziel in die Praxis umsetzen kannst. Wie kannst du dein formuliertes Ziel im Alltag umsetzen? Was wäre ein erster kleiner Schritt?“*.

In der letzten Phase des Lerncoachings verabschieden sich beide Parteien und der bisherige Verlauf des Coachings wird zusammengefasst. Für diesen Zweck wird GPT-3.5 wieder verwendet, um den kompletten Verlauf zusammenzufassen.

Insgesamt sind Phase eins und drei standardisiert und bei jeder Person gleich. Die restlichen Phasen werden von GPT-4 mithilfe von Few-Shot-Beispielen beantwortet.

5.2 Forschungsdesign

Nach Döring [2023] kann das Untersuchungsdesign in neun Kennzeichen unterteilt werden, die nun im Folgenden in Bezug auf diese Arbeit dargestellt werden.

Die Studie kennzeichnet sich als eine quantitative, experimentelle Feldforschung mit mehrfacher Messung, die sich auf die Grundlagenforschung konzentriert. Es handelt

sich hierbei um eine empirische Originalstudie, die aus diesem Grund auf einer Primäranalyse basiert. Da in dieser Arbeit eine spezifische Forschungsfrage überprüft wird, hat sie einen erklärungsorientierten, explanativen Charakter. Im Rahmen der Studie erfolgt eine Gruppenanalyse, wobei eine ausgewählte Stichprobe zur Datenerhebung herangezogen wird. Diese methodologische Herangehensweise bildet das Fundament, auf dem die Forschungsarbeit aufbaut.

5.2.1 Vorstudie

Im ersten Schritt wurde eine Vorstudie durchgeführt. Dabei ist ein Wizard-of-Oz-Design zum Einsatz gekommen. Dies ist eine Simulationstechnik, bei der ein sogenannter „Wizard“ Teile eines Systems manuell steuert, um die Funktionalität eines computergetriebenen Dialogsystems zu imitieren [63]. In der Vorstudie lag der Fokus auf der Sammlung von Daten zur Nutzerinteraktion mit einem LC und der Validierung des Studiendesigns. Diese Datenerhebung war von fundamentaler Bedeutung, da bis dato keine vergleichbaren Studien zu Mensch-Computer-Interaktionen im Kontext des Lerncoachings existierten. Die Besonderheit der Interaktion zwischen Mensch und Maschine, die sich deutlich von menschlichen Konversationen unterscheidet, unterstrich die Notwendigkeit einer sorgfältigen Erfassung [63, 64]. Die fünf Teilnehmer*innen der Vorstudie wurden mithilfe eines Schneeballverfahrens rekrutiert. Sie führten das Gespräch mit dem Wizard über Telegramm durch, wobei sie einen Fragebogen vor und nach der Studie ausfüllten. Der Fragebogen wurde in der Vorstudie verwendet, um den Ablauf der gesamten Studie zu simulieren. Genaue Details zu dem Schneeballverfahren und zu dem verwendeten Fragebogen folgen in weiteren Kapiteln. Der Wizard ist der Forschungsleitende dieser Arbeit. Er besitzt eine Coaching Ausbildung zertifiziert durch die Industrie- und Handelskammer (IHK) Akademie in Ostbayern [45]. Mithilfe der Lerncoaching Literatur und der Coaching-Ausbildung konnte der Wizard im Kontext dieser Arbeit adäquat die Nachrichten der teilnehmenden Personen beantworten. Die Teilnehmer*innen wurden nach der Studie über das Design aufgeklärt, umso den ethischen Bedenken eines Wizard-of-Oz Experiments gerecht zu werden [63]. Da die Ergebnisse des Wizard-of-Oz Experiment den Aufbau und Ablauf der Hauptstudie beeinflussen, werden diese vorgezogen und präsentiert, um im Nachgang ein besseres Verständnis für die Hauptstudie zu erlangen.

	Durchschnittliche Anzahl an Nachrichten	Standardabweichung der Anzahl der Nachrichten	Durchschnittliche Dauer in Minuten	Standardabweichung in Minuten
Phase 1	1	0	0	0
Phase 2	25,4	5,2	14,8	5,0
Phase 3	16,4	3,8	9,2	3,2
Phase 4	10	2,7	4,8	1,1
Phase 5	4,6	3,6	2	2,8
Gesamt	57,4		30,8	

Tabelle 2 Auswertung der Vorstudie nach den Phasen des Lerncoachings

Wie in der Lerncoaching Literatur beschrieben wurde die Interaktion der Teilnehmenden und dem Wizard in fünf Phasen aufgeteilt. Die Ergebnisse können in Tabelle 2 beobachtet werden. Eine kurze Beschreibung des Aufbaus der Vorstudie inklusive eines Chatbeispiels können im Anhang A betrachtet werden. Phase eins, drei und fünf waren vor der Studie verfasst worden. Phase eins ist die Begrüßung und beinhaltet somit nur die Eingangsnachricht. In der dritten Phase wurde ein Ziel nach der Literatur verfasst. In der letzten Phase wurde sich von den Teilnehmer*innen verabschiedet. Leichte Unterschiede in diesen Phasen sind dadurch entstanden, dass der Wizard in manchen Situationen nachfragen gestellt hat und somit leicht vom Skript abgewichen ist. Phase zwei und vier konnten kaum vorbereitet werden, da diese für jede teilnehmende Person unterschiedlich sind. Phase zwei ist der längste Abschnitt der Studie. In diesem Teil wird mit den Teilnehmer*innen erörtert welches Thema sie besprechen wollen. Zusätzlich werden zu der Thematik passende Informationen gesammelt, um ein Ziel formulieren zu können.

Diese Ergebnisse bekräftigen die Aussagen aus dem Abschnitt 5.1 Phase zwei und vier unterscheiden sich im Inhalt zwischen den Teilnehmer*innen stark, somit können hier keine vorgefertigten Chatbot Nachrichten verwendet werden. In diesen Abschnitten wird GPT-4 verwendet. Eine weitere Erkenntnis der Vorstudie ist die Nachrichtenanzahl der Phasen. Mithilfe dieser Ergebnisse wurde die Länge der Phasen in der Studie angepasst. Zusätzlich konnten Ausschnitte der Chats für das Few-Shot-Learning verwendet werden. Die Prompts an GPT-4 in Kombination mit den Beispielen der Vorstudie können in Anhang B betrachtet werden.

5.2.2 Hauptstudie

Abbildung 2 zeigt den Aufbau der Hauptstudie. Um die Effekte des Lerncoaching Chatbots zu untersuchen, kam ein Between-Group Design zum Einsatz. Dieses Verfahren teilt die gesammelte Stichprobe in zwei separate Gruppen, die jeweils unterschiedliche Systeme nutzen. Die erste Gruppe verwendet den LC, während die zweite Gruppe mit einem NC interagiert. Sobald die Studie startet, werden beide Gruppen von den Chatbots aufgefordert den Fragebogen der Studie auszufüllen. Daraufhin startet die erste Session. LC durchläuft dabei die fünf Phasen des Lerncoachings und verwendet die Few-Shot-Beispiele der Vorstudie. Die Nachrichtenanzahl der Phasen wurde mit den Ergebnissen der Vorstudie bestimmt. NC verwendet keine Few-Shot-Beispiele. Der Chatbot bekommt neutrale Anweisungen. Die Länge des NC richtet sich nach der Nachrichtenanzahl der Phasen zwei und vier des LC. Sobald die Chatbots sich von den Teilnehmer*innen verabschieden, wird ihnen der Posttest zugesendet. Zusätzlich können sich alle Teilnehmer*innen für eine zweite Session anmelden. Wie in Abschnitt 2.2 beschrieben, sind in Mensch-Mensch Coachings normalerweise mehrere Sitzungen üblich, um an einem Ziel zu arbeiten. Aufgrund von Rekrutierungsschwierigkeiten für mehrere Sitzungen bot man als Kompromiss eine freiwillige zweite Sitzung an. Die Teilnehmenden konnten wählen, ob sie an einer oder beiden Sitzungen teilnehmen möchten. Eine Woche nach der ersten Session war die Durchführung der zweiten Session möglich. Der LC durchlief erneut die fünf Phasen des Lerncoachings, wobei die Prompts an die vorherigen Ergebnisse der Nutzer*innen angepasst wurden. Der NC blieb in der zweiten Session unverändert zur ersten. Nach Abschluss wurde den Teilnehmenden der Posttest zugesendet. Der erste Posttest dient dabei als Baseline, während der zweite als Follow-Up betrachtet wird. Chatbeispiele und die Few-Shot-Beispiele können in Anhang B betrachtet werden.

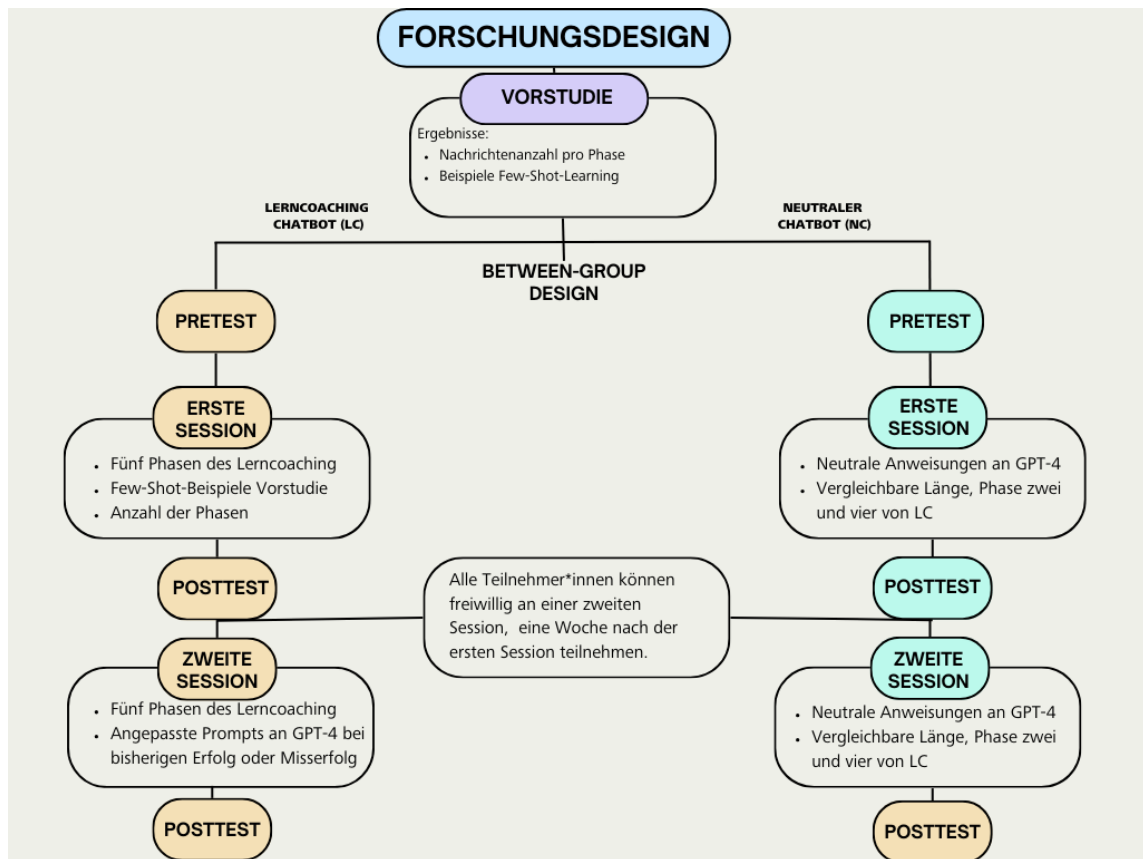


Abbildung 2: Ablauf der Studie

5.3 Stichprobenziehung

In der Vorstudie kam das Schneeballverfahren [26] zur Rekrutierung der Teilnehmenden zum Einsatz. Dieser Ansatz wurde gewählt, da die Vorstudie explorative Ziele verfolgte. Hauptziel war die Gewinnung von Daten für das Few-Shot-Learning in der Hauptstudie. Zusätzlich diente die Vorstudie der Überprüfung des Studiendesigns. Die Entscheidung für das Schneeballverfahren fiel insbesondere aufgrund der damit verbundenen Ressourceneinsparungen.

Für die Hauptstudie wurde eine Kombination aus Gelegenheitsstichprobe [18] und Schneeballverfahren [26] zur Rekrutierung der Teilnehmenden genutzt. Ziel der Studie ist es, die Wirksamkeit eines LC bei der Unterstützung von Studierenden zu untersuchen. Da die Teilnahme an der Studie nicht in Universitätsvorlesungen verpflichtend gemacht werden konnte, war man auf die freiwillige Teilnahme der Proband*innen angewiesen. Daher erfolgte die Rekrutierung in jenen Vorlesungen, deren Leitungen bereit waren, das Projekt zu unterstützen, indem die Studie dort vorgestellt wurde. Um eine ausreichende Anzahl an Teilnehmenden zu gewährleisten, wurde darüber hinaus das Schneeballverfahren in der Hauptstudie eingesetzt.

5.3.1 Zielgruppe und Auswahlkriterien

Die Zielgruppe für die Vor- und Hauptstudie, abgeleitet aus der Forschungsfrage, umfasste Studierende ohne spezifische Einschränkungen hinsichtlich Alter, Studiengang, Fachsemester oder Geschlecht. Ein entscheidendes Auswahlkriterium war jedoch, dass die Teilnehmenden in Kursen eingeschrieben sein sollten, die regelmäßiges Lernen fördern. Dies umfasste Kurse, die wöchentliche Arbeitsblätter verlangen, regelmäßige Vokabeltests abhalten oder kontinuierliche Prüfungen in Form von Studienleistungen durchführen. Diese Kriterien gewährleisteten, dass die Studienteilnehmer*innen in einem Lernumfeld aktiv sind, das die effektive Nutzung und Bewertung des LC ermöglicht.

5.3.2 Rekrutierung der Teilnehmenden

Die Rekrutierungsmethode für die Gelegenheitsstichprobe der Hauptstudie erfolgte durch die Vorstellung der Studie in insgesamt 16 Vorlesungen und Seminaren der Universität Regensburg. Diese umfassten sowohl geistes- als auch naturwissenschaftliche Kurse. Bei der Präsentation wurden die Vorteile der Studie hervorgehoben, wobei darauf geachtet wurde, nicht zu detailliert zu werden, um eine Beeinflussung der potenziellen Proband*innen zu vermeiden. Die Studierenden hatten die Möglichkeit, sich über eine interne Webseite der Universität für die Studie anzumelden. Insgesamt wurde die Studie vor 1346 Studierenden präsentiert, von denen sich 133 für eine Teilnahme entschieden. Nach der Anmeldung war es erforderlich, dass die Studierenden einen Termin auswählen um die Studie durchzuführen. Die finale Stichprobe der Studie wird im nächsten Punkt erläutert.

5.3.3 Größe und Struktur der Stichprobe

Beschreibung	Wert
Geschlecht: männlich	40.00%
Geschlecht: weiblich	60.00%
Anzahl der Teilnehmer*innen	5
Durchschnittsalter	23.8 Jahre
Standardabweichung Alter	2.14 Jahre
Coaching Erfahrung: Nein	100%
Coaching Erfahrung: Ja	0%
Chatbot Erfahrung: Nein	60%
Chatbot Erfahrung: Ja	40%

Tabelle 3: Alter, Geschlecht und Studienfach der Teilnehmenden der Vorstudie

In der Vorstudie, die in Tabelle 3 dargestellt ist, die fünf Teilnehmende umfasste, lag der Anteil der weiblichen Teilnehmerinnen bei 60% und der der männlichen Teilnehmer bei 40%, mit einem Durchschnittsalter von 23,8 Jahren. Die Erfahrungen mit Coaching und Chatbots waren dabei unterschiedlich verteilt, wobei keiner der Teilnehmenden Coaching-Erfahrung hatte und 40% Chatbot-Erfahrung aufwiesen.

Die Hauptstudie, deren Daten in Tabelle 4 zusammengefasst sind, die 30 Teilneh-

Beschreibung	Wert
Geschlecht: weiblich	70.00%
Geschlecht: männlich	30.00%
Anzahl der Teilnehmer*innen	30
Durchschnittsalter	23.04 Jahre
Standardabweichung Alter	3.54 Jahre
Coaching Erfahrung: Nein	76.66%
Coaching Erfahrung: Ja	23.34%
Chatbot Erfahrung: Nein	50.00%
Chatbot Erfahrung: Ja	50.00%

Tabelle 4: Demografische Daten der Teilnehmenden der Hauptstudie mende ein-

schloss, waren die demografischen Daten tendenziell ähnlich: 70.00% weiblich, 30.00% männlich, mit einem leicht niedrigeren Durchschnittsalter von 23.04 Jahren. Die Erfahrungen der Teilnehmenden in den Bereichen Coaching und Chatbots waren ebenfalls vergleichbar, mit 76.66% ohne Coaching-Erfahrung und einer nahezu ausgeglichenen Verteilung in der Chatbot-Erfahrung. Diese Ähnlichkeiten in den demografischen Daten

und Vorerfahrungen zwischen Vor- und Hauptstudie unterstreichen die Eignung der Vorstudie zur Vorbereitung der Hauptstudie. 13 Teilnehmer*innen, die die Studie nicht vollständig abgeschlossen haben, sei es durch das Nichtausfüllen der Fragebögen oder eine nicht adäquate Durchführung, wurden von der Studie ausgeschlossen. 13 Versuchsteilnehmer*innen hatten sich für die zweite Session angemeldet, jedoch führten nur 4 diese Session durch. Somit konnte für den Follow-Up Test nicht genügend Daten gesammelt werden, um eine Analyse durchzuführen.

5.4 Datenerhebung

Für die Datenerhebung sowohl des Pretests als auch des Posttests kam der Motivated Strategies for Learning Questionnaire (MSLQ) zum Einsatz [77].

Der MSLQ ist ein Selbstauskunftsinstrument, mit dem die motivationalen Orientierungen von Hochschulstudierenden und die Art und Weise, wie sie verschiedene Lernmethoden in einem bestimmten Kurs einsetzen, bewertet werden [77]. Der Fragebogen ist in zwei Hauptabschnitte unterteilt: einen zur Motivation und einen zu den Lerntechniken. In dieser Arbeit liegt der Schwerpunkt auf dem ersten Teil des Fragebogens, da das Ziel ist, Veränderungen in Motivation und Selbstwirksamkeit zu messen. Gemäß dem Handbuch des MSLQ können die verschiedenen Abschnitte des Fragebogens entsprechend der spezifischen Forschungsfragen separat eingesetzt werden [77]. Somit wird die Motivation mit folgenden Skalen gemessen: Intrinsic Goal Orientation, Extrinsic Goal Orientation, Task Value, Control of Learning Beliefs und Self-Efficacy welches für Selbstwirksamkeit steht. Der Motivationsteil umfasst insgesamt 31 Items, die auf einer 7-Punkte-Likert-Skala basieren.

Bei der Erstellung dieser Arbeit wurde das Manual des MSLQ bereits 6515 Mal zitiert. Zusätzlich hat die Metastudie von Marcus Credé and L. Alison Phillips [2011] gezeigt, dass der MSLQ ein zuverlässiges Maß für Konstrukte ist, von denen einige bedeutungsvolle Zusammenhänge mit der akademischen Leistung aufweisen.

Obwohl es generell bevorzugt wird, vorhandene deutsche Fragebögen zu nutzen, um Übersetzungsproblematiken zu vermeiden [18], wurde in diesem Fall eine Ausnahme gemacht. Für diese Studie wurde der MSLQ, der bisher keine offizielle deutsche Version vorweist, ins Deutsche übersetzt. Diese Entscheidung basierte auf der hohen thematischen Eignung des MSLQ, seiner umfassenden Erprobung [53] und seinem

erfolgreichen Einsatz in zahlreichen Studien mit Studierenden [53]. Die spezifischen Merkmale und die bewährte Testmethodik des MSLQ überwogen in diesem Fall die üblichen Bedenken gegenüber der Übersetzung, weshalb die Anpassung ins Deutsche als der beste Ansatz angesehen wurde. Der übersetzte Fragebogen ist in Anhang C zu finden.

5.5 Software der Chatbots

Zuerst wird das Front-End der Chatbots erörtert, um daraufhin im zweiten Teil das Back-End zu erklären. Zur Veranschaulichung der Zusammenhänge und Funktionsweisen der Chatbot-Komponenten dient das in Abbildung 3 dargestellte Flussdiagramm. Bei der Programmierung der Bots wurde mit der Programmiersprache Python (Version 3.11.5) gearbeitet. Die wichtigsten Pakete, die bei der Herstellung der Bots verwendet wurden, werden in den nächsten beiden Abschnitten erläutert

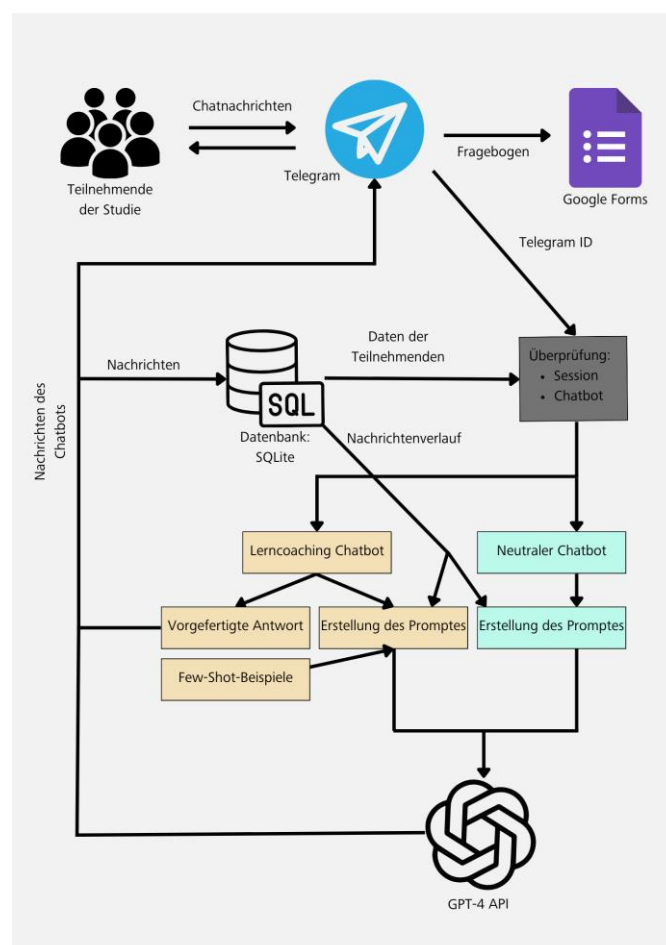


Abbildung 3: Flussdiagramm der Chatbots

5.5.1 Front-End: Telegram

Die Studienteilnehmenden interagierten über die Benutzeroberfläche von Telegram mit den Chatbots. Es wurde ihnen die Möglichkeit geboten, entweder über Telegram Web, also mittels eines Webbrowsers, oder durch die mobile Anwendung auf ihren Handys mit den Bots zu kommunizieren. Für die Benutzeroberfläche kam das Python-Paket *aiogram* [2] zum Einsatz. Aiogram ist ein asynchrones Framework, das für die Telegram Bot API entwickelt wurde, und es nutzt *asyncio* sowie *aiohttp* [2]. Im ersten Schritt wird mittels der Telegram-Anwendung *BotFather* [24] ein Bot-Token generiert. Im zweiten Schritt können dann mithilfe dieses Tokens und *aiogram* Nachrichten verarbeitet werden, um sie anschließend an die Nutzer*innen weiterzuleiten.

Sobald die Studie gestartet wird, erhalten die Teilnehmenden den Pretest. Dieser wird über Google Formulare [27] bereitgestellt. Mithilfe der anonymen TelegramID die automatisch in das Google Dokument übertragen wird, können die Fragebögen zugeordnet werden. Nach Beenden der Studie erhalten die Personen den Posttest.

5.5.2 Back-End: GPT-4, SQLite-Datenbank

Wenn aiogram eine Nachricht einfängt, wird mithilfe der Datenbank überprüft, in welcher Session sich die Teilnehmenden befinden und welchem Chatbot sie zugeordnet sind. Als Datenbanklösung wurde SQLite mit dem Python Paket *sqlite3* [78] gewählt. Nachdem die Zuordnung zu den Chatbots erfolgt ist, wird geprüft, in welcher Phase sich die Nutzer*innen aktuell befinden und wie viele Nachrichten sie bereits gesendet haben. Je nach diesen Informationen sendet entweder der LC eine vorbereitete Nachricht, oder es wird ein Prompt mit dem bisherigen Chatverlauf erstellt und an die GPT-4 API weitergeleitet. Für die Erstellung des Prompts beim LC werden zusätzlich Few-Shot-Beispiele verwendet. Die Integration der GPT-4 API erfolgt durch das OpenAI Paket [67].

5.6 Datenauswertung

Für die Berechnung der Unterschiede zwischen den Gruppen kam ein unabhängiger t-Test zum Einsatz. Hierbei wurde die Differenz aus den einzelnen Punkten des Pre- und Posttests gebildet, und im Folgenden mithilfe des t-Tests zwischen den Gruppen verglichen. Diese Analyse erfolgte mit Python (Version 3.11.5), unter Verwendung des SciPy Pakets [92].

Zur Ermittlung der Unterschiede innerhalb der Gruppen wurde ein gepaarter t-Test verwendet. Nachdem im ersten Schritt der Unterschied zwischen den Gruppen berechnet wurde, betrachtete man in diesem Schritt beide Gruppen als eine Einheit. Auch hier kam das gleiche Python Paket zum Einsatz.

Die statistischen Tests wurden auf jede Subskala des MSLQ angewandt, was insgesamt 12 t-Tests ergab. Aufgrund der gleichzeitigen Anwendung mehrerer Tests wurde die Bonferroni-Korrektur eingesetzt [9]. Diese Korrektur ist essenziell, um das Risiko von irrtümlich positiven Ergebnissen zu verringern, die entstehen können, wenn mehrere statistische Tests simultan durchgeführt werden [9].

Da die Stichprobengröße bei 30 liegt, kann von einer Normalverteilung ausgegangen werden. Zusätzlich sind t-Tests robust gegenüber der Normalverteilung [89, 114].

6 Ergebnisse

<i>Strategien</i>	LC Pre- test:Mean,SD	LC Post- test:Mean,SD	NC Pre- test:Mean,SD	NC Post- test:Mean,SD
<i>Motivation</i>				
<i>Intrinsic Goal Orientation</i>	4.15, 1.18	4.52, 1.23	3.78, 1.42	4.20, 1.53
<i>Extrinsic Goal Orientation</i>	4.28, 1.14	4.27, 1.11	3.78, 1.36	3.98, 1.10
<i>Task Value</i>	4.87, 1.36	5.00, 1.38	4.72, 1.12	4.80, 1.25
<i>Control of Learning Beliefs</i>	5.90, 0.88	5.75, 0.89	5.35, 0.72	5.60, 0.75
<i>Self-Efficacy</i>	4.67, 1.10	5.24, 1.03	4.26, 1.01	4.63, 1.13
<i>Test Anxiety</i>	4.27, 1.15	4.31, 1.39	4.57, 1.58	4.29, 1.66

In **Tabelle 5: Mittelwerte und Standardabweichungen der beiden Gruppen**

die-

sem Abschnitt der Arbeit werden die Ergebnisse der Studie dargestellt. Zur Kontrolle der Fehlerwahrscheinlichkeit erster Art wurde eine Bonferroni-Korrektur angewendet, wodurch das Signifikanzniveau für jeden einzelnen Test auf 0.0042 angepasst wurde.

Die Mittelwerte und die Standardabweichungen der LC-Gruppe und der NC-Gruppe werden in Tabelle 5 dargestellt. Hervorzuheben sind die Verbesserungen bei Self-Efficacy und Intrinsic Goal Orientation. Bei der Self-Efficacy zeigte sich eine Verbesserung im LC mit einem Anstieg des Mittelwerts von 4.67 (SD = 1.10) auf 5.24 (SD = 1.03) und im NC von 4.26 (SD = 1.01) auf 4.63 (SD = 1.13). Ebenso verbesserte sich die Intrinsic Goal Orientation im LC von 4.15 (SD = 1.18) auf 4.52 (SD = 1.23) und im NC von 3.78 (SD = 1.42) auf 4.20 (SD = 1.53), was auf eine Zunahme der intrinsischen Motivation in beiden Gruppen hinweist.

Bei den anderen Konstrukten ergab sich ein gemischtes Bild. Der Task Value verbesserte sich im LC (von 4.87, SD = 1.36 auf 5.00, SD = 1.38) und im NC (von 4.72, SD = 1.12 auf 4.80, SD = 1.25). Die Extrinsic Goal Orientation (EGO) zeigte im LC einen minimalen Rückgang (von 4.28, SD = 1.14 auf 4.27, SD = 1.11), während im NC eine leichte Verbesserung (von 3.78, SD = 1.36 auf 3.98, SD = 1.10) zu beobachten war. Die Test Anxiety änderte sich leicht im LC (von 4.27, SD = 1.15 auf 4.31, SD = 1.39) und verringerte sich im NC (von 4.57, SD = 1.58 auf 4.29, SD = 1.66).

Die Resultate bezüglich der Control of Learning Beliefs verdienen eine besondere Beachtung. Hier zeigte die NC-Gruppe, die als Kontrollgruppe diente, eine Verbesserung. Während im LC ein Rückgang von 5.90 (SD = 0.88) auf 5.75 (SD = 0.89) festzustellen

war, stieg der Mittelwert im NC von 5.35 (SD = 0.72) auf 5.60 (SD = 0.75). Dies deutet darauf hin, dass der neutrale Chatbot in Bezug auf die Förderung der Überzeugungen zur Kontrolle des eigenen Lernens positive Effekte hatte.

Strategien	t-Wert	p-Wert	df
<i>Motivation</i>			
<i>Intrinsic Goal Orientation</i>	-0,17	0,868	28
<i>Extrinsic Goal Orientation</i>	-0,94	0,356	28
<i>Task Value</i>	0,30	0,768	28
<i>Control of Learning Beliefs</i>	-2,51	0,018	28
<i>Self-Efficacy</i>	1,03	0,311	28
<i>Test Anxiety</i>	1,21	0,236	28

Tabelle 6: Unterschiede der Motivation zwischen den beiden Gruppen

Tabelle 6 berichtet über den unabhängigen t-Test zwischen den beiden Gruppen. Es wurde getestet, ob ein signifikanter Unterschied in den Motivationsskalen erkannt werden kann. Bei den untersuchten Konstrukten wurden keine signifikanten Unterschiede zwischen den Gruppen festgestellt. Bei der Intrinsic Goal Orientation zeigte sich kein signifikanter Unterschied, $t(28) = -0,17$, $p = 0,868$. Ähnlich war es bei der Extrinsic Goal Orientation, wo kein signifikanter Unterschied zu verzeichnen war, $t(28) = -0,94$, $p = 0,356$. Beim Task Value zeigte sich ebenfalls kein signifikanter Unterschied zwischen den Gruppen, $t(28) = 0,30$, $p = 0,768$. In Bezug auf Self-Efficacy wurde ebenfalls kein signifikanter Unterschied festgestellt, $t(28) = 1,03$, $p = 0,311$. Auch bei der Test Anxiety gab es keinen signifikanten Unterschied, $t(28) = 1,21$, $p = 0,236$.

Hervorzuheben sind jedoch die Ergebnisse bei den Control of Learning Beliefs. Hier wurde ein t-Wert von -2,51 ermittelt, der p-Wert lag jedoch bei 0,018, was über dem durch die Bonferroni-Korrektur festgelegten Signifikanzniveau liegt. Obwohl dieser Befund statistisch nicht als signifikant eingestuft werden kann, ist es bemerkenswert, dass bei diesem Konstrukt der größte Unterschied in den Mittelwerten zwischen den beiden Gruppen zu verzeichnen war. Dies könnte auf einen potenziell relevanten Effekt hindeuten, der weiterer Untersuchungen bedarf. Insgesamt bestätigen die Ergebnisse des t-Tests die geringen Veränderungen zwischen den beiden Gruppen in den Mittelwerten, die in der Studie beobachtet wurden.

Strategien	t-Wert	p-Wert	df
<i>Motivation</i>			
<i>Intrinsic Goal Orientation</i>	-2,68	0,012	29
<i>Extrinsic Goal Orientation</i>	-0,80	0,433	29
<i>Task Value</i>	-1,15	0,261	29
<i>Control of Learning Beliefs</i>	-0,58	0,569	29
<i>Self-Efficacy</i>	-5,06	<0,001 ^a	29
<i>Test Anxiety</i>	0,90	0,375	29

Tabelle 7: Unterschiede der Motivation innerhalb der beiden Gruppen

^ap < 0,0042

In Tabelle 7 der Studie, die die Unterschiede der Motivation innerhalb der beiden Gruppen untersucht, wurden folgende Ergebnisse ermittelt. Zunächst ist der Befund bei der Self-Efficacy hervorzuheben. Hier wurde ein signifikanter Unterschied festgestellt, $t(29) = -5,06$, $p < 0,001$, was deutlich unter dem durch die Bonferroni-Korrektur festgelegten Signifikanzniveau von 0,0042 liegt. Dies zeigt eine bedeutende Veränderung in der Selbstwirksamkeit innerhalb der Gruppen.

Im Gegensatz dazu wurde bei der Intrinsic Goal Orientation kein signifikanter Unterschied festgestellt, $t(29) = -2,68$, $p = 0,012$, da der p-Wert über dem durch die Bonferroni-Korrektur festgelegten Signifikanzniveau liegt.

Bei den anderen Konstrukten ergaben sich ebenfalls keine signifikanten Unterschiede. Die Extrinsic Goal Orientation zeigte keinen signifikanten Unterschied, $t(29) = -0,80$, $p = 0,433$. Beim Task Value wurde kein signifikanter Unterschied beobachtet, $t(29) = -1,15$, $p = 0,261$. Ähnlich verhält es sich mit den Control of Learning Beliefs, wo sich ebenfalls kein signifikanter Unterschied ergab, $t(29) = -0,58$, $p = 0,569$. Bei der Test Anxiety zeigten die Ergebnisse keinen signifikanten Unterschied, $t(29) = 0,90$, $p = 0,375$.

Zusammenfassend lässt sich sagen, dass die Bereiche der Intrinsic Goal Orientation und Self-Efficacy die größten positiven Veränderungen in den Mittelwerten verzeichnet haben, obwohl bei ersterer keine statistische Signifikanz gefunden wurde. Wie schon beschrieben konnten für die zweite Session nicht genügend Daten gesammelt werden.

7 Diskussion

Die vorliegenden Ergebnisse liefern wichtige Einsichten in die Dynamik der Motivationsveränderungen und der Veränderung der Selbstwirksamkeit innerhalb und zwischen den untersuchten Gruppen.

7.1 Ergebnisse zwischen den beiden Gruppen

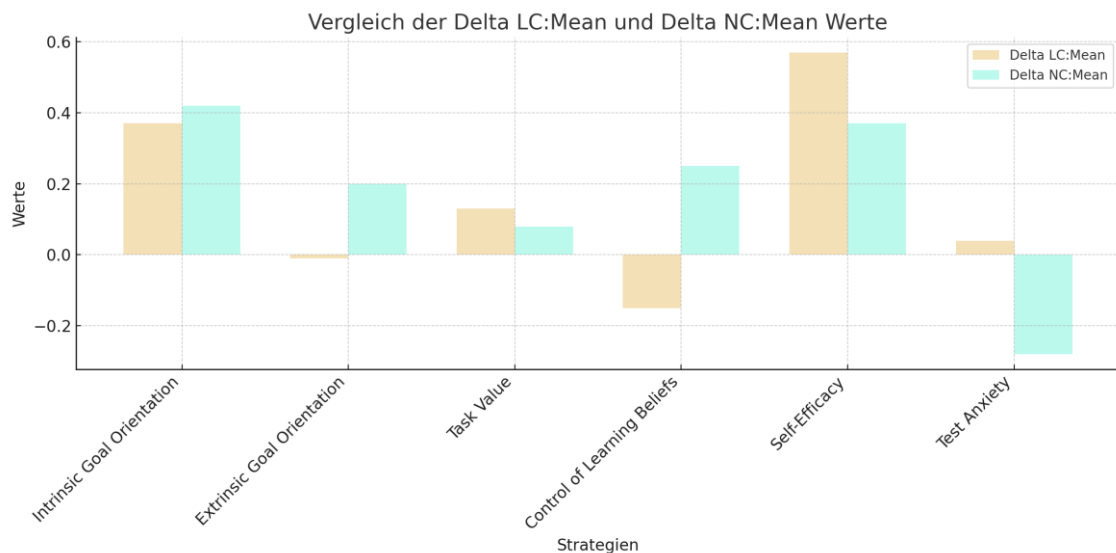


Abbildung 4: Die Differenz und statistische Werte für jede Strategie

Abbildung 4 präsentiert ein Balkendiagramm, das den Vergleich der Werte von Delta LC:Mean und Delta NC:Mean über die Subskalen der Motivation illustriert. Die dargestellten Subskalen umfassen Intrinsic Goal Orientation, Extrinsic Goal Orientation, Task Value, Control of Learning Beliefs, Self-Efficacy und Test Anxiety. Die Höhe und Richtung jedes Balkens veranschaulichen das Ausmaß der Veränderung innerhalb der jeweiligen Subskala, wodurch die Unterschiede der beiden Gruppen deutlich werden.

Intrinsic bzw. Extrinsic Goal Orientation, Control of Learning Beliefs und Test Anxiety haben in der NC-Gruppe eine größere Veränderungen erlebt als in der LC-Gruppe. Jedoch sind alle Ergebnisse zwischen den beiden Gruppen nach der Bonferroni-Korrektur nicht signifikant. Der größte Unterschied zwischen den beiden Gruppen ist bei Control of Learning Beliefs zu finden, da sich die LC-Gruppe im Gegensatz zu NC verschlechtert hat. Der P-Wert ist dabei unter 0.05, jedoch über der Bonferroni-Korrektur. Somit lässt sich ein möglicher Trend erkennen, welcher jedoch weiter untersucht werden muss. Gemäß Pintrich et al. (1991) beziehen sich die Control of Learning Beliefs auf die Annahme der Lernenden, dass sie ihre Lernergebnisse und Leistungen durch

eigene Anstrengungen und Fähigkeiten beeinflussen können. Diese Überzeugung, dass interne Faktoren wie Anstrengung und Fähigkeit – und nicht externe Faktoren wie Aufgabenschwierigkeit oder Zufall – eine entscheidende Rolle spielen, ist von zentraler Bedeutung. Eine mögliche Erklärung für dieses Ergebnis könnte sein, dass der NC-Chatbot direkte Lösungsvorschläge geben kann und somit der Fokus auf die Lösung und den eigenen Fähigkeiten gelegt wird. Ein Beispiel dafür ist diese Nachricht des NC-Chatbots: *„Hast du schon mal überlegt, Mindmaps oder Konzeptdiagramme zu erstellen, um das Gelernte mehr zu visualisieren? Welche Aspekte könntest du dabei betonen, um den Überblick über die Zusammenhänge und Schlüsselideen zu behalten?“*. Der LC-Chatbot dagegen versucht die Teilnehmer*innen zu einer Reflektion der Aussagen zu bewegen, um einen Fortschritt zu erzielen. Dies geschieht zum Beispiel mit einer zirkulären Frage nach dem Lerncoaching: *„Das klingt nach einer Herausforderung. Wie haben es andere Leute, die du kennst, geschafft, sich in ähnlichen Situationen zu motivieren?“*. Dies kann jedoch dazu führen, dass äußere Faktoren mehr miteinbezogen werden.

Erwähnenswert ist die Test Anxiety. Bei diesem Score ist ein geringerer Wert erwünscht [77]. Hierbei sieht man in der NC-Gruppe eine Verbesserung, die jedoch auch nicht zwischen den Gruppen statistisch signifikant ist und somit noch weitere Untersuchungen benötigt.

Task Value und Self-Efficacy besitzen eine größere Steigerung in der LC-Gruppe. Jedoch ist auch hier keine statistische Signifikanz zwischen den Gruppen zu erkennen. Self-Efficacy also die Selbstwirksamkeit ist ein wichtiger Teil des Lerncoachings. Somit könnte die Veränderung zwischen den beiden Gruppen erläutert werden, jedoch ohne Signifikanz.

Insgesamt ist jedoch kein signifikanter Unterschied zwischen den beiden Gruppen zu erkennen.

7.2 Ergebnisse innerhalb der Gruppen

Nach der Intervention zeigen sich die deutlichsten Veränderungen in den Bereichen Intrinsic Goal Orientation und Self-Efficacy, wie in Abbildung 4 illustriert. Andere Bereiche wie Extrinsic Goal Orientation, Task Value, Control of Learning Beliefs und Test Anxiety verzeichnen keine starken Veränderungen. Diese Ergebnisse spiegeln sich auch in den Resultaten des gepaarten t-Tests wider: Self-Efficacy zeigt signifikante

Veränderungen, jedoch während die Intrinsic Goal Orientation nach der Anwendung der Bonferroni-Korrektur nicht mehr als statistisch signifikant gilt. Die übrigen Konstrukte weisen keine signifikanten Unterschiede auf.

Daraus lässt sich schließen, dass GPT-4, selbst ohne Few-Shot-Learning und mit nur minimalen Anweisungen, ausreichend ist, um eine Veränderung in der Self-Efficacy der Teilnehmer*innen zu bewirken. Self-Efficacy bezieht sich auf die Überzeugung, eine Aufgabe erfolgreich bewältigen zu können [77]. Sie umfasst auch das Vertrauen in die eigenen Fähigkeiten, diese Aufgabe auszuführen [77]. In beiden Gruppen nutzen Teilnehmende häufig die Gelegenheit, um mit den Chatbots spezifische Probleme in ihren Studienfächern zu diskutieren. Dadurch rückt das Erlernen von Strategien zur Bewältigung solcher Aufgaben in den Vordergrund. Ein Beispiel hierfür beim LC-Bot ist folgender Dialog: Eine*r der Teilnehmenden fragt: *„Ist es schwierig, die Theorien in Erziehungswissenschaft zu erlernen?“* Darauf antwortet der LC-Bot: *„Als Lerncoach möchte ich dir dabei helfen, die für dich effektivsten Lernmethoden zu finden. Hast du schon Grundkenntnisse in Erziehungswissenschaft, oder beginnst du gerade erst, dich mit diesem Fachgebiet auseinanderzusetzen?“*. Der neutrale Chatbot fokussiert sich stärker darauf, direkt einen Lösungsweg zu finden, wobei auch hier das Lernen und das Lösen von Problemen im Mittelpunkt stehen. Ein Dialogbeispiel ist: Eine*r der Teilnehmenden äußert: *„Ich habe Schwierigkeiten, mich längere Zeit auf das Lernen zu konzentrieren.“* Daraufhin fragt der NC-Bot: *„Könntest du bitte beschreiben, wie deine aktuelle Lernumgebung aussieht?“*. Mit diesen Methoden schaffen es beide Chatbots die Self-Efficacy zu steigern. Es ist überraschend, dass zwischen den beiden Bots kein signifikanter Unterschied in der Wirksamkeit festgestellt wurde. Dies ist besonders bemerkenswert, da Selbstwirksamkeit ein wichtiger Bestandteil des Lerncoachings ist, ein Bereich, in dem der LC-Bot spezialisiert ist, während der NC-Bot über kein entsprechendes Framework in diesem Bereich verfügt. Dennoch konnte durch die Fokussierung des NC-Bots auf das direkte Lösen von studienbezogenen Problemen ebenfalls eine Steigerung der Selbstwirksamkeit erreicht werden.

7.3 Analyse Chatlogs

Gruppe	Dauer (Min)	Mean Nach- richten Mensch und Chatbot	Durchschn. Länge Mensch	Std. Abw. Mensch	Durchschn. Länge Chat- bot	Std. Abw. Chat- bot
LC	25.71	874	10.31	10.19	28.89	26.08
NC	14.57	504	8.82	9.20	24.35	19.23

Tabelle 8: Menschliche/Chatbot Durchschnittliche Dauer & Nachrichtenlänge

In Tabelle 8 lässt sich die durchschnittliche Dauer einer Session für beide Gruppen betrachten. Es ist ersichtlich, dass Teilnehmer*innen in der LC-Gruppe durchschnittlich 11,14 Minuten länger mit dem Chatbot interagiert haben. Zusätzlich wird deutlich, dass LC-Nutzer*innen mehr Nachrichten verfasst haben als jene in der Kontrollgruppe. Dies ist auf das Lerncoachingframework zurückzuführen. Die Zielerfassung, also Phase 3 des Lerncoachings, kam ausschließlich in der LC-Gruppe vor. Zudem konnten LC-Nutzer*innen aufgrund einer Skalierungsfrage vor Phase 3 mehr Nachrichten mit dem Bot austauschen. Somit lässt sich die erhöhte Dauer und Anzahl der Nachrichten erklären. Die Ergebnisse bestätigen die Studie von Cannon-Bowers et al. [2023] in der die Dauer des Coachings keinen Einfluss auf die Resultate hat. Somit konnten die NC Teilnehmer*innen in einer geringeren Zeit ihre Self-Efficacy steigern. Dies stellt einen Vorteil dar. Wenn der Bot Studierenden frei angeboten würde und diese ihn häufiger nutzen würden, könnte eine Zeitersparnis bei mehreren Sitzungen noch deutlicher erkennbar sein.

Die durchschnittliche Nachrichtenlänge bei den Teilnehmer*innen der LC-Gruppe ist länger als in der Kontrollgruppe. Auch der LC-Chatbot antwortet mit längeren Nachrichten. Daher lässt sich annehmen, dass der LC-Bot die Konversation besser anregt als der NC-Bot. Jedoch besteht eine hohe Standardabweichung, die sich durch die Nachrichten der Teilnehmer*innen erklären lässt. Einige Teilnehmerinnen schreiben nur kurze Nachrichten, während andere längere, ausführlichere verfassen. Die Bots antworten meist entsprechend der Länge der Nachrichten der Nutzer*innen.

Mithilfe der Chats der LC- und NC-Gruppe wurde eine Sentiment Analyse und ein Topic Modelling durchgeführt. Für die Sentiment Analyse wird ein Score zwischen 1

(positiv) und -1 (negativ) erstellt. Beide Gruppen besitzen neutrale Scores: LC = 0.15 und NC = 0.1. Somit kann davon ausgegangen werden, dass keine schwerwiegenden Probleme besprochen wurden. Dies steht im Einklang mit dem Lerncoaching, welches keine Therapieform darstellt, sondern sich auf das Thema Lernen fokussiert. Dies spiegelt sich auch im Topic Modelling wider, wobei vor allem Lernen als Thema erkannt wird. Die Durchführung der Analysen kann im GitHub Repository¹ betrachtet werden.



Abbildung 5: Wordcloud LC-Gruppe links, NC-Gruppe rechts

Mithilfe der Wordclouds aus Abbildung 5 lassen sich die am häufigsten verwendeten Wörter erkennen. Wörter wie „lernen“ und „Zeit“ treten in beiden Gruppen auf, ebenso wie die Antworten „Ja“ und „Nein“. Da beide Bots in der Regel pro Nachricht eine Frage stellten, gab es Teilnehmer*innen, die häufig mit Ja oder Nein auf die Antworten der Bots reagierten. Jedoch ist das Wort „Nein“ weniger in der LC-Gruppe zu finden, welches auf eine bessere Kooperation hindeuten kann. Noch zu erwähnen ist das Wort „Motivation“. Dies tritt vor allem in der LC-Gruppe auf. Ansonsten überschneiden sich viele Wörter und Themen welche weniger oft vorkommen. Somit ist ein geringer Unterschied zwischen den beiden Gruppen zu erkennen.

7.4 Praktische Implikationen

Es wurde kein signifikanter Unterschied in den Ergebnissen zwischen den beiden Gruppen festgestellt. Interessanterweise konnte der NC-Bot, ohne den Einsatz von Few-Shot-Learning und mit nur minimalen Anweisungen, die gleichen Resultate erzielen wie der LC-Bot. Dies deutet darauf hin, dass auch weniger spezialisierte Ansätze effektiv sein können. Somit ist die Erstellung solcher Chatbot Assistenten erleichtert, da kein Datensatz nötig ist, um ein System zu trainieren oder Few-Shot-Learning zu verwenden. Diese

¹ <https://github.com/AlexVike/Masterarbeit>

Erkenntnis vereinfacht die praktische Umsetzung solcher Systeme. Zusätzlich ist bemerkenswert, dass von allen untersuchten Subskalen die Selbstwirksamkeit die stärkste Korrelation mit der Endnote eines Kurses aufweist [77]. Dies unterstreicht die Bedeutung der Selbstwirksamkeit. Aus diesem Grund könnte es sich lohnen, ein solches System als Unterstützung in das Curriculum von Universitäten einzubauen.

8 Abschluss

In diesem Abschnitt wird ein Ausblick gegeben, der auf den bisher erzielten Ergebnissen basiert. Zusätzlich werden die Limitationen der Studie dargelegt, um ein abschließendes Fazit zu ziehen.

8.1 Ausblick

In dieser Arbeit wurde ein Online-Coaching Chatbot der Stufe eins erstellt, da dieser nur mithilfe eines Chats mit den Teilnehmer*innen kommuniziert [7]. Für zukünftige Entwicklungen könnte die Implementierung eines fortgeschritteneren Chatbots, der über diese erste Stufe hinausgeht, von Vorteil sein. Dazu wäre der Aufbau einer erweiterten Plattform erforderlich, die den Nutzer*innen eine interaktive Kommunikation mit dem Chatbot ermöglicht. Ein erster Schritt wäre die Einführung einer auditiven Kommunikation. Technologien wie Whisper [81] könnten dabei eingesetzt werden, um Sprachnachrichten der Nutzer*innen zu transkribieren und somit eine effektive Interaktion mit dem GPT-4-Modell zu ermöglichen. Ferner könnten Funktionen zur Sprachsynthese integriert werden, wodurch der Chatbot in der Lage wäre, auch auditiv mit den Klient*innen zu kommunizieren. Diese Erweiterung würde die Kommunikationsfähigkeiten des Bots signifikant verbessern und ihn auf eine höhere Stufe heben. Eine weitere Möglichkeit wäre die Einführung eines Avatars, um den Chatbot als virtuellen Coach auch visuell zu repräsentieren. Um Stufe vier für den Bot zu ermöglichen, müsste zusätzlich eine Bilderkennung eingeführt werden. Modelle wie GPT-4 können Bilder als Input bekommen [66]. Dies würde es ermöglichen, Coaching-Tools wie diagnostische Instrumente [7] einzusetzen oder Übungen durchzuführen, bei denen die Klient*innen Zeichnungen anfertigen oder mit Objekten interagieren, und dies in Interaktion mit dem Chatbot. Eine zusätzliche Erweiterung könnte die Analyse der Emotionen der Klient*innen durch Videoaufnahmen umfassen, um auf Mimik und Gestik reagieren zu können. Dies ist ein entscheidender Aspekt im Coaching-Prozess, da nonverbale Kommunikation wesentliche Einblicke in die Gefühlswelt der Klient*innen bietet [30, 61, 62, 87]. Hierfür könnte Facial Expression Recognition (FER) verwendet werden [51]. Mit diesen Methoden wäre es möglich eine Plattform zu erstellen, welche Stufe vier des Online-Coachings erreicht und somit sehr nahe an ein reales Coaching knüpft.

Ein weiterer Ansatzpunkt ist die Coaching-Methode. In einer zukünftigen Arbeit könnte eine andere Methodik verwendet werden, um neue Ergebnisse zu überprüfen. Zusätzlich könnte es sinnvoll sein das Studiendesign anzupassen. Coaching wird freiwillig bei einer Problematik aufgesucht [30, 61, 62, 87]. In diesem Studiensetting ist dies nicht gegeben. Aus diesem Grund könnte es sich lohnen eine Feldstudie zu betreiben. Dies könnte mithilfe einer longitudinalen Studie realisiert werden, indem der Chatbot als Unterstützung neben den Vorlesungen angeboten wird. Sobald ein Problem auftaucht, kann eine studierende Person auf den Bot zugreifen. So ist ein realistischeres Setting möglich und ein Grundprinzip des Coachings wird abgedeckt. Zusätzlich könnte es sich lohnen zu überprüfen ob die Noten mit den Ergebnissen der Studie korrelieren. Somit kann ein zusätzliches Merkmal überprüft werden, ob sich Chatbots lohnen in der Lehre einzusetzen.

Ein weiterer Ansatz könnte die Durchführung der Studie mit einer größeren Stichprobe sein, um die in dieser Untersuchung beobachteten Trends zu überprüfen und zu validieren.

Abschließend sollen andere technologische Umsetzungen angesprochen werden. GPT-4 ist zum Zeitpunkt dieser Arbeit das State-of-the-Art Modell welches zugänglich ist [66]. Jedoch wurde mittlerweile Gemini Ultra von Google angekündigt [105]. Dieses Modell ist ein Teil der Gemini-Familie und zeichnet sich durch die Fähigkeit aus, in 30 von 32 Benchmarks State-of-the-Art Leistungen zu erzielen, einschließlich einer Leistung auf dem Niveau menschlicher Experten beim MMLU-Benchmark [31]. Gemini Ultra demonstriert außergewöhnliche multimodale Fähigkeiten in der Verarbeitung von Bild-, Audio-, Video- und Textdaten [105]. Somit könnte es sich lohnen, die angesprochenen Veränderungsmöglichkeiten für einen Stufe vier Chatbot mit Gemini Ultra zu realisieren. Jedoch ist das Modell noch nicht frei zugänglich und kann nicht unabhängig überprüft werden. Da Google ein wirtschaftliches Interesse daran hat, ihr eigenes Modell gegenüber GPT-4 als überlegen darzustellen, sollten die Ergebnisse mit einer gewissen Vorsicht betrachtet werden.

8.2 Limitationen

Die Implementierung des Chatbots über Telegram brachte sowohl Vorteile als auch Herausforderungen mit sich. Die Analyse der TelegramIDs in der Datenbank deutet darauf

hin, dass die Mehrheit der Studierenden Telegram vor Beginn der Studie nicht verwendete. Dies stellte eine wesentliche Barriere für die Teilnahme an der Studie dar. Es ist anzunehmen, dass eine größere Anzahl an Teilnehmenden hätte erreicht werden können, wenn der Chatbot über eine Vielfalt an Benutzeroberflächen zugänglich gemacht worden wäre, um unterschiedliche Präferenzen und Vorerfahrungen der Studierenden zu berücksichtigen.

Eine weitere Limitation ist in der Coaching-Methode zu finden. Das Lerncoaching besitzt alle grundlegenden Elemente des Coachings. Somit kann man davon ausgehen, dass die Effekte im Coaching auch im Lerncoaching stattfinden. Jedoch wäre es ratsam dies in einer separaten Studie zu überprüfen. Zuletzt ist zu erwähnen, dass die Analyse der Chatprotokolle des Lerncoaching-Chatbots mit GPT-4 in dieser Studie fehlt. Zukünftige Forschungen sollten untersuchen, ob GPT-4 effektiv als Lerncoach agiert. Dafür sind zusätzliche Annotator*innen in einem neuen Projekt nötig, die in dieser Arbeit nicht verfügbar waren.

8.3 Fazit

Das Hauptziel dieser Arbeit war zu untersuchen, ob GPT-4 durch Few-Shot-Learning effektiv als Lerncoach eingesetzt werden kann, um die Motivation und Selbstwirksamkeit der Lernenden zu erhöhen. Dieser Ansatz zielte darauf ab, eine bestehende Forschungslücke im aktuellen wissenschaftlichen Diskurs zu adressieren. Die Bedeutung dieser Fragestellung wurde im Kontext des ausgearbeiteten theoretischen Rahmens dargestellt. In der Zusammenfassung lässt sich festhalten, dass die Forschungsfrage teilweise beantwortet werden konnte. Es wurde festgestellt, dass GPT-4, mit nur wenigen Anweisungen und einem minimalen Rahmen, die Selbstwirksamkeit von Studierenden steigern kann. Je nach theoretischer Grundlage ist die Selbstwirksamkeit ein Teil der Motivation.

Diese Arbeit konnte einen Teil dazu leisten die Forschungslücke zu schließen. Darüber hinaus wurden neue Perspektiven für zukünftige Forschungsarbeiten aufgezeigt, die im Ausblick der Studie diskutiert wurden.

9 References²


- [1] Martin Adam, Michael Wessel, and Alexander Benlian. 2021. AI-based chatbots in customer service and their effects on user compliance. *Electron Markets* 31, 2, 427–445. DOI: <https://doi.org/10.1007/s12525-020-00414-7>.
- [2] 2023. *aiogram 3.2.0 documentation* (November 2023). Retrieved November 21, 2023 from <https://docs.aiogram.dev/en/latest/>.
- [3] Shweta Arora, Shaurya Tiwari, Niharika Negi, Saurabh Pargaien, and Abhishek Misra. 2023. The Role of Artificial Intelligence in Mentoring Students. In *2023 1st International Conference on Circuits, Power and Intelligent Systems (CCPIS)*, 1–6. DOI: <https://doi.org/10.1109/CCPIS59145.2023.10291479>.
- [4] Amanda Askill, Yuntao Bai, Anna Chen, Dawn Drain, Deep Ganguli, Tom Henighan, Andy Jones, Nicholas Joseph, Ben Mann, Nova DasSarma, Nelson Elhage, Zac Hatfield-Dodds, Danny Hernandez, Jackson Kernion, Kamal Ndousse, Catherine Olsson, Dario Amodei, Tom Brown, Jack Clark, Sam McCandlish, Chris Olah, and Jared Kaplan. 2021. *A General Language Assistant as a Laboratory for Alignment*. DOI: <https://doi.org/typos>.
- [5] Tatiana Bachkirova and Carol Kauffman. 2008. Many ways of knowing: how to make sense of different research perspectives in studies of coaching. *Coaching: An International Journal of Theory, Research and Practice* 1, 2, 107–113. DOI: <https://doi.org/10.1080/17521880802328186>.
- [6] Tatiana Bachkirova and Carol Kauffman. 2009. The blind men and the elephant: using criteria of universality and uniqueness in evaluating our attempts to define coaching. *Coaching: An International Journal of Theory, Research and Practice* 2, 2, 95–105. DOI: <https://doi.org/10.1080/17521880903102381>.
- [7] Elke Berninger-Schäfer. 2018. *Online-Coaching*. Springer, Wiesbaden.
- [8] Abdullah Bin Sawad, Bhuva Narayan, Ahlam Alnefaie, Ashwaq Maqbool, Indra Mckie, Jemma Smith, Berkan Yuksel, Deepak Puthal, Mukesh Prasad, and A. B. Kocaballi. 2022. A Systematic Review on Healthcare Artificial Intelligent

² In dieser Masterarbeit wurde das Sprachmodell GPT-4 von OpenAI nach Genehmigung des Erstgutachters als Hilfsmittel zur sprachlichen Überarbeitung eingesetzt. Keine Texte wurden direkt von GPT-4 erstellt; das Tool diente lediglich der Verbesserung von Formulierung, Grammatik und Stil.

- Conversational Agents for Chronic Conditions. *Sensors (Basel, Switzerland)* 22, 7. DOI: <https://doi.org/10.3390/s22072625>.
- [9] 2023. *Bonferroni Correction -- from Wolfram MathWorld* (November 2023). Retrieved November 25, 2023 from <https://mathworld.wolfram.com/BonferroniCorrection.html>.
- [10] Josip Bozic and Franz Wotawa. 2018. Security Testing for Chatbots. In *Testing Software and Systems. 30th IFIP WG 6.1 International Conference, ICTSS 2018, Cadiz, Spain, October 1-3, 2018, Proceedings*, Inmaculada Medina-Bulo, Mercedes G. Merayo and Robert Hierons, Eds. Lecture Notes in Computer Science. Springer, Cham, 33–38. DOI: https://doi.org/10.1007/978-3-319-99927-2_3.
- [11] Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel M. Ziegler, Jeffrey Wu, Clemens Winter, Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. 2020. *Language Models are Few-Shot Learners*. DOI: <https://doi.org/pages>.
- [12] Daniel Burt and Zenobia Talati. 2017. The unsolved value of executive coaching: a meta-analysis of outcomes using randomised control trial studies. DOI: <https://doi.org/10.24384/000248>.
- [13] Janis A. Cannon-Bowers, Clint A. Bowers, Catherine E. Carlson, Shannon L. Doherty, Jocelyne Evans, and Julie Hall. 2023. Workplace coaching: a meta-analysis and recommendations for advancing the science of coaching. *Frontiers in psychology* 14, 1204166. DOI: <https://doi.org/10.3389/fpsyg.2023.1204166>.
- [14] François Chollet. 2017. *Deep Learning with Python*. Manning Publications, New York.
- [15] Christopher Rauen. 2021. Coaching-Prozesse. In *Handbuch Coaching*, Christopher Rauen, Ed. Hogrefe Verlag GmbH & Co. KG, 521–548.
- [16] Christopher Rauen. 2021. Tools im Coaching. In *Handbuch Coaching*, Christopher Rauen, Ed. Hogrefe Verlag GmbH & Co. KG, 549–576.

- [17] Christopher Rauen. 2021. Varianten des Coachings. In *Handbuch Coaching*, Christopher Rauen, Ed. Hogrefe Verlag GmbH & Co. KG, 37–65.
- [18] Nicola Döring. 2023. *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. (6th ed.). Springer Berlin; Springer, Berlin.
- [19] Edward W. Felten, Manav Raj, and Robert C. Seamans. 2023. How will Language Modelers like ChatGPT Affect Occupations and Industries? *SSRN Electronic Journal*.
- [20] Eleni Adamopoulou and Lefteris Moussiades. 2020. Chatbots: History, technology, and applications. *Machine Learning with Applications* 2, 100006. DOI: <https://doi.org/10.1016/j.mlwa.2020.100006>.
- [21] Tyna Eloundou, Sam Manning, Pamela Mishkin, and Daniel Rock. 2023. *GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models*.
- [22] E. Engelhardt. 2013. Qualitätsmerkmale guter Onlineberatung-Aktuelle Anforderungen an Forschung und Praxis. *Zeitschrift für systemische Therapie und Beratung* 31, 3, 111–115.
- [23] Emily M. Engelhardt. 2019. Onlineberatung - Digitales Beratungsangebot für Alle? In *Digital Diversity: Bildung und Lernen im Kontext gesellschaftlicher Transformationen*, Holger Angenent, Birte Heidkamp and David Kergel, Eds. Springer Fachmedien Wiesbaden, Wiesbaden, 161–173. DOI: https://doi.org/10.1007/978-3-658-26753-7_10.
- [24] 2023. *From BotFather to 'Hello World'* (November 2023). Retrieved November 20, 2023 from <https://core.telegram.org/bots/tutorial>.
- [25] GITNEX. 2023. *Must-Know Chat GPT-4 Statistics [Recent Analysis]* • Gitnux (December 2023). Retrieved December 25, 2023 from <https://gitnux.org/chat-gpt-statistics/>.
- [26] Leo A. Goodman. 1961. Snowball Sampling. *The Annals of Mathematical Statistics* 32, 1, 148–170.
- [27] 2023. *Google Formulare: App zum Erstellen von Onlineformularen* | Google Workspace (November 2023). Retrieved November 21, 2023 from <https://www.google.com/forms/about/>.
- [28] Siegfried Greif, Heidi Möller, and Wolfgang Scholl, Eds. 2016. *Handbuch Schlüsselkonzepte im Coaching*. Springer Berlin Heidelberg, Berlin, Heidelberg.

- [29] Biyang Guo, Xin Zhang, Ziyuan Wang, Minqi Jiang, Jinran Nie, Yuxuan Ding, Jianwei Yue, and Yupeng Wu. 2023. *How Close is ChatGPT to Human Experts? Comparison Corpus, Evaluation, and Detection*.
- [30] Hanna Hardeland. 2023. *Lerncoaching und Lernberatung: Lernende in ihrem Lernprozess wirksam begleiten und unterstützen : ein Buch zur (Weiter-)Entwicklung der theoretischen und praktischen (Lern-)Coachingkompetenz*. (9th ed.). Schneider Verlag Hohengrehren GmbH, Baltmannsweiler.
- [31] Dan Hendrycks, Collin Burns, Steven Basart, Andy Zou, Mantas Mazeika, Dawn Song, and Jacob Steinhardt. 2020. *Measuring Massive Multitask Language Understanding*. DOI: <https://doi.org/10.48550/arXiv.2009.03300>.
- [32] Miri Heo and Kyoung J. Lee. 2018. Chatbot as a New Business Communication Tool: The Case of Naver TalkTalk. *Bus. Commun. Res. Pract.* 1, 1, 41–45. DOI: <https://doi.org/10.22682/bcrp.2018.1.1.41>.
- [33] Sebastian Hobert. 2019. *How Are You, Chatbot? Evaluating Chatbots in Educational Settings – Results of a Literature Review*. DOI: https://doi.org/10.18420/delfi2019_289.
- [34] Sebastian Hobert and Raphael Meyer von Wolff. 2019. Say hello to your new automated tutor-a structured literature review on pedagogical conversational agents.
- [35] Redaktion B. L. Internetredaktion. 2023. *Digitalisierung im Bildungsbereich - BMBF Empirische Bildungsforschung* (2023). Retrieved December 25, 2023 from <https://www.empirische-bildungsforschung-bmbf.de/de/Digitalisierung-im-Bildungsbereich-1804.html>.
- [36] Katharina Jeblick, Balthasar Schachtner, Jakob Dextl, Andreas Mittermeier, Anna T. Stüber, Johanna Topalis, Tobias Weber, Philipp Wesp, Bastian Sabel, Jens Rieke, and Michael Ingrisch. 2022. *ChatGPT Makes Medicine Easy to Swallow: An Exploratory Case Study on Simplified Radiology Reports*.
- [37] Joao Luis Zeni Montenegro, Cristiano André da Costa, and Rodrigo da Rosa Righi. 2019. Survey of conversational agents in health. *Expert Systems with Applications* 129, 56–67. DOI: <https://doi.org/10.1016/j.eswa.2019.03.054>.
- [38] Florian Johannsen, Susanne Leist, Daniel Konadl, and Michael Basche. 2018. Comparison of commercial chatbot solutions for supporting customer interaction.

- [39] Rebecca J. Jones, Stephen A. Woods, and Yves R. F. Guillaume. 2016. The effectiveness of workplace coaching: A meta-analysis of learning and performance outcomes from coaching. *J Occupat & Organ Psyc* 89, 2, 249–277. DOI: <https://doi.org/10.1111/joop.12119>.
- [40] Dan Jurafsky and James H. Martin. 2023. *Speech and Language Processing* (March 2023). Retrieved March 16, 2023 from <https://web.stanford.edu/~jurafsky/slp3/>.
- [41] Jürgen Rudolph, Samson Tan, and Shannon Tan. 2023. ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *JALT* 6, 1. DOI: <https://doi.org/10.37074/jalt.2023.6.1.9>.
- [42] Richard R. Kilburg. 1996. Toward a conceptual understanding and definition of executive coaching. *Consulting Psychology Journal: Practice and Research* 48, 2, 134–144. DOI: <https://doi.org/10.1037/1061-4087.48.2.134>.
- [43] Helen Knauf. 2023. Lernen auf Distanz während der Coronapandemie. *Zeitschrift für Grundschulforschung* 16, 2, 465–479. DOI: <https://doi.org/10.1007/s42278-023-00168-z>.
- [44] Ahmet B. Kocaballi, Shlomo Berkovsky, Juan C. Quiroz, Liliana Laranjo, Huong L. Tong, Dana Rezazadegan, Agustina Briatore, and Enrico Coiera. 2019. The Personalization of Conversational Agents in Health Care: Systematic Review. *Journal of medical Internet research* 21, 11, e15360. DOI: <https://doi.org/10.2196/15360>.
- [45] 2023.  Kurse | IHK-Akademie in Ostbayern (November 2023). Retrieved November 13, 2023 from <https://www.ihk-wissen.de/kurse/kurs/Systemischer-Coach-IHK-mwd/18535MA050>.
- [46] Sen-Tarnng Lai, Fang-Yie Leu, and Jeng-Wei Lin. 2019. A Banking Chatbot Security Control Procedure for Protecting User Data Security and Privacy. In *Advances on Broadband and Wireless Computing, Communication and Applications. Proceedings of the 13th International Conference on Broadband and Wireless Computing, Communication and Applications (BWCCA-2018)*, Leonard Barolli, Fang-Yie Leu, Tomoya Enokido and Hsing-Chung Chen, Eds. Lecture Notes on Data Engineering and Communications Technologies, 25. Springer International Publishing, Cham, 561–571. DOI: https://doi.org/10.1007/978-3-030-02613-4_50.
- [47] Liliana Laranjo, Adam G. Dunn, Huong L. Tong, Ahmet B. Kocaballi, Jessica Chen, Rabia Bashir, Didi Surian, Blanca Gallego, Farah Magrabi, Annie Y. S. Lau,

- and Enrico Coiera. 2018. Conversational agents in healthcare: a systematic review. *Journal of the American Medical Informatics Association : JAMIA* 25, 9, 1248–1258. DOI: <https://doi.org/10.1093/jamia/ocy072>.
- [48] Matthew J. Leach. 2005. Rapport: a key to treatment success. *Complementary therapies in clinical practice* 11, 4, 262–265. DOI: <https://doi.org/10.1016/j.ctcp.2005.05.005>.
- [49] Stephanie Lin, Jacob Hilton, and Owain Evans. TruthfulQA: Measuring How Models Mimic Human Falsehoods. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Association for Computational Linguistics, Stroudsburg, PA, USA, 3214–3252. DOI: <https://doi.org/10.18653/v1/2022.acl-long.229>.
- [50] Edwin A. Locke and Gary P. Latham. 1990. *A theory of goal setting & task performance*. A theory of goal setting & task performance. Prentice-Hall, Inc, Englewood Cliffs, NJ, US.
- [51] M Anand and Dr. S. Babu. 2022. A Comprehensive Investigation on Emotional Detection in Deep Learning. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*.
- [52] Vanessa Mai, Caterina Neef, and Anja Richert. 2022. Developing an AI-based Coaching Chatbot: A Study on Disclosure as Effectiveness Factor in Human-Machine-Coaching. In *Human Interaction & Emerging Technologies (IHET-AI 2022): Artificial Intelligence & Future Applications*. AHFE International. AHFE International. DOI: <https://doi.org/10.54941/ahfe100917>.
- [53] Marcus Credé and L. Alison Phillips. 2011. A meta-analytic review of the Motivated Strategies for Learning Questionnaire. *Learning and Individual Differences* 21, 4, 337–346. DOI: <https://doi.org/10.1016/j.lindif.2011.03.002>.
- [54] Wari Maroengsit, Thanarath Piyakulpinyo, Korawat Phonyiam, Suporn Pongnumkul, Pimwadee Chaovalit, and Thanaruk Theeramunkong. 2019. A Survey on Evaluation Methods for Chatbots. In , 111–119. DOI: <https://doi.org/10.1145/3323771.3323824>.
- [55] W. J. McKeachie, P. R. Pintrich, Y. G. Lin, and D. Smith. 1987. Teaching and Learning in the College Classroom. A Review of the Research Literature (1986) and November 1987 Supplement.

- [56] Kenneth P. de Meuse, Guangrong Dai, and Robert J. Lee. 2009. Evaluating the effectiveness of executive coaching: beyond ROI? *Coaching: An International Journal of Theory, Research and Practice* 2, 2, 117–134. DOI: <https://doi.org/10.1080/17521880902882413>.
- [57] Madison Milne-Ives, Caroline de Cock, Ernest Lim, Melissa H. Shehadeh, Nick de Pennington, Guy Mole, Eduardo Normando, and Edward Meinert. 2020. The Effectiveness of Artificial Intelligence Conversational Agents in Health Care: Systematic Review. *Journal of medical Internet research* 22, 10, e20346. DOI: <https://doi.org/10.2196/20346>.
- [58] Quim Motger, Xavier Franch, and Jordi Marco. 2023. Software-Based Dialogue Systems: Survey, Taxonomy, and Challenges. *ACM Comput. Surv.* 55, 5, 1–42. DOI: <https://doi.org/10.1145/3527450>.
- [59] Quim Motger, Xavier Franch, and Jordi Marco. 2023. Software-Based Dialogue Systems: Survey, Taxonomy, and Challenges. *ACM Comput. Surv.* 55, 5, 1–42. DOI: <https://doi.org/10.1145/3527450>.
- [60] Alexander T. Neumann, Tamar Arndt, Laura Köbis, Roy Meissner, Anne Martin, Peter de Lange, Norbert Pengel, Ralf Klamma, and Heinz-Werner Wollersheim. 2021. Chatbots as a Tool to Scale Mentoring Processes: Individually Supporting Self-Study in Higher Education. *Frontiers in artificial intelligence* 4, 668220. DOI: <https://doi.org/10.3389/frai.2021.668220>.
- [61] Torsten Nicolaisen. 2017. *Einführung in das systemische Lerncoaching*. . Carl-Auer Compact. Carl-Auer Verlag, Heidelberg.
- [62] Torsten Nicolaisen. 2017. *Lerncoaching-Praxis: Coaching in pädagogischen Arbeitsfeldern*. (2nd ed.). Pädagogisches Training. Beltz Juventa, Weinheim.
- [63] Nils Dahlbäck, Arne Jönsson, and Lars Ahrenberg. 1993. Wizard of Oz studies: why and how. *Knowl. Based Syst.* 6, 258–266.
- [64] Norman M. Fraser and G.Nigel Gilbert. 1991. Simulating speech systems. *Computer Speech & Language* 5, 1, 81–99. DOI: [https://doi.org/10.1016/0885-2308\(91\)90019-M](https://doi.org/10.1016/0885-2308(91)90019-M).
- [65] OECD. 2023. *Changes in performance and equity in education between 2018 and 2022*.
- [66] OpenAI. 2023. *GPT-4 Technical Report*. DOI: <https://doi.org/pages>.

- [67] 2023. *OpenAI API* (November 2023). Retrieved November 21, 2023 from <https://openai.com/blog/openai-api>.
- [68] Elvis Ortega-Ochoa, Marta Arguedas, and Thanasis Daradoumis. 2023. Empathic pedagogical conversational agents: A systematic literature review. *Brit J Educational Tech*. DOI: <https://doi.org/10.1111/bjet.13413>.
- [69] Long Ouyang, Jeff Wu, Xu Jiang, Diogo Almeida, Carroll L. Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, John Schulman, Jacob Hilton, Fraser Kelton, Luke Miller, Maddie Simens, Amanda Aspell, Peter Welinder, Paul Christiano, Jan Leike, and Ryan Lowe. 2022. *Training language models to follow instructions with human feedback*.
- [70] Long Ouyang, Jeff Wu, Xu Jiang, Diogo Almeida, Carroll L. Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, John Schulman, Jacob Hilton, Fraser Kelton, Luke Miller, Maddie Simens, Amanda Aspell, Peter Welinder, Paul Christiano, Jan Leike, and Ryan Lowe. 2022. *Training language models to follow instructions with human feedback*.
- [71] Waldemar Pallasch and Uwe Hameyer. 2012. *Lerncoaching: Theoretische Grundlagen und Praxisbeispiele zu einer didaktischen Herausforderung*. (2nd ed.). Pädagogisches Training. Juventa Verlag, Weinheim, München.
- [72] Sajan B. Patel and Kyle Lam. 2023. ChatGPT: the future of discharge summaries? *The Lancet Digital Health* 5, 3, e107-e108. DOI: [https://doi.org/10.1016/s2589-7500\(23\)00021-3](https://doi.org/10.1016/s2589-7500(23)00021-3).
- [73] José Q. Pérez, Thanasis Daradoumis, and Joan M. M. Puig. 2020. Rediscovering the use of chatbots in education: A systematic literature review. *Comput Appl Eng Educ* 28, 6, 1549–1565. DOI: <https://doi.org/10.1002/cae.22326>.
- [74] Monika Perkhofer-Czapek and Renate Potzmann. 2016. (Lern-)Coach und (Lern-)Coaching. In *Begleiten, beraten und coachen. Der Lehrberuf im Wandel*, Monika Perkhofer-Czapek and Renate Potzmann, Eds. Research. Springer VS, Wiesbaden, 143–199. DOI: https://doi.org/10.1007/978-3-658-12828-9_6.
- [75] Paul R. Pintrich. 1988. A process-oriented view of student motivation and cognition. *New Drctns for Instit Rsrch* 1988, 57, 65–79. DOI: <https://doi.org/10.1002/ir.37019885707>.

- [76] Paul R. Pintrich and Elisabeth V. de Groot. 1990. Motivational and self-regulated learning components of classroom academic performance. *Journal of educational psychology* 82, 1, 33–40. DOI: <https://doi.org/10.1037/0022-0663.82.1.33>.
- [77] Paul R. Pintrich and others. 1991. A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ).
- [78] Python documentation. 2023. *sqlite3 — DB-API 2.0 interface for SQLite databases* (November 2023). Retrieved November 21, 2023 from <https://docs.python.org/3/library/sqlite3.html>.
- [79] Junaid Qadir. 2022. *Engineering Education in the Era of ChatGPT: Promise and Pitfalls of Generative AI for Education*. Institute of Electrical and Electronics Engineers (IEEE).
- [80] Qin Chen, Jinfeng Ge, Huaqing Xie, Xingcheng Xu, and Yanqing Yang. 2023. Large Language Models at Work in China’s Labor Market. *ArXiv* abs/2308.08776.
- [81] Alec Radford, Jong W. Kim, Tao Xu, Greg Brockman, Christine McLeavey, and Ilya Sutskever. 2022. *Robust Speech Recognition via Large-Scale Weak Supervision*. DOI: <https://doi.org/10.48550/arXiv.2212.04356>.
- [82] Alec Radford, Karthik Narasimhan, Tim Salimans, and Ilya Sutskever. 2018. *Improving language understanding by generative pre-training*.
- [83] Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever. 2019. *Language models are unsupervised multitask learners*.
- [84] Marco Ramponi. 2022. How ChatGPT actually works. *News, Tutorials, AI Research* (Dec. 2022).
- [85] Rasa Technologies Inc. 2023. *How to Build a Contextual AI Assistant* (December 2023). Retrieved December 24, 2023 from <https://info.rasa.com/how-to-build-contextual-assistant>.
- [86] Christopher Rauen. 2014. *Coaching*. (3rd ed.). Praxis der Personalpsychologie, Bd. 2. Hogrefe, Göttingen, Bern, Wien, Paris, Oxford, Prag, Toronto, Boston, Mass., Amsterdam, Kopenhagen, Stockholm, Florenz, Helsinki.
- [87] Christopher Rauen, Ed. 2021. *Handbuch Coaching*. Hogrefe Verlag GmbH & Co. KG. DOI: <https://doi.org/10.1026/02259-000>.
- [88] Anne Ribbers and Alexander Waringa. 2015. *E-coaching: Theory and practice for a new online approach to coaching*. Routledge, Abingdon, Oxon, New York, N.Y.

- [89] Roberto Pagano. 1981. Understanding Statistics in the Behavioral Sciences. In.
- [90] Malik Sallam. 2023. ChatGPT Utility in Healthcare Education, Research, and Practice: Systematic Review on the Promising Perspectives and Valid Concerns. *Healthcare* 11, 6, 887. DOI: <https://doi.org/10.3390/healthcare11060887>.
- [91] Theresa Schachner, Roman Keller, and Florian V Wangenheim. 2020. Artificial Intelligence-Based Conversational Agents for Chronic Conditions: Systematic Literature Review. *Journal of medical Internet research* 22, 9, e20701. DOI: <https://doi.org/10.2196/20701>.
- [92] 2023. *SciPy documentation — SciPy v1.11.4 Manual* (November 2023). Retrieved November 25, 2023 from <https://docs.scipy.org/doc/scipy/index.html>.
- [93] Ashish Sharma, Kevin Rushton, Inna Lin, David Wadden, Khendra Lucas, Adam Miner, Theresa Nguyen, and Tim Althoff. 2023. Cognitive Reframing of Negative Thoughts through Human-Language Model Interaction. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Association for Computational Linguistics, Stroudsburg, PA, USA, 9977–10000. DOI: <https://doi.org/10.18653/v1/2023.acl-long.555>.
- [94] Muskan Sharma, Anita Kumari, and Jyotsna. 2023. AI-Based Deep Learning Chatbot for Career and Personal Mentorship. In *2023 IEEE 3rd International Conference on Technology, Engineering, Management for Societal impact using Marketing, Entrepreneurship and Talent (TEMSMET)*, 1–6. DOI: <https://doi.org/10.1109/TEMSMET56707.2023.10149907>.
- [95] Shirley C. Sonesh, Chris W. Coultas, Christina N. Lacerenza, Shannon L. Marlow, Lauren E. Benishek, and Eduardo Salas. 2015. The power of coaching: a meta-analytic investigation. *Coaching: An International Journal of Theory, Research and Practice* 8, 2, 73–95. DOI: <https://doi.org/10.1080/17521882.2015.1071418>.
- [96] Heung-yeung Shum, Xiao-dong He, and Di Li. 2018. From Eliza to XiaoIce: challenges and opportunities with social chatbots. *Frontiers Inf Technol Electronic Eng* 19, 1, 10–26. DOI: <https://doi.org/10.1631/FITEE.1700826>.
- [97] Renzo Siegrist. 2006. Online-Coaching. In *Coaching: Angewandte Psychologie für die Beratungspraxis*, Eric Lippmann, Ed. Springer Berlin Heidelberg, Berlin, Heidelberg, 304–314. DOI: https://doi.org/10.1007/3-540-28451-6_24.

- [98] Siqi Shen, Charles F Welch, Rada Mihalcea, and Verónica Pérez-Rosas. 2020. Counseling-Style Reflection Generation Using Generative Pretrained Transformers with Augmented Context. In *SIGDIAL Conferences*.
- [99] Pavel Smutny and Petra Schreiberova. 2020. Chatbots for learning: A review of educational chatbots for the Facebook Messenger. *Computers & Education* 151, 103862. DOI: <https://doi.org/10.1016/j.compedu.2020.103862>.
- [100] Stella Kanatouri. 2021. Digitales Coaching. In *Handbuch Coaching*, Christopher Rauen, Ed. Hogrefe Verlag GmbH & Co. KG, 131–162.
- [101] Taylor N. Stephens, Angela Joerin, Michiel Rauws, and Lloyd N. Werk. 2019. Feasibility of pediatric obesity and prediabetes treatment support through Tess, the AI behavioral coaching chatbot. *Translational behavioral medicine* 9, 3, 440–447. DOI: <https://doi.org/10.1093/tbm/ibz043>.
- [102] Maja Storch. 2009. Motto-Ziele, S.M.A.R.T.-Ziele und Motivation. In *Coachingwissen: Denn sie wissen nicht, was sie tun?*, Bernd Birgmeier, Ed. VS Verlag für Sozialwissenschaften, Wiesbaden, 183–205. DOI: https://doi.org/10.1007/978-3-531-91766-5_12.
- [103] Maja Storch and Frank Krause. 2007. *Selbstmanagement - ressourcenorientiert: Grundlagen und Trainingsmanual für die Arbeit mit dem Zürcher Ressourcen Modell (ZRM)*. (4th ed.). Psychologie Praxis. Huber, Bern.
- [104] Teo Susnjak. 2022. *ChatGPT: The End of Online Exam Integrity?*
- [105] Gemini Team. 2023. *Gemini: A Family of Highly Capable Multimodal Models*. DOI: <https://doi.org/10.48550/arXiv.2312.11805>.
- [106] Tim Theeboom, Bianca Beersma, and Annelies E. van Vianen. 2014. Does coaching work? A meta-analysis on the effects of coaching on individual level outcomes in an organizational context. *The Journal of Positive Psychology* 9, 1, 1–18. DOI: <https://doi.org/10.1080/17439760.2013.837499>.
- [107] N. T. Thomas. 2016. An e-business chatbot using AIML and LSA. In *2016 International Conference on Advances in Computing, Communications and Informatics (I-CACCI)*. IEEE. DOI: <https://doi.org/10.1109/icacci.2016.7732476>.
- [108] Lorainne Tudor Car, Dhakshenya A. Dhinakaran, Bhone M. Kyaw, Tobias Kowatsch, Shafiq Joty, Yin-Leng Theng, and Rifat Atun. 2020. Conversational

- Agents in Health Care: Scoping Review and Conceptual Analysis. *Journal of medical Internet research* 22, 8, e17158. DOI: <https://doi.org/10.2196/17158>.
- [109] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. 2017. *Attention Is All You Need*.
- [110] Qing Wang, Yi-Ling Lai, Xiaobo Xu, and Almuth McDowall. 2022. The effectiveness of workplace coaching: a meta-analysis of contemporary psychologically informed coaching approaches. *Journal of Work-Applied Management* 14, 1, 77–101. DOI: <https://doi.org/10.1108/JWAM-04-2021-0030>.
- [111] Robert Wegener. 2018. *Coaching-Prozessforschung: Forschung und Praxis im Dialog – Fachbuch*. . Vandenhoeck et Ruprecht, Göttingen.
- [112] Robert Wegener, Silvia Deplazes, Melanie Hasenbein, Hansjörg Künzli, Annamari Ryter, and Beat Uebelhart, Eds. 2016. *Coaching als individuelle antwort auf gesellschaftliche entwicklungen*. Vs Verlag Fur Sozialwisse, [Place of publication not identified]. DOI: <https://doi.org/10.1007/978-3-658-12854-8>.
- [113] Bernard Weiner. 1989. *Human motivation*. L. Erlbaum, Hillsdale, N.J.
- [114] Rand Wilcox. 2012. Introduction to robust estimation and hypothesis testing. 3rd ed. *Introduction to Robust Estimation and Hypothesis Testing* 93. DOI: <https://doi.org/10.2307/2669876>.
- [115] Sebastian Wollny, Jan Schneider, Daniele Di Mitri, Joshua Weidlich, Marc Rittberger, and Hendrik Drachsler. 2021. Are We There Yet? - A Systematic Literature Review on Chatbots in Education. *Frontiers in artificial intelligence* 4, 654924. DOI: <https://doi.org/10.3389/frai.2021.654924>.
- [116] Anbang Xu, Zhe Liu, Yufan Guo, Vibha Sinha, and Rama Akkiraju. 2017. A New Chatbot for Customer Service on Social Media. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, 3506–3510. DOI: <https://doi.org/10.1145/3025453.3025496>.

Anhang A: Aufbau der Vorstudie

Die Teilnehmenden wurden zu einem Zoom-Meeting eingeladen, in dem die Studie und der Ablauf der Session erläutert wurden. Ihnen wurde eine Datenschutzerklärung zur Unterzeichnung vorgelegt. Es wurde ihnen auch mitgeteilt, dass der Bot ein Prototyp ist, was zu eventuell längeren Antwortzeiten führen könnte. Somit wurden eventuell längere Pausen erläutert, da der Wizard Zeit benötigt, um Antworten zu bilden.

A1 Beispielchat einer Vorstudie

Im Folgenden wird ein Beispielchat aus der Vorstudie präsentiert. Die Phasen des Lerncoachings sind farblich markiert. Der Forschende wird mit F abgekürzt, die Teilnehmenden mit T.

- 1. Einstieg/Kontakt herstellen
- 2. Thema/Anliegen erfassen
- 3. Ziel erfassen
- 4. Lösungen
- 5. Abschluss

Message

F	Hallo und herzlich willkommen! Ich bin Mila, dein Lerncoaching Chatbot. Es ist mir eine Freude, dich auf deiner individuellen Lernreise zu unterstützen.
T	Hausarbeiten schreiben unter Zeitdruck
F	Verstehe ich das richtig, du möchtest über das Thema Hausarbeiten unter Zeitdruck sprechen? Bist du bereits in einer solchen Situation gewesen?
T	Ja. Jedes Semester
F	Was tust du normalerweise in solchen Situationen?
T	Stress, Frust darüber nicht früher angefangen zu haben.
F	Verstehe, du Ärgerst dich über dich selbst und musst länger arbeiten.
T	Genau so ist es
F	Hattest du schon einmal eine Situation, in der du dem Problem entkommen konntest?
T	öfters schon, ja
F	Was hat in einer solchen Situation den Unterschied gemacht?
T	Verlängerungen der Abgabefrist, exzessive Arbeitssessions oder einfach nicht Abgabe der Arbeit
F	Verstehe. Wie hast du dich gefühlt, wenn du solche Lösungen angewendet hast?
T	Je nach Dozent*in unangenehm beim erfragen nach einer Verlängerung.
F	Kann man das auch als Stärke betrachten, wenn man sich traut nachzufragen und dies als Ressource nutzt?
T	Ja, doch, ich glaube schon

F	Wenn morgen eine gute Fee erscheinen würde und du einen Wunsch bezüglich deiner Hausarbeiten hättest, was würdest du ändern?
T	Strukturierteres Arbeiten und besseres Zeitmanagement
F	Ich verstehe, für dich ist strukturiertes Arbeiten und Zeitmanagement wichtig. Auf einer Skala von 1 bis 10, wie wichtig ist dieses Thema für dich?
T	9
F	Also eine 9. Wie ich sehe, ist das Thema wichtig für dich. Ist es für dich in Ordnung, wenn wir auf Basis dieses Themas ein konkretes Ziel für dich formulieren?
T	Klar, gerne
F	Ich möchte nun mit dir ein Ziel formulieren. Zielformulierung im Lerncoaching bietet Klarheit, fördert die Motivation und ermöglicht eine messbare Überwachung des Fortschritts.
T	Ja, voll
F	Super! Um ein Ziel genau formulieren zu können möchte ich mit dir die SMART Regel verwenden. SMART ist ein Akronym und steht für: Spezifisch, Messbar, Anspruchsvoll/Attraktiv, Realistisch und Terminiert. Lass uns die einzelnen Schritte einmal gemeinsam durchgehen.
F	1. Was genau möchtest du in deinen eigenen Worten erreichen?
T	Ich möchte früher damit beginnen, meine Hausarbeiten anzufertigen und diese nacheinander und ohne Zeitdruck finalisieren. Zeitdruck soll nicht als nötige Motivation dienen
F	2. Woran würdest du merken, dass du dein Ziel erreicht hast?
T	Weniger Stress, Frust und schlechtes Gewissen. Außerdem wäre ich dann körperlich und geistig nicht so ausgelaugt
F	Ich verstehe, spielen diese Faktoren auch eine Rolle bei der Bedeutung des Ziels?
T	Ja
F	5.
T	Bis zur nächsten Prüfungsphase
F	Vielen Dank für deine Antworten! Lass uns nun die gesammelten Informationen in einem oder maximal zwei Sätzen zusammenfassen. Hier ist ein Beispiel für dich:
T	Bis zum Ende des nächsten Semesters will ich strukturierte Arbeitsweisen mir aneignen um für die nächsten Hausarbeiten durch eine bessere Vorplanung und rechtzeitige Durchführung nicht mehr Zeitdruck als Motivation zum schreiben meine Hausarbeiten zu brauchen.
F	Prima! Ich danke dir für deine aktive Mitarbeit. Zum Abschluss besprechen wir gemeinsam, wie du dein definiertes Ziel in die Praxis umsetzen kannst.
T	Feste Zeitslots für Arbeitssessions
F	Möchtest du also feste Zeitslots in deinem Zeitplan einrichten?
T	Ja
F	Ist das zurzeit für dich realistisch? Wie fühlt sich die erste Umsetzung dieser Idee an?
T	An die Bibliothek fahren zum Arbeiten um nicht davon abgelenkt zu werden
F	Wenn etwas dazwischenkommt und du die Zeitslots nicht einhalten kannst, welche Maßnahmen könntest du ergreifen, um damit umzugehen?
T	An anderen Tagen Zeit den Slot erfüllen oder mich damit abfinden, einen Slot nicht einhalten zu können

F	Wenn du in die Zukunft blickst und dir vorstellst, wie du eine Hausarbeit ohne Stress abgibst, wie würde sich das für dich anfühlen?
T	Absolut grandios
F	Das freut mich zu hören. Abschließend eine Frage: Wie kannst du das Ziel im Kopf behalten?
T	Visualisierung der Zeitplanung, zum Beispiel durch einen Kalender.
F	Vielen Dank für deine Lösungsvorschläge. Ich wünsche dir viel Erfolg bei der Umsetzung deines Zieles!
T	Diese Übung hat geholfen, Punkte die ich eigentlich bereits weis und meistens mit mir selber ausmache zu visualisieren und dadurch mit einer anderen Perspektive darauf zu schauen
F	Vielen Dank für deine Teilnahme an dieser Studie! Wende dich nun an den Betreuer.

Tabelle 9: Beispielchat einer Vorstudie

Anhang B: Prompt- und Chatbeispiele

In diesem Abschnitt werden die Prompts vorgestellt, die für die Chatbots verwendet wurden. Dies umfasst sowohl die Verhaltensaufforderungen an GPT-4 als auch die Few-Shot-Beispiele, die zum Einsatz kamen. Zusätzlich werden Chatbeispiele beider Chatbots präsentiert.

B1 Prompts an GPT-4

Bei dem Lerncoaching Chatbot sind unterschiedliche Prompts passend zur jeweiligen Phase verwendet worden. Zusätzlich veränderten sich die Prompts nach einer vordefinierten Anzahl an Turns innerhalb der Phasen. In Phase zwei und vier werden jeweils drei unterschiedliche Prompts verwendet.

Phase zwei:

1. Aufgabenstellung:

- Verhalte dich wie ein Lerncoach. Fokus sehr wichtig: Gebe Keine Lösungen vor. Hilfe dem Klienten eine Lösung zu finden. Stelle pro Turn eine Frage. Keine Ratschläge oder Richtungen. Am Anfang ist es wichtig Beispiele zu erfragen und den Klienten zu verstehen.

1. Few-Shot-Beispiel:

- Klient: Hausarbeiten schreiben unter Zeitdruck
Lerncoach: Verstehe ich das richtig, du möchtest über das Thema Hausarbeiten unter Zeitdruck sprechen? Bist du bereits in einer solchen Situation gewesen?
Klient: Ja. Jedes Semester
Lerncoach: Was tust du normalerweise in solchen Situationen?
Klient: Stress, Frust darüber nicht früher anfangen zu haben.
Lerncoach: Verstehe, du ärgerst dich über dich selbst und musst länger arbeiten.
Klient: Genau so ist es
Lerncoach: Hattest du schon einmal eine Situation, in der du dem Problem entgegenkommen konntest?"
Klient: Öfters schon, ja"

2. Aufgabenstellung:

- Verhalte dich wie ein Lerncoach. Fokus sehr wichtig: Gebe Keine Lösungen vor. Hilfe dem Klienten eine Lösung zu finden. Stelle pro Turn eine Frage. Keine Rat-schläge oder Richtungen. Versuche das Problem des Klienten zu verstehen.

2. Few-Shot-Beispiel:

- Klient: Der innere Druck wird zu hoch und ich fange an zu lernen, weil ich weiß, dass es dann etwas besser wird.

Lerncoach: Der innere Druck treibt dich also dazu an zu lernen. Kannst du be-schreiben, wie sich das genau anfühlt?

Klient: Es fühlt sich an wie ein Gewicht auf meiner Brust

Lerncoach: Das scheint eine Menge Energie zu erfordern."

Klient: Das stimmt

Lerncoach: Wenn morgen eine gute Fee auftauchen würde und dir bei deinem Problem helfen könnte, welche Veränderungen würdest du erwarten?

Klient: Ich würde hoffen, dass sie mir zu mehr Antrieb verhilft und ich mich bes-ser konzentrieren kann.

3. Aufgabenstellung:

- Verhalte dich wie ein Lerncoach. Fokus sehr wichtig: Gebe Keine Lösungen vor. Hilfe dem Klienten eine Lösung zu finden. Stelle pro Turn eine Frage. Keine Rat-schläge oder Richtungen. Versuche das Thema des Klienten zu konkretisieren.

3. Few-Shot-Beispiel:

Klient: Hattest du schon einmal eine Prüfung, bei der du dachtest, dass du Schwierigkeiten haben würdest, aber du konntest sie dennoch erfolgreich bewäl-tigen?

Lerncoach: Ja das ist auch schon vorgekommen.

Klient: Was war in dieser Situation der Schlüssel zum Erfolg, und was kannst du daraus für die nächste Prüfung lernen?

Lerncoach: Naja, ich habe mich durchgebissen und hatte auch eventuell glück bei der Prüfung.

Klient: Kann man das Durchbeißen auch als eine Form von Stärke und Ressource betrachten?

Lerncoach: Ja, finde ich schon. Solang das Ergebnis zufriedenstellend ist und man sich auf dem weg nicht verliert.

Klient: Vielen Dank für die umfangreichen Informationen und die ersten Vorschläge zu deinem Thema.

Phase vier:

1. Aufgabenstellung:

- Verhalte dich wie ein Lerncoach. Fokus sehr wichtig: Gebe Keine Lösungen vor. Hilfe dem Klienten eine Lösung zu finden. Stelle pro Turn eine Frage. Keine Ratschläge oder Richtungen. Hilfe dem Klienten sein Thema im Alltag umzusetzen.

2. Few-Shot-Beispiel:

- Lerncoach: Prima! Ich danke dir für deine aktive Mitarbeit. Zum Abschluss besprechen wir gemeinsam, wie du dein definiertes Ziel in die Praxis umsetzen kannst. Wie kannst du dein formuliertes Ziel im Alltag umsetzen? Was wäre ein erster kleiner Schritt?"

Klient: Mich durch Kalendereinträge ans lernen zu erinnern

Lerncoach: Wie genau würde das funktionieren?

Klient: Durch gewissenhafte Pflege meines Kalenders

Lerncoach: Das hört sich schon sehr gut an. Wann möchtest du damit anfangen?

Klient: Sobald das Semester beginnt.

2. Aufgabenstellung:

- Verhalte dich wie ein Lerncoach. Fokus sehr wichtig: Gebe Keine Lösungen vor. Hilfe dem Klienten eine Lösung zu finden. Stelle pro Turn eine Frage. Keine Ratschläge oder Richtungen. Hilfe dem Klienten sein Thema im Alltag umzusetzen.
(Identisch zur letzten Aufgabenstellung)

2. Few-Shot-Beispiel:

- Klient: Bereits während der Vorlesungszeit jedes abgeschlossene Kapitel weitestgehend zusammenzufassen

Lerncoach: Ich verstehe, du möchtest Kapitel bereits während der Vorlesungszeit zusammenfassen.

Klient: Ja, um damit Zeit im gleichen Zuge Zeit zu sparen

Lerncoach: Gibt es noch Möglichkeiten, wie du die Zusammenfassungen effizienter gestalten könntest, um noch mehr Zeit zu sparen?

Klient: Die wichtigsten Punkte der Kapitel möglichst in Stichpunkten, aber verständlich aufzuschreiben

Lerncoach: Stell dir vor, der Kurs wäre vorbei, und du hättest erfolgreich den Stoff strukturiert. Wie fühlst du dich dabei?

3. Aufgabenstellung:

- Verhalte dich wie ein Lerncoach. Fokus sehr wichtig: Gebe Keine Lösungen vor. Hilfe dem Klienten eine Lösung zu finden. Stelle pro Turn eine Frage. Keine Ratschläge oder Richtungen. Hilfe dem Klienten sein Thema im Alltag umzusetzen, konkretisiere die Vorschläge des Klienten.

3. Few-Shot-Beispiel:

- Lerncoach: Wenn etwas dazwischenkommt und du die Zeitslots nicht einhalten kannst, welche Maßnahmen könntest du ergreifen, um damit umzugehen?
Klient: An anderen Tagen Zeit den Slot erfüllen oder mich damit abfinden, einen Slot nicht einhalten zu können
Lerncoach: Wenn du in die Zukunft blickst und dir vorstellst, wie du eine Hausarbeit ohne Stress abgibst, wie würde sich das für dich anfühlen?
Klient: Absolut grandios
Lerncoach: Das freut mich zu hören. Abschließend eine Frage: Wie kannst du das Ziel im Kopf behalten?
Klient: Visualisierung der Zeitplanung, zum Beispiel durch einen Kalender.
Lerncoach: Vielen Dank für deine Lösungsvorschläge. Ich wünsche dir viel Erfolg bei der Umsetzung deines Zieles!

Bei dem Neutralen Chatbot wurde eine spezifische Aufgabenstellung implementiert, ohne den Einsatz von Few-Shot-Beispielen. Die Entscheidung, GPT-4 nicht ohne eine vorgegebene Aufgabenstellung zu verwenden, erfolgte nach dem Testen des Systems. Es stellte sich heraus, dass die Antworten von GPT-4 ohne klare Aufgabenstellung oft zu lang waren und somit das Token-Limit gefährdeten. Zudem regten diese Antworten keine Konversation an, da GPT-4 meist mit konkreten Lösungsvorschlägen oder Fakten antwortete. Da der Lerncoaching Chatbot versucht eine Konversation zu führen, war es notwendig, dass auch der neutrale Chatbot dies tut, um die Vergleichbarkeit zu

gewährleisten. Nach mehrmaligen Tests verschiedener Prompts wurde eine Aufgabenstellung entwickelt, die neutrale, jedoch gesprächsanregende Antworten ermöglicht.

Aufgabenstellung:

- Verhalte dich wie ein neutraler Assistent. Hilfe dem Klienten eine Lösung zu finden. Versuche den Klienten zu einem aktiven Gespräch zu animieren, frage nach Kontext. Stelle pro Turn eine Frage. Halte deinen Redeanteil kurz.

B2 Chatbeispiele

Nun sollen zwei Chatausschnitte der Chatbots präsentiert werden.

Lerncoaching Chatbot:

Aussage des Teilnehmenden	Antwort des Chatbots
Ich fühle mich sehr ausgelastet mit meinem Privatleben und dem Studium.	Verstehe ich das richtig, dass es dir gerade schwerfällt, dein Privatleben und das Studium unter einen Hut zu bekommen? Kannst du mir ein Beispiel dafür geben?
Ich arbeite 20 Stunden in der Woche um für meine Familie zu sorgen und unsere Ausgaben zu decken und studiere nebenbei Vollzeit und versuche so viel wie möglich zu schaffen.	Das klingt nach einer sehr anspruchsvollen Situation. Wie organisierst du bislang deine Zeit, um sowohl deinen Arbeitsverpflichtungen als auch dem Studium gerecht zu werden?
Ich arbeite meist 2 Stunden in der Früh, bevor ich in die Uni muss. Dann habe ich dort ein bis drei Veranstaltungen, komme dann wieder nach Hause um zu arbeiten und fahre danach erneut in die Uni, um weitere Kurse zu absolvieren. Da bleibt mir wenig Zeit zu lernen oder bei meiner Frau zu sein.	Es klingt so, als hättest du einen sehr vollen Terminkalender, und ich kann verstehen, dass das herausfordernd sein kann. Was hast du bisher unternommen, um eine Balance zwischen Lernen, Arbeit und Zeit mit deiner Frau zu finden?
Ich versuche nach 18:30 Uhr nichts mehr für die Arbeit oder die Uni zu tun,	Es ist toll, dass du einen klaren Cut-off-Zeitpunkt für die Arbeit und das

sondern mit meiner ganzen Aufmerksamkeit daheim zu sein.	Studium festgelegt hast, um bei deiner Familie präsent zu sein. Wie gut funktioniert das bisher und fühlst du dich mit dieser Herangehensweise ausgewogen?
Es funktioniert meistens gut. An einem Tag habe ich bis 20 Uhr Vorlesungen.	Es ist schön zu hören, dass es an den meisten Tagen gut funktioniert. Wie bewältigst du dann diesen einen Tag, an dem du bis 20 Uhr Vorlesungen hast?
Ich habe 4 Stunden Pause zwischen meiner Übung und der letzten Vorlesung, die um 18 Uhr startet. In dieser Zeit arbeite ich. Wenn ich dann um 20:15 zu Hause bin, esse ich und bin ausgelastet.	Es klingt, als würdest du diesen Tag gut für Arbeit nutzen und trotz des späten Unterrichtsendes schnell zu Hause sein. Wie fühlst du dich am Ende eines solchen Tages und wie beeinflusst das deine Energie und Motivation für den nächsten Tag?
Zurzeit fühle ich mich noch gut. Da mir das Studium generell viel Spaß macht, fehlt es nicht an Motivation. Aber ich bin auch erst in der 2. Woche. Ich hoffe, dass ich mich im Laufe des Semesters nicht überfordere.	Es ist großartig, dass du in dieser Phase noch motiviert bist und das Studium Spaß macht. Was sind deine bisherigen Erfahrungen oder Strategien, um sicherzustellen, dass du dich während des Semesters nicht überforderst?
Ich versuche mir einen Tag der Woche nicht zu voll mit Arbeit zu machen, damit ich Zeit habe, mich gut für das Studium vorzubereiten und auch Zeit zum Lernen zu finden, die sehr knapp werden kann in einer vollen Woche. Den Samstag halte ich mir immer frei, um Zeit mit meiner schwangeren Frau zu verbringen und abschalten zu können.	Es klingt, als hättest du eine gute Strategie, um Zeit für das Studium und persönliche Belange zu schaffen. Wie effektiv ist diese Herangehensweise bisher für dich gewesen und wie konntest du Stress oder Überforderungen damit bewältigen?
Bisher habe ich nicht viel Zeit gefunden, mich auf das Studium vorzubereiten, da	Es ist schön zu hören, dass das freie Wochenende dir hilft, zu entspannen und

es oftmals zu kurzfristigen Änderungen in meiner täglichen Routine kam. Das freie Wochenende hilft mir aber immer sehr, mich zu entspannen und Energie zu sammeln.	Energie zu sammeln. Was könntest du möglicherweise tun, um trotz kurzfristiger Änderungen in deiner täglichen Routine genügend Zeit für die Studiovorbereitung zu finden?
Hast du eine Idee für mich?	Welche Möglichkeiten siehst du, um die Qualität deiner Lernzeit zu verbessern oder effektiver zu nutzen, selbst wenn du weniger Zeit zur Verfügung hast, aufgrund von kurzfristigen Änderungen?
Ich könnte mir eine To-do-Liste erstellen, geordnet nach Wichtigkeit. Dort könnte ich Aufgaben anführen, die zu erledigen sind bis zu einem bestimmten Zeitpunkt und auch Lerninhalte, die wichtig wären, sofort zu vertiefen.	Eine To-do-Liste klingt wie ein praktischer Ansatz, um deine Aufgaben und Lerninhalte zu organisieren. Wie denkst du, würde diese Liste dir helfen, fokussierter zu arbeiten und deine Lernzeit effektiver zu gestalten?

Tabelle 10: Beispiel Lerncoaching-Chatbot

Neutraler Chatbot:

Aussage des Teilnehmenden	Antwort des Chatbots
Mir fällt es oft schwer, mich für Aufgaben zu motivieren und mich nicht ablenken zu lassen.	Welche Art von Aufgaben fällt dir besonders schwer zu erledigen?
Texte lesen	Gibt es bestimmte Faktoren, die das Lesen von Texten für dich besonders unangenehm oder schwierig machen?
Wenn sie auf Englisch sind oder sehr technisch	In welchen Situationen musst du solche Texte lesen und wie wichtig sind sie für deine persönlichen oder beruflichen Zielsetzungen?
Für Literaturrecherche für meine Masterarbeit	Wie organisierst du derzeit deinen Arbeitsprozess im Rahmen der Literaturrecherche für die

	Masterarbeit und welche Zeitpläne hast du bereits festgelegt?
Ich überlege mir jeweils, wie viel ich an einem Tag schaffen möchte, und versuche, möglichst strukturiert vorzugehen und meine Ergebnisse in einem Dokument festzuhalten.	Welche Ressourcen oder Strategien nutzt du derzeit, um technische Texte oder englischsprachige Inhalte besser zu verstehen und effektiv in deine Arbeit zu integrieren?
Schwierige Textpassagen übersetze ich mir ins Deutsche oder google schwierige Begriffe.	Welche Tools oder Techniken hast du bisher zur Verbesserung deiner Englischkenntnisse verwendet, oder welche würdest du gerne ausprobieren, um dich langfristig beim Verstehen von technischen Texten auf Englisch wohlfühlen?
Ich möchte Englisch gern mehr in meinen Alltag integrieren, z.B. englische Bücher lesen, Serien auf Englisch schauen.	Welches Genre von Büchern und Serien interessiert dich am meisten und welche Titel hast du bereits auf deiner Liste, um dein Englisch im Alltag zu verbessern?
Ich mag gerne Romane. Aber ich habe keine Liste mit Büchern, die ich lesen möchte.	Wie wäre es, wenn du dir eine Liste von Autoren oder Buchempfehlungen zu deinem bevorzugten Genre ansiehst, um daraus geeignete englischsprachige Bücher für deinen Leseplan auszuwählen?
Das ist eine gute Idee.	Würdest du dabei die Klassiker bevorzugen oder bist du offen für aktuelle Bestseller und weniger bekannte Autoren?
Ich denke, ich bin offen. Aktuelle Bücher sind mir sogar lieber.	Möchtest du lieber Bücher von britischen oder amerikanischen Autoren lesen, um dich mit verschiedenen Varianten der englischen Sprache vertraut zu machen, oder hast du keine bestimmte Präferenz?

Hier habe ich keine bestimmte Präferenz, finde ich beides gut.	In welchem Genre suchst du vorrangig nach englischsprachigen Büchern? Zum Beispiel Krimi, Drama, Fantasy, Science-Fiction oder Historisch?
--	--

Tabelle 11: Beispiel Neutraler-Chatbot

Anhang C: Überstzung, Motivated Strategies for Learning Questionnaire

In diesem Abschnitt wird der Fragebogen präsentiert.

Deutsche Aussage	Englische Übersetzung
In einem Kurs wie diesem ziehe ich Kursmaterial vor, das mich wirklich herausfordert, damit ich Neues lernen kann.	In a class like this, I prefer course material that really challenges me so I can learn new things.
Wenn ich auf angemessene Weise lerne, dann werde ich in der Lage sein, den Stoff dieses Kurses zu verstehen.	If I study in appropriate ways, then I will be able to learn the material in this course.
Wenn ich einen Test schreibe, denke ich darüber nach, wie schlecht ich im Vergleich zu anderen Studierenden abschneide.	When I take a test I think about how poorly I am doing compared with other students.
Ich denke, dass ich das, was ich in diesem Kurs lerne, auch in anderen Kursen verwenden kann.	I think I will be able to use what I learn in this course in other courses.
Ich glaube, dass ich in diesem Kurs eine ausgezeichnete Note bekommen werde.	I believe I will receive an excellent grade in this class.
Ich bin sicher, dass ich den schwierigsten Stoff, der in der Literatur für diesen Kurs vorgestellt wird, verstehen kann.	I'm certain I can understand the most difficult material presented in the readings for this course.
Eine gute Note in diesem Kurs zu bekommen, ist für mich gerade das Befriedigendste.	Getting a good grade in this class is the most satisfying thing for me right now.
Wenn ich einen Test schreibe, denke ich an Fragen in anderen Teilen des Tests, die ich nicht beantworten kann.	When I take a test I think about items on other parts of the test I can't answer.
Es ist meine eigene Schuld, wenn ich den Stoff dieses Kurses nicht lerne.	It is my own fault if I don't learn the material in this course.
Es ist wichtig für mich, den Stoff dieses Kurses zu lernen.	It is important for me to learn the course material in this class.

Das Wichtigste für mich gerade ist, meinen Gesamtnotendurchschnitt zu verbessern. Daher ist meine Hauptbesorgnis in diesem Kurs, eine gute Note zu bekommen.	The most important thing for me right now is improving my overall grade point average, so my main concern in this class is getting a good grade.
Ich bin zuversichtlich, die Grundkonzepte, die in diesem Kurs gelehrt werden, zu lernen.	I'm confident I can learn the basic concepts taught in this course.
Wenn möglich, möchte ich in diesem Kurs bessere Noten bekommen als die meisten anderen Studierenden.	If I can, I want to get better grades in this class than most of the other students.
Wenn ich Tests mache, denke ich an die Konsequenzen des Durchfallens.	When I take tests I think of the consequences of failing.
Ich bin zuversichtlich, dass ich den komplexesten Stoff, der von dem Dozenten/der Dozentin in diesem Kurs präsentiert wird, verstehen kann.	I'm confident I can understand the most complex material presented by the instructor in this course.
In einem Kurs wie diesem ziehe ich Kursmaterial vor, das meine Neugier weckt, auch wenn es schwer zu lernen ist.	In a class like this, I prefer course material that arouses my curiosity, even if it is difficult to learn.
Ich bin sehr am Inhalt dieses Kurses interessiert.	I am very interested in the content area of this course.
Wenn ich mich genug anstrengende, werde ich den Kursstoff verstehen.	If I try hard enough, then I will understand the course material.
Ich fühle mich unruhig und aufgeregt, wenn ich eine Prüfung schreibe.	I have an uneasy, upset feeling when I take an exam.
Ich bin zuversichtlich, dass ich bei den Aufgaben und Tests in diesem Kurs hervorragende Arbeit leisten kann.	I'm confident I can do an excellent job on the assignments and tests in this course.
Ich erwarte, dass ich in diesem Kurs gut abschneide.	I expect to do well in this class.

Das Befriedigendste für mich in diesem Kurs ist, den Inhalt so gründlich wie möglich zu verstehen.	The most satisfying thing for me in this course is trying to understand the content as thoroughly as possible.
Ich denke, dass der Kursstoff in diesem Kurs nützlich für mich ist zu lernen.	I think the course material in this class is useful for me to learn.
Wenn ich die Möglichkeit habe, wähle ich in diesem Kurs Aufgaben aus, von denen ich lernen kann, auch wenn sie keine gute Note garantieren.	When I have the opportunity in this class, I choose course assignments that I can learn from even if they don't guarantee a good grade.
Wenn ich den Kursstoff nicht verstehe, liegt es daran, dass ich mich nicht genug angestrengt habe.	If I don't understand the course material, it is because I didn't try hard enough.
Ich mag den Stoff dieses Kurses.	I like the subject matter of this course.
Es ist mir sehr wichtig, den Stoff dieses Kurses zu verstehen.	Understanding the subject matter of this course is very important to me.
Ich spüre, wie mein Herz schnell schlägt, wenn ich eine Prüfung schreibe.	I feel my heart beating fast when I take an exam.
Ich bin sicher, dass ich die Fähigkeiten, die in diesem Kurs gelehrt werden, beherrschen kann.	I'm certain I can master the skills being taught in this class.
Ich möchte in diesem Kurs gut abschneiden, weil es wichtig ist, meine Fähigkeiten meiner Familie, meinen Freunden, meinem Arbeitgeber oder anderen zu zeigen.	I want to do well in this class because it is important to show my ability to my family, friends, employer, or others.
Wenn ich die Schwierigkeit dieses Kurses, den Dozenten/ die Dozentin und meine Fähigkeiten berücksichtige, denke ich, dass ich in diesem Kurs gut abschneiden werde.	Considering the difficulty of this course, the teacher, and my skills, I think I will do well in this class.

Tabelle 12: Übersetzung Motivated Strategies for Learning Questionnaire

Erklärung zur Urheberschaft

Die vorgelegten Druckexemplare sowie die vorgelegte digitale Version der Arbeit sind identisch. Ich habe die Arbeit selbständig verfasst, keine anderen als die angegebenen Quellen und Hilfsmittel benutzt und die Arbeit nicht bereits an einer anderen Hochschule zur Erlangung eines akademischen Grades eingereicht. Von den zu § 26 Abs. 6 der Prüfungsordnung vorgesehenen Rechtsfolgen habe ich Kenntnis.