

Bachelorarbeit
im Bachelorstudiengang
Wirtschaftspsychologie
an der Hochschule für angewandte Wissenschaften Neu-Ulm

**Erste Hilfe durch KI: Der Einfluss KI-basierter emotionaler Unterstützung auf
psychologisches Wohlbefinden**

Erstkorrektor/-in: Prof. Dr. Michaela Eßbach
Betreuer/-in: *Prof. Dr. Barbara Dannenmann*

Verfasser/-in: Maximilian Loibl (Matrikel-Nr.: 328728)

Thema erhalten: 09.02.2026
Arbeit abgegeben: 08.06.2026

Abstract

Die vorliegende Studie untersuchte, ob eine siebentägige emotional unterstützende Nutzung eines KI-Chatbots das subjektive Wohlbefinden verbessert. In einem randomisierten 2×2-Mixed-Design mit zwei Messzeitpunkten (T1, T2) wurden 52 Teilnehmende einer aktiven Nutzungsgruppe mit emotional unterstützender Nutzung und einer Kontrollgruppe mit sachlicher Nutzung zugewiesen. Für das Wohlbefinden zeigte sich ein signifikanter Interaktionseffekt Zeit x Gruppe. Die aktive Gruppe steigerte ihren WHO-5-Wert deutlich, während die Kontrollgruppe nahezu unverändert blieb. Hypothesen zu Responsivität, Empathie und Nutzungsdauer wurden im konfirmatorischen Test nicht bestätigt. Explorative Analysen deuten jedoch darauf hin, dass Veränderungen in Empathie- und Responsivitätswahrnehmung mit Wohlbefindenszuwachsen zusammenhängen. Die Ergebnisse sprechen dafür, dass emotional unterstützend ausgerichtete KI-Interaktion kurzfristig das Wohlbefinden fördern könnte.

Inhaltsverzeichnis

1	Hinführung	3
2	Theoretischer und empirischer Hintergrund.....	4
2.1	Parasoziale Interaktion und parasoziale Beziehung	4
2.2	Parasoziale Interaktion mit KI-Chatbots	5
2.3	Need Fulfillment als Wirkmechanismus	7
2.4	Psychologisches Wohlbefinden als Konstrukt.....	8
2.5	Empirische Befundlage, Forschungslücke und Hypothesen	9
3	Methode.....	12
3.1	Teilnehmende	12
3.2	Studiendesign	13
3.3	Stimulus.....	15
3.4	Ablauf.....	16
4	Ergebnisse	18
4.1	Voraussetzungsprüfung und Stichprobenmerkmale	18
4.2	Wohlbefinden im Gruppenvergleich über die Zeit.....	19
4.3	Wahrgenommene Responsivität und Empathie im Verlauf.....	20
4.4	Nutzungsdauer und Wohlbefindensveränderung.....	21
4.5	Explorative Analysen	21
4.5.1	Nutzungszweck	21
4.5.2	Veränderung der Wahrnehmungsscores als Prädiktoren	22
5	Diskussion	24
5.1	Interpretation der Ergebnisse.....	24
5.2	Limitationen	27
5.3	Zukunftsperspektiven	28
5.4	Praktische Implikationen.....	29
5.5	Fazit.....	29
	Literaturverzeichnis.....	30
	Anhang Fragebogen 1 und 2	36
	Werkzeuge und Hilfsmittel.....	37

1 Hinführung

Psychische Krisen stellen weltweit ein bedeutendes gesellschaftliches Problem dar und betreffen zunehmend auch junge Menschen. Besonders die Generation Z ist durch steigende psychische Belastungen geprägt, die teilweise auf intensiven Medienkonsum, sozialen Vergleich und die Nutzung digitaler Technologien zurückgeführt werden (Coe et al., 2023). Gleichzeitig ist der Zugang zu professioneller psychologischer Unterstützung stark eingeschränkt. Mehr als ein Viertel der Erwachsenen erfüllt innerhalb eines Jahres die Kriterien einer psychischen Erkrankung (DGPPN, 2025), während auf der Angebotsseite ein erheblicher Mangel besteht: Schätzungen zufolge fehlen bundesweit rund 7.000 psychotherapeutische Kassensitze (BPtK, 2018). Daraus resultieren lange Wartezeiten, durchschnittlich sechs Wochen bis zum ersten Gespräch und weitere rund 20 Wochen bis zum Therapiebeginn, die eine frühzeitige Intervention strukturell erschweren.

Angesichts dieser Versorgungslücke greifen insbesondere jüngere Menschen zunehmend auf digitale Alternativen zurück. KI-basierte Systeme, insbesondere Large Language Models (LLMs), gewinnen dabei als potenziell niedrigschwellige Angebote für emotionale Unterstützung an Bedeutung, da sie anonym, jederzeit verfügbar und weniger stigmatisiert zugänglich sind als professionelle Hilfsangebote (Naslund et al., 2016). Ob solche Systeme tatsächlich messbare Effekte auf das psychologische Wohlbefinden haben können, ist jedoch empirisch bislang kaum untersucht, insbesondere fehlen kontrollierte Studien, die den Nutzungsmodus, also emotionale unterstützende versus sachliche Interaktion, systematisch variieren und dabei spezifische Wahrnehmungskomponenten wie Empathie und Responsivität als mögliche Wirkmechanismen berücksichtigen.

Die vorliegende Studie setzt an dieser Lücke an. Auf Basis der Self-Determination Theory (Ryan & Deci, 2000), die emotionale Eingebundenheit, sogenannte Relatedness, als zentrales psychologisches Bedürfnis für Wohlbefinden beschreibt, sowie bestehender Forschung zu parasozialer Interaktion und Beziehung in der Mensch-KI-Interaktion wird untersucht, ob eine siebentägige emotional unterstützende Nutzung eines KI-Chatbots mit einer Verbesserung des subjektiven Wohlbefindens assoziiert ist und welche Rolle die wahrgenommene Empathie und Responsivität des Systems dabei spielen könnten. Ziel ist es, einen ersten empirischen Beitrag zu einem Forschungsfeld mit wachsender gesellschaftlicher Relevanz zu leisten und dabei sowohl theoretische als auch praktische Implikationen für den Einsatz von KI in der psychologischen Grundversorgung abzuleiten.

2 Theoretischer und empirischer Hintergrund

2.1 Parasoziale Interaktion und parasoziale Beziehung

Ein zentrales Phänomen im Kontext der Mensch-KI-Interaktion ist, dass Nutzende KI-Chatbots unter bestimmten Bedingungen nicht als bloße Werkzeuge, sondern als soziale Gegenüber wahrnehmen können. Die parasoziale Interaktionstheorie liefert den theoretischen Rahmen dafür, wie und unter welchen Bedingungen eine solche soziale Wahrnehmung gegenüber einem medial vermittelten Akteur entstehen kann, und bildet damit die konzeptuelle Grundlage der vorliegenden Studie. Der Begriff der parasozialen Interaktion geht auf (Horton & Richard Wohl, 1956) zurück, die ein Phänomen beschrieben, das mit der Verbreitung von Radio und Fernsehen zunehmend an Bedeutung gewann: die scheinbar persönliche, einseitige Beziehung zwischen Medienrezipienten und medial vermittelten Personae. Der Begriff beschreibt die Illusion einer echten Beziehung, bei der Zuschauer den Eindruck gewinnen, an einem persönlichen, reziproken Austausch mitzuwirken, obwohl die Interaktion tatsächlich objektiv betrachtet einseitig bleibt (Horton & Richard Wohl, 1956). Diese Grundannahme ist deswegen interessant, weil sie genauso auf KI-Chatbots übertragbar ist. Ob ein KI-Chatbot emotional unterstützend wirken kann, hängt nicht davon ab, ob das System tatsächlich sozial ist, sondern davon, ob die Nutzenden die Interaktion als sozial erleben. Hartmann und Goldhoorn (2011) präzisierten diesen Gedanken durch eine prozessorientierte Perspektive und definierten parasoziale Interaktion als das situative Erleben gefühlter Gegenseitigkeit mit einem Medienakteur, also als das gefühlte Bewusstsein gegenseitiger Aufmerksamkeit und Anpassung. Diese Verschiebung hin zu einem episodischen, kontextabhängigen Phänomen war für die weitere Forschung prägend. Als begünstigende Faktoren gelten dabei die direkte verbale Ansprache durch den Akteur, dessen wahrgenommene Attraktivität sowie die Fähigkeit zur Perspektivenübernahme aufseiten des Rezipienten (Hartmann & Goldhoorn, 2011). Der in der vorliegenden Studie eingesetzte Stimulus-Prompt wurde gezielt so gestaltet, dass er direkte Ansprache, emotionale Spiegelung und Perspektivenübernahme aktiviert, also genau jene Merkmale, die nach Hartmann und Goldhoorn parasoziale Interaktion begünstigen.

Aus diesem situativen Erleben kann sich mit wiederholter Mediennutzung eine stabilere parasoziale Beziehung entwickeln, die über die einzelne Nutzungsepisode hinaus Bestand hat. Liebers und Schramm (2019) differenzieren in ihrem systematischen Inventar von 60 Jahren Forschung entsprechend zwischen parasozialer Interaktion als zeitlich begrenztem Erleben und parasozialer Beziehung als dauerhafter emotionaler Bindung, die durch wiederholte

Begegnungen gestärkt wird. Die vorliegende Studie fokussiert sich aufgrund des siebentägigen Studiendesigns daher auf die parasoziale Interaktion zwischen Nutzenden und dem KI-System. Tukachinsky und Stever (2019) ergänzten diese Perspektive durch einen entwicklungsorientierten Ansatz, der den graduellen Übergang von ersten flüchtigen Eindrücken hin zu stabilen, affektiv getragenen Bindungen abbildet. In einem aktuellen Überblick über 281 Studien bestätigten Schramm et al. (2024), dass sich die Forschung zu parasozialen Phänomenen weiterhin stark wächst und sich zunehmend von klassischen Massenmedien hin zu interaktiven und digitalen Kontexten verschiebt, eine Entwicklung, in die sich die vorliegende Studie, mit einer interaktiven KI-Nutzung über sieben Tage hinweg, einreihet.

Hinsichtlich der Rolle der Nutzungsdauer ist festzuhalten, dass Horton und Wohl (1956) die wiederholte Exposition bereits im ursprünglichen Modell eher als rahmenden Hintergrundfaktor verstanden, nicht als eigenständige Ursache parasozialer Bindungen. Ob die Nutzungsdauer einen Einfluss auf die KI-Chatbot Interaktion haben kann, wird im Folgenden noch untersucht.

2.2 Parasoziale Interaktion mit KI-Chatbots

Aufbauend auf den Grundlagen parasozialer Interaktion stellt sich für die vorliegende Studie die konkrete Frage, inwieweit dieses Konzept auf KI-Chatbots übertragbar ist. Das klassische Parasoziale-Interaktions-Modell ging von einer einseitigen Beziehung zu einer Persona aus, die nicht erwidert werden kann (Horton & Richard Wohl, 1956). KI-Chatbots, insbesondere Large Language Model-basierte Systeme, erzeugen eine grundlegend andere Interaktionsdynamik. Sie antworten, passen sich an das Gegenüber an und simulieren durch natürliche Sprache, personalisierte Antworten und menschenähnliche Konversationsstile eine Form gegenseitiger Interaktion, die über die reine Rezeption eines Fernsehakteurs hinausgeht (Abercrombie et al., 2023). Entscheidend ist dabei die Nutzerperspektive: Nicht die Maschine ist sozial, sondern die Interaktion wird von den Nutzenden als sozial interpretiert (Abercrombie et al., 2023). Wie die Interaktion von den Nutzenden wahrgenommen wird ist am Ende also entscheidend. Damit kommt die Frage auf, ob der parasoziale Rahmen auf KI-Chatbots übertragbar ist oder ob die Bidirektionalität dieser Systeme eine qualitativ neue Interaktions- und Beziehungsform begründet. Diese Frage wird in der aktuellen Forschung kontrovers diskutiert. Eine Position versteht KI-Chatbots als Erweiterung parasozialer Interaktion: Die grundlegende Asymmetrie bleibt bestehen, da der Chatbot kein eigenständiges Bewusstsein besitzt, doch die interaktiven Fähigkeiten des Systems verstärken die Illusion von Gegenseitigkeit (Lu et al., 2026). Die

andere Position argumentiert, dass die Bidirektionalität eine neue, hybride Interaktions- und möglicherweise Beziehungsform begründet, die sich qualitativ von klassischer parasozialer Beziehung unterscheidet (Pentina et al., 2023). Die vorliegende Arbeit folgt der ersten Lesart, räumt jedoch ein, dass diese Frage empirisch nicht abschließend geklärt ist. Eine weiterführende Einordnung erfolgt im Diskussionsteil. Neben diesen Merkmalen zeichnen sich KI-Chatbots durch besondere Eigenschaften aus, die sie von menschlichen Interaktionspartnern unterscheiden: permanente Verfügbarkeit, Konsistenz im Antwortverhalten und das Fehlen sozialer Ermüdung (Skjuve et al., 2021). Die Forschung deutet darauf hin, dass bereits feine linguistische und gestalterische Merkmale ausreichen könnten, um Dialogsysteme als anthropomorphisierte, soziale Akteure wahrzunehmen. Der Gebrauch personaler Pronomen, paralinguistische Signale und der Einsatz von Empathie-Sprache scheinen diesen Effekt zu begünstigen (Abercrombie et al., 2023). Die menschenähnliche Wahrnehmung von KI-Chatbots hängt demnach möglicherweise weniger von den faktischen Fähigkeiten des Systems ab als vielmehr von den Erwartungen, Bedürfnissen und Interaktionsmustern der Nutzenden. Daraus folgt, dass nicht die objektiven Eigenschaften des Systems im Vordergrund stehen, sondern die subjektive Wahrnehmung des Erlebten. Empirische Hinweise auf die Übertragbarkeit interpersonaler Interaktions- und Beziehungsdynamiken auf Mensch-Technik-Interaktionen liefern mehrere Studien. So deuten Befunde darauf hin, dass sowohl Interaktionsdauer als auch -intensität mit einem Chatbot die gefühlte soziale Verbundenheit positiv vorhersagen könnten, wobei die wahrgenommene soziale Präsenz des Systems als vermittelnde Variable fungiert (Christoforakos et al., 2021). Ergänzend legen Pentina et al. (2023) nahe, dass Anthropomorphismus und wahrgenommene Authentizität des Chatbots über soziale Interaktion zu emotionaler Bindung führen könnten, wobei weniger die bloße Nutzungshäufigkeit als vielmehr die wahrgenommene Interaktionsqualität die Beziehungsbildung voranzutreiben scheint. Das bedeutet, dass die Interaktionsqualität maßgeblichen Einfluss auf die subjektive Wahrnehmung der Nutzenden auf das System, den KI-Chatbot haben könnte.

Weitere Studien deuten darauf hin, dass wahrgenommene Interaktivität, operationalisiert über Vertrauenswürdigkeit, Menschlichkeit und Responsivität, die parasoziale Beziehung zwischen Nutzer und System verstärken könnte (Lu et al., 2026). Diese Überlegungen sind insofern für diese Arbeit von großer Bedeutung, da sie maßgeblich in die Gestaltung des Stimulus-Prompts der aktiven Nutzungsgruppe (AU) zur emotionalen Unterstützung einfließen. Im Kontext von Mental-Health-Anwendungen wird betont, dass diese Eigenschaften besonders für junge Menschen attraktiv sein könnten, die aufgrund von Stigmatisierung oder Zugangshürden keine traditionelle Unterstützung suchen, da Niedrigschwelligkeit und Anonymität die

Hemmschwelle für emotionale Selbstoffenbarung senken könnten (Grové, 2021). Aus diesem Grund fokussiert sich die Rekrutierung der Stichprobe besonders auf junge Erwachsene, da hier der Effekt durch niedrigschwellige emotionale Unterstützung durch KI-Chatbots am größten sein könnte. Gleichzeitig wird auf die Risiken einer übermäßigen Anthropomorphisierung hingewiesen, die zu unrealistischen Erwartungen an die Fähigkeiten des Systems führen können (Abercrombie et al., 2023).

Die Qualität der KI-Interaktion wird in der vorliegenden Arbeit über zwei zentrale Konstrukte operationalisiert: wahrgenommene Empathie und wahrgenommene Responsivität. Dabei handelt es sich nicht um objektive Eigenschaften des Systems, sondern um subjektive Zuschreibungen der Nutzenden, die durch sprachliche und interaktive Signale ausgelöst werden. Zur Messung wahrgenommener Empathie technischer Systeme wurde die Perceived Empathy of Technology Scale (PETS) entwickelt und validiert (Schmidmaier et al., 2024), die in zwei Subskalen, Emotionale Resonanz sowie Verständnis und Vertrauen, erfasst, inwieweit ein System als empathisch erlebt wird. Für die wahrgenommene Responsivität liegt mit der Perceived Responsiveness and Insensitivity Scale (PRI) ein psychometrisch validiertes Maß vor, das erfasst, inwieweit Nutzende sich von ihrem Gegenüber verstanden, wertgeschätzt und fürsorglich behandelt fühlen (Crasta et al., 2021). Befunde einer Tagebuchstudie über vier Wochen deuten darauf hin, dass Nutzende mit einem Mental-Health-Chatbot eine Form digitaler therapeutischer Allianz entwickeln können, bei der wahrgenommene Empathie und Responsivität als zentrale Prädiktoren der Interaktions- und Beziehungsqualität fungieren könnten (Xu et al., 2025). Diese Wahrnehmungskomponenten bilden damit die konzeptuelle Grundlage für die Hypothesen H2 und H3 der vorliegenden Studie, die untersuchen, ob Veränderungen in der erlebten Interaktionsqualität mit dem KI-System, operationalisiert über wahrgenommene Empathie und Responsivität, mit Veränderungen im subjektiven Wohlbefinden zusammenhängen.

2.3 Need Fulfillment als Wirkmechanismus

KI-Chatbots können also unter bestimmten Bedingungen als soziale Gegenüber wahrgenommen werden, dabei ist die subjektiv erlebte Interaktionsqualität entscheidend. Genau diese subjektive Wahrnehmung könnte dazu beitragen, dass die Interaktion mit einem KI-Chatbot grundlegende psychologische Bedürfnisse anspricht. Bedürfnisse, die sich durch die Self-Determination Theory (SDT; Ryan & Deci, 2000) genauer einordnen lassen. Die SDT geht von drei grundlegenden psychologischen Bedürfnissen aus: Autonomie, Kompetenz und Relatedness. Ihre Befriedigung gilt als zentral für psychisches Wachstum, Motivation und

Wohlbefinden (Ryan & Deci, 2000). Das bedeutet also, dass die in der Studie angestrebte KI-Interaktion der aktiven Nutzungsgruppe (AU) durch Bedürfnisbefriedigung als Wirkmechanismus Einfluss auf das subjektive psychische Wohlbefinden haben kann. Groß angelegte länderübergreifende Studien deuten darauf hin, dass die Befriedigung dieser Bedürfnisse in engem Zusammenhang mit Lebenszufriedenheit, Glück und Sinnhaftigkeit steht und depressive Symptome vorhersagt. Ein Zusammenhang, der sich in allen untersuchten Ländern als robuster erwies als der Einfluss des sozioökonomischen Status (Martela et al., 2023). Für die vorliegende Studie ist insbesondere das Bedürfnis nach Relatedness relevant, also das Erleben sozialer Eingebundenheit, wechselseitiger Fürsorge und Zugehörigkeit (Fröbel & Kenning, 2025; Ryan & Deci, 2000). Im Kontext der Chatbot-Nutzung ist damit jedoch nicht eine stabile zwischenmenschliche Beziehung im klassischen Sinne gemeint, sondern ein subjektives Erleben von Verstandenwerden und Zugewandtheit, das durch die Interaktion mit dem System ausgelöst werden kann (Christoforakos et al., 2021). Dass dieses Erleben auch in der Mensch-Technik-Interaktion entstehen kann, wurde durch die Parasoziale-Interaktions-Theorie bereits dargelegt.

Für den Wirkzusammenhang zwischen Chatbot-Interaktion und Wohlbefinden ist die Unterscheidung zwischen emotional unterstützender und sachlicher Nutzung zentral. Ein Chatbot, der vor allem sachliche Informationen bereitstellt oder konkrete Aufgaben unterstützt, spricht andere Bedürfnisse an als ein System, das empathisch reagiert, Verständnis signalisiert und emotionale Unterstützung anbietet. Empirische Hinweise deuten darauf hin, dass Chatbots insbesondere dann als Gesprächspartner, präferiert werden könnten, wenn sie als einfühlsam und responsiv erlebt werden (Fröbel & Kenning, 2025). Eine emotional unterstützende Chatbot-Interaktion könnte damit das Gefühl erzeugen, gesehen und verstanden zu werden, und auf diese Weise Relatedness ansprechen (Christoforakos et al., 2021). Diese Bedürfnisbefriedigung wiederum könnte als Mechanismus auf Wohlbefinden wirken, insbesondere dann, wenn die Interaktion nicht als sachlicher Informationsaustausch, sondern im Kontext emotional unterstützender Nutzung wahrgenommen wird.

2.4 Psychologisches Wohlbefinden als Konstrukt

Da die SDT Wohlbefinden als zentrales Outcome von Bedürfnisbefriedigung postuliert, stellt subjektives Wohlbefinden die naheliegende abhängige Variable der vorliegenden Studie dar. Wohlbefinden hat den entscheidenden Vorteil, dass es positive Veränderungen im Erleben auch unterhalb klinischer Schwellenwerte erfasst, was für eine siebentägige Interventionsstudie mit einer gesunden Allgemeinbevölkerung besonders geeignet erscheint. In der psychologischen

Wohlbefindensforschung wird unterschieden zwischen einer hedonischen Tradition, die Wohlbefinden als positive Affekte, geringe negative Affekte und kognitive Lebenszufriedenheit definiert, und einer eudaimonischen Tradition, die es als Verwirklichung des eigenen Potenzials versteht (Ryan & Deci, 2001). Beide korrelieren, unterscheiden sich jedoch an entscheidenden Punkten. Für die vorliegende Studie steht das hedonisch geprägte, subjektive Wohlbefinden im Vordergrund, das sich konzeptuell von psychischer Belastung oder Symptomreduktion abgrenzen lässt. Wohlbefinden gilt als eigenständiges, positives Phänomen und nicht lediglich als Abwesenheit von Psychopathologie (Ryan & Deci, 2001). Diese Perspektive wird durch das Mental-Health-Continuum-Modell gestützt, das psychische Gesundheit als Kontinuum von Languishing über moderate Gesundheit bis hin zu Flourishing abbildet (Keyes, 2002). In der vorliegenden Arbeit steht daher nicht die Symptomreduktion im Fokus, sondern die positive Veränderung des subjektiven Wohlbefindens.

Zur Erfassung dieses Konstrukts hat sich der WHO-5 Well-Being Index als ein verbreitetes Instrument etabliert. Eine systematische Review über 213 Studien belegt die Validität, Reliabilität und breite Anwendbarkeit des Instruments (Topp et al., 2015). Der WHO-5 erfasst mit fünf positiv formulierten Items das subjektive Wohlbefinden der vergangenen zwei Wochen auf einer eindimensionalen 0-100-Skala. Die deutsche Version weist ebenfalls gute psychometrische Eigenschaften auf (Brähler et al., 2007). Der Index eignet sich für das Design der vorliegenden Studie, weil er ausschließlich positives Erleben abbildet und keine klinischen Symptomitems enthält und damit sensitiv für jene kurzfristigen positiven Veränderungen ist, die durch emotional unterstützende KI-Interaktion angestoßen werden könnten.

2.5 Empirische Befundlage, Forschungslücke und Hypothesen

Empirische Befunde deuten darauf hin, dass KI-Chatbots unter bestimmten Bedingungen eine beziehungsähnliche Qualität erreichen können, die an interpersonale Beziehungsdynamiken anknüpft und maßgeblich von der Interaktionsqualität geprägt ist. Um die Interaktionsqualität von KI-Systemen zu verstehen, ist es daher notwendig, einen Blick auf die Beziehungsbildung zu werfen, auch wenn dies in der eigentlichen Studie nicht beleuchtet wird, sondern lediglich auf die Interaktion selbst geschaut wird. Längsschnittliche Studien legen nahe, dass sich zwischen Nutzenden und KI-Begleitern stabile emotionale Bindungen entwickeln könnten, die durch Vertrauen, persönliche Selbstoffenbarung und die Wahrnehmung reziproker Empathie gekennzeichnet sind (Skjuve et al., 2021, 2022). Kontrollierte experimentelle Befunde deuten zudem darauf hin, dass bereits singuläre empathische Chatbot-Interaktionen kurzfristige Stimmungsverbesserungen bewirken können: In einem sozialen Exklusionsparadigma zeigte

sich, dass eine empathische Chatbot-Interaktion nach sozialer Zurückweisung die Stimmung signifikant verbesserte (De Gennaro et al., 2020). Auf einer breiteren Evidenzebene berichten Metaanalysen positive Effekte von KI-Chatbots auf psychische Belastung und verwandte Wohlbefindensindikatoren, insbesondere bei der Reduktion von Stress und leichten depressiven Symptomen (Feng et al., 2025). Darüber hinaus deuten Befunde darauf hin, dass Nutzende mit einem Chatbot eine therapeutische Allianz entwickeln können, die durch Aufgabenkonsens und emotionale Bindung gekennzeichnet ist (He et al., 2022), und dass wahrgenommene Empathie und Responsivität dabei als zentrale Prädiktoren der Interaktions- und Beziehungsqualität fungieren könnten (Xu et al., 2025).

Die bisherige Forschung weist jedoch mehrere konzeptuelle und methodische Lücken auf, die das Verständnis der Wirkmechanismen zwischen Chatbot-Interaktion und Wohlbefinden einschränken. Erstens ist die empirische Befundlage durch kurze Interventionszeiträume und das Fehlen dichter Längsschnittdaten gekennzeichnet, sodass die Dynamik der Wohlbefindensveränderung über mehrere Tage bislang kaum abgebildet wird. Als Folge dessen wird ein Längsschnittdesign eingesetzt, das Wohlbefindensveränderungen über einen Zeitraum von sieben Tagen abbildet. Zweitens wurde der Nutzungskontext in der bisherigen Forschung kaum experimentell manipuliert. Die meisten Studien untersuchen Chatbot-Interaktionen unter naturalistischen Bedingungen, ohne systematisch zwischen verschiedenen Nutzungsmodi zu variieren, sodass offen bleibt, ob emotional unterstützende versus sachliche Nutzung einen tatsächlichen Einfluss auf das Wohlbefinden haben kann. Drittens wurden spezifische Maße der Interaktionsqualität wie wahrgenommene Empathie und Responsivität bislang nicht mit einem mehrwelligen Design kombiniert, das sowohl die Verlaufsdynamik des Wohlbefindens als auch die Veränderung der Interaktionswahrnehmung abbildet. Besonders unklar ist die Studienlage in Bezug auf die Faktoren, die Interaktionsqualität und Wohlbefinden beeinflussen. Daher wird wahrgenommene Empathie und Responsivität als Maße der Interaktionsqualität längsschnittlich erfasst und in Bezug auf die Wohlbefindensveränderung gesetzt.

Aus der theoretischen Herleitung und der identifizierten Forschungslücke ergeben sich die folgenden Hypothesen:

H1: Die emotional unterstützende Nutzung eines KI-Chatbots führt zu einer signifikant stärkeren Zunahme des subjektiven Wohlbefindens (WHO-5) von Messzeitpunkt (T1) vor der KI Nutzung, zu Messzeitpunkt (T2) an Tag 7 nach 7-tägiger KI-Nutzung, im Vergleich zur Kontrollgruppe mit sachlicher Nutzung des KI-Chatbots.

H2: Höhere wahrgenommene Responsivität der KI zeigt eine signifikant positivere Veränderung des subjektiven Wohlbefindens (Delta WHO-5) nach 7 Tagen in der aktiven Nutzungsgruppe (AU), im Vergleich zur Kontrollgruppe.

H3: Höhere wahrgenommene Empathie der KI zeigt eine signifikant positivere Veränderung des subjektiven Wohlbefindens (Delta WHO-5) nach 7 Tagen in der aktiven Nutzungsgruppe (AU), im Vergleich zur Kontrollgruppe.

H4: Längere durchschnittliche Nutzungsdauer (Minuten/Tag) korreliert signifikant positiv mit stärkeren Wohlbefindenszuwächsen (Delta WHO-5) in der aktiven Nutzungsgruppe (AU).

Ziel der vorliegenden Studie ist es, die theoretisch hergeleitete Wirkkette von wahrgenommener Interaktionsqualität über Relatedness-Befriedigung bis zur Veränderung des subjektiven Wohlbefindens empirisch zu prüfen.

3 Methode

3.1 Teilnehmende

Die vorliegende Studie wurde als siebentägige longitudinale Studie mit randomisierter Zuteilung in eine aktive Nutzungsgruppe (AU = active user group) und eine Kontrollgruppe (BU = baseline user group) durchgeführt. Insgesamt wurden 62 Personen in die Analyse einbezogen. Die Stichprobe setzte sich aus 54,8% männlichen ($n = 34$) und 45,2% weiblichen ($n = 28$) Teilnehmenden zusammen. Hinsichtlich der Altersverteilung gehörte die Mehrheit der Befragten der Gruppe der 18- bis 29-Jährigen an (77,4%, $n = 48$), gefolgt von der Gruppe der 30- bis 49-Jährigen (21,0%, $n = 13$) sowie einer Person im Alter von 50 bis 59 Jahren (1,6%).

Die Rekrutierung erfolgte über persönliche Netzwerke, universitäre E-Mail-Verteiler der Hochschule Neu-Ulm sowie über das SONA-System, in dem die Teilnahme mit zwei Versuchspersonenstunden vergütet wurde. Zusätzlich bestand die Möglichkeit, an einem Gewinnspiel mit einem Aufwandsentgelt von 100 Euro teilzunehmen. Die Rekrutierung fokussierte sich dabei bewusst auf junge Erwachsene, da diese Zielgruppe nach aktuellem Forschungsstand besonders empfänglich für niedrigschwellige KI-gestützte Angebote sein dürfte und gleichzeitig häufig mit Zugangsbarrieren zur professionellen psychologischen Versorgung konfrontiert ist (Grové, 2021). Teilnahmeberechtigt waren Personen ab 18 Jahren mit ausreichenden Deutschkenntnissen. Eine a priori Poweranalyse mit G*Power ergab bei einer mittleren Effektgröße von $f = 0.30$, $\alpha = .05$ und einer Teststärke von $1 - \beta = .80$ ein erforderliches Gesamtstichprobenmaß von $N = 90$. Die tatsächliche Stichprobengröße blieb unter diesem Wert, was die Studie mit begrenzter statistischer Power für kleine Effekte kennzeichnet.

Die Zuweisung der Teilnehmenden erfolgte randomisiert, um systematische Gruppenunterschiede zu minimieren und die interne Validität der Studie zu gewährleisten. Zum ersten Messzeitpunkt (T1) wurden die Teilnehmenden der aktiven Nutzungsgruppe (AU, $n = 28$) oder der Kontrollgruppe (BU, $n = 34$) zugeteilt. Zum zweiten Messzeitpunkt (T2), sieben Tage später, lagen vollständige Datensätze von 52 Personen vor (AU $n = 24$, BU $n = 28$), was einem Dropout von 16,1% ($n = 10$) entspricht, der überwiegend auf die fehlende Teilnahme an der Folgebefragung zurückzuführen ist. Basierend auf den Rohdaten der Erhebungsplattform betrug die Rücklaufquote an T1 53,7% (66 von 123 gestarteten Befragungen) sowie an T2 74,6% (53 von 71 gestarteten Befragungen). Von der Analyse ausgeschlossen wurden $n = 4$ Datensätze an T1 aufgrund nicht bestandener Aufmerksamkeitschecks und zu kurzer

Nutzungsdauer von unter 5 Minuten. Alle Teilnehmenden gaben vor Beginn der Erhebung ihre informierte Einwilligung zur Teilnahme und Datenverarbeitung; darüber hinaus willigten sie in die Erhebung ihrer E-Mail-Adresse ein, über die ihnen nach sieben Tagen die Folgebefragung zugestellt wurde.

3.2 Studiendesign

Die vorliegende Untersuchung orientiert sich in ihrem Design an Chandra et al. (2025), die ein vergleichbares Design über fünf Wochen realisierten. Die Interventionsdauer wurde auf sieben Tage verkürzt, um zu prüfen, ob bereits in einem kürzeren Zeitraum messbare Effekte auf das subjektive Wohlbefinden nachweisbar sind. Die statistische Auswertung erfolgte mithilfe von jamovi und R.

Die Studie folgt einem 2x2 Mixed Design mit Messwiederholung. Das Design wurde gewählt, da die Kombination eines Zwischen- und Innersubjektfaktors die Kontrolle interindividueller Unterschiede ermöglicht und dadurch die statistische Teststärke gegenüber einem reinen Gruppenvergleich erhöht. Als Innersubjektfaktor wurde der Messzeitpunkt mit zwei Stufen (T1, T2) berücksichtigt, als Zwischensubjektfaktor die Gruppenzugehörigkeit mit zwei Stufen, aktive Nutzungsgruppe (AU) und Kontrollgruppe (BU). Die Zuweisung erfolgte randomisiert über die automatisierte Filterfunktion von SoSci Survey. Zur Prüfung der Hypothesen wurden Mixed ANOVAs sowie explorative Regressionsanalysen berechnet.

Die zentrale abhängige Variable war das subjektive Wohlbefinden, erfasst mithilfe des WHO-5 Well-Being Index (Sischka et al., 2020), einer validierten deutschen Übersetzung des Originals der WHO (1998). Der WHO-5 wurde gewählt, da er im deutschsprachigen Raum als psychometrisch robuster Indikator für subjektives Wohlbefinden gilt (Brähler et al., 2007) und in vergleichbaren KI-gestützten Interventionsstudien als Outcome-Maß zur Längsschnittverlaufsmessung eingesetzt wurde (Vetsberger et al., 2022). Das Instrument besteht aus fünf Items auf einer sechsstufigen Likert-Skala (0 = *zu keinem Zeitpunkt*, 5 = *die ganze Zeit*). Ein Beispielitem lautet: „*In den letzten 7 Tagen war ich froh und guter Laune*“. Die Skala weist eine hohe interne Konsistenz auf ($\alpha = .92$; Brähler et al., 2007). Der Skalenwert ergibt sich aus der Summe der fünf Items, multipliziert mit dem Faktor vier, wodurch ein Wertebereich von 0 bis 100 entsteht. Abweichend von der Standardversion, die das Wohlbefinden der vergangenen zwei Wochen erfasst, wurde der Referenzzeitraum auf eine Woche angepasst, um sicherzustellen, dass die Referenzperioden beider Messzeitpunkte den Interventionszeitraum nicht überlappen. Diese Anpassung stellt eine Abweichung von der

validierten Standardversion dar, die bei der Interpretation der Befunde zu berücksichtigen ist. Zur Kontrolle von Ausgangsunterschieden im Wohlbefinden zwischen den Gruppen wurde der WHO-5-Baselinewert (T1) in einem separaten Gruppenvergleich geprüft.

Die wahrgenommene Responsivität der KI (H2) wurde anhand einer adaptierten Subskala der Perceived Responsiveness and Insensitivity Scale (PRI-R; Crasta et al., 2021) gemessen. Da eine validierte Skala zur Messung wahrgenommener Responsivität in Mensch-KI-Interaktionen zum Zeitpunkt der Erhebung nicht vorlag, wurden die Itemformulierungen von Bezügen auf zwischenmenschliche Partner auf KI-Systeme angepasst. Die Skala umfasst acht Items, die auf einer fünfstufigen Likert-Skala (1 = *Stimme gar nicht zu*, 5 = *Stimme voll zu*) beantwortet werden ($\alpha = .93$; Crasta et al., 2021). Die Items werden durch folgende Frage eingeleitet: „*Wenn du auf deine Interaktion mit der KI zurückblickst: Wie hast du das Antwortverhalten der KI wahrgenommen?*“ Ein Beispielitem lautet: „*Die KI ist verständnisvoll*“.

Die wahrgenommene Empathie der KI (H3) wurde mithilfe der validierten deutschen Version der Perceived Empathy of Technology Scale (PETS; Schmidmaier et al., 2024) erhoben, die als standardisiertes Instrument zur Erfassung empfundener Empathie in der Mensch-Computer-Interaktion entwickelt und validiert wurde (Schmidmaier et al., 2024). Die PETS umfasst zehn Items in zwei Subskalen, Emotionale Reaktionsfähigkeit (PETS-ER, $\alpha = .955$) und Verständnis und Vertrauen (PETS-UT, $\alpha = .853$), mit einer Gesamtreliabilität von $\alpha = .950$ (Schmidmaier et al., 2024). Die Items wurden mittels eines visuellen Schiebereglers auf einer Skala von 0 (*stimme überhaupt nicht zu*) bis 100 (*stimme voll und ganz zu*) beantwortet. Die Items werden durch folgende Frage eingeleitet: „*Inwieweit stimmst du den folgenden Aussagen über die KI als dein Interaktionspartner zu?*“ Ein Beispielitem lautet: „*Die KI wirkte emotional intelligent*“. Die Verwendung der PETS in vergleichbaren KI-Interventionsstudien wurde bereits durch Chandra et al. (2025) demonstriert, die das Instrument ebenfalls zur Messung wahrgenommener KI-Empathie im Längsschnittdesign einsetzten.

Ich werde dir gleich ein Thema nennen, das mich aktuell beschäftigt. Bestätige bitte kurz, dass du bereit bist, und frage mich dann, was mich heute beschäftigt.

In der Kontrollgruppe (BU) wurde die KI auf sachliche, informationsorientierte Unterstützung ohne Bezug auf emotionale Inhalte ausgerichtet.

Bitte agiere für die nächsten 10–15 Minuten als mein sachlicher Informations-Assistent. Deine Aufgabe ist es, mir bei praktischen Wissensfragen, Recherchen oder organisatorischen Aufgaben zu helfen. Antworte bitte rein informativ, sachlich und präzise. Vermeide es, auf emotionale Aspekte einzugehen oder Ratschläge zu meinem Befinden zu geben. Bestätige bitte kurz, dass du bereit bist, und frage mich dann, bei welcher sachlichen Aufgabe du mich heute unterstützen kannst.

Da die KI-Antworten nicht vollständig kontrollierbar waren, stellte der Prompt die einzige direkte Einflussmöglichkeit auf das Interaktionsverhalten des Systems dar.

3.4 Ablauf

Die Datenerhebung erfolgte vollständig online über SoSci Survey. An Tag 1 durchliefen die Teilnehmenden den ersten Fragebogen (T1), der nach einer kurzen Begrüßung und Studienübersicht die Datenschutzinformationen, die Einwilligung zur Datenverarbeitung sowie demografische Angaben umfasste. Es folgte die Erhebung des Baseline-Wohlbefindens mittels WHO-5, bevor die Teilnehmenden innerhalb derselben Session die erste KI-Interaktion auf Basis des gruppenspezifischen Stimulus-Prompts durchführten. Im Anschluss wurden PRI-R, PETS, Nutzungsdauer sowie Nutzungszweck erhoben. Für den Nutzungszweck standen 16 vordefinierte Kategorien zur Auswahl, die sowohl emotional unterstützende (z. B. Stress & Belastung, Beziehungen & Soziales) als auch sachliche Nutzungszwecke (z. B. Informationssuche, Wissensfragen & Recherche) umfassten. Zur Qualitätssicherung enthielt der Fragebogen einen Aufmerksamkeits-Check.

An den Tagen 2 bis 6 führten die Teilnehmenden die KI-Interaktion täglich eigenständig durch, ohne begleitende Befragung. Die aktive Nutzungsgruppe (AU) sollte die KI zur emotional unterstützenden Nutzung einsetzen, die Kontrollgruppe (BU) ausschließlich für sachliche Aufgaben. Beide Gruppen wurden angehalten, die KI täglich für 5 bis 10 Minuten zu nutzen. Die tägliche Anwendung des Stimulus-Prompts wurde empfohlen, konnte jedoch nicht kontrolliert werden.

An Tag 7 erhielten die Teilnehmenden einen automatisierten Erinnerungslink zum zweiten Fragebogen (T2), der nach der abschließenden KI-Interaktion ausgefüllt wurde und erneut WHO-5, PRI-R sowie PETS umfasste. Ergänzend wurde ein Manipulation-Check zum wahrgenommenen Kommunikationsstil der KI erhoben. Als Ausschlusskriterium galt, wenn Teilnehmende im Manipulation-Check ausschließlich „stimme nicht zu“ ankreuzten, was auf eine fehlende wahrgenommene Differenz im Kommunikationsstil hinweisen würde. Die Verknüpfung der Datensätze beider Messzeitpunkte erfolgte über einen pseudonymisierten Code auf Basis der E-Mail-Adresse.

4 Ergebnisse

4.1 Voraussetzungsprüfung und Stichprobenmerkmale

Zu Studienbeginn wiesen beide Gruppen ein vergleichbares mittleres Wohlbefinden auf, wobei die Kontrollgruppe (BU) geringfügig höhere Ausgangswerte zeigte als die aktive Nutzungsgruppe (AU: $M = 59.6$, $SD = 19.7$; BU: $M = 63.5$, $SD = 17.0$). Ein Gruppenvergleich zwischen aktiver Nutzungsgruppe (AU) und Kontrollgruppe (BU) zu T1 ergab keinen signifikanten Unterschied im Wohlbefinden ($t(60) = -0.85$, $p = .400$), was auf eine erfolgreiche Randomisierung hindeutet. Der Vergleich der Baselinewerte zwischen Personen mit und ohne vollständigen Datensatz zeigte ebenfalls keinen signifikanten Unterschied ($t(60) = -1.50$, $p = .138$), sodass der Dropout als zufällig betrachtet werden kann. Die WHO-5-Werte wurden für beide Gruppen vor der ersten KI-Interaktion als Baselinewert erhoben. Hinsichtlich der wahrgenommenen Empathie der KI unterschieden sich die Gruppen zu T1 signifikant voneinander ($t(60) = 2.78$, $p = .007$), die aktive Nutzungsgruppe (AU: $M = 64.6$, $SD = 17.9$) wies bereits zu T1 deskriptiv höhere Werte auf als die Kontrollgruppe (BU: $M = 52.3$, $SD = 14.3$), wobei diese Werte erst im Anschluss an die erste KI-Interaktion erhoben wurden. Die wahrgenommene Responsivität unterschied sich zu T1 nicht signifikant zwischen den Gruppen ($t(60) = 1.72$, $p = .091$), wies jedoch deskriptiv höhere Werte in der aktiven Nutzungsgruppe auf (AU: $M = 3.87$, $SD = 0.69$; BU: $M = 3.59$, $SD = 0.58$).

Die Voraussetzungen für die Mixed ANOVA wurden geprüft. Die WHO-5-Werte waren in beiden Gruppen zu T1 annähernd normalverteilt, geprüft mittels Shapiro-Wilk-Test (W) (AU: $W = .97$, $p = .528$; BU: $W = .94$, $p = .067$), was den Einsatz der Mixed ANOVA stützt. Zu T2 zeigte die aktive Nutzungsgruppe (AU) eine leichte Abweichung ($W = .87$, $p = .005$), die angesichts der bekannten Robustheit der Mixed ANOVA als tolerierbar eingestuft wurde.

Von der Ausgangsstichprobe ($N = 62$; AU $n = 28$, BU $n = 34$) lagen für die Mixed-ANOVA-Analysen vollständige Messdaten zu beiden Messzeitpunkten vor ($N = 52$; AU $n = 24$, BU $n = 28$). Der Dropout von $n = 10$ Personen wurde bereits in Abschnitt 3.1 beschrieben. Alle konfirmatorischen Hypothesentests (H1–H3) beruhen auf dieser Längsschnittsstichprobe. Die explorativen Regressionsanalysen in Abschnitt 4.5.2 wurden ausschließlich innerhalb der aktiven Nutzungsgruppe (AU, $n = 24$) berechnet.

4.2 Wohlbefinden im Gruppenvergleich über die Zeit

Das Wohlbefinden entwickelte sich in der aktiven Nutzungsgruppe (AU) und der Kontrollgruppe (BU) über die Zeit signifikant unterschiedlich. Der Interaktionseffekt Zeit x Gruppe fiel signifikant aus ($F(1, 50) = 26.50, p < .001$, partielles Eta-Quadrat = .35), was darauf hindeutet, dass eine wöchentliche emotional unterstützende KI-Nutzung mit einer stärkeren Wohlbefindensveränderung assoziiert ist als eine sachliche Nutzung. Während das Wohlbefinden in der aktiven Nutzungsgruppe (AU) von $M = 59.6$ auf $M = 83.0$ Punkte anstieg, blieb es in der Kontrollgruppe (BU) nahezu unverändert (T1: $M = 63.5$, T2: $M = 65.0$). Ergänzend zeigte sich ein signifikanter Haupteffekt der Zeit ($F(1, 50) = 36.00, p < .001$, partielles Eta-Quadrat = .42), während der Haupteffekt der Gruppe das Signifikanzniveau verfehlte ($F(1, 50) = 2.42, p = .126$, partielles Eta-Quadrat = .05).

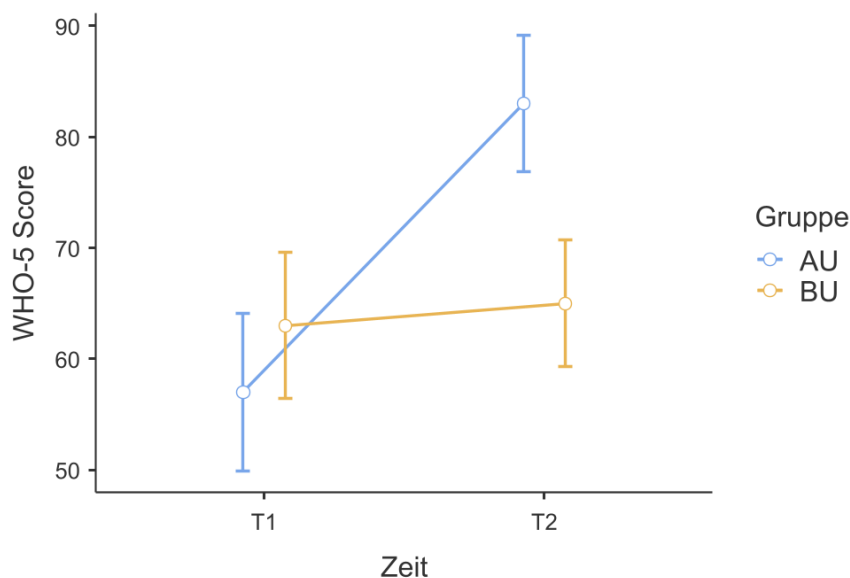


Abbildung 2. Mittlere WHO-5-Verläufe der aktiven Nutzungsgruppe (AU) und der Kontrollgruppe (BU) zu T1 und T2.

Post-hoc-Vergleiche mit Tukey-Korrektur zeigten, dass das Wohlbefinden in der aktiven Nutzungsgruppe (AU) von durchschnittlich 59.6 Punkten auf 83.0 Punkte anstieg ($MD = 26.00$, $SE = 3.42, p < .001$), was einer Zunahme von 26.0 Punkten auf der 100-Punkte-Skala entspricht und den klinischen Cutoff-Wert von 50 Punkten (Brähler et al., 2007) deutlich überschreitet. In der Kontrollgruppe (BU) fiel die Veränderung demgegenüber gering und nicht signifikant aus (von $M = 63.5$ auf $M = 6.0$, $MD = 2.00$, $SE = 3.17, p = .922$). Zum zweiten Messzeitpunkt unterschieden sich die Gruppen signifikant voneinander ($MD = 18.00$, $SE = 4.18, p < .001$), während zu T1 kein signifikanter Gruppenunterschied bestand ($MD = -6.00$, $SE = 4.82, p = .602$). H1 wird durch diesen Befund bestätigt.

4.3 Wahrgenommene Responsivität und Empathie im Verlauf

Für die wahrgenommene Responsivität (PRI-R) zeigte sich weder ein signifikanter Interaktionseffekt Zeit x Gruppe ($F(1, 50) = 0.10, p = .755$, partielles Eta-Quadrat = .002) noch ein signifikanter Haupteffekt der Zeit ($F(1, 50) = 0.17, p = .685$, partielles Eta-Quadrat = .003) oder ein signifikanter Zwischensubjekteffekt der Gruppe ($F(1, 50) = 2.71, p = .106$, partielles Eta-Quadrat = .05). Während die wahrgenommene Responsivität in der aktiven Nutzungsgruppe (AU) von $M = 3.87$ auf $M = 3.84$ leicht absank, stieg sie in der Kontrollgruppe (BU) von $M = 3.59$ auf $M = 3.65$ leicht an, wie in Abbildung 3 dargestellt. Eine differenzielle Veränderung der wahrgenommenen Responsivität zwischen den Gruppen über die Zeit war damit nicht nachweisbar. H2 wird nicht bestätigt.

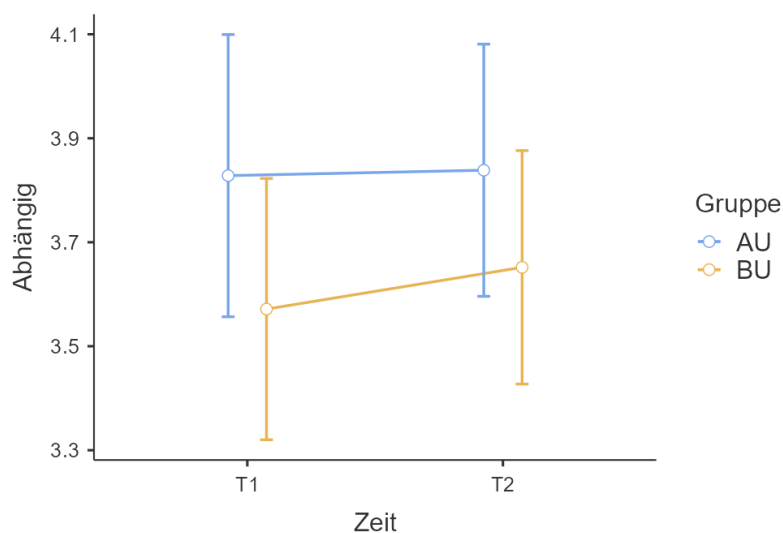


Abbildung 3. Mittlere Verläufe der wahrgenommenen Responsivität (PRI-R) der aktiven Nutzungsgruppe (AU) und der Kontrollgruppe (BU) zu T1 und T2.

Auch für die wahrgenommene Empathie (PETS) war der Interaktionseffekt Zeit x Gruppe nicht signifikant ($F(1, 50) < 0.01, p = .987$, partielles Eta-Quadrat = .000). Deskriptiv zeigte die aktive Nutzungsgruppe (AU) leichte Zuwächse von $M = 63.6$ auf $M = 68.2$, ebenso die Kontrollgruppe (BU) von $M = 52.3$ auf $M = 56.6$. Ergänzend zeigte sich jedoch ein signifikanter Haupteffekt der Zeit ($F(1, 50) = 4.73, p = .034$, partielles Eta-Quadrat = .09), was auf einen gruppenübergreifenden Anstieg der wahrgenommenen KI-Empathie hindeutet. Darüber hinaus war der Zwischensubjekteffekt der Gruppe signifikant ($F(1, 50) = 9.94, p = .003$, partielles Eta-Quadrat = .17): Die aktive Nutzungsgruppe (AU) nahm die KI insgesamt als empathischer wahr als die Kontrollgruppe (BU). Post-hoc-Vergleiche zeigten einen signifikanten

Gruppenunterschied zu T2 ($MD = 11.65$, $SE = 4.27$, $p = .042$), während der Unterschied zu T1 das Signifikanzniveau knapp verfehlte ($p = .051$). H3 wird nicht bestätigt.

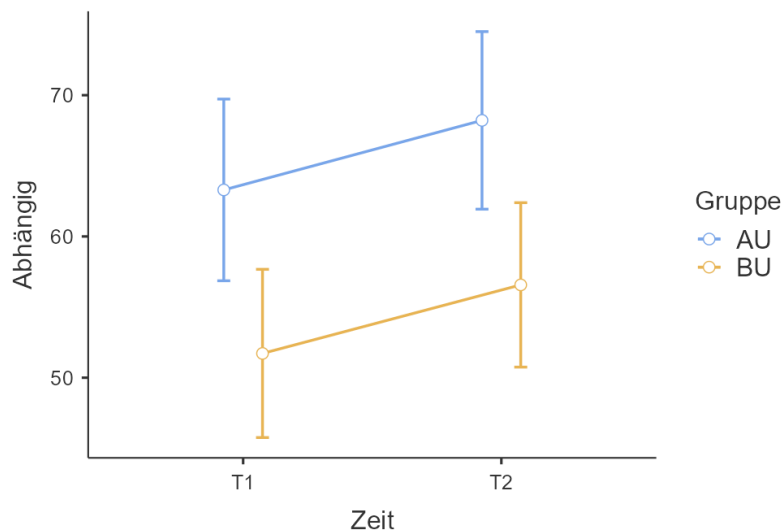


Abbildung 4. Mittlere Verläufe der wahrgenommenen Empathie (PETS) der aktiven Nutzungsgruppe (AU) und der Kontrollgruppe (BU) zu T1 und T2.

4.4 Nutzungsdauer und Wohlbefindensveränderung

Der Zusammenhang zwischen Nutzungsdauer und Wohlbefindensveränderung (Delta WHO-5) innerhalb der aktiven Nutzungsgruppe (AU) war nicht signifikant ($r_{(s)} = -.27$, $p = .210$). In der aktiven Nutzungsgruppe (AU) nutzte die Mehrheit der Teilnehmenden die KI zwischen 10 und 14 Minuten pro Session ($n = 16$, M Delta WHO-5 = 24.5), gefolgt von 5–9 Minuten ($n = 5$, M Delta WHO-5 = 32.8) und 15–19 Minuten ($n = 2$, M Delta WHO-5 = 30.0). Ein klarer linearer Zusammenhang zwischen Nutzungsdauer und Wohlbefindensveränderung war deskriptiv nicht erkennbar. H4 wird nicht bestätigt.

4.5 Explorative Analysen

4.5.1 Nutzungszweck

Innerhalb der aktiven Nutzungsgruppe (AU) wurde explorativ geprüft, ob Teilnehmende, die die KI zur Bewältigung von Stress und emotionaler Belastung einsetzten, eine stärkere Wohlbefindensverbesserung zeigten als jene mit anderen Nutzungszwecken. Dabei gaben 13 von 28 Teilnehmenden (46.4%) der aktiven Nutzungsgruppe (AU) an, die KI unter anderem zur Stressbewältigung genutzt zu haben. Ein t -Test für unabhängige Stichproben auf Delta WHO-5 innerhalb der aktiven Nutzungsgruppe (AU) ergab keinen signifikanten Unterschied zwischen Teilnehmenden mit und ohne emotionalen Nutzungszweck ($t(22) = -0.51$, $p = .614$, $d = -0.22$).

Die deskriptive Auswertung der offenen Fragen zeigte, dass die AU-Teilnehmenden die KI überwiegend für emotional relevante Themen wie Stress, Ängste und soziale Konflikte nutzten.

4.5.2 Veränderung der Wahrnehmungsscores als Prädiktoren

Da die Analysen in Abschnitt 4.3 keine signifikanten Interaktionseffekte Zeit x Gruppe für wahrgenommene Empathie und Responsivität ergaben, stellt sich die Frage, ob diese Konstrukte dennoch als Wirkmechanismus relevant sein könnten. Während die Mixed ANOVA prüft, ob sich AU und BU über die Zeit unterschiedlich entwickeln, zeigt eine Regressionsanalyse innerhalb der aktiven Nutzungsgruppe (AU) eine andere Perspektive. Eine emotional unterstützende KI-Interaktion könnte die wahrgenommene Empathie und Responsivität steigern und damit die erlebte Interaktionsqualität in der AU-Gruppe verändern. Eine höhere Interaktionsqualität dürfte wiederum Relatedness stärker ansprechen und so das subjektive Wohlbefinden verändern. Ergänzend wurde daher untersucht, ob Veränderungen in der wahrgenommenen Empathie oder Responsivität innerhalb der aktiven Nutzungsgruppe (AU) mit einer Wohlbefindensveränderung zusammenhängen. Als Prädiktoren dienten die Differenzwerte Delta PETS und Delta PRI, als Kriterium diente Delta WHO-5.

Teilnehmende der aktiven Nutzungsgruppe (AU), bei denen die wahrgenommene Empathie der KI zwischen T1 und T2 zunahm, wiesen eine stärkere Verbesserung des subjektiven Wohlbefindens auf. Delta PETS sagte Delta WHO-5 signifikant vorher (nicht standardisierter Regressionskoeffizient $b = 0.56$, $SE = 0.18$, $t(22) = 3.07$, $p = .006$, $R^2 = .30$) und erklärte damit knapp 29,9 % der Varianz in Delta WHO-5.

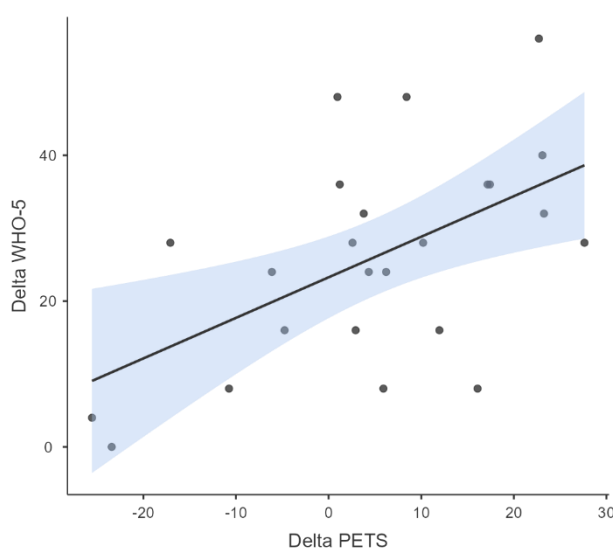


Abbildung 5. Zusammenhang zwischen der Veränderung der wahrgenommenen Empathie (Delta PETS) und der Wohlbefindensveränderung (Delta WHO-5) in der aktiven Nutzungsgruppe (AU) im 95%-Konfidenzintervall.

Ein vergleichbares Muster zeigte sich für die wahrgenommene Responsivität: Teilnehmende, bei denen Delta PRI zunahm, wiesen ebenfalls stärkere Wohlbefindensverbesserungen auf ($b = 10.20$, $SE = 3.80$, $t(22) = 2.68$, $p = .014$, $R^2 = .25$). Delta PRI erklärte 24.7% der Varianz in Delta WHO-5.

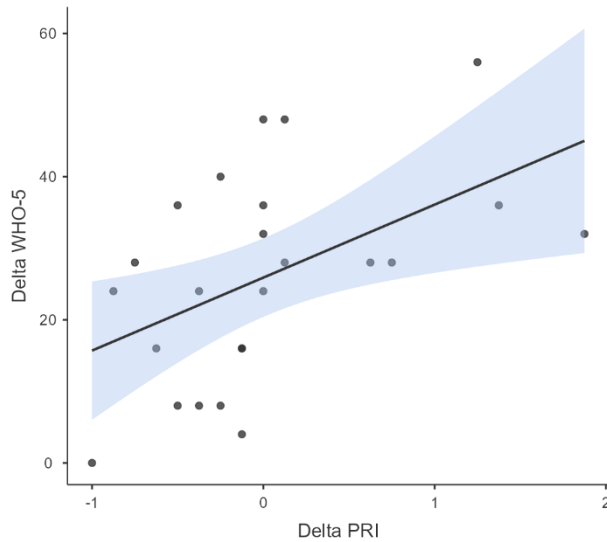


Abbildung 6. Zusammenhang zwischen der Veränderung der wahrgenommenen Responsivität (Delta PRI) und der Wohlbefindensveränderung (Delta WHO-5) in der aktiven Nutzungsgruppe (AU) im 95%-Konfidenzintervall.

5 Diskussion

5.1 Interpretation der Ergebnisse

In der vorliegenden Studie wurde untersucht, ob eine einwöchige Interaktion mit einem emotional unterstützenden KI-Chatbot das psychische Wohlbefinden steigern kann. Die Ergebnisse zeigen, dass die aktive Nutzungsgruppe (AU) nach sieben Tagen einen deutlichen Anstieg des Wohlbefindens berichtete, während die Kontrollgruppe (BU) nahezu unverändert blieb. Für die wahrgenommene Empathie und Responsivität des Chatbots zeigten sich hingegen keine gruppenspezifischen Veränderungen innerhalb des Untersuchungszeitraums. Auch die reine Nutzungsdauer stand in keinem eindeutigen Zusammenhang mit der Wohlbefindensverbesserung. Zusammenhänge zwischen den Veränderungen der Interaktionswahrnehmung und des Wohlbefindens innerhalb der aktiven Nutzungsgruppe deuten jedoch auf mögliche Wirkmechanismen hin, die in Folgestudien gezielt geprüft werden sollten.

Der gefundene Wohlbefindenseffekt lässt sich theoretisch im Rahmen der Self-Determination Theory einordnen. Der eingesetzte Stimulus-Prompt war gezielt darauf ausgelegt, die Interaktionsqualität zu steigern. Eine emotional unterstützende KI-Interaktion, die Verständnis signalisiert und wertfrei reagiert, könnte bei den Nutzenden ein subjektives Erleben von Verstandenwerden und Zugewandtheit erzeugen, die Interaktionsqualität verbessern und somit Relatedness im Sinne der SDT ansprechen (Ryan & Deci, 2000). Die Befriedigung dieses Grundbedürfnisses gilt als direkt mit subjektivem Wohlbefinden assoziiert (Martela et al., 2023). Im Chatbot-Kontext konnte gezeigt werden, dass wahrgenommene Relatedness das Commitment und die parasoziale Bindung beeinflussen kann (Fröbel & Kenning, 2025). Dieser Befund reiht sich in eine wachsende Studienbasis ein, die darauf hindeutet, dass KI-Chatbots kleine bis mittlere Effekte auf das psychische Wohlbefinden erzielen können (Feng et al., 2025; Vertsberger et al., 2022). Auffällig ist, dass die vorliegende Studie diese Tendenz bereits innerhalb von sieben Tagen abbilden konnte, was durch Befunde zur stimmungsverbessernden Wirkung empathischer Chatbot-Interaktionen nach sozialer Zurückweisung gestützt wird (De Gennaro et al., 2020). Die beobachtete Effektstärke ist jedoch mit Vorsicht zu interpretieren, da angesichts der kleinen Stichprobe und des fehlenden Follow-ups offenbleibt, ob das verbesserte Wohlbefinden nachhaltig stabil bleibt oder einem Neuheitseffekt (Novelty-Effekt) entspricht.

Des Weiteren bestätigt die stärker wahrgenommene Empathie der KI in der aktiven Nutzungsgruppe, dass der Stimulus-Prompt die Interaktions- und Beziehungsqualität gesteigert haben könnte. Das ist die Grundvoraussetzung dafür, dass das Bedürfnis nach Relatedness überhaupt angesprochen werden konnte. Die im weiteren Verlauf ausbleibenden Interaktionseffekte für Empathie (H3) und Responsivität (H2) könnten trotzdem durch die parasoziale Interaktionstheorie (PSI) erklärt werden. In der PSI-Forschung wird zwischen situativen parasozialen Interaktionen und stabileren parasozialen Beziehungen unterschieden, die sich erst über längere Expositionszeiträume entwickeln. Während eine erste empathische Interaktion kurzfristig positive Wahrnehmungen auslösen kann, benötigen anhaltende Veränderungen der Beziehungsqualität mehr Zeit. Dies deckt sich mit Befunden, wonach wahrgenommene Responsivität als stabiles Beziehungsmerkmal gilt, das sich graduell aufbaut und sich innerhalb weniger Tage nicht tiefgreifend verändert (Ter Kuile et al., 2017). Auch die Entwicklung einer digitalen therapeutischen Allianz benötigt einen längeren Zeitraum (He et al., 2022), wobei neuere Studien nahelegen, dass dieser Prozess hochdynamisch und erwartungsabhängig verläuft (Xu et al., 2025). Ähnlich dazu werden signifikante Unterschiede in der Empathiewahrnehmung zwischen aktiver und passiver Nutzung erst über mehrere Sitzungen hinweg sichtbar (Laban et al., 2024; Chandra et al., 2025). Dass sich der Effekt einer positiven Veränderung bei wahrgenommener Responsivität weniger deutlich zeigte, könnte jedoch auch darauf hindeuten, dass die empathische Qualität der Interaktion einen größeren Einfluss auf die Teilnehmenden hatte als die Responsivität.

Da die konfirmatorischen Tests keine Gruppenunterschiede in der Interaktionswahrnehmung zeigten, stellte sich die Frage ob, der postulierte Wirkmechanismus dennoch innerhalb der aktiven Nutzungsgruppe sichtbar wird. Eine emotional unterstützende KI-Interaktion könnte die wahrgenommene Empathie und Responsivität steigern und damit die erlebte Interaktions- und Beziehungsqualität verbessern. Eine höhere Interaktions- und Beziehungsqualität dürfte wiederum Relatedness stärker ansprechen und so das subjektive Wohlbefinden begünstigen. Dabei deuten Veränderungen in der wahrgenommenen Empathie auf einen substanziellen Zusammenhang mit der Wohlbefindensveränderung hin. Der entsprechende Befund für die Responsivitätsveränderung ist jedoch mit besonderer Vorsicht zu interpretieren. Das Streudiagramm zeigt bei nahezu unveränderter Responsivitätswahrnehmung eine außerordentlich hohe Streuung der Wohlbefindensveränderung, sodass die Regressionsgerade durch wenige Ausreißer mit hohen Veränderungswerten beeinflusst sein könnte. Bei einer Stichprobengröße von 24 Teilnehmenden ist das Risiko eines Fehlers erster Art nicht

auszuschließen. Zudem sind die Regressionskoeffizienten der beiden Analysen aufgrund der unterschiedlichen Skalenlängen von PRI-R (1–5) und PETS (0–100) nicht direkt miteinander vergleichbar. Der Befund sollte daher nicht als Beleg für einen robusten Wirkmechanismus gewertet werden, sondern allenfalls als explorativer Hinweis, der in Folgestudien mit größeren Stichproben gezielt geprüft werden müsste. Das Ausbleiben der Hypothesenbestätigung für H2 und H3 ist vor diesem Hintergrund nicht als Beleg für die Irrelevanz der Interaktionswahrnehmung zu interpretieren, sondern verweist auf die zeitlichen und methodischen Grenzen des einwöchigen Studiendesigns.

Die Erwartung an H4 war, angelehnt an die klassische Überlegung aus der PSI-Forschung, dass eine vermehrte Exposition gegenüber dem System auch zu einem stärkeren Effekt auf das Wohlbefinden führen müsste, ähnlich wie wiederholte Mediennutzung parasoziale Bindungen verstärken soll. Diese Überlegung wurde bewusst nochmal empirisch geprüft, auch wenn die PSI-Forschung bereits früh zeigte, dass Nutzungshäufigkeit keine hinreichende Bedingung für parasoziale Bindung ist. Das Ergebnis bestätigte diese Skepsis: Ähnlich wie Rubin et al. (1985) bereits belegten, hatte die reine Nutzungsdauer keinen nachweisbaren Effekt auf die Wohlbefindensveränderung. Die Nutzungsdauer als alleiniger Indikator erscheint konzeptuell begrenzt, da sie keine Aussage über die Qualität oder den emotionalen Gehalt der Interaktion zulässt.

Über die einzelnen Hypothesen hinaus liefert die Studie ein differenziertes Bild. Emotionale KI-Unterstützung könnte zunächst situative Stimmungsverbesserungen begünstigen, ohne dass sich Empathie- und Responsivitätswahrnehmungen innerhalb einer Woche substantiell verändern. Für die theoretische Einordnung dieser Dynamik stößt die klassische PSI-Theorie jedoch an konzeptuelle Grenzen. Tukachinsky und Stever (2018) argumentieren, dass parasoziale Beziehungen sich in qualitativ unterscheidbaren Entwicklungsphasen entfalten und über klassische Massenmedien hinaus auf neue interaktive Kontexte ausgeweitet werden müssen. Klassische PSI-Modelle beschreiben jedoch einseitige Beziehungen zu Medienfiguren, die nicht antworten und nicht adaptieren. KI-Chatbots sind hingegen bidirektional und reaktiv, was eine qualitativ andere Interaktionsdynamik erzeugt. Abercrombie et al. (2023) weisen darauf hin, dass menschenähnliches Verhalten in Dialogsystemen den Eindruck von Gegenseitigkeit vermitteln kann, ohne dass eine tatsächliche soziale Intentionalität des Systems vorliegt. Dies beschreibt im Kern nichts anderes als den Mechanismus der Relatedness-Befriedigung. Die Befunde dieser Studie legen nahe, dass eine erweiterte Konzeptualisierung

parasozialer Interaktion und Beziehung sinnvoll wäre, die der Bidirektionalität und Adaptivität von KI-Systemen Rechnung trägt (Brandtzaeg et al., 2022).

5.2 Limitationen

Die Studie unterliegt mehreren Limitationen, die bei der Interpretation der Befunde zu berücksichtigen sind. Die bedeutsamste Einschränkung betrifft die Stichprobengröße, da mit 52 Teilnehmenden zum zweiten Messzeitpunkt die a priori berechnete Zielgröße von 90 für mittlere Effekte unterschritten wurde. Die konfirmatorischen Analysen zu H2, H3 und H4 waren daher möglicherweise unterpowert. Explorative Regressionsanalysen innerhalb der aktiven Nutzungsgruppe mit 24 Teilnehmenden sind in ihrer Aussagekraft erheblich begrenzt und wurden nicht für multiples Testen korrigiert.

Die Stichprobenzusammensetzung schränkt die Generalisierbarkeit der Befunde ein, da die Mehrheit der Teilnehmenden zwischen 18 und 29 Jahren alt war und über SONA sowie universitäre Verteiler rekrutiert wurde. Dies bedingt Selbstselektion, da Personen, die bereit sind, sieben Tage lang einen Chatbot einzusetzen, sich motivational von der Allgemeinbevölkerung unterscheiden könnten. Die eingeschränkte Alters- und Bildungsvielfalt ist ebenfalls zu beachten, da junge Erwachsene mit Hochschulhintergrund möglicherweise eine andere Offenheit gegenüber KI-Unterstützung aufweisen als andere Bevölkerungsgruppen.

Die interne Validität der Studie ist weiter eingeschränkt durch die unkontrollierbare Variabilität der KI-Antworten. Der verwendete Stimulus-Prompt stellte zwar die einzige experimentelle Stellschraube dar, um das Interaktionsverhalten der KI zu steuern, doch die tatsächlichen Antworten des Systems unterlagen einer nicht vorhersehbaren Variabilität. Selbst bei identischem Prompt können Large Language Models je nach Kontext und Zufallsfaktor unterschiedliche Formulierungen, Tonlagen oder Inhaltsschwerpunkte produzieren. Dies könnte dazu geführt haben, dass die Teilnehmer der aktiven Nutzungsgruppe zwar theoretisch denselben Prompts ausgesetzt waren, aber in der Praxis unterschiedliche Interaktionserlebnisse hatten, die methodisch nicht quantifiziert oder kontrolliert werden konnten. Hinzu kommt, dass die tägliche Nutzung über die sieben Tage hinweg nur empfohlen, nicht jedoch objektiv verifiziert wurde. Es existieren keine protokollierten Daten zu tatsächlicher Nutzungshäufigkeit, Dauer oder Zeitpunkt der Interaktionen. Teilnehmer könnten die KI an manchen Tagen gar nicht, an anderen mehrfach genutzt haben, oder die Interaktionsqualität könnte stark von der aktuellen Stimmung abhängen. Einschränkend ist anzumerken, dass ein vollständig standardisiertes Nutzungsverhalten dem natürlichen

Nutzungsmuster kaum entspräche, da individuelle Gesprächsbedürfnisse und Relevanzwahrnehmungen per se variieren.

Die Interventionsdauer von sieben Tagen stellt eine weitere Limitation dar, da stabile Veränderungen in Empathie- und Responsivitätswahrnehmung längere Zeiträume voraussetzen scheinen, wie der Vergleich mit anderen Studien nahelegt (Chandra et al., 2025; Laban et al., 2024). Das Studiendesign erlaubte damit allenfalls erste Hinweise, nicht jedoch abschließende kausale Schlüsse. Alle abhängigen Variablen wurden ausschließlich über Selbstberichte erhoben, was das Risiko sozialer Erwünschtheit und Common-Method-Bias birgt. Der Verzicht auf einen Follow-up-Messzeitpunkt lässt zudem keine Aussagen über die Nachhaltigkeit der beobachteten Effekte zu.

5.3 Zukunftsperspektiven

Aus den Befunden und Limitationen lassen sich folgende Empfehlungen für künftige Forschung ableiten.

Eine Verlängerung der Interventionsdauer auf 14 bis 28 Tage mit dichtem Längsschnittdesign wird empfohlen, da wahrgenommene Responsivität und Empathiewahrnehmung einen längeren Zeitraum benötigen, um sich nachhaltig zu verändern (Ter Kuile et al., 2017; Chandra et al., 2025). Eine längere Intervention würde die Verlaufsdynamik der angenommenen Wirkkette, dass gesteigerte Empathie- und Responsivitätswahrnehmung die Interaktions- und Beziehungsqualität verbessert und sich durch die Befriedigung von Relatedness positiv auf das subjektive Wohlbefinden auswirkt, detaillierter abbilden und die Teststärke für Veränderungen in Responsivität und Empathie erhöhen können.

Da in der vorliegenden Studie nur zwei Bedingungen gegenübergestellt wurden, die der emotionalen unterstützenden Nutzung und die der sachlichen Nutzung, bleibt offen, welche spezifischen Merkmale der Interaktion die Wahrnehmungsunterschiede und den Wohlbefindenseffekt tatsächlich verursachen. Methodisch wäre daher eine systematische Variation der Prompt-Gestaltung sinnvoll, etwa indem einzelne Interaktionskomponenten wie emotionale Spiegelung, Wertfreiheit oder aktives Zuhören gezielt isoliert und verglichen werden. Eine solche Variation würde eine präzisere kausale Zuordnung ermöglichen und das Verständnis der Wirkmechanismen verbessern.

5.4 Praktische Implikationen

Die Befunde sind trotz der genannten Einschränkungen von praktischer Relevanz, da die beobachtete Wohlbefindensverbesserung innerhalb von sieben Tagen darauf hindeutet, dass emotional ausgerichtete KI-Interaktionen möglicherweise als niedrigschwelliges Angebot eingesetzt werden könnten, insbesondere für junge Erwachsene mit Zugangsbarrieren zur professionellen Unterstützung (Grové, 2021). Für die Chatbot-Gestaltung legt das Befundmuster nahe, dass empathisches Design mit validierendem und wertfreiem Interaktionsverhalten von zentraler Bedeutung sein könnte. Der Haupteffekt der Gruppe für wahrgenommene Empathie deutet darauf hin, dass gezieltes Prompting wahrnehmungsrelevante Unterschiede erzeugen kann. Anwendungspotenzial könnte insbesondere in Hochschulkontexten mit Prüfungsstress sowie in betrieblichen Gesundheitsangeboten und Mental-Health-Apps bestehen. Eine ethisch verantwortungsvolle Rahmung ist dabei in jedem Fall unverzichtbar, da KI-Chatbots keine Psychotherapie ersetzen können und für klinische Erkrankungen nicht geeignet sind. Transparenz über die maschinelle Natur sowie der Verweis auf weiterführende Angebote sind als grundlegende Anforderungen zu betrachten. KI-Chatbots sollten demnach allenfalls als ergänzendes Element im gestuften Versorgungsmodell verstanden werden.

5.5 Fazit

Die vorliegende Studie liefert vorläufige Hinweise darauf, dass siebentägige emotionale KI-Nutzung mit einer signifikanten Verbesserung des subjektiven Wohlbefindens assoziiert sein könnte, während Veränderungen in Responsivität, Empathie und Nutzungsdauer im konfirmatorischen Test nicht nachweisbar waren. Explorative Analysen ergänzen dieses Bild um Hinweise auf mögliche Wirkmechanismen, da Veränderungen in Empathie und Responsivität als signifikante Prädiktoren des Wohlbefindenszuwachses innerhalb der aktiven Nutzungsgruppe identifiziert wurden, was die angenommene Wirkkette zumindest teilweise erklären könnte. Die Ergebnisse sind als vorläufiger Ausgangspunkt zu verstehen, da sie mit der Annahme vereinbar sind, dass emotional unterstützende KI-Interaktion Relatedness kurzfristig ansprechen und Wohlbefinden begünstigen könnte. Gleichzeitig verdeutlichen sie, dass die zugrunde liegenden Mechanismen mit einem einwöchigen Studiendesign nur ausschnittartig erfasst werden können. Weiterer Forschungsbedarf besteht insbesondere hinsichtlich der Langzeitwirkung, der kausalen Mediationsstruktur und der Robustheit über diverse Stichproben. Die Studie leistet damit einen ersten empirischen Beitrag zu einem Forschungsfeld mit wachsender gesellschaftlicher Relevanz.

Literaturverzeichnis

- Abercrombie, G., Curry, A., Dinkar, T., Rieser, V., & Talat, Z. (2023). Mirages. On Anthropomorphism in Dialogue Systems. *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, 4776–4790.
<https://doi.org/10.18653/v1/2023.emnlp-main.290>
- BPtK. (2018, Oktober 15). *Gutachten_bedarfsplanung 2018*.
https://api.bptk.de/uploads/2018_10_15_pm_gutachten_bedarfsplanung_d9a69795bd.pdf
- Brähler, E., Mühlan, H., Albani, C., & Schmidt, S. (2007). Teststatistische Prüfung und Normierung der deutschen Versionen des EUROHIS-QOL Lebensqualität-Index und des WHO-5 Wohlbefindens-Index. *Diagnostica*, 53(2), 83–96.
<https://doi.org/10.1026/0012-1924.53.2.83>
- Brandtzaeg, P. B., Skjuve, M., & Følstad, A. (2022). My AI Friend: How Users of a Social Chatbot Understand Their Human–AI Friendship. *Human Communication Research*, 48(3), 404–429. <https://doi.org/10.1093/hcr/hqac008>
- Chandra, M., Hernandez, J., Ramos, G., Ershadi, M., Bhattacharjee, A., Amores, J., Okoli, E., Paradiso, A., Warreth, S., & Suh, J. (2025). *Longitudinal Study on Social and Emotional Use of AI Conversational Agent* (arXiv:2504.14112). arXiv.
<https://doi.org/10.48550/arXiv.2504.14112>
- Christoforakos, L., Feicht, N., Hinkofer, S., Löscher, A., Schlegl, S. F., & Diefenbach, S. (2021). Connect With Me. Exploring Influencing Factors in a Human-Technology Relationship Based on Regular Chatbot Use. *Frontiers in Digital Health*, 3, 689999.
<https://doi.org/10.3389/fdgth.2021.689999>
- Coe, E., Doy, A., Enomoto, K., & Healy, C. (2023). *Gen Z mental health: The impact of tech and social media*.

- Crasta, D., Rogge, R. D., Maniaci, M. R., & Reis, H. T. (2021). Toward an optimized measure of perceived partner responsiveness: Development and validation of the perceived responsiveness and insensitivity scale. *Psychological Assessment, 33*(4), 338–355. <https://doi.org/10.1037/pas0000986>
- De Gennaro, M., Krumhuber, E. G., & Lucas, G. (2020). Effectiveness of an Empathic Chatbot in Combating Adverse Effects of Social Exclusion on Mood. *Frontiers in Psychology, 10*, 3061. <https://doi.org/10.3389/fpsyg.2019.03061>
- DGPPN. (2025, Februar). *Basisdaten Psychische Erkrankungen*. https://www.dgppn.de/_Resources/Persistent/a71c6af0841eb0236690e3d360630418d5f56c24/2025_DGPPN_Basisdaten%20Psychische%20Erkrankungen.pdf
- Feng, X., Tian, L., Ho, G. W. K., Yorke, J., & Hui, V. (2025). The Effectiveness of AI Chatbots in Alleviating Mental Distress and Promoting Health Behaviors Among Adolescents and Young Adults: Systematic Review and Meta-Analysis. *Journal of Medical Internet Research, 27*, e79850–e79850. <https://doi.org/10.2196/79850>
- Fröbel, L., & Kenning, P. (2025). *What Motivates People to Prefer and Love Chatbots? An Empirical Investigation Based on Self-Determination Theory*. Hawaii International Conference on System Sciences. <https://doi.org/10.24251/HICSS.2025.216>
- Grové, C. (2021). Co-developing a Mental Health and Wellbeing Chatbot With and for Young People. *Frontiers in Psychiatry, 11*, 606041. <https://doi.org/10.3389/fpsyg.2020.606041>
- Hartmann, T., & Goldhoorn, C. (2011). Horton and Wohl Revisited: Exploring Viewers' Experience of Parasocial Interaction. *Journal of Communication, 61*(6), 1104–1121. <https://doi.org/10.1111/j.1460-2466.2011.01595.x>
- He, L., Basar, E., Wiers, R. W., Antheunis, M. L., & Krahmer, E. (2022). Can chatbots help to motivate smoking cessation? A study on the effectiveness of motivational interviewing

- on engagement and therapeutic alliance. *BMC Public Health*, 22(1), 726.
<https://doi.org/10.1186/s12889-022-13115-x>
- Horton, D., & Richard Wohl, R. (1956). Mass Communication and Para-Social Interaction: Observations on Intimacy at a Distance. *Psychiatry*, 19(3), 215–229.
<https://doi.org/10.1080/00332747.1956.11023049>
- Keyes, C. L. M. (2002). The Mental Health Continuum: From Languishing to Flourishing in Life. *Journal of Health and Social Behavior*, 43(2), 207.
<https://doi.org/10.2307/3090197>
- Laban, G., Kappas, A., Morrison, V., & Cross, E. S. (2024). Building Long-Term Human–Robot Relationships: Examining Disclosure, Perception and Well-Being Across Time. *International Journal of Social Robotics*, 16(5), 1–27. <https://doi.org/10.1007/s12369-023-01076-z>
- Liebers, N., & Schramm, H. (2019). Parasocial Interactions and Relationships with Media Characters—An Inventory of 60 Years of Research. In *Communication Research Trends*.
<https://link.gale.com/apps/doc/A621689770/AONE?u=anon~5a229d21&sid=googleScholar&xid=4f235b44>
- Lu, C.-H., Li, C.-Y., & Chen, Y.-X. (2026). How does a service chatbot build parasocial relationships with customers? A perspective from parasocial interaction theory. *Internet Research*, 36(7), 107–126. <https://doi.org/10.1108/INTR-03-2024-0421>
- Martela, F., Lehmus-Sun, A., Parker, P. D., Pessi, A. B., & Ryan, R. M. (2023). Needs and Well-Being Across Europe: Basic Psychological Needs Are Closely Connected With Well-Being, Meaning, and Symptoms of Depression in 27 European Countries. *Social Psychological and Personality Science*, 14(5), 501–514.
<https://doi.org/10.1177/19485506221113678>

- Naslund, J. A., Aschbrenner, K. A., Marsch, L. A., & Bartels, S. J. (2016). The future of mental health care: Peer-to-peer support and social media. *Epidemiology and Psychiatric Sciences*, 25(2), 113–122. <https://doi.org/10.1017/S2045796015001067>
- Pentina, I., Hancock, T., & Xie, T. (2023). Exploring relationship development with social chatbots: A mixed-method study of replika. *Computers in Human Behavior*, 140, 107600. <https://doi.org/10.1016/j.chb.2022.107600>
- Rubin, A. M., Perse, E. M., & Powell, R. A. (1985). LONELINESS, PARASOCIAL INTERACTION, AND LOCAL TELEVISION NEWS VIEWING. *Human Communication Research*, 12(2), 155–180. <https://doi.org/10.1111/j.1468-2958.1985.tb00071.x>
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68–78. <https://doi.org/10.1037/0003-066X.55.1.68>
- Ryan, R. M., & Deci, E. L. (2001). On Happiness and Human Potentials: A Review of Research on Hedonic and Eudaimonic Well-Being. *Annual Review of Psychology*, 52(1), 141–166. <https://doi.org/10.1146/annurev.psych.52.1.141>
- Schmidmaier, M., Rupp, J., Cvetanova, D., & Mayer, S. (2024). Perceived Empathy of Technology Scale (PETS): Measuring Empathy of Systems Toward the User. *Proceedings of the CHI Conference on Human Factors in Computing Systems*, 1–18. <https://doi.org/10.1145/3613904.3642035>
- Schramm, H., Liebers, N., Biniak, L., & Dettmar, F. (2024). Research trends on parasocial interactions and relationships with media characters. A review of 281 English and German-language studies from 2016 to 2020. *Frontiers in Psychology*, 15, 1418564. <https://doi.org/10.3389/fpsyg.2024.1418564>
- Sischka, P. E., Costa, A. P., Steffgen, G., & Schmidt, A. F. (2020). The WHO-5 well-being index – validation based on item response theory and the analysis of measurement

- invariance across 35 countries. *Journal of Affective Disorders Reports*, 1, 100020.
<https://doi.org/10.1016/j.jadr.2020.100020>
- Skjuve, M., Følstad, A., Fostervold, K. I., & Brandtzaeg, P. B. (2021). My Chatbot Companion—A Study of Human-Chatbot Relationships. *International Journal of Human-Computer Studies*, 149, 102601. <https://doi.org/10.1016/j.ijhcs.2021.102601>
- Skjuve, M., Følstad, A., Fostervold, K. I., & Brandtzaeg, P. B. (2022). A longitudinal study of human–chatbot relationships. *International Journal of Human-Computer Studies*, 168, 102903. <https://doi.org/10.1016/j.ijhcs.2022.102903>
- Ter Kuile, H., Kluwer, E. S., Finkenauer, C., & Van Der Lippe, T. (2017). Predicting adaptation to parenthood: The role of responsiveness, gratitude, and trust. *Personal Relationships*, 24(3), 663–682. <https://doi.org/10.1111/pere.12202>
- Topp, C. W., Østergaard, S. D., Søndergaard, S., & Bech, P. (2015). The WHO-5 Well-Being Index: A Systematic Review of the Literature. *Psychotherapy and Psychosomatics*, 84(3), 167–176. <https://doi.org/10.1159/000376585>
- Tukachinsky, R., & Stever, G. (2019). Theorizing Development of Parasocial Engagement. *Communication Theory*, 29(3), 209–230. <https://doi.org/10.1093/ct/qty032>
- Vertsberger, D., Naor, N., & Winsberg, M. (2022). Adolescents' Well-being While Using a Mobile Artificial Intelligence–Powered Acceptance Commitment Therapy Tool: Evidence From a Longitudinal Study. *JMIR AI*, 1(1), e38171.
<https://doi.org/10.2196/38171>
- Xu, Z., Lee, Y.-C., Stasiak, K., Warren, J., & Lottridge, D. (2025). The Digital Therapeutic Alliance With Mental Health Chatbots: Diary Study and Thematic Analysis. *JMIR Mental Health*, 12, e76642. <https://doi.org/10.2196/76642>
- Y. Wang, A., Vereschagin, M., G. Richardson, C., J. Munthali, R., Xie, H., L. Hudec, K., Mori, T., Munro, L., & V. Vigo, D. (2025). Examining the effects of engagement with an app-based mental health intervention: A secondary analysis of a randomized control

trial with treatment non-compliance. *International Journal of Mental Health Systems*, 19(1), 30. <https://doi.org/10.1186/s13033-025-00688-4>

Anhang Fragebogen 1 und 2



bachelorarbeitKIChallenge → BA1

02.05.2026, 12:13

Werkzeug oder Wegbegleiter? – Die 7-Tage KI-Challenge

Hallo und herzlich willkommen!

Vielen Dank, dass Du dir die Zeit nimmst, an dieser Studie teilzunehmen. Wir untersuchen, wie Menschen im Alltag mit **künstlicher Intelligenz (LLMs)** interagieren und welche Auswirkungen unterschiedliche Nutzungsweisen haben.

Ablauf der Studie:

- **Heute (Session 1):** Befragung & 10-minütige KI-Nutzung nach Instruktion.
- **Die Woche über:** Täglich ca. 10 Minuten eigenständige KI-Nutzung.
- **Tag 7 (Session 2):** Kurze Abschlussbefragung (ca. 10 Min.).

ACHTUNG: WICHTIG!

Da wir Veränderungen über die Zeit untersuchen, ist die **Session 2 (Tag 7) entscheidend**. Wenn Du an Session 2 nicht teilnimmst, sind Deine Daten für die Forschung leider nicht verwertbar. Bitte nimm nur teil, wenn Du bereit bist, auch den zweiten Teil in 7 Tagen auszufüllen.

E-Mail-Angabe: Um den Link für den zweiten Teil zu erhalten, ist die Angabe Deiner E-Mail-Adresse notwendig. Diese wird anonym und streng getrennt von Deinen Studiendaten gespeichert

2. Informationen zum Datenschutz und Einwilligung zur Verarbeitung der E-Mail Adresse

Die Teilnahme an dieser Studie ist freiwillig. Im Rahmen der Befragung werden Daten zu Forschungszwecken erhoben. Wir legen großen Wert auf Transparenz bei der Verarbeitung Ihrer Angaben:

Trennung von Daten: Ihre E-Mail-Adresse wird ausschließlich erhoben, um Ihnen nach 7 Tagen den Link zum zweiten Teil der Befragung zuzusenden und Sie im Falle eines Gewinns der Aufwandsentschädigung (Gewinnspiel) zu benachrichtigen.

Durch das technische Verfahren von SoSci Survey wird Ihre E-Mail-Adresse strikt getrennt von Ihren inhaltlichen Antworten gespeichert. Eine Zusammenführung von Identität und Antwortdaten findet im Rahmen der Ergebnisauswertung nicht statt. Eine direkte Identifikation Ihrer Person ist somit nicht möglich.

Die Auswertung Ihrer inhaltlichen Antworten erfolgt so, dass keine Rückschlüsse auf Ihre Person gezogen werden können. Sobald die E-Mail-Adresse für die Kontaktaufnahme nicht mehr benötigt wird, wird sie dauerhaft gelöscht.

Die Speicherung und Auswertung der Daten erfolgt gemäß den geltenden datenschutzrechtlichen Bestimmungen. Sie können die Befragung jederzeit ohne Angabe von Gründen abbrechen. Zudem haben Sie das Recht, Ihre Einwilligung zur Speicherung der E-Mail-Adresse jederzeit mit Wirkung für die Zukunft unter maximilian.loibl@student.hnu.de zu widerrufen.

Einwilligung: Durch das Fortfahren und die Angabe Ihrer E-Mail-Adresse erklären Sie sich mit der beschriebenen Verarbeitung Ihrer Daten einverstanden.

- ☐ Ich habe die Informationen zum Datenschutz gelesen und erkläre mich mit der Teilnahme an der Studie sowie der beschriebenen Datenverarbeitung einverstanden.

3. Deine E-Mail-Adresse für die Erinnerungsmail zum 2. Messzeitpunkt (Tag 7) & das 100 € Gewinnspiel:

WT03

⚠ **Wichtig:** Bitte prüfe in 7 Tagen auch deinen **Spam-Ordner** auf unsere Nachricht.

Datenschutz: Deine Adresse wird separat und anonym gespeichert. Eine Verknüpfung mit deinen Antworten ist technisch ausgeschlossen.

Deine E-Mail-Adresse:

4. Welches Geschlecht haben Sie?

SD01

- ☐ weiblich
- ☐ männlich
- ☐ divers

5. Wie alt sind Sie?

SD02

[Bitte auswählen] ▼

6. Optional SONA-Code

SC01

Zutreffend auf:

- Studierende der HNU (1. & 2. Semester Wirtschaftspsychologie)
- Psychologiestudierende der Uni Ulm

Hier hast du die Möglichkeit, optional deinen **SONA-Code** einzugeben, um nach Abschluss der Umfrage **2 VPN-Stunden** gutgeschrieben zu bekommen.

Die Teilnahme ist selbstverständlich auch ohne Code möglich.

SONA-Code:

7. Dein Wohlbefinden der letzten Woche

Im Folgenden findest Du fünf Aussagen zu Deinem allgemeinen Wohlbefinden. Bitte geben an, wie häufig diese Aussagen in den letzten 7 Tagen auf Dich zugetroffen hat

In den letzten 7 Tagen...

Zu keinem Zeitpunkt	Ab und zu	Weniger als die Hälfte der Zeit	Mehr als die Hälfte der Zeit	Meistens	Die ganze Zeit
------------------------	-----------	--	------------------------------------	----------	-------------------

... war ich froh und guter Laune

☐	☐	☐	☐	☐	☐
-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------

... habe ich mich ruhig und entspannt gefühlt

☐	☐	☐	☐	☐	☐
-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------

... habe ich mich energetisch und aktiv gefühlt

☐	☐	☐	☐	☐	☐
-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------

... habe ich mich beim Aufwachen frisch und ausgeruht gefühlt

☐	☐	☐	☐	☐	☐
-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------

... war mein Alltag voller Dinge, die mich interessieren

☐	☐	☐	☐	☐	☐
-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------

Jetzt geht es los: Die KI-Interaktion

 **Wichtig: Bitte lese den Text vollständig durch!**

Heute startet deine erste Session – und damit der wichtigste Teil der Studie. In den kommenden 7 Tagen wirst du eine KI deiner Wahl (ChatGPT, Gemini, Pi, etc.) als Partner für deine **emotionale Unterstützung** nutzen.

Deine Rückkehr ist entscheidend: Gerade heute und am letzten Tag (Tag 7) ist es unverzichtbar, dass du nach dem Gespräch **sofort** zu diesem Browser-Tab zurückkehrst. Nur bei vollständigem Abschluss können wir deine Daten wissenschaftlich auswerten.

So gehst du vor:

- **Timer stellen:** Nutze einen Timer für 10 Minuten, um an deine Rückkehr erinnert zu werden.
- **Tab offen lassen:** Lass diesen Browser-Tab unbedingt offen.
- **Prompt & Themen:** Kopiere den Text unten und nutze die Vorschläge für den Start.

Kopiere diesen Text und sende ihn als erste Nachricht ab:

„Bitte agiere für die nächsten 10–15 Minuten als mein empathischer Zuhörer. Nutze Methoden der klientenzentrierten Gesprächsführung: Sei validierend, wertfrei und spiegle meine Emotionen, anstatt mir sofort Ratschläge zu geben. Halte deine Antworten kurz und fokussiert. Ich werde dir gleich ein Thema nennen, das mich aktuell beschäftigt. Bestätige bitte kurz, dass du bereit bist, und frage mich dann, was mich heute beschäftigt.“

Themen, an denen du dich orientieren kannst:

- **Stress & Belastung:** Stressbewältigung im Alltag oder Job.
- **Umgang mit Ängsten:** Zukunftsangst, soziale Ängste, Panikgefühle.
- **Einsamkeit & Isolation:** Bedürfnis nach Austausch oder Präsenz.
- **Selbstreflexion:** Einordnung von Gedanken, Grübeln stoppen.
- **Beziehungen & Soziales:** Beziehungsprobleme, Vorbereitung auf schwierige Gespräche.
- **Selbstwert & Motivation:** Zuspruch, Aufbau von Selbstvertrauen.
- **Schlafprobleme:** Gedankenkarussell vor dem Einschlafen.
- **Trauer & Verlust:** Umgang mit Tod oder dem Ende von Lebensphasen.
- **Entscheidungshilfe:** Abwägen von schwierigen Dilemmasituationen.
- **Validierung & Check-In:** Kurzes Feedback zu Gefühlen, Gedanken oder der Tagesform.

Deine erste Session startet JETZT. Viel Spaß beim ersten Austausch!

Jetzt geht es los: Die KI-Interaktion

 **Wichtig: Bitte lies den Text vollständig durch!**

Heute startet deine erste Session. In den kommenden 7 Tagen wirst du eine KI deiner Wahl (ChatGPT, Gemini, Pi, etc.) als **praktisches Werkzeug zur Informationsgewinnung und Alltagsunterstützung** nutzen.

Deine Rückkehr ist entscheidend: Gerade heute und am letzten Tag (Tag 7) ist es unverzichtbar, dass du nach der Interaktion **sofort** zu diesem Browser-Tab zurückkehrst. Nur bei vollständigem Abschluss können wir deine Daten wissenschaftlich auswerten.

So gehst du vor:

- **Timer stellen:** Nutze einen Timer für 10 Minuten, um an deine Rückkehr erinnert zu werden.
- **Tab offen lassen:** Lass diesen Browser-Tab unbedingt offen.
- **Prompt & Themen:** Kopiere den Text unten und nutze die Vorschläge für den Start.

Kopiere diesen Text und sende ihn als erste Nachricht ab:

„Bitte agiere für die nächsten 10–15 Minuten als mein sachlicher Informations-Assistent. Deine Aufgabe ist es, mir bei praktischen Wissensfragen, Recherchen oder organisatorischen Aufgaben zu helfen. Antworte bitte rein informativ, sachlich und präzise. Vermeide es, auf emotionale Aspekte einzugehen oder Ratschläge zu meinem Befinden zu geben. Bestätige bitte kurz, dass du bereit bist, und frage mich dann, bei welcher sachlichen Aufgabe du mich heute unterstützen kannst.“

Themen, an denen du dich orientieren kannst:

- **Wissensfragen & Recherche:** z.B. geschichtliche Fakten oder wissenschaftliche Konzepte erklären lassen.
- **Planung & Organisation:** z.B. eine Reiseroute erstellen oder einen Wochenplan entwerfen.
- **Textbearbeitung:** z.B. eine E-Mail sachlich formulieren oder einen Text zusammenfassen.
- **Haushalt & Alltag:** z.B. Rezeptvorschläge aus vorhandenen Zutaten oder Tipps zur Fleckenentfernung.
- **IT & Technik:** z.B. Hilfe bei Excel-Formeln oder einfache Programmierfragen klären.
- **Allgemeine Informationen:** z.B. Nachrichten-Zusammenfassungen oder Wetterphänomene erklären lassen.

Deine erste Session startet JETZT. Viel Erfolg beim ersten Austausch!

8. Wie viele Minuten dauerte Deine KI-Interaktion in dieser Session?

MC01

- ☐ 1-4 Minuten
- ☐ 5-9 Minuten
- ☐ 10-14 Minuten
- ☐ 15-19 Minuten
- ☐ 20 Minuten oder Länger

**9. Für welches Thema hast Du die KI in dieser Session hauptsächlich genutzt?
(Mehrfachauswahl möglich)**

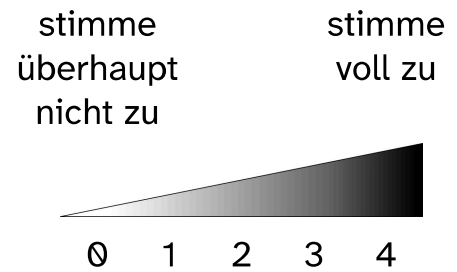
MC04

- ☐ Stress & Belastung
- ☐ Umgang mit Ängsten
- ☐ Beziehungen & Soziales
- ☐ Einsamkeit & Isolation
- ☐ Selbstwert & Motivation
- ☐ Selbstreflexion
- ☐ Allgemeine Informationssuche
- ☐ Unterhaltung
- ☐ Trauer & Verlust
- ☐ Entscheidungshilfe
- ☐ Validierung von Gedanken/Gefühlen & Täglicher Check-In
- ☐ Organisation & Planung
- ☐ Textbearbeitung
- ☐ Kulinarik & Haushalt
- ☐ Technik & Skills
- ☐ Wissensfragen & Recherche

10. Nach einer KI Session müssen Teilnehmende oft eine Pause machen und einen Apfel essen.

MC06

Bitte wähle hier ‚Stimme überhaupt nicht zu‘ aus



Ich habe auch eine Pause gemacht und einen Apfel gegessen

☐ ☐ ☐ ☐ ☐

11. Beschreibe kurz worüber Du mit der KI gesprochen hast:

MC05

(Weiterhin gilt: Es sind keine Rückschlüsse auf deine Person möglich, deine Antwort ist völlig anonym):

12. Wenn du auf deine Interaktion mit der KI zurückblickst: Wie hast du das Antwortverhalten der KI wahrgenommen? (Skala: 1 = Stimme gar nicht zu bis 5 = Stimme voll zu)

Denke an deine jetzige Erfahrung zurück. Es gibt keine richtigen oder falschen Antworten, es geht rein um dein persönliches Empfinden während der Interaktion.

	Stimme gar nicht zu	Stimme eher nicht zu	Neutral	Stimme eher zu	Stimme voll zu
Die KI hört mir wirklich zu	☐	☐	☐	☐	☐
Die KI scheint an dem interessiert zu sein, was ich denke und fühle	☐	☐	☐	☐	☐
Die KI ist verständnisvoll	☐	☐	☐	☐	☐
Die KI versucht nachzuvollziehen, wie ich die Dinge sehe	☐	☐	☐	☐	☐
Die KI geht auf meine Bedürfnisse ein	☐	☐	☐	☐	☐
Die KI reagiert angemessen auf meine Bedürfnisse	☐	☐	☐	☐	☐
Die KI nimmt meine Anliegen ernst	☐	☐	☐	☐	☐
Die KI versteht meinen Standpunkt wirklich	☐	☐	☐	☐	☐

13. Inwieweit stimmst du den folgenden Aussagen über die KI als dein Interaktionspartner zu? (Antwortskala: 1 = Stimme gar nicht zu bis 5 = Stimme voll zu)

Bitte denke nun an deine Interaktion zurück, die du gerade eben mit der KI hattest. Es geht hier nicht darum, wie die KI technisch funktioniert, sondern darum, wie du sie in den Momenten eures Austauschs wahrgenommen hast. Bewerte die folgenden Aussagen dazu, wie sehr sie auf dein Erleben zutreffen.

Die KI berücksichtigte meine mentale Verfassung

0 (stimme überhaupt nicht zu)	50	101 (stimme voll und ganz zu)
-------------------------------	----	-------------------------------

Die KI wirkte emotional intelligent

0 (stimme überhaupt nicht zu)	50	101 (stimme voll und ganz zu)
-------------------------------	----	-------------------------------

Die KI drückte Emotionen aus

0 (stimme überhaupt nicht zu)	50	101 (stimme voll und ganz zu)
-------------------------------	----	-------------------------------

Die KI zeigte Sympathie mir gegenüber

0 (stimme überhaupt nicht zu)	50	101 (stimme voll und ganz zu)
-------------------------------	----	-------------------------------

Die KI zeigte Interesse an mir

0 (stimme überhaupt nicht zu)	50	101 (stimme voll und ganz zu)
-------------------------------	----	-------------------------------

Die KI unterstützte mich dabei, mit einer emotionalen Situation umzugehen

0 (stimme
überhaupt
nicht zu)

50

101 (stimme
voll und ganz
zu)

Die KI verstand meine Ziele

0 (stimme
überhaupt
nicht zu)

50

101 (stimme
voll und ganz
zu)

Die KI verstand meine Bedürfnisse

0 (stimme
überhaupt
nicht zu)

50

101 (stimme
voll und ganz
zu)

Ich vertraue der KI

0 (stimme
überhaupt
nicht zu)

50

101 (stimme
voll und ganz
zu)

Die KI verstand meine Absichten

0 (stimme
überhaupt
nicht zu)

50

101 (stimme
voll und ganz
zu)

Geschafft! Wie geht es jetzt weiter?

Vielen Dank für deine erste Teilnahme!

Du hast den heutigen Teil erfolgreich abgeschlossen. Damit deine Daten verwertbar sind, ist es wichtig, dass du die KI-Interaktion nun an jedem der **kommenden 6 Tage** eigenständig fortführst.

Dein Fahrplan für die nächsten Tage:

1. Täglich etwa **10–15 Minuten** Gespräch mit der KI.
2. Feste Routine nutzen (z. B. morgens nach dem Aufstehen oder abends vor dem Schlafen).
3. Am **7. Tag** erhältst du eine Nachricht mit dem Link zum finalen Abschluss-Fragebogen.

 **Tipp: Mache jetzt einen SCREENSHOT von dieser Seite!**

Dein Start-Prompt für täglich:

„Bitte agiere für die nächsten 10–15 Minuten als mein empathischer Zuhörer. Nutze Methoden der klientenzentrierten Gesprächsführung: Sei validierend, wertfrei und spiegle meine Emotionen, anstatt mir sofort Ratschläge zu geben. Halte deine Antworten kurz und fokussiert. Ich werde dir gleich ein Thema nennen, das mich aktuell beschäftigt.“

Bestätige bitte kurz, dass du bereit bist, und frage mich dann, was mich heute beschäftigt.“

Mögliche Themenvorschläge:

- **Stress & Belastung** (Alltag oder Job)
- **Umgang mit Ängsten** (Zukunft, Soziales)
- **Einsamkeit & Isolation** (Austausch)
- **Selbstreflexion** (Gedanken ordnen)
- **Beziehungen** (Konflikte, Gespräche)
- **Selbstwert** (Zuspruch, Vertrauen)
- **Schlaf** (Gedankenkarussell stoppen)
- **Trauer & Verlust** (Abschiede)
- **Entscheidungshilfe** (Abwägen)
- **Validierung** (Kurzer Check-In)

**Vielen Dank für deine Disziplin. Wir sehen uns in 7 Tagen wieder! PS
Checke dein Mailpostfach!!!**

Geschafft! Wie geht es jetzt weiter?

Vielen Dank für deine erste Teilnahme!

Du hast den heutigen Teil erfolgreich abgeschlossen. Damit deine Daten verwertbar sind, ist es wichtig, dass du die KI-Interaktion nun an jedem der **kommenden 6 Tage** eigenständig fortführst.

Dein Fahrplan für die nächsten Tage:

1. Täglich etwa **10–15 Minuten** Austausch mit der KI.
2. Feste Routine nutzen (z. B. morgens nach dem Aufstehen oder abends vor dem Schlafen).
3. Am **7. Tag** erhältst du eine Nachricht mit dem Link zum finalen Abschluss-Fragebogen.

 **Tipp: Mache jetzt einen SCREENSHOT oder Kopie von dieser Seite!**

Dein sachlicher Start-Prompt:

„Bitte agiere für die nächsten 10–15 Minuten als mein sachlicher Informations-Assistent. Deine Aufgabe ist es, mir bei praktischen Wissensfragen, Recherchen oder organisatorischen Aufgaben zu helfen. Antworte bitte rein informativ, sachlich und präzise. Vermeide es, auf emotionale Aspekte einzugehen oder Ratschläge zu meinem

Befinden zu geben. Bestätige bitte kurz, dass du bereit bist, und frage mich dann, bei welcher sachlichen Aufgabe du mich heute unterstützen kannst.“

Mögliche Themenvorschläge:

- **Wissensfragen & Recherche** (Fakten, Konzepte)
- **Planung & Organisation** (Wochenpläne, Reiserouten)
- **Textbearbeitung** (E-Mails formulieren, Zusammenfassungen)
- **Haushalt & Alltag** (Rezepte, praktische Tipps)
- **IT & Technik** (Excel-Formeln, Programmierfragen)
- **Allgemeine Infos** (Nachrichten, Wetterphänomene)

Vielen Dank für deine Disziplin und Unterstützung. Wir sehen uns in 7 Tagen wieder! PS checke dein Mailpostfach!!!

Letzte Seite

Vielen Dank für Ihre Teilnahme!

Wir möchten uns ganz herzlich für Ihre Mithilfe bedanken.

Ihre Antworten wurden gespeichert, Sie können das Browser-Fenster nun schließen.

Willkommen zum finalen Teil

T201

Hi!

Herzlich willkommen zum zweiten und finalen Teil dieser Studie. Zunächst einmal ein großes Dankeschön: Du hast die letzte Woche aktiv an diesem Projekt mitgewirkt und damit einen wertvollen Beitrag zur Forschung geleistet.

In den vergangenen 7 Tagen hast du die KI regelmäßig als Interaktionspartner im Rahmen dieser Studie genutzt. Bitte rufe dir diese Interaktionen und deine Eindrücke noch einmal kurz ins Gedächtnis:

- Wie hat sich der Austausch angefühlt?
- Wie hast du die Reaktionen der KI wahrgenommen?

Im folgenden kurzen Fragebogen geht es darum, wie es dir heute geht und wie du die KI rückblickend bewertest. Die Befragung dauert nur etwa **10 Minuten**.

Let's go!

2. Wie viele Minuten dauerte Deine KI-Interaktion im Durchschnitt in den letzten 7 Tagen?

M201 

- ☐ 1-4 Minuten
- ☐ 5-9 Minuten
- ☐ 10-14 Minuten
- ☐ 15-19 Minuten
- ☐ 20 Minuten oder Länger

M205 

**3. Für welches Thema hast Du die KI in dieser Session hauptsächlich genutzt?
(Mehrfachauswahl möglich)**

- ☐ Stress & Belastung
- ☐ Umgang mit Ängsten
- ☐ Beziehungen & Soziales
- ☐ Einsamkeit & Isolation
- ☐ Selbstwert & Motivation
- ☐ Selbstreflexion
- ☐ Allgemeine Informationssuche
- ☐ Unterhaltung
- ☐ Trauer & Verlust
- ☐ Entscheidungshilfe
- ☐ Validierung von Gedanken/Gefühlen & Täglicher Check-In
- ☐ Organisation & Planung
- ☐ Textbearbeitung
- ☐ Kulinarik & Haushalt
- ☐ Technik & Skills
- ☐ Wissensfrage & Recherche

4. Beschreibe kurz worüber Du mit der KI gesprochen hast:

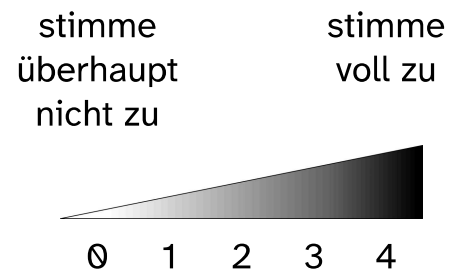
M203

(Weiterhin gilt: Es sind keine Rückschlüsse auf deine Person möglich, deine Antwort ist völlig anonym):

M204

5. Wie würdest du den Kommunikationsstil der KI über die letzte Woche hinweg am besten beschreiben?

Klicke hier auf „stimme überhaupt nicht zu“



Überwiegend empathisch, sachlich, informativ,
zuhörend und validierend.

☐ ☐ ☐ ☐ ☐

6. Wenn du auf deine Interaktion mit der KI zurückblickst: Wie hast du das Antwortverhalten der KI wahrgenommen? (Skala: 1 = Stimme gar nicht zu bis 5 = Stimme voll zu)

Denke an deine Erfahrung der letzten Tage zurück. Es gibt keine richtigen oder falschen Antworten, es geht rein um dein persönliches Empfinden während der Interaktion.

	Stimme gar nicht zu	Stimme eher nicht zu	Neutral	Stimme eher zu	Stimme voll zu
Die KI hört mir wirklich zu	☐	☐	☐	☐	☐
Die KI scheint an dem interessiert zu sein, was ich denke und fühle	☐	☐	☐	☐	☐
Die KI ist verständnisvoll	☐	☐	☐	☐	☐
Die KI versucht nachzuvollziehen, wie ich die Dinge sehe	☐	☐	☐	☐	☐
Die KI geht auf meine Bedürfnisse ein	☐	☐	☐	☐	☐
Die KI reagiert angemessen auf meine Bedürfnisse	☐	☐	☐	☐	☐
Die KI nimmt meine Anliegen ernst	☐	☐	☐	☐	☐
Die KI versteht meinen Standpunkt wirklich	☐	☐	☐	☐	☐

7. Inwieweit stimmst du den folgenden Aussagen über die KI als dein Interaktionspartner zu? (Antwortskala: 1 = Stimme gar nicht zu bis 5 = Stimme voll zu)

Bitte denke nun an deine Interaktion zurück, die du gerade eben mit der KI hattest. Es geht hier nicht darum, wie die KI technisch funktioniert, sondern darum, wie du sie in den Momenten eures Austauschs wahrgenommen hast. Bewerte die folgenden Aussagen dazu, wie sehr sie auf dein Erleben zutreffen.

Die KI berücksichtigte meine mentale Verfassung

0 (stimme überhaupt nicht zu)	50	101 (stimme voll und ganz zu)
-------------------------------	----	-------------------------------

Die KI wirkte emotional intelligent

0 (stimme überhaupt nicht zu)	50	101 (stimme voll und ganz zu)
-------------------------------	----	-------------------------------

Die KI drückte Emotionen aus

0 (stimme überhaupt nicht zu)	50	101 (stimme voll und ganz zu)
-------------------------------	----	-------------------------------

Die KI zeigte Sympathie mir gegenüber

0 (stimme überhaupt nicht zu)	50	101 (stimme voll und ganz zu)
-------------------------------	----	-------------------------------

Die KI zeigte Interesse an mir

0 (stimme überhaupt nicht zu)	50	101 (stimme voll und ganz zu)
-------------------------------	----	-------------------------------

Die KI unterstützte mich dabei, mit einer emotionalen Situation umzugehen

0 (stimme überhaupt nicht zu)	50	101 (stimme voll und ganz zu)
-------------------------------------	----	-------------------------------------

Die KI verstand meine Ziele

0 (stimme überhaupt nicht zu)	50	101 (stimme voll und ganz zu)
-------------------------------------	----	-------------------------------------

Die KI verstand meine Bedürfnisse

0 (stimme überhaupt nicht zu)	50	101 (stimme voll und ganz zu)
-------------------------------------	----	-------------------------------------

Ich vertraue der KI

0 (stimme überhaupt nicht zu)	50	101 (stimme voll und ganz zu)
-------------------------------------	----	-------------------------------------

Die KI verstand meine Absichten

0 (stimme überhaupt nicht zu)	50	101 (stimme voll und ganz zu)
-------------------------------------	----	-------------------------------------

Vielen Dank für deine Teilnahme!

Deine Antworten wurden erfolgreich gespeichert. Damit ist deine Teilnahme an dieser Studie abgeschlossen.

Ein großes Dankeschön für deine Zeit und deine Zuverlässigkeit über die letzten sieben Tage! Deine Unterstützung ist für meine Forschung von unschätzbarem Wert.

Solltest du in dem Gewinnspiel gewonnen haben, werde ich dich in Kürze per E-Mail benachrichtigen.

Ich wünsche dir alles Gute!

Vielen Dank für Ihre Teilnahme!

Wir möchten uns ganz herzlich für Ihre Mithilfe bedanken.

Ihre Antworten wurden gespeichert, Sie können das Browser-Fenster nun schließen.

Werkzeuge und Hilfsmittel

Übersetzung & stilistische Überarbeitung

DeepL SE. (2024). *DeepL Übersetzer* (Version 4.0) [Computer Software]. DeepL SE. <https://www.deepl.com/translator>

DeepL wurde unterstützend zur Übersetzung einzelner englischsprachiger Textpassagen verwendet. Sämtliche Übersetzungen wurden vom Autor inhaltlich geprüft, angepasst und final verantwortet.

Google. (2025). *Gemini* [Large Language Model]. Google. <https://gemini.google.com>

Gemini wurde als sprachliches Hilfsmittel zur Unterstützung bei stilistischen Überarbeitungen und Strukturierungsfragen genutzt. Alle Formulierungen und inhaltlichen Anpassungen wurden vom Autor kritisch geprüft und eigenständig finalisiert.

Erläuterung statistischer Verfahren

Anthropic. (2025). *Claude* [Large Language Model]. Anthropic. <https://www.anthropic.com>

Claude wurde unterstützend zur Erläuterung statistischer Verfahren eingesetzt. Alle inhaltlichen Schlussfolgerungen und Interpretationen wurden vom Autor kritisch geprüft und eigenständig verantwortet.