



HOCHSCHULE RUHR WEST
BWL - INTERNATIONALES HANDELSMANAGEMENT & LOGISTIK
MÜLHEIM A. D. RUHR

ABSCHLUSSARBEIT ZUR ERLANGUNG DES
AKADEMISCHEN TITELS BACHELOR OF ARTS (B.A.)

AUSWIRKUNGEN VON KI-GESTÜTZTEN CHATBOTS AUF
DIE CUSTOMER EXPERIENCE IM E-COMMERCE

Erstprüferin: Prof. Dr. rer. pol. Ellen Roemer
Zweitprüferin: Prof. Dr. Julia Thalmann-Ulshöfer
Eingereicht am: 10.01.2024

Mike Trynczyk
Matrikelnummer: 10013652

Gender-Hinweis

Zur besseren Lesbarkeit wird in dieser Arbeit das generische Maskulinum verwendet. Dabei sind andere Geschlechteridentitäten ausdrücklich mitgemeint, sofern es für die Aussage erforderlich ist.

Zusammenfassung

Aufgrund der rasanten Entwicklung von Künstlicher Intelligenz (KI) in den letzten Jahren werden im Electronic Commerce (E-Commerce) neue Potentiale geboten, um das Kundenerlebnis grundlegend zu verändern. Im Fokus stehen moderne Chatbots, die auf KI basieren und in der Lage sind, den menschlichen Mitarbeiter aus dem stationären Handel widerzuspiegeln sowie die Besucher eines Onlineshops in unterschiedlichen Anliegen zu unterstützen.

In dieser Arbeit werden die Auswirkungen von KI-gestützten Chatbots auf die Customer Experience im E-Commerce untersucht. Das Forschungsziel besteht darin, zu verstehen, welchen Einfluss ein textbasierter Chatbot eines Onlineshops auf die einzelnen Dimensionen der Kundenerfahrung hat. Hierbei liegt der Fokus auf den kognitiven, affektiven, sozialen und physischen beziehungsweise sensorischen Reaktionen der Kunden. Mithilfe eines Mixed-Methods-Designs in Form eines Laborexperiments werden Daten durch Eye-Tracking, einer Hautreaktionsmessung und einem begleitenden Fragebogen gesammelt, analysiert und interpretiert. Die Ergebnisse der Gruppe, die einen Chatbot zur Absolvierung der Aufgaben nutzt, werden einer Kontrollgruppe ohne Chatbot-Interaktion gegenübergestellt.

Die Auswertung der retrospektiven subjektiven Wahrnehmung der Versuchspersonen zeigt, dass signifikante Unterschiede zwischen den Gruppen feststellbar sind und der Chatbot in allen vier Dimensionen einen starken Effekt hat. Jede Erfahrungsdimension wurde von Chatbot-Nutzern höher und positiver bewertet im Vergleich zur Kontrollgruppe. Sowohl in den Hautreaktionsmesswerten als auch in den Eye-Tracking-Aufzeichnungen sind unterschiedliche Werte zwischen den Gruppen beobachtbar, jedoch kann kein signifikanter Unterschied in den ausgewählten Metriken festgestellt werden. Daher sind die Ergebnisse mit Vorsicht zu behandeln. Eine qualitative Analyse der Aufzeichnungen präsentiert deutlich, dass der Chatbot die Aufmerksamkeit der Nutzer auf sich lenkt und eine intensivere kognitive Verarbeitung auslöst. Außerdem können eindeutige Anstiege in den Hautleitwerten und somit in der emotionalen Reaktion der Nutzer während der Interaktion mit dem digitalen Assistenten festgestellt werden. Die Erkenntnisse tragen wesentlich zur Erforschung neuer Technologien im E-Commerce bei, indem sie aufzeigen, dass moderne Chatbots bedeutende Auswirkungen auf die Kundenerfahrung haben können. Diese werden von den Nutzern sowohl als sozialer und physisch angenehmer empfunden und lösen zugleich intensivere kognitive und emotionale Reaktionen aus im Vergleich zu Besuchern eines Onlineshops, die nicht mit einem Chatbot interagieren.

Inhaltsverzeichnis

Zusammenfassung	I
Abbildungsverzeichnis	IV
Tabellenverzeichnis.....	V
Abkürzungsverzeichnis	VI
Symbolverzeichnis	VII
1 Einleitung.....	1
1.1 Problemstellung und Herleitung der Forschungsfrage	1
1.2 Aufbau der Arbeit.....	3
2 Chatbots	3
2.1 Definition von Chatbots	3
2.2 Geschichte von Chatbots	4
2.3 KI-gestützte Chatbots	5
3 Forschungsüberblick	6
3.1 Aktueller Forschungsstand	6
3.2 Forschungslücke	11
4 Customer Experience.....	13
4.1 Kognitive Erfahrung.....	15
4.2 Affektive Erfahrung	16
4.3 Soziale Erfahrung	17
4.4 Physische Erfahrung.....	18
5 Empirische Analyse: Nutzung eines Chatbots im E-Commerce.....	19
5.1 Forschungsdesign	19
5.2 Aufbau und Durchführung der Untersuchung.....	20
5.3 Erhebungsmethoden	23
5.3.1 Neurophysiologische Messung	24
5.3.2 Eye-Tracking.....	25
5.3.3 Begleitender Fragebogen	26
6 Darstellung und Auswertung der Ergebnisse	28
6.1 Beschreibung der Stichprobe.....	28
6.2 Ergebnisse des Laborexperiments	29
6.2.1 Auswertung des Fragebogens	29

6.2.2	Auswertung der GSR-Messungen.....	32
6.2.3	Auswertung der Eye-Tracking-Aufzeichnungen	34
6.3	Hypothesentests.....	35
7	Diskussion der Ergebnisse.....	41
7.1	Interpretation der Ergebnisse.....	41
7.2	Kritische Diskussion der Methode und weiterer Forschungsbedarf	47
8	Fazit.....	48
	Literaturverzeichnis	51
	Anhang.....	57

Abbildungsverzeichnis

Abbildung 1: Dimensionen der Customer Experience	14
Abbildung 2: Darstellung Onlineshop	20
Abbildung 3: Digitaler Assistent (Chatbot)	21
Abbildung 4: Versuchsablauf	23
Abbildung 5: Mittelwerte – (<i>KO01</i>) und (<i>AF01</i>)	30
Abbildung 6: Mittelwerte – (<i>SO01</i>) und (<i>PH01</i>)	31
Abbildung 7: Skalenmittelwerte der Variablen <i>KO</i> , <i>AF</i> , <i>SO</i> , <i>PH</i>	36
Abbildung 8: Heatmaps (links mit und rechts ohne Chatbot)	42
Abbildung 9: Verlauf des Hautleitwerts (Proband 22: mit Chatbot)	44

Tabellenverzeichnis

Tabelle 1: Ausgewählte Studien zu Chatbots und Customer Experience	9
Tabelle 2: Items des Fragebogens	27
Tabelle 3: Finale Stichprobe.....	28
Tabelle 4: Hautreaktion - Messwerte.....	32
Tabelle 5: Hautreaktion - Mann-Whitney-U-Test.....	33
Tabelle 6: Eye-Tracking - ausgewählte Metriken	34
Tabelle 7: Eye-Tracking - Mann-Whitney-U-Test.....	35
Tabelle 8: T-Test Variable <i>KO</i>	37
Tabelle 9: T-Test Variable <i>AF</i>	38
Tabelle 10: T-Test Variable <i>SO</i>	39
Tabelle 11: T-Test Variable <i>PH</i>	40

Abkürzungsverzeichnis

AI	Artificial Intelligence
AIML	Artificial Intelligence Markup Language
ALICE	Artificial Linguistic Internet Computer Entity
API	Application Programming Interface (Programmierschnittstelle)
AR	Augmented Reality
BITKOM	Bundesverband Informationswirtschaft, Telekommunikation und neue Medien e.V.
CX	Customer Experience
DL	Deep Learning
E-Commerce	Electronic Commerce
f.	folgende
ff.	fortfolgende
GSR	Galvanic Skin Response (Galvanische Hautreaktion)
H	Hypothese
HDE	Handelsverband Deutschland
IFH	Institut für Handelsforschung
ISCX	In-store Customer Experience
KI	Künstliche Intelligenz
K-S-T	Kolmogorov-Smirnov-Test
LLMs	Large Language Models (große Sprachmodelle)
ML	Maschinelles Lernen
NLP	Natural Language Processing (natürliche Sprachverarbeitung)
S.	Seite
Sig.	Signifikanz
Std.	Standard
SEMs	Strategische Erlebnismodule
SST	Self-Service-Technologie
S-W-T	Shapiro-Wilk-Test
vgl.	vergleiche
z.B.	zum Beispiel

Symbolverzeichnis

α	Alpha
H_0	Nullhypothese
H_1	Alternativhypothese
mm	Millimeter
ms	Millisekunden
n	Stichprobenumfang
p	Signifikanzniveau
μS	Mikrosiemens
<	kleiner als
>	größer als

1 Einleitung

1.1 Problemstellung und Herleitung der Forschungsfrage

Über die letzten Jahre hat sich der Electronic Commerce (E-Commerce) zu einem unverzichtbaren Bestandteil der globalen Wirtschaft entwickelt. Die digitale Transformation und die Zunahme der Internetnutzung haben den E-Commerce gleichermaßen zu einem Schlüsselement für Unternehmen und Verbraucher gemacht (vgl. Basalisco et al. 2021, S. 9). Mit der COVID-19-Pandemie hat der E-Commerce zusätzlich an Bedeutung gewonnen, da immer mehr Verbraucher auf Online-Plattformen zurückgreifen, um ihre Einkäufe zu erledigen und Dienstleistungen in Anspruch zu nehmen (vgl. HDE 2021). Laut Prognosen wird der weltweite Umsatz im Markt für E-Commerce zwischen 2023 und 2027 weiterhin um 1,8 Billionen Euro (+53.71%) steigen und im Jahr 2027 einen neuen Höchststand von 5,2 Billionen Euro erreichen (vgl. Statista Market Insights 2023). Auch in Deutschland ist der Umsatz durch E-Commerce in den letzten Jahren rasant angestiegen und durch die pandemiebedingten Schließungen im Einzelhandel weiterhin angeregt worden. Während im Jahr 2015 der jährliche B2C-E-Commerce-Umsatz noch unter 40 Milliarden Euro lag, hat er sich 2022 mit etwa 84 Milliarden Euro mehr als verdoppelt (vgl. HDE 2023).

Des Weiteren sehen sich Unternehmen zunehmend mit der Herausforderung konfrontiert, innovative Technologien zu nutzen, um die steigenden Erwartungen ihrer Kunden zu erfüllen und das Kundenerlebnis zu optimieren. Dadurch differenzieren sie sich von anderen Wettbewerbern und verschaffen sich einen Wettbewerbsvorteil. Die Integration von Künstlicher Intelligenz (KI) im E-Commerce bietet Potential, die Kundenerfahrung zu transformieren und eine Wertsteigerung zu generieren. Dennoch müssen Unternehmen sicherstellen, dass diese Technologie effektiv genutzt wird, ohne das Kundenerlebnis zu beeinträchtigen (vgl. Bernhard/Mühling 2020, S. 5 f.; Hoyer et al. 2020, S. 57 f.).

Aufgrund der verstärkten Entwicklung in den letzten Jahren, ausgelöst durch Big Data, Cloud Computing und fortgeschrittene maschinelle Lernalgorithmen, hat KI als wegweisende Technologie das Potential, die Customer Experience (CX) im E-Commerce grundlegend zu verändern. Unter den vielfältigen Einsatzmöglichkeiten von KI im Onlinehandel rücken vor allem moderne KI-gestützte Chatbots in den Fokus, die eine direkte Interaktion mit dem Kunden ermöglichen und zu einem

wichtigen Schlüsselement werden können (vgl. Bawack et al. 2022, S. 298; Bernhard/Mühling 2020, S. 11 f.).

Jedoch stellen Chatbots keine neuartige Innovation dar. Sie werden bereits seit einigen Jahren im E-Commerce eingesetzt (vgl. Bogula 2022, S. 157). Eine repräsentative Befragung aus dem Jahr 2016 zeigt, dass die Mehrheit der Befragten bisher an der Zuverlässigkeit von herkömmlichen Chatbots zweifelte. Zudem war der Großteil der Onlineshop-Besucher der Meinung, dass die Technologie noch nicht ausgereift sei und sie nicht mit einem Computer kommunizieren wollen. Einige der Befragungspersonen zeigten sich hingegen zuversichtlich. Jeder Vierte gab an, in Zukunft Chatbots nutzen zu wollen (vgl. Bitkom 2017). Dies führte bislang zu einer zurückhaltenden Einstellung von Kunden und Unternehmen zur Nutzung dieser Technologie. Zahlreiche aktuelle Studien, die den Fokus auf innovativere KI-gestützte Chatbots legen, präsentieren dagegen, dass Kunden menschliche Servicemitarbeiter zwar immer noch als kompetenter einstufen, Chatbots die Kundenanfragen jedoch mindestens genau so zuverlässig und bis zu viermal effektiver beantworten können als menschliche Mitarbeiter (vgl. Lou et al. 2022; Luo et al. 2019). Zudem ist sich eine Vielzahl von Forschern darüber einig, dass KI-gestützte Chatbots zu den Technologien zählen, die den größten Einfluss auf die CX haben werden (vgl. Hoyer et al. 2020, S. 58 ff.). Auch in Deutschland haben dies bereits einige Unternehmen erkannt und testen seit 2023 verschiedene KI-Assistenten in ihren Onlineshops (vgl. Gondorf 2023). Immer mehr Kunden entdecken auch das Potential von Chatbots und deren zahlreichen Vorteile (vgl. EOS-Gruppe 2021). Es stellt sich jedoch die Frage, ob KI-gestützte Chatbots tatsächlich eine Bereicherung für die CX im E-Commerce darstellen und inwiefern sie sich auf die einzelnen Erfahrungsdimensionen auswirken (vgl. Bakkouri et al. 2022, S. 435 f.).

Das Ziel der vorliegenden Arbeit ist es, diese Forschungsfrage anhand der gewonnenen Daten einer selbstdurchgeführten empirischen Studie in Form eines Laborexperiments zu beantworten. Dabei werden die unterschiedlichen Dimensionen der CX herausgearbeitet und Hypothesen formuliert. Diese Hypothesen werden getestet und die Ergebnisse werden im Anschluss interpretiert.

1.2 Aufbau der Arbeit

Die vorliegende Arbeit ist in einen theoretischen und einen empirischen Teil gegliedert. Zudem ist sie in acht wesentliche Kapitel strukturiert. Das erste Kapitel beschreibt die aktuelle Situation und das vorherrschende Problem, woraus die Forschungsfrage und Zielsetzung der empirischen Studie hergeleitet werden. Im zweiten Kapitel wird der Begriff Chatbot definiert, die Geschichte von Chatbots wird beschrieben und deren Funktionsweise wird erläutert. Ein besonderes Augenmerk wird auf KI-gestützte Chatbots gelegt, die im Fokus dieser Arbeit stehen. Im Forschungsüberblick wird zunächst der aktuelle Forschungsstand thematisiert, bevor die Forschungslücke herausgearbeitet wird. Auf Basis einer selbstständigen Literaturrecherche werden aktuelle relevante Studien mit den Schwerpunkten Chatbot und E-Commerce aufgelistet und wesentlichen Erkenntnisse werden erläutert.

In Kapitel 4 wird der theoretische Rahmen der Customer Experience für die empirische Forschung festgelegt. Demnach wird der Begriff zunächst definiert und Möglichkeiten der Messung werden aufgezeigt. Der Fokus liegt auf den vier Erfahrungsdimensionen nach Schmitt (1999), deren Untersuchung und dem Zusammenhang mit der Chatbot-Nutzung. Es werden vier Hypothesen aufgestellt. Das darauffolgende Kapitel widmet sich der empirischen Analyse. Das Forschungsdesign sowie der Aufbau und die Durchführung des Laborexperiments werden erläutert und die drei ausgewählten Erhebungsmethoden werden vorgestellt.

Den zentralen Teil der Arbeit bilden Kapitel 6 und 7, in denen die Ergebnisse zunächst dargestellt und mittels statistischer Verfahren ausgewertet werden, um sie anschließend zu interpretieren. In der Fortsetzung wird die Methode einer kritischen Betrachtung unterzogen. Zum Schluss folgt ein Fazit, in welchem die Ergebnisse kurz zusammengefasst werden und die Forschungsfrage beantwortet wird, woraus abschließend Handlungsempfehlungen abgeleitet werden.

2 Chatbots

2.1 Definition von Chatbots

Künstliche Intelligenz integriert sich zunehmend in unterschiedlichsten Formen in den Alltag der Menschen. Der Anwendungsbereich erstreckt sich von einfachen Tätigkeiten bis hin zum Lösen von komplexen Aufgaben. Chatbots sind ein grundlegendes Beispiel für KI-Systeme und die meistgenannte Technologie in der

Mensch-Computer-Interaktion. Es ist ein von Menschen künstlich geschaffenes Konstrukt, um mit anderen Menschen zu interagieren (vgl. Adamopoulou/Moussiades 2020, S. 373). In der aktuellen Fachliteratur wird *Chatbot* folgendermaßen definiert:

„a computer program, which responds like a smart entity when conversed with through text or voice and understands one or more human languages by Natural Language Processing“ (Adamopoulou/Moussiades 2020, S. 373).

Demnach können Chatbots nicht nur textbasiert (z.B. Chat GPT), sondern auch sprachbasiert (z.B. Alexa, Siri) sein und auf vielen verschiedenen Sprachen kommunizieren. Chatbots sind in der Lage, menschliche Konversationen zu führen und dessen Benutzer zu unterhalten. Zudem finden sie in verschiedenen Bereichen Anwendung. Sie werden in der Informationsbeschaffung, im Bildungswesen als auch im E-Commerce genutzt (vgl. Abu Shawar/Atwell 2007, S. 35). Im digitalen Handel werden Chatbots als neue Evolutionsstufe betrachtet. Sie bieten das Potential, den Kundenservice grundlegend zu verändern und sind sowohl für Unternehmen als auch deren Kunden ein Gewinn, denn Unternehmen können ihre Kommunikation effizienter gestalten und Kunden erhalten schneller Antworten auf ihre Fragen und Anliegen (vgl. Heinemann 2018, S. 169; Hoyer et al. 2020, S. 60).

Darüber hinaus werden viele Synonyme für den Begriff Chatbot verwendet, dazu zählen unter anderem *Chatterbots, smarte Bots, interaktive Agenten, virtuelle, digitale* oder *intelligente Assistenten* (vgl. Adamopoulou/Moussiades 2020, S. 373; Bakkouri et al. 2022, S. 435).

2.2 Geschichte von Chatbots

Chatbots sind keine neue Innovation der letzten Jahre. Die Geschichte der Chatbot-Entwicklung reicht zurück in die 1960er Jahre, in denen der erste digitale Assistent entwickelt wurde. Im Jahr 1966 erstellte Joseph Weizenbaum ELIZA. Dieser hatte die Aufgabe, auf Eingaben zu reagieren und einfache psychotherapeutische Dialoge zu führen. Obwohl er nur begrenzte Fähigkeiten hatte, legte der regelbasierte Chatbot einen Grundstein für die spätere Entwicklung (vgl. Adamopoulou/Moussiades 2020, S. 374). Der nächste große Schritt war ALICE (Artificial Linguistic Internet Computer Entity), ein Chatbot, welcher 1995 von Richard Wallace entwickelt wurde und auf einem Musterabgleichsalgorithmus, der AIML (Artificial Intelligence Markup Language), basiert. Dies ist eine wichtige Sprache

in der technischen Entwicklung, die es ermöglicht, Regeln und Muster für Konversationen zu definieren (vgl. Abu Shawar/Atwell 2007, S. 30 ff.; Adamopoulou/Moussiades 2020, S. 376).

Mit dem Aufkommen von maschinellem Lernen (ML) und Künstlicher Intelligenz in den letzten Jahren haben Chatbots einen bedeutenden Fortschritt gemacht. Moderne Chatbots, sogenannte KI-gestützte Chatbots, nutzen Deep Learning (DL), um aus großen Datensätzen zu lernen und natürliche Sprache besser zu verstehen. Auch im digitalen Handel wird die Entwicklung von Chatbots vorangetrieben. Große Entwicklerunternehmen wie OpenAI, Meta, Microsoft oder Google sowie kleinere lokale Anbieter, wie beispielweise moinAI, melibo oder Userlike, bieten ihre Chatbot-Lösungen für digitale Handelsunternehmen an. Die Integration eines Chatbots in den Onlineshop ist so einfach wie nie zuvor. Mittels Programmierschnittstellen (API) können die Anwendungen in kurzer Zeit in die Internetseiten, Messenger-Dienste und mobile Apps implementiert und mit Daten trainiert werden (vgl. Gentsch 2019, S. 96; Heinemann 2018, S. 169 f.).

2.3 KI-gestützte Chatbots

Grundsätzlich wird zwischen zwei Arten beziehungsweise Funktionsweisen von digitalen Assistenten differenziert. Einerseits gibt es regelbasierte Chatbots, die mit vorab festgelegten Mustern arbeiten, andererseits gewinnen moderne Assistenten, die auf KI basieren, an Bedeutung (vgl. Adamopoulou/Moussiades 2020, S. 374; Gentsch 2019, S. 31 ff.). KI-gestützte Chatbots sind ein Produkt der jüngsten Fortschritte in den Bereichen des maschinellen Lernens, natürlicher Sprachverarbeitung (NLP – Natural Language Processing) und Datenanalyse. Sie sind intelligente Systeme, die in der Lage sind, den geschriebenen oder gesprochenen Input mithilfe von KI zu analysieren und den passenden Output wiederzugeben (vgl. Adamopoulou/Moussiades 2020, S. 377 f.).

Dabei stellt NLP den wichtigsten technischen Bestandteil dar. Es handelt sich um eine Disziplin der Computerlinguistik, die sich vor allem mit der Semantik, also mit der Interpretation des Geschriebenen, beschäftigt. Da diese nicht in jeder Sprache einheitlich ist und von kulturellen Unterschieden, Grammatik und Intention abhängt, stellt sie derzeit eine der größten Herausforderungen dar. Der Fortschritt

in dieser Disziplin wird jedoch durch schnellere Hardware und Big Data¹ begünstigt, indem solche Chatbot-Systeme mit einer Vielzahl an Daten in einer vergleichsweise kurzen Zeit trainiert werden können. Chatbots, die auf KI basieren, können größtenteils selbstständig mithilfe von Algorithmen aus riesigen Datenmengen lernen und sich stetig verbessern (vgl. Gentsch 2019, S. 31 f.).

Im E-Commerce können derartige Chatbots verschiedene Funktionen übernehmen. Am häufigsten werden sie als digitale Assistenten eingesetzt, die den Mitarbeiter im stationären Handel widerspiegeln sollen. Sie helfen bei der Beantwortung von Produktfragen, empfehlen Produkte und unterstützen bei der Kaufabwicklung. Die digitalen Assistenten basieren zum einen auf allgemeinen Sprachmodellen, die es ermöglichen, die Sprache zu verstehen. Zum anderen analysieren sie große Mengen an Kundendaten, wie beispielsweise Präferenzen der Kunden oder Suchverläufe im Onlineshop, um spezifische Fragen besser zu verstehen und relevante Informationen bereitzustellen (vgl. Gentsch 2019, S. 97 ff.; Hoyer et al. 2020, S. 58).

3 Forschungsüberblick

3.1 Aktueller Forschungsstand

In den letzten drei Jahrzehnten wurde eine Vielzahl von wissenschaftlichen Studien zum Thema Künstlicher Intelligenz im E-Commerce durchgeführt. Während es 2013 noch 160 Veröffentlichungen waren, steigt seitdem die Anzahl der weltweiten Publikationen in diesem Forschungsgebiet mit einer durchschnittlichen jährlichen Wachstumsrate von 10,45% auf 522 Veröffentlichungen im Jahr 2019. Es wird erwartet, dass dieser Trend weiter anhält. Zu den Vorreitern, gemessen an der Anzahl an Publikationen, gehören überwiegend China (4261) gefolgt von USA (1481), Indien (1397) und Südkorea (379). Dies kann damit begründet werden, dass dort die Entwicklung und die Bereitschaft zur Nutzung von KI fortgeschrittener ist als in Europa. Das größte Interesse in diesem Forschungsgebiet liegt auf KI-gestützten Empfehlungssystemen jeglicher Art. Jedoch wurde ebenfalls eine Vielzahl an Studien zum Einsatz von Chatbots im E-Commerce durchgeführt (vgl. Bakkouri et al. 2022, S. 434 f.; Bawack et al. 2022, S. 300 ff.; Bitkom 2022).

¹ Big Data umfasst große und komplexe Datenmengen, die traditionelle Software überfordern. Für die Auswertung und Analyse von Big Data werden vielzählige KI-Methoden verwendet (vgl. Gentsch 2019, S. 8 f.).

In Tabelle 1 werden relevante Studien aus den letzten vier Jahren aufgelistet, in denen der Schwerpunkt auf Chatbots und E-Commerce liegt. Die Studien sind nach angewandter Forschungsmethode sortiert. Die Tabelle basiert auf einer selbstständigen Literaturrecherche, die mit den Suchmaschinen Google Scholar, EBSCO und ScienceDirect von Elsevier durchgeführt wurde. Die Suche wurde durch die Nutzung der KI-gestützten Suchmaschinen Semantic Scholar und Connected Papers unterstützt. Es wurde nach den Schlüsselwörtern „Chatbot“ in Kombination mit „Customer Experience“ gesucht. Es kamen nur Studien in Betracht, die in renommierten Fachzeitschriften veröffentlicht, von Fachleuten begutachtet und in englischer oder deutscher Sprache verfasst worden sind. Jedoch ist anzumerken, dass keine in deutscher Sprache verfasste Studie gefunden wurde. Zudem wurden alle Studien, die vor dem Jahr 2019 veröffentlicht worden sind, aussortiert, um die Aktualität zu gewährleisten. Zusätzlich wurden nach dem Lesen des Abstracts weitere Arbeiten aussortiert, in denen Chatbots oder E-Commerce keine wesentlichen Bestandteile der Studie waren. Nach dem Lesen der gesamten Studien wurde eine nach Relevanz festgelegte Auswahl in die Tabelle aufgenommen (vgl. Bakkouri et al. 2022, S. 433 f.; Bawack et al. 2022, S. 298 f.).

Im Folgenden werden die Forschungsziele und Erkenntnisse der ausgewählten Studien zusammengefasst. Bislang wurde die Auswirkung der allgemeinen Einstellung zu neuen Technologien auf Unternehmen und die Marke untersucht. Es wurde festgestellt, dass die Nutzung eines Chatbots abhängig von der allgemeinen Einstellung zu neuen Technologien ist. Jedoch konnte in dieser Feldstudie kein Zusammenhang zwischen der Chatbot-Nutzung und Markenimage festgestellt werden (vgl. Moriuchi et al. 2021). Andere Studien haben sich mit weiteren Aspekten beschäftigt, darunter der Einfluss der Aufgabenschwierigkeit und die Offenlegung von Chatbots auf das Vertrauen der Verbraucher (vgl. Li/Wang 2023). Außerdem wurden die Auswirkungen des Sprachstils hinsichtlich Empathie und Freundlichkeit auf das Nutzererlebnis untersucht. Cheng et al. (2022) haben festgestellt, dass sich einerseits die Empathie und Freundlichkeit von Chatbots positiv auf das Vertrauen der Nutzer auswirkt. Andererseits beeinflussen Aufgabenschwierigkeit und Offenlegung von Chatbots das Vertrauen der Kunden negativ. Dies haben Xu et al. (2020) ebenfalls bestätigen können. Bei Aufgaben mit einer geringeren Komplexität wird die Problemlösungsfähigkeit eines Chatbots von den Kunden als größer

betrachtet als die des menschlichen Kundenservices. In dem Fall tendieren Kunden eher zur Nutzung der KI. Jedoch bewerteten Kunden den Nutzen des menschlichen Kundenservices höher bei komplexeren Anfragen (vgl. Xu et al. 2020). Li und Wang (2023) kamen zu der Erkenntnis, dass zwischen dem Sprachstil eines Chatbots und der Absicht der Kunden zur Fortsetzung der Nutzung eines Chatbots ein signifikanter Zusammenhang besteht. Das Ergebnis zeigt, dass die fortgesetzte Nutzung der Kunden zunimmt, wenn Chatbots einen informellen Sprachstil annehmen. Es konnte ebenfalls ein positiver Effekt auf die Einstellung gegenüber der Marke nachgewiesen werden. Liebrecht et al. (2021) untersuchten in einer Studie die Auswirkungen von informalen Chatbot-Antworten auf die Bewertung der Nutzer bezüglich der sozialen Präsenz, Angemessenheit, Markeneinstellung und Qualität der Interaktion. Sie konnten den positiven Effekt auf das von Kunden bewertete Markenimage bestätigen. Zudem konnte ein indirekter Effekt auf die soziale Präsenz gefunden werden. In einer weiteren Studie konnte Trivedi (2019) zusätzlich untermauern, dass Unternehmen ihre Beziehung zwischen den Verbrauchern und der Marke festigen können, wenn digitale Assistenten, welche die Erwartungen der Nutzer erfüllen, genutzt werden. Außerdem untersuchten Lo Presti et al. (2021), wie sich die Nutzung von Chatbots auf den Kaufprozess und die Kaufabsicht der Nutzer auswirkt. Es wurde festgestellt, dass Chatbots in der Lage sind, die Kaufabsicht der Kunden positiv zu beeinflussen. Luo et al. (2019) erforschten, inwiefern die Offenlegung des Chatbots die Kaufabsicht der Kunden beeinflusst. Nach einer Analyse von über 6.200 Feldexperimentdaten kamen sie zu dem Entschluss, dass nicht offengelegte sprachgesteuerte Chatbots in Verkaufsgesprächen genau so effektiv sind wie menschliche Verkaufsmitarbeiter und vielmal effektiver als unerfahrene Mitarbeiter. Jedoch führt eine Offenlegung des Chatbots zu einer erheblichen Reduktion der Verkaufsraten. Zudem haben Lou et al. (2022) herausgefunden, dass Kunden sich während einer Chatbot-Interaktion anders benehmen als bei der Interaktion mit einem menschlichen Mitarbeiter. Im Zusammenhang mit der Kontingenz und dem Mitgefühl des Gegenübers (Mensch oder Chatbot) werden Menschen von den Kunden als kompetenter eingestuft.

Tabelle 1: Ausgewählte Studien zu Chatbots und Customer Experience

Auto(en), Jahr	Quelle	Titel/Thema	Forschungsmethode	Unabhängige Variable	Abhängige Variable
Moriuchi et al. 2021	Journal of Strategic Marketing	„Engagement with chatbots versus augmented reality interactive technology in e-commerce“	Empirisch (Feldstudie)	Nutzung Chatbots & Augmented Reality (AR)	Markenimage
Luo et al. 2019	Marketing Science	„Frontiers: Machines vs. Humans: The Impact of Artificial Intelligence Chatbot Disclosure on Customer Purchases“	Empirisch (Feldstudie)	Nutzung eines Chatbots (sprachgesteuert)	Kaufabsicht der Kunden
Lo Presti et al. 2021	Italian Journal of Marketing	„The role of the chatbot on customer purchase intention: towards digital relational sales“	Empirisch (Laborversuch)	Nutzung eines Chatbots	Kaufprozess und Kaufabsicht
Li/Wang 2023	Journal of Retailing and Consumer Services	„Chatbots in e-commerce: The effect of chatbot language style on customers' continuance usage intention and attitude toward brand“	Empirisch (Laborversuch)	Sprachstil eines Chatbots	Nutzungsabsicht und Einstellung gegenüber der Marke
Liebrecht et al. 2021	Chatbot Research and Design	„Too Informal? How a Chat-bot's Communication Style Affects Brand Attitude and Quality of Interaction“	Empirisch (Online-Experiment)	Sprachstil eines Chatbots	Markeneinstellung und Qualität der Interaktion

Fortsetzung auf nächster Seite

Tabelle 1: Fortsetzung

Auto(en), Jahr	Quelle	Titel/Thema	Forschungsmethode	Unabhängige Variable	Abhängige Variable
Lou et al. 2022	International Journal of Advertising	„Bots vs. humans: how schema congruity, contingency-based interactivity, and sympathy influence consumer perceptions and patronage intentions“	Empirisch (Online-Experiment)	Kontingenz und Mitgefühl von Chatbots	Kompetenzbewertung der Kunden
Xu et al. 2020	Australasian Marketing Journal	„AI Customer Service: Task Complexity, Problem-Solving Ability, and Usage Intention“	Empirisch (Online-Feldexperiment)	Aufgabenkomplexität	Nutzungsabsicht der Kunden
Trivedi 2019	Journal of Internet Commerce	„Examining the Customer Experience of Using Banking Chatbots and Its Impact on Brand Love: The Moderating Role of Perceived Risk“	Empirisch (Quantitative Umfrage)	Nutzung eines Chatbots	Einstellung gegenüber der Marke und Markenimage
Cheng et al. 2022	Internet Research	„Exploring consumers' response to text-based chatbots in e-commerce: the moderating role of task complexity and chatbot disclosure“	Empirisch (Quantitative Umfrage)	Aufgabenschwierigkeit und Offenlegung von Chatbots	Vertrauen der Verbraucher

Quelle: Eigene Darstellung basierend auf eigener Erhebung

Wie Tabelle 1 widerspiegelt, basiert der Großteil der Erkenntnisse auf empirischen Forschungsmethoden. Sowohl in qualitativen Laborversuchen als auch in quantitativen (Online-)Umfragen wurde untersucht, wie sich die Nutzung und verschiedene Aspekte eines digitalen Assistenten auf dessen Nutzer beziehungsweise auf die Kundenerfahrung auswirkt. Ein weiteres Mittel der Wahl ist vor allem ein Mixed-Methods-Design in Form von Feldstudien und Experimenten mit einer höheren Stichprobenanzahl. Auf diese Weise werden qualitative und quantitative Daten kombiniert, um optimale Ergebnisse zu erzielen. Im Gegensatz dazu lassen sich nur wenige Veröffentlichungen finden, die auf einer Literaturanalyse basieren. Dies kann auf die Aktualität des Forschungsthemas zurückzuführen sein. Angesichts des aktuellen Aufschwungs der Themen KI und Chatbots, insbesondere durch die Veröffentlichung bedeutender Sprachmodelle (LLMs – Large Language Models) wie zum Beispiel ChatGPT von Open AI, ist zu erwarten, dass die Forschung in diesem Gebiet steigen wird (vgl. OpenAI 2023). Aufgrund der rasanten Entwicklung und des Fortschritts in diesem Bereich bedarf es einer fortführenden Auswertung der Situation (vgl. Bakkouri et al. 2022, S. 436).

3.2 Forschungslücke

Es ist nicht verwunderlich, dass der Großteil der wissenschaftlichen Veröffentlichungen aus Asien und den USA stammen (vgl. Bawack et al. 2022, S. 301). Unternehmen wie Amazon Inc. (USA), JD.com (China), Alibaba Group Holding Ltd. (China), Wayfair (USA), eBay (USA) gehören, am Umsatz gemessen, zu den größten E-Commerce-Unternehmen weltweit und haben großes Interesse an der Weiterentwicklung dieser Technologien (vgl. Bawack et al. 2022, S. 307 f.). Damit auch europäische beziehungsweise deutsche Unternehmen wettbewerbsfähig bleiben, sollte in Europa mehr Forschung in diesem Gebiet geleistet werden. Ein weiterer wichtiger Aspekt, der berücksichtigt werden sollte, ist, dass sich die europäischen Märkte in ihren kulturellen und wirtschaftlichen Rahmenbedingungen deutlich von anderen Teilen der Welt unterscheiden. Deshalb ist es wichtig, die Auswirkungen von KI-gestützten Technologien spezifisch für den europäischen beziehungsweise den deutschen Markt zu untersuchen (vgl. Bawack et al. 2022, S. 308).

Es empfiehlt sich, den Fokus auf den Einsatz aktueller fortgeschrittener KI-Technologien im E-Commerce zu legen. Dazu zählen unter anderem Chatbots, die

mit dem Nutzer in natürlicher Sprache kommunizieren und als neue „Gesprächsempfehlungssysteme“ eingesetzt werden können (vgl. Bawack et al. 2022, S. 308; Iovine et al. 2020, S. 1; Ngai et al. 2021, S. 12). Angesichts des gegenwärtigen technischen Fortschritts von KI-gestützten Chatbots haben immer mehr Unternehmen ein zunehmendes Interesse daran, diese Technologie als Touchpoint² in die Customer Journey³ zu implementieren und somit die CX zu optimieren. Denn diese Technik kann Unternehmensabläufe nicht nur effizienter machen und zur Kosteneinsparungen führen, sondern kann sich auch direkt auf die Kunden auswirken und deren Einkaufserlebnis erheblich beeinflussen (vgl. Bakkouri et al. 2022, S. 435 f.).

Während in den letzten Jahren im Wesentlichen experimentelle Studien durchgeführt worden sind, in denen insbesondere der Einfluss von Chatbots auf den Kaufprozess, die Kaufabsicht und das Markenimage untersucht wurde (vgl. Tabelle 1), fehlen umfassende Analysen der Auswirkungen von KI-gestützten Chatbots auf die direkte Kundenerfahrung. Zwar wurden vereinzelt Studien zum Einfluss auf einzelne Aspekte durchgeführt, jedoch mangelt es an Studien, die die holistische CX der Kunden betrachten (vgl. Bakkouri et al. 2022, S. 434 ff.; Verhulst et al. 2020, S. 2). Diese Arbeit soll helfen, diese Forschungslücke zu minimieren und wichtige Erkenntnisse liefern, inwiefern KI-gestützte Chatbots die CX im E-Commerce beeinflussen.

Praktisch betrachtet können die gewonnenen Erkenntnisse aus der Schließung dieser Forschungslücke wichtige Erkenntnisse für Unternehmen mit sich bringen, die ihre Wettbewerbsfähigkeit stärken wollen. Im heutigen (Online-)Handel ist die CX ein Schlüsselement, um sich von anderen Wettbewerbern zu unterscheiden und Kunden an sich zu binden (vgl. Bustamante/Rubio 2017, S. 885). Ein tieferes Verständnis von KI-gestützten Chatbots und deren Auswirkungen kann dazu dienen, gezielte Strategien zu entwickeln, um personalisierte Einkaufserlebnisse zu schaffen und die Effizienz im Unternehmen zu erhöhen. Darüber hinaus wird durch die Analyse der einzelnen Dimensionen aufgeschlüsselt, in welchem Maße be-

² Touchpoints (Berührungspunkte) sind die verschiedenen Interaktionspunkte zwischen einem Kunden und einer Marke entlang der Customer Journey (vgl. Lemon/Verhoef 2016, S. 76 f.).

³ Als Customer Journey (Kundenreise) wird der gesamte Prozess eines Kunden mit einem Unternehmen oder einer Marke bezeichnet (vgl. Lemon/Verhoef 2016, S. 74 ff.).

stimmte Bereiche von digitalen Assistenten beeinflusst werden und in welchen Fällen eine Ergänzung mithilfe anderer Touchpoints ratsam wäre. Es wird die Chance geboten, die Kundenzufriedenheit zu steigern und somit eine engere Kundenbindung, aufgrund eines optimierten Kundenerlebnisses, zu schaffen (vgl. Bernhard/Mühling 2020, S. 5 f.; Hoyer et al. 2020, S. 57 f.).

4 Customer Experience

Die Customer Experience, auch Kundenerfahrung oder Kundenerlebnis genannt, wird in der einschlägigen Literatur auf unterschiedliche Weise definiert. Zusammenfassend kann gesagt werden, dass die CX eine ganzheitliche Betrachtung aller internen und subjektiven Reaktionen der Kunden auf jeden direkten oder indirekten Kontakt mit einem Unternehmen ist (vgl. Bustamante/Rubio 2017, S. 885; Lemon/Verhoef 2016, S. 70). Verhoef et al. (2009) erweitern den Gedanken und präzisieren Kundenerfahrung folgendermaßen:

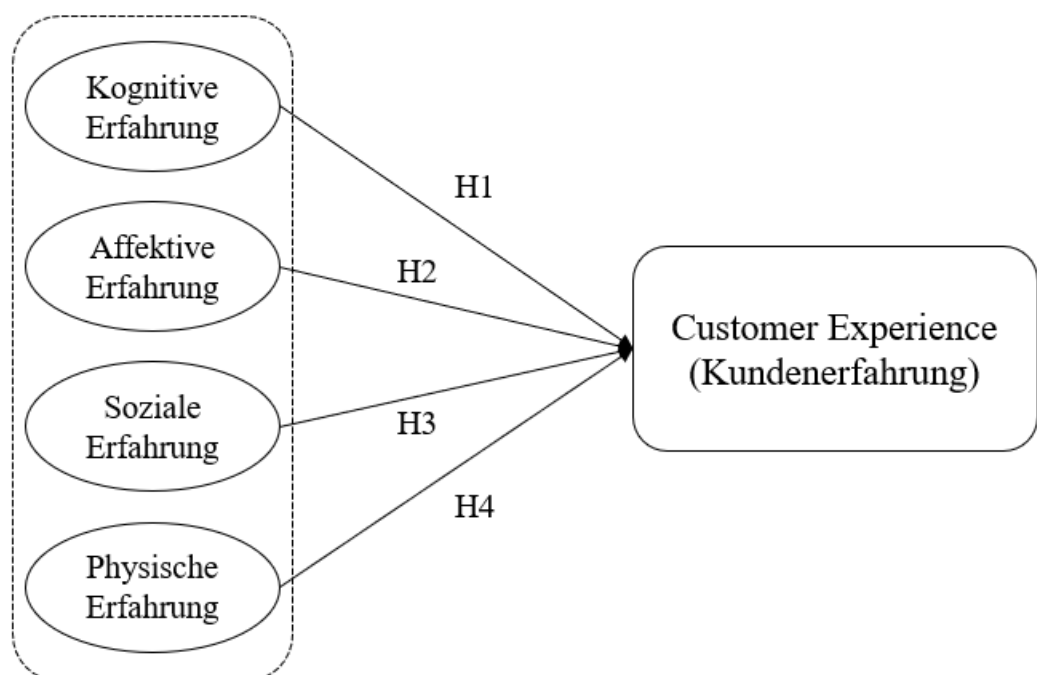
“[...] the customer experience construct is holistic in nature and involves the customer’s cognitive, affective, emotional, social and physical responses to the retailer.” (Verhoef et al. 2009, S. 32).

Weiterhin kann Erfahrung als Zustand definiert werden, der bei einer Person als Reaktion auf einen Reiz erzeugt wird. Sie ist nicht nur eine Wahrnehmung von Reizen, sondern auch eine Interpretation dieser (vgl. Poulsson/Kale 2004, S. 270 ff.). Daher setzt sich das Kundenerlebnis aus allen kognitiven, emotionalen, sozialen und physischen Reaktionen eines Kunden mit den zahlreichen Touchpoints einer Customer Journey zusammen. Demnach können die Berührungspunkte als Reize betrachtet werden (vgl. Lemon/Verhoef 2016, S. 70; Meyer/Schwager 2007, S. 117 ff.).

Ein wesentliches Vorgehen, um die Customer Experience zu messen, ist die Möglichkeit, die Reaktionen der Kunden zu erfassen. Seit Jahren versuchen Forscher und Unternehmen optimale Messansätze zu finden und Modelle zu entwickeln, die idealerweise die gesamte Kundenerfahrung und das Erlebnis in der jeweiligen Phase der Customer Journey und für alle Touchpoints ermitteln (vgl. Bustamante/Rubio 2017, S. 885 ff.; Lemon/Verhoef 2016, S. 71 / 80 f.). Ein weit verbreiteter Ansatz zur Messung der CX basiert auf dem Konzept von Schmitt

(1999), der die unterschiedlichen Dimensionen der Reaktionen als *strategische Erlebnismodule* (SEMs) definiert hat. Er betrachtet die Kundenerfahrung als Konstrukt, das unter anderem aus den Komponenten der sowohl kognitiven und affektiven Erfahrung als auch sozialer und physischer Erlebnisse besteht (vgl. Bustamante/Rubio 2017, S. 889 ff.; Schmitt 1999, S. 60 ff.). In der Forschung fand dieser Ansatz der vier Dimensionen bereits zahlreiche Anwendungen (vgl. Hermes/Riedl 2021, S. 76 ff.). Beispielweise nutzten Bleier et al. (2019) das Konzept, um den Zusammenhang der Customer Experience mit der Produktseite zu messen. Ebenfalls analysierten Bagdare und Jain (2013) die Customer Experience im stationären Handel mithilfe der SEMs, jedoch lag der wesentliche Fokus dieser Untersuchung auf den affektiven Erlebnissen der Kunden. Auch Bustamante und Rubio (2017) übernahmen Schmitts strategische Erlebnismodule als theoretischen Rahmen zur Entwicklung eines Instruments zur Messung der Kundenerfahrung im stationären Handel (ISCX). Sie verfolgten einen phänomenologischen Ansatz, um die CX als Ergebnis eines aktiven Erlebens zu definieren. Die Kundenzufriedenheit und Loyalität können demnach die Ergebnisse eines positiven aktiven Erlebens sein (vgl. Bustamante/Rubio 2017, S. 886 ff.; Lemon/Verhoef 2016, S. 71; Schmitt 1999, S. 61 f.).

Abbildung 1: Dimensionen der Customer Experience



Quelle: Eigene Darstellung in Anlehnung an Roy et al. (2022), S. 46

Basierend auf der Literatur analysiert die Studie im Rahmen dieser Arbeit die CX eines KI-gestützten Chatbots als gesamteinheitliches Konstrukt, welches alle vier zuvor genannten Dimensionen miteinbezieht, die in Abbildung 1 visualisiert werden. Als Basis werden die von Schmitt (1999) vorgestellten SEMs genutzt und es wird sich auf die vier wesentlichen Module fokussiert. In der Studie werden verschiedene Erhebungsmethoden genutzt, um die internen Prozesse bezüglich der einzelnen Dimensionen in Abhängigkeit der Chatbot-Nutzung zu messen. Die Datenerhebungsverfahren und deren Anwendung werden in Kapitel 5.3 genauer erläutert. In den folgenden Kapiteln 4.1 bis 4.4 werden die vier Erfahrungsdimensionen genauer definiert und jeweilige Hypothesen (H1 bis H4) aufgestellt. Die Hypothesen dienen dazu, das Forschungsthema zu untersuchen. Sie leiten sich aus der Forschungsfrage ab und sollen dabei helfen, diese zu beantworten. Mit den erhobenen Daten aus dem Laborversuch werden die Hypothesen im weiteren Verlauf der Arbeit getestet.

4.1 Kognitive Erfahrung

Kognition ist die Fähigkeit, Informationen wahrzunehmen, Daten zu verarbeiten, Lösungen zu finden und Entscheidungen zu treffen. Die kognitive Erfahrungsebene der CX bezieht sich auf die mentalen Prozesse, die ein Kunde erlebt, wenn er sich mit den Dienstleistungen, Produkten oder anderen Aspekten eines Unternehmens oder einer Marke auseinandersetzt (vgl. Da Silva/Syed 2006, S. 293 f.). Eine Vielzahl von Forschern ist sich darüber einig, dass Kunden in erster Linie eine kognitive Erfahrung machen, bevor es zu einer affektiven Erfahrung kommt, da die wahrgenommenen Informationen, beispielweise eines Services in Form eines Chatbots, zuerst verarbeitet werden müssen, bevor sich Einstellung und Emotionen im Kunden bilden können (vgl. Bleier et al. 2019, S. 99; Bustamante/Rubio 2017, S. 889 f.).

Die kognitive Erfahrung ist viel mehr als der Erwerb von Wissen. Sie besteht aus mentalen Reaktionen, welche durch Reize hervorgerufen werden. Diese versuchen, Kunden auf verschiedene Weise anzusprechen (vgl. Schmitt 1999, S. 61 f.). Einer dieser Reize können KI-gestützte Chatbots im E-Commerce sein. Sie bieten die Möglichkeit, informative und personalisierte Interaktionen mit Kunden zu füh-

ren. Indem sie in der Lage sind, Fragen gezielt zu beantworten, wesentliche Informationen bereitzustellen und Kunden bei Entscheidungen zu unterstützen, bieten sie den Nutzern einen kognitiven Mehrwert (vgl. Hoyer et al. 2020, S. 63 f.). Infolgedessen kann die Annahme getroffen werden, dass die Besucher eines Onlineshops, die mit einem Chatbot interagieren, aktivere Teilnehmer am Informationsaustausch sind als Besucher ohne Chatbot-Interaktion, was sich wiederum in intensiveren kognitiven Reaktionen und Erfahrungen widerspiegeln kann (vgl. Lou et al. 2022, S. 673 ff.; Xu et al. 2020, S. 190 f.). Daraus abgeleitet, lautet die erste Forschungshypothese (H1):

H1: Besucher eines Onlineshops, die mit einem Chatbot interagieren, weisen im Vergleich zu Besuchern ohne Chatbot-Interaktion ein höheres Maß an kognitiven Erfahrungen auf.

4.2 Affektive Erfahrung

Die affektive Erfahrung kann als Gefühlszustand definiert werden, der bei einem Kunden entsteht, wenn dieser mit dem Unternehmen oder der Marke interagiert (vgl. Erevelles 1998, S. 199). Die Affektivität umfasst nicht nur positive und negative Emotionen, sondern auch Stimmungslagen. Dennoch ist es bei der Erforschung der Kundenerfahrung effektiver, sich auf die hervorgerufenen Emotionen während eines Erlebnisses zu konzentrieren, da diese in der Regel intensiver sind. Um solche Erfahrungen zu schaffen, ist es erforderlich zu erforschen, wie bestimmte Reize die Intensität von Emotionen beeinflussen (vgl. Bustamante/Rubio 2017, S. 890; Schmitt 1999, S. 61).

Im Zusammenhang mit Chatbots spielen Empathie, Verständnis und die Fähigkeit, eine positive emotionale Bindung zum Kunden zu schaffen, eine wesentliche Rolle. Sie werden üblicherweise als benutzerfreundliche Werkzeuge konzipiert, die Kunden das Gefühl von individueller Unterstützung und Aufmerksamkeit bieten können (vgl. Bleier et al. 2019, S. 99; Bogula 2022, S. 158 f.). Einerseits kann sich diese persönlichere Interaktion auf eine positive emotionale Bindung auswirken. Andererseits können KI-gestützte Chatbots eine positivere Affektivität hervorrufen oder zumindest Unzufriedenheit verhindern, indem sie Anfragen schneller und zu-

friedenstellender bearbeiten als es der Kunde selbst oder mithilfe von Dritten erledigen kann (vgl. Bogula 2022, S. 160; Hoyer et al. 2020, S. 63; Jenneboer et al. 2022, S. 224 f.). Dies führt zu einer zweiten Forschungshypothese (H2):

H2: Besucher eines Onlineshops, die mit einem Chatbot interagieren, weisen im Vergleich zu Besuchern ohne Chatbot-Interaktion ein höheres Maß an positiver Affektivität auf.

4.3 Soziale Erfahrung

Die soziale Erfahrung wird in der sozialen Identitätstheorie definiert und umfasst sowohl die Beziehung zu anderen Menschen als auch zu sich selbst. Beziehungen entstehen durch komplexe kognitive, emotionale und bewertende Prozesse. Demnach enthält die soziale Identität Bestandteile der kognitiven, affektiven und physischen Dimensionen (vgl. Tajfel 2010, S. 17 ff.). Die soziale Erfahrung kann auch als soziale Präsenz definiert werden (vgl. Gefen/Straub 2003, S. 9 f.). In der Literatur werden Erlebnisse unter anderem als individuell, ereignisspezifisch und kontextspezifisch betrachtet. Infolgedessen entstehen soziale Erfahrungen durch Erlebnisse, die im sozialen Kontext mit anderen Menschen geschaffen werden (vgl. Bustamante/Rubio 2017, S. 890 f.; Schmitt 1999, S. 62 f.). Der stationäre Handel bietet eine Vielzahl von solchen Möglichkeiten, um mit anderen Menschen zu interagieren. Es wird der Aufbau einer Beziehung zu den Menschen und zu dem Geschäft ermöglicht. Im Onlinehandel wird sich bemüht, ein ähnliches Gefühl der sozialen Präsenz zu schaffen, denn in Forschungen wurde gezeigt, dass die soziale Präsenz einer Internetseite die wahrgenommene Greifbarkeit und Nähe zu einem Produkt erhöhen kann. Außerdem kann das Vergnügen, die Kaufabsichten und die Loyalität durch eine höhere soziale Präsenz gesteigert werden (vgl. Bleier et al. 2019, S. 101).

Die Integration von KI-gestützten Chatbots stellt eine Lösung dar, um dieses Ziel zu realisieren. Sie können menschenähnliche Interaktionen simulieren und sind in der Lage, den Kunden dadurch ein ähnliches Gefühl zu vermitteln wie ein Servicemitarbeiter. Ein Dialog zwischen Menschen und Computern kann entstehen, indem sie auf direkte Fragen eingehen, Rückfragen stellen und von sich aus neue Themen eröffnen können (vgl. Bogula 2022, S. 158 f.; Hoyer et al. 2020, S. 63).

Dabei kann in Kunden das Gefühl des menschlichen Kontakts, Wärme und Sensibilität ausgelöst werden (vgl. Bleier et al. 2019, S. 99 f.; Gefen/Straub 2003, S. 13). Daher wird angenommen, dass Besucher eines Onlineshops, die während des Einkaufsprozesses auf der Webseite mit einem Chatbot interagieren, ein höheres soziales Empfinden aufbauen als Besucher ohne derartigen Austausch. Daraus hergeleitet, wird folgende Forschungshypothese (H3) aufgestellt:

H3: Besucher eines Onlineshops, die mit einem Chatbot interagieren, haben im Vergleich zu Besuchern ohne Chatbot-Interaktion ein höheres soziales Empfinden.

4.4 Physische Erfahrung

Die physische beziehungsweise sensorische Erfahrung umfasst unter anderem Aspekte, die den Seh-, Hör-, Geruchs-, Geschmacks- oder Tastsinn stimulieren. Sie bezieht sich auf die physiologischen Reaktionen eines Kunden während der Interaktion mit seiner Umgebung oder im E-Commerce mit dem Onlineshop und seinen Komponenten (vgl. Bleier et al. 2019, S. 101). Diese Reaktionen können als Zustand des Wohlbefindens und der Entspannung oder als Mangel an Energie beziehungsweise als Ermüdungszustand differenziert werden. Es herrscht ein Zusammenhang zur affektiven Dimension, denn die Interaktion mit der Umgebung kann auch einen Einfluss auf die Emotionen und Stimmung des Kunden ausüben (vgl. Bustamante/Rubio 2017, S. 891). Im Onlinehandel ist der Umfang der sensorischen Erlebnisse eingeschränkt. Allerdings kann die Wahrnehmung durch das Design, das Bild- und Videomaterial oder andere Berührungspunkte, wie Chatbots, beeinflusst werden. Studien haben offenbart, dass sich das physische Gefühl sowohl auf die Wahrnehmung eines Produkts als auch die Kaufabsichten auswirken kann (vgl. Bleier et al. 2019, S. 101).

KI-gestützte Chatbots können zu physischem Wohlbefinden führen, indem sie dazu beitragen, sowohl Stress als auch Unsicherheiten bei dessen Benutzern zu reduzieren. Durch die Möglichkeit, sofort helfen zu können und relevante Informationen bereitzustellen, können sie die Anfragen der Kunden schnell lösen und infolgedessen eine reibungslosere und stressfreiere Customer Journey gewährleisten. Jedoch ist zu erwähnen, dass Chatbots sich ebenfalls negativ auswirken und bei Nut-

zern Unbehagen auslösen können, beispielsweise aufgrund von technischen Problemen (vgl. Lou et al. 2022, S. 673 ff.; Röder 2023). In der Studie wird ein funktionsfähiger Chatbot angenommen. Demnach wird die folgende und vierte Forschungshypothese (H4) formuliert:

H4: Besucher eines Onlineshops, die mit einem Chatbot interagieren, weisen im Vergleich zu Besuchern ohne Chatbot-Interaktion ein höheres Maß an physischem Wohlbefinden auf.

5 Empirische Analyse: Nutzung eines Chatbots im E-Commerce

5.1 Forschungsdesign

Der zentrale Bestandteil dieser Arbeit umfasst eine empirische Untersuchung in Form eines Experiments im Labor mit dem Ziel der Messung der CX während der Chatbot-Nutzung. Es wurde untersucht, wie sich die Nutzung eines KI-gestützten Chatbots auf die Kundenerfahrung auswirkt. Hierzu wurden sowohl die kognitiven, affektiven, sensorischen beziehungsweise physischen Reaktionen als auch das soziale Empfinden der Nutzer eines Chatbots in einem Onlineshop erfasst. Weiterhin wurden diese den Reaktionen der Besucher eines Onlineshops ohne Chatbot-Nutzung gegenübergestellt. Entsprechend der Vorgehensweise zahlreicher anderer Studien in diesem Forschungsgebiet, ist für die vorliegende Untersuchung eine empirische Forschungsmethode ausgewählt worden. Die Wahl fällt auf die Form des Laborversuchs, um einerseits externe Faktoren, die ebenfalls einen Einfluss auf die CX haben können, so gut wie möglich auszuschließen (vgl. Li/Wang 2023; Lo Presti et al. 2021; Verhulst et al. 2020). Andererseits sind KI-gestützte Chatbots im E-Commerce im deutschsprachigen Raum noch nicht genug verbreitet, um eine repräsentative quantitative Befragung durchzuführen, wie bereits in Kapitel 3.2 dargestellt wird. Als Orientierung für den experimentellen Aufbau und die Durchführung dieser empirischen Forschung dient im Wesentlichen die Studie von Verhulst et al. (2020), in der die affektive Dimension der CX, während einer Self-Service-Technologie-Interaktion (SST), in einem Laborexperiment gemessen wurde.

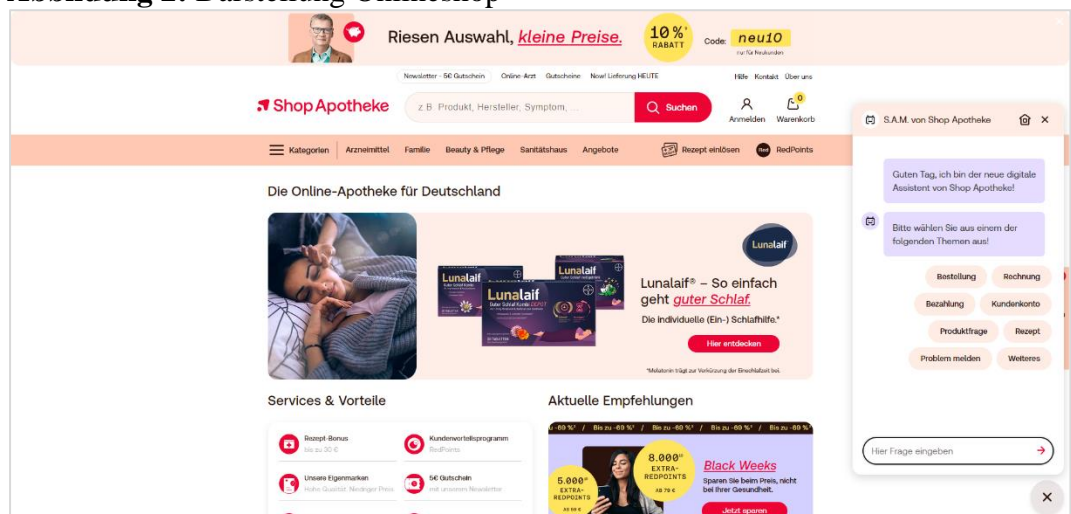
Die Datenerhebung erfolgte sowohl während des Experiments durch die Anwendung einer neurophysiologischen Messmethode wie der Hautreaktionsmes-

sung, einer Eye-Tracking-Messung, als auch unmittelbar im Anschluss an das Experiment mithilfe eines begleitenden Fragebogens, den die Versuchspersonen beantworteten. Die Hautreaktionsmessung wurde gewählt, da eine reine retrospektive Befragung der Affektivität Einschränkungen mit sich bringt. Die Eye-Tracking-Messung ermöglicht eine genauere Analyse der Kognition und der begleitende Fragebogen deckt retrospektiv alle vier Dimensionen ab. Somit wird eine möglichst genaue Erfassung der gesamten Kundenerfahrung gewährleistet (vgl. Bojko 2013, S. 97 f.; Hermes/Riedl 2021, S. 82 f.).

5.2 Aufbau und Durchführung der Untersuchung

Der experimentelle Laborversuch wurde in einem Between-Subject-Design durchgeführt. Dadurch waren die Versuchspersonen jeweils nur einer Bedingung ausgesetzt, entweder mit oder ohne Chatbot-Nutzung. Damit wurden Carryover-Effekte vermieden, bei denen die Erfahrungen aus einer Bedingung in die andere übertragen werden und möglicherweise Ergebnisse verfälschen (vgl. Charness et al. 2012, S. 4). Zuerst wurde ein Pilotversuch (n = 4) durchgeführt. Eine Auswahl von zwei Teilnehmern pro Gruppe war ausreichend, um sicherzustellen, dass der Versuchsablauf und die Instrumente zur Messung der neurophysiologischen und Eye-Tracking Aktivitäten ordnungsgemäß funktionieren (vgl. Verhulst et al. 2020, S. 4).

Abbildung 2: Darstellung Onlineshop

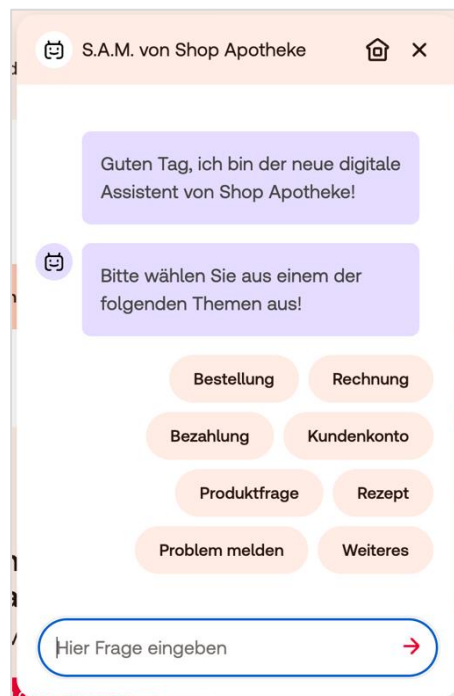


Quelle: Eigene Bildschirmaufnahme

In der Hauptstudie wurden die Probanden randomisiert zwei verschiedenen Gruppen zugeordnet. Dabei waren der Versuchsaufbau, der Ablauf und das Ziel in

beiden Gruppen identisch, um die interne Validität so gut wie möglich zu gewährleisten (vgl. Kuß et al. 2014, S. 186 f.). Die Anzahl der Versuchspersonen je Gruppe ist gleich, um einen bestmöglichen Vergleich sicherzustellen (vgl. Charness et al. 2012, S. 2; Verhulst et al. 2020, S. 4). Die Geschlechterverteilung in beiden Gruppen ist identisch, da sich Frauen und Männer in bestimmten biologischen Reaktionen unterscheiden und auf diese Weise geschlechterspezifische Unterschiede minimiert werden können (vgl. Bianchin/Angrilli 2012, S. 929). Die Szenarien unterscheiden sich lediglich in der unabhängigen Variable. In der ersten Gruppe wurde ein KI-gestützter Chatbot zur Absolvierung der Aufgabe genutzt. Die Kontrollgruppe löste die Aufgabe, indem sie die benötigten Informationen im Onlineshop selbstständig suchten und verwendeten. Die Nutzung des digitalen Assistenten war in dem Fall nicht erlaubt.

Abbildung 3: Digitaler Assistent (Chatbot)



Quelle: Eigene Bildschirmaufnahme

Beide Szenarien wurden im gleichen Apotheken-Onlineshop (www.shop-apotheke.com) durchgeführt, um weitere Faktoren, wie beispielsweise das Design und Layout, die einen Einfluss auf die CX haben können, auszuschließen (vgl. Charness et al. 2012, S. 4). Die Startseite des Onlineshops wird in Abbildung 2 dargestellt. Shop-Apotheke bietet seinen Besuchern einen digitalen Assistenten als Hilfe an, der direkt auf der Startseite in der unteren rechten Ecke des Bildschirms zu finden

ist. Zum einen kann dieser den Nutzern bei einfachen Fragen zur Bestellung und zum Kundenkonto helfen, zum anderen ist er in der Lage, nützliche Produktinformationen bereitzustellen oder passende Produkte zu empfehlen. Wie in Abbildung 3 verdeutlicht wird, können Kunden eigene Fragen und Schlüsselwörter eintippen oder eines der vordefinierten Themen auswählen (vgl. Shop-Apotheke 2023).

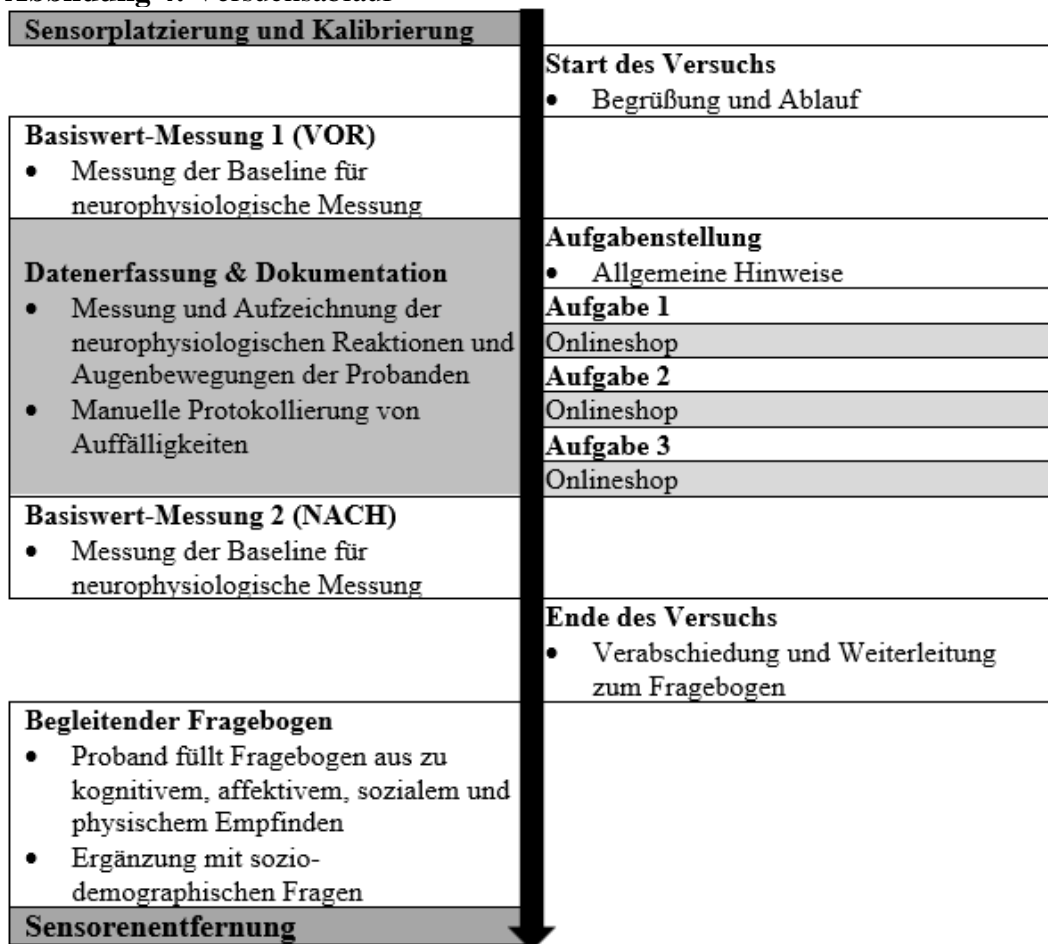
In beiden Gruppen erhielten die Probanden eine identische Aufgabenstellung. Um eine Überforderung der Probanden zu vermeiden und die Dauer des Experiments angemessen zu gestalten, wurde die Studie auf drei wesentliche Aufgaben beschränkt (Ermittlung von Produktinformationen, Recherche von Bezaharten, Feststellung von Versandkosten). Diese sollten nacheinander und separat voneinander bearbeitet werden. In der Chatbot-Gruppe wurde die Aufgabenstellung mit einem Hinweis ergänzt, indem die Aufgabe mithilfe eines Chatbots absolviert werden soll („Digitaler Assistent“ auf www.shop-apotheke.com). In der Kontrollgruppe erhielten die Teilnehmer einen Hinweis mit der Bitte, den Chatbot nicht zu benutzen. Die genaue Aufgabenstellung und Darstellung sind in Anhang B zu finden.

In Abbildung 4 wird der Versuchsablauf veranschaulicht. Nachdem die Versuchsperson den Raum betreten hat, von dem Versuchsleiter begrüßt wurde und eine Einverständniserklärung unterzeichnet hat, setzte sie sich auf einen Stuhl gegenüber von einem Computer. Der Versuchsleiter erteilte daraufhin eine kurze Einführung zum Ablauf und zu den Besonderheiten (siehe Anhang A-4). Die Einführung war für alle Versuchsteilnehmer identisch. Es wurden die Sensoren für die Hautwiderstandmessung angebracht und das Eye-Tracking wurde kalibriert. Nach einer Basiswert-Messung, die für die Hautreaktionsmessung notwendig ist, begann das eigentliche Experiment (vgl. Verhulst et al. 2020, S. 5).

Nach einer allgemeinen Erklärung der Aufgaben in Textform auf dem Monitor, folgten jeweils drei unterschiedliche Aufgaben und eine dazugehörige Weiterleitung zum Onlineshop. Die Versuchsperson sollte dort die Aufgaben absolvieren. Eine Zeitbegrenzung gab es nicht. Alle Teilaufgaben haben unterschiedliche Schwerpunkte (Produkt, Versandkosten, Zahlarten), die gewählt worden sind, um den Teilnehmern eine vielseitige Interaktion mit dem Onlineshop und dem Chatbot zu ermöglichen (vgl. Bleier et al. 2019, S. 101 f.; Lemon/Verhoef 2016, S. 76 f.). Die Aufgabenstellung war sowohl mit den Informationen im Onlineshop als auch

mit dem digitalen Assistenten lösbar. Wenn die Probanden der Meinung waren, die jeweilige Aufgabe absolviert zu haben, wurden sie aus dem Onlineshop zurückgeleitet. Nach den drei Aufgaben folgte eine Verabschiedung mit einer anschließenden zweiten Basiswert-Messung und die Weiterleitung zum begleitenden Fragebogen. Nach dem Ausfüllen des Fragebogens wurden die Sensoren entfernt und der Versuch war beendet.

Abbildung 4: Versuchsablauf



Quelle: Eigene Darstellung in Anlehnung an Verhulst et al. (2020), S. 4

5.3 Erhebungsmethoden

Während des Versuchs wurden die unterschiedlichen Reaktionen der Probanden gemessen. Zum einen wurde eine neurophysiologische Messung in Form der elektrodermalen Aktivität beziehungsweise Hautreaktionsmessung (GSR – Galvanic Skin Response) vorrangig zur Messung der Affektivität durchgeführt, um das Erregungslevel der Emotionen während des Versuchs zu ermitteln (vgl. Christop-

oulos et al. 2019, S. 403). Zum anderen wurde Eye-Tracking genutzt, um insbesondere die kognitive und sensorische Aktivität der Probanden zu messen (vgl. Bojko 2013, S. 122 ff.).

Zudem wurden die erhobenen Daten mit einem begleitenden Fragebogen ergänzt, in dem die subjektive Wahrnehmung der Teilnehmer zu allen vier Dimensionen der CX retrospektiv erfasst wurde. Auf diese Weise kann das gemessene Erregungslevel aus der GSR-Messung mit der Valenz aus dem Fragebogen verglichen werden (vgl. Verhulst et al. 2020, S. 3). Darüber hinaus kann auch die physische und kognitive Wahrnehmung der Versuchspersonen den Eye-Tracking-Aufzeichnungen gegenübergestellt werden. Eine Befragung über einen Fragebogen allein kann zu verzerrten Ergebnissen führen, da die Selbstbewertung von Probanden aus unterschiedlichen Gründen befangen sein kann (vgl. Clore et al. 2001, S. 41 f.; Lemon/Verhoef 2016, S. 88). Die Daten sind zueinander und zu den jeweiligen Probanden über eine jeweilige Probanden-Nummer zuordenbar. Alle Namen sind anonymisiert. Die Datenerhebung und Synchronisation der GSR- und Eye-Tracking-Messungen erfolgte über die Software Tobii Pro Lab sowie über einen Eye-Tracking Monitor von Tobii und einen Hautreaktionsmesssensor von Shimmer (siehe Anhang A-1 und A-2). Die Datenerhebung über den Fragebogen erfolgte über die webbasierte Plattform SosciSurvey. Diese Kombination der methodischen Ansätze ermöglicht eine optimale Auswertung und Interpretation der Ergebnisse (vgl. Bojko 2013, S. 97 ff.; Hermes/Riedl 2021, S. 82 f.; Lemon/Verhoef 2016, S. 88 f.).

5.3.1 Neurophysiologische Messung

Die GSR-Messung dient wesentlich zur Erfassung der affektiven Erfahrung, da die diesbezügliche Selbstauskunft von Probanden aus unterschiedlichen Gründen verzerrt sein kann. Ihre Einstellung oder Emotionen können ihnen (noch) nicht bewusst sein oder sie wollen diese nicht wahrhaben. Daher kann die Messung der elektrodermalen Aktivität als physiologischer Indikator der psychischen Aktivierung genutzt werden. Die Veränderungen des elektrischen Widerstands der Haut der Probanden können Auskunft über den Grad der Aktivierung und das Aktivierungspotential der Reize, zum Beispiel des Chatbots, geben. Jedoch ist die Valenz der Reaktion nicht messbar, das heißt ob der Reiz als negativ oder positiv empfunden wird (vgl. Boucsein 2013, S. 310 ff.). Aus diesem Grund ist es empfehlenswert

begleitende Daten mit einem Fragebogen zu erheben, um die Reaktionen zu interpretieren (vgl. Christopoulos et al. 2019, S. 403; Wübbenhorst 2018).

Im durchgeführten Laborversuch wird die Hautreaktion mit dem Hautleitwert wiedergegeben. Dies ist der Kehrwert des Hautwiderstandes und wird in Mikrosiemens (μS) ausgegeben. Der Hautleitwert ist demnach ein direkter Indikator für die physische Anspannung und ermöglicht, im Vergleich zum Hautwiderstand, eine reibungslosere Auswertung und Interpretation ohne Inversion der Werte (vgl. Boucsein 2013, S. 47 f.). In dem Laborversuch wurde der Hautwiderstand in drei wesentlichen Phasen gemessen. Einerseits während der Aufgabenabsolvierung im Onlineshop und andererseits wurde unmittelbar vor und nach den Aufgaben jeweils eine 15-sekündige Basiswert-Messung durchgeführt, um mögliche Veränderungen im Verlauf erkennen zu können. Die Messung erfolgte über zwei Sensoren an Zeige- und Mittelfinger an der nicht-dominanten Hand. Diese durfte während des Versuchs nicht genutzt werden. Die Teilnehmer setzten ausschließlich eine Hand zum Bedienen der Tastatur und Maus ein. Damit wurden Messungenauigkeiten aufgrund von Bewegungen vermieden (vgl. Verhulst et al. 2020, S. 5).

5.3.2 *Eye-Tracking*

Das Eye-Tracking ermöglicht sowohl die Erfassung physischer als auch kognitiver und affektiver Erfahrungen, indem es die visuellen Reaktionen der Teilnehmer durch die Verfolgung ihrer Augenbewegungen erfasst. Diese Augenbewegungen sind in der Regel reflexartige Reaktionen und verlaufen größtenteils außerhalb des bewussten Einflusses der Teilnehmer. Diese Eigenschaft des Eye-Trackings bietet die Möglichkeit, objektive Daten zu erheben, die die subjektiven Angaben der Teilnehmer ergänzen. Darüber hinaus können die Daten automatisch analysiert, aggregiert und visuell dargestellt werden und ermöglichen eine effiziente Auswertung der Ergebnisse (vgl. Bojko 2013, S. 6 ff.; Roemer et al. 2022, S. 162 ff.).

Im Rahmen der Studie wird Eye-Tracking verwendet, um annähernd alle Dimensionen der CX zu analysieren. Eye-Tracking bietet eine Vielzahl von Metriken und sowohl qualitativen als auch quantitativen Auswertungsmöglichkeiten. Der Fokus der Analyse liegt allerdings auf ausgewählten Messgrößen, da im Rahmen dieser Arbeit nicht das komplette Auswertungspotential ausgeschöpft werden kann. Um die Auswirkungen von Chatbots auf die CX zu analysieren, werden die Anzahl der Fixationen, die durchschnittliche Dauer einer Fixation (Blickdauer) und die

durchschnittliche Pupillengröße als Indikatoren gewählt. Diese Metriken werden aufgrund ihrer Fähigkeit ausgewählt, um Einblicke in verschiedene Aspekte, wie die vertiefte kognitive Verarbeitung, Aufmerksamkeit, emotionale Erregung und physische Anstrengung, zu ermöglichen. Die Messgrößen können zwischen den Gruppen verglichen werden, wodurch Erkenntnisse zum Einfluss von KI-gestützten Chatbots auf die Erfahrungsdimensionen der Versuchspersonen gewonnen werden können (vgl. Bojko 2013, S. 123 ff.; Holmqvist 2011, S. 439 ff.; Roemer et al. 2022, S. 162 ff.).

5.3.3 *Begleitender Fragebogen*

Der Fragebogen basiert auf den theoretischen Grundlagen der CX. Die Basis der Itembatterien bilden die strategischen Erlebnismodule von Schmitt (1999). Die Items wurden aus früheren Forschungsarbeiten übernommen und ins Deutsche übersetzt. Die Items zur Messung der kognitiven, affektiven und physischen Erfahrung wurden aus der ISCX-Messskala entnommen, die Bustamante und Rubio (2017) entwickelt haben und zur Messung der CX im Einzelhandel einsetzten. Die Items zur Messung der sozialen Erfahrung stammen von Bleier et al. (2019), die in einer Studie die Auswirkungen des Designs einer Produktseite auf das digitale Kundenerlebnis untersuchten. Die Entscheidung, diese Items zur Messung der sozialen Erfahrung zu verwenden, wurde getroffen, da sie sich, im Gegensatz zu den Items von Bustamante und Rubio, spezifisch auf die sozialen Merkmale eines Onlineshops beziehen, anstatt die Erfahrung durch die sozialen Interaktionen mit Mitarbeitern und anderen Kunden abzufragen (vgl. Bleier et al. 2019, S. 103 ff.).

Die essenzielle Struktur der Fragen wurde beibehalten, jedoch wurden einzelne Formulierungen entsprechend den Anforderungen des Laborversuchs im Wortlaut angepasst, beispielsweise wurde der Begriff „Geschäft“ in „Onlineshop“ umformuliert. Alle Items werden auf einer siebenstufigen Likert-Skala quantifiziert und weisen zufriedenstellende Werte ($\alpha > 0,70$) für das Cronbach-Alpha auf, was auf eine hohe interne Konsistenz hinweist und zeigt, dass die Items jeweils homogene Konstrukte messen. Außerdem wurde die hohe Validität der Konstrukte mittels Faktorenanalysen in den Studien bestätigt (vgl. Bleier et al. 2019, S. 104 ff.; Bustamante/Rubio 2017, S. 897). Darüber hinaus wurde der Fragebogen durch das Einbeziehen von soziodemographischen Abfragen zu Geschlecht, Alter und beruf-

lichem Status erweitert. Zusätzlich wurden Vorfragen gestellt, welche die Probanden befragten, ob sie den digitalen Assistenten während der Aufgabenabsolvierung genutzt haben oder nicht. Zudem wurde nach der Probanden-Nummer gefragt, um die erhobenen Daten zueinander zuordnen zu können. Es sei angemerkt, dass sämtliche Fragen, die von den Teilnehmern in beiden untersuchten Bedingungen beantwortet wurden, identisch sind. Eine detaillierte Aufschlüsselung der Items und deren Anordnung wird in Tabelle 2 veranschaulicht. Die Codierung der Items und die Darstellung im digitalen Fragebogen ist in Anhang C zu finden.

Tabelle 2: Items des Fragebogens

Itembatterien	Items
Kognitiv (Bustamante/Rubio 2017; Schmitt 1999)	1) Die Darstellung, der Service und die Interaktion mit dem Onlineshop... a) lehren mich interessante Dinge. b) wecken meine Kreativität. c) bringen mich auf interessante Ideen.
Affektiv (Bustamante/Rubio 2017; Schmitt 1999)	2) Die Darstellung, der Service und die Interaktion mit dem Onlineshop machen mich... a) schlecht gelaunt – gut gelaunt b) unglücklich – glücklich c) pessimistisch – optimistisch
Sozial (Bleier et al. 2019; Gefen/Straub 2003; Schmitt 1999)	3) Die Darstellung, der Service und die Interaktion mit dem Onlineshop... a) geben mir ein Gefühl des menschlichen Kontakts. b) geben mir ein Gefühl der menschlichen Wärme. c) geben mir ein Gefühl der menschlichen Sensibilität.
Physisch (Bitner 1992; Bustamante/Rubio 2017; Schmitt 1999)	4) Die Darstellung, der Service und die Interaktion mit dem Onlineshop haben... a) mich überfordert – mich entspannt b) mir Energie geraubt – mir Energie gegeben c) mich ermüdet – mich vitalisiert

Quelle: Eigene Darstellung

Die kognitiven und sozialen Erfahrungsdimensionen werden anhand der Zustimmung zu den Aussagen der jeweiligen Items von den Probanden bewertet. Hierbei steht der Wert 1 für „stimme gar nicht zu“ und repräsentiert gleichzeitig den geringsten Wert, während der Wert 7 „stimme voll und ganz zu“ darstellt und gleichzeitig den höchsten Wert widerspiegelt. Die affektiven und physischen Erfahrungsdimensionen werden ebenfalls auf siebenstufigen Skalen erfasst. Jedoch

wird die Affektivität auf drei bipolaren Skalen abgefragt, bei denen die gegenüberliegenden Pole gegensätzliche Bewertungen darstellen. Dabei repräsentiert der Wert 1 die negativste Bewertung (schlecht gelaunt; unglücklich; pessimistisch), während der Wert 7 die positivste Bewertung (gut gelaunt; glücklich; optimistisch) bedeutet. In Bezug auf die physische Erfahrungsdimension verhält es sich ähnlich. Das physische Unbehagen (Überforderung, Energieverlust, Ermüdung) wird durch den niedrigsten Wert 1 widerspiegelt und das physische Wohlbefinden (Entspannung, Energiezufuhr, Vitalisierung) wird durch den höchsten Wert 7 dargestellt (vgl. Bleier et al. 2019, S. 104 ff.; Bustamante/Rubio 2017, S. 897).

6 Darstellung und Auswertung der Ergebnisse

6.1 Beschreibung der Stichprobe

Im Folgenden werden die Ergebnisse der empirischen Untersuchung vorgestellt. Die Daten des Laborversuchs wurden mithilfe von Tobii Pro Lab, SosciSurvey, SPSS und Excel erfasst und im Anschluss ausgewertet. Es haben insgesamt 37 Probanden an dem Experiment teilgenommen. Diese wurden durch persönliche Ansprachen im engeren Familien- und Freundeskreis sowie durch gezielte Rekrutierung an der Hochschule gewonnen. Auf die Studie wurde zudem durch Plakate in den Gebäuden und Beiträge in Onlineforen der Hochschule aufmerksam gemacht.

Die Aufzeichnungen wurden manuell auf Vollständigkeit geprüft. Es mussten sieben Ergebnisse aufgrund von fehlenden GSR-Daten verworfen werden. Der Grund dafür waren Verbindungsprobleme zwischen Sensoren und Computer während der Aufzeichnungen. Diese Schwierigkeiten führten dazu, dass GSR-Messdaten teilweise oder vollständig fehlten. Der Fragebogen wurde von jedem Teilnehmer vollständig beantwortet. Zudem waren die Eye-Tracking-Aufzeichnungen bei jeder Versuchsperson erfolgreich.

Tabelle 3: Finale Stichprobe

	MIT	OHNE	Gesamt
Teilnehmer	15	15	30
davon Männer	8	8	16
davon Frauen	7	7	14
Durchschnittsalter	24,93	25,20	25,07

Quelle: Eigene Darstellung basierend auf eigener Erhebung

Die finale Stichprobe besteht somit aus 30 verwertbaren Aufzeichnungen. Davon haben 15 Versuchspersonen den Chatbot zur Aufgabenabsolvierung genutzt (MIT) und 15 weitere Teilnehmer gehören zur Kontrollgruppe und haben die Aufgaben ohne Chatbot gelöst (OHNE). Die wesentlichen sozio-demographischen Ergebnisse werden in Tabelle 3 dargestellt. Insgesamt haben 14 Frauen (46,67%) und 16 Männer (53.33%) teilgenommen. Je Bedingung haben 8 Männer und 7 Frauen an der Studie partizipiert. Das Durchschnittsalter aller Probanden beträgt 25,07 Jahre. Im Durchschnitt sind die Teilnehmer mit Chatbot-Nutzung 24,93 Jahre alt. Der Median in dieser Gruppe liegt bei 23 Jahren. Das Alter in der Kontrollgruppe ist annähernd gleich und weist keinen signifikanten Unterschied auf. Das Durchschnittsalter in der zweiten Gruppe ist 25,20 Jahre. Der Median beträgt 22 Jahre. Der Unterschied zwischen Mittelwert und Median deutet auf Ausreißer in den Stichproben hin. Dies lässt sich bestätigen, da in beiden Gruppen jeweils auch Versuchspersonen, die über 40 Jahre alt sind, teilgenommen haben (vgl. Kuß et al. 2014, S. 218). Der derzeitige berufliche Status wurde ebenfalls abgefragt. Es gaben 24 Teilnehmer an, dass sie hauptsächlich studieren. Außerdem haben vier Angestellte, ein Proband in der Ausbildung und eine verbeamtete Person teilgenommen.

6.2 Ergebnisse des Laborexperiments

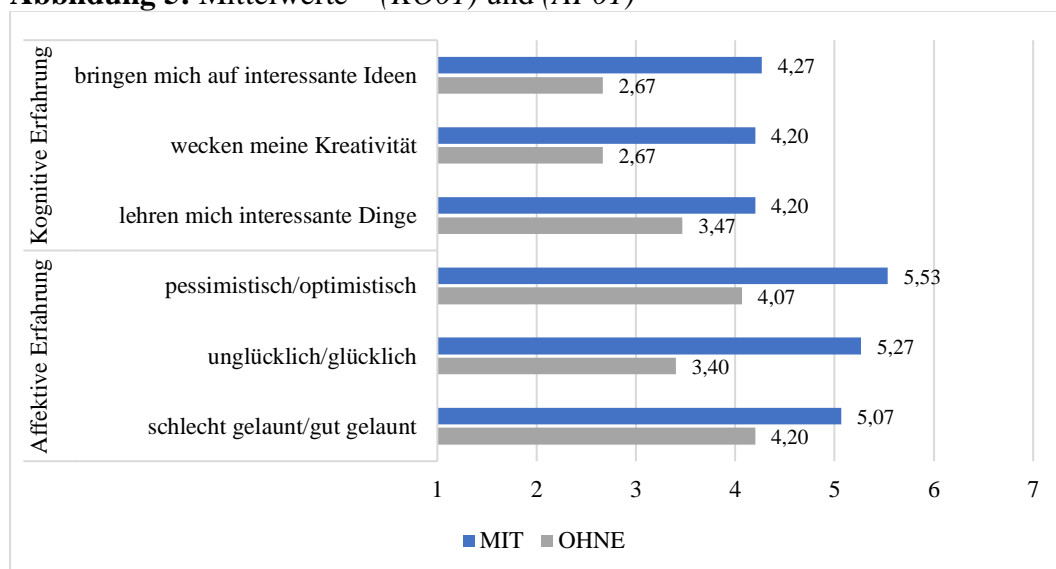
6.2.1 Auswertung des Fragebogens

Im folgenden Abschnitt werden die erhobenen Daten der jeweiligen Erhebungsmethoden dargestellt. Zuerst werden die gewonnenen Ergebnisse aus dem Fragebogen beschrieben, bevor die Aufzeichnungen der GSR- und Eye-Tracking-Messungen ausgewertet werden. Es werden ausschließlich relevante Kennzahlen präsentiert, während eine umfassende Aufschlüsselung aller Ergebnisse im Anhang D dieser Arbeit zu finden ist. In den Ergebnissen liegen unabhängige Stichproben in beiden Gruppen vor, da die Teilnehmer den Versuch nur unter einer Bedingung und unabhängig voneinander durchgeführt haben (vgl. Bühl 2016, S. 175). In der begleitenden Umfrage wurden alle Items auf einer siebenstufigen Likert-Skala quantifiziert. Die Ratingskalen zeichnen sich durch gleichmäßige Abstände zwischen den Punkten aus und können daher als intervallskaliert betrachtet werden. Somit erfüllen sie die Kriterien für die Auswertung metrischer Skalen (vgl. Kuß et al. 2014, S. 208 f.). Für die Darstellung der Ergebnisse wird zunächst das arithmetische

Mittel für jedes Item pro Gruppe gebildet. Die Werte werden in Abbildung 5 und Abbildung 6 gegenübergestellt und im weiteren Verlauf verglichen.

Die kognitive Erfahrung (*KO01*) während der Interaktion mit dem Onlineshop wurde von der Gruppe mit Chatbot-Nutzung im Fragebogen anhand aller drei Items höher bewertet als ohne Nutzung des Chatbots. Die Aussagen „lehren mich interessante Dinge“ und „wecken meine Kreativität“ erhielten von den Chatbot-Nutzern mit einem durchschnittlichen Wert von jeweils 4,20 eine höhere Zustimmung, während die Kontrollgruppe diese Items mit 3,47 und 2,67 bewertet hat. Zudem wurde das Item „bringen mich auf interessante Ideen“ von der ersten Gruppe mit durchschnittlich 4,27 am höchsten bewertet, während die zweite Gruppe die Aussage mit 2,67 bewertete.

Abbildung 5: Mittelwerte – (*KO01*) und (*AF01*)



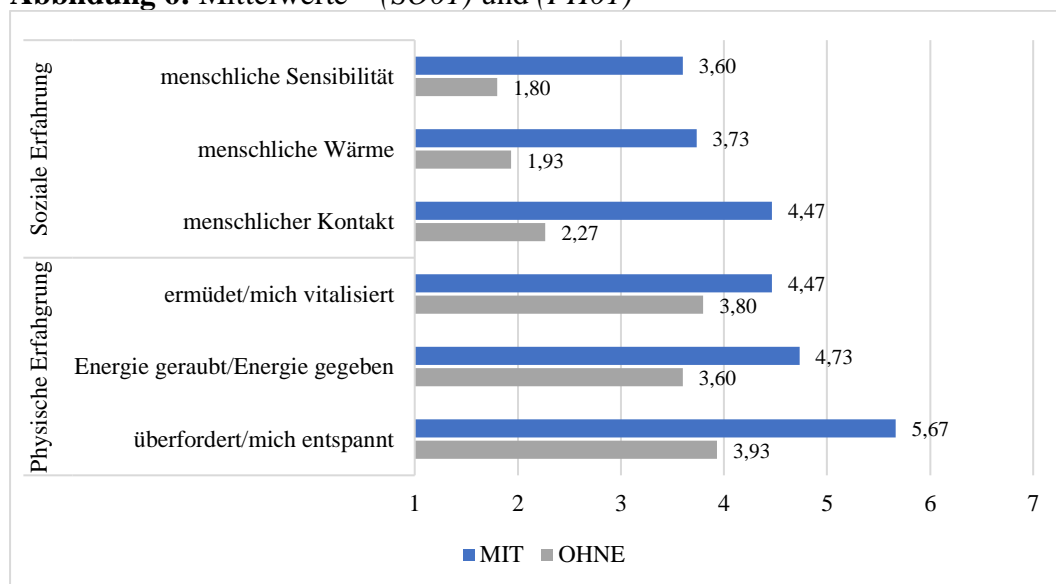
Quelle: Eigene Darstellung basierend auf eigener Erhebung

Die erlebte Affektivität (*AF01*) wurde von der Gruppe, die den digitalen Assistenten nutzte, ebenfalls höher bewertet als von der Kontrollgruppe ohne Chatbot. Demnach gaben die Chatbot-Nutzer an, während des Besuchs des Onlineshops eher gut gelaunt gewesen zu sein (5,07) im Vergleich zu den Probanden ohne Chatbot (4,20). Diese haben ebenfalls angegeben, glücklicher gewesen zu sein (5,27) als die Versuchspersonen, die keinen Chatbot genutzt haben (3,40). Dies spiegelt sich auch im Optimismus wider. Die Teilnehmer der ersten Gruppe haben ihre Erfahrung mit

einem Wert von 5,53 als optimistischer eingestuft, während die zweite Gruppe das Item im Durchschnitt mit 4,07 bewertet hat.

Im Folgenden wird das soziale Empfinden (*SO01*) der Probanden beschrieben. Die Darstellung, der Service und die Interaktion mit dem Onlineshop wurde von den Probanden, die den KI-gestützten Chatbot nutzten, als sozialer eingestuft im Vergleich zu den Teilnehmern, die keinen Chatbot verwendeten. Insbesondere das Item „Gefühl des menschlichen Kontakts“ erhielt in beiden Gruppen die höchste Zustimmung, wobei die Werte ebenso den größten Unterschied aufweisen. Die Chatbot-Nutzer bewerteten den menschlichen Kontakt während des Experiments durchschnittlich mit 4,47, während die Kontrollgruppe eine Bewertung von 2,27 abgab. Auch die anderen beiden Items, die das Gefühl der menschlichen Wärme und Sensibilität betreffen, zeigen deutliche Unterschiede in der Zustimmung zwischen den Gruppen. Besucher mit Chatbot-Nutzung vergaben durchschnittlich Werte von 3,73 für die menschliche Wärme und 3,60 für die menschliche Sensibilität, während die Teilnehmer, die die Aufgaben ohne Chatbot lösten, die Items der sozialen Dimension im Durchschnitt mit 1,93 und 1,80 bewerteten.

Abbildung 6: Mittelwerte – (*SO01*) und (*PH01*)



Quelle: Eigene Darstellung basierend auf eigener Erhebung

Zuletzt wird die physische Wahrnehmung (*PH01*) dargestellt. Das physische Wohlbefinden der Chatbot-Nutzer während der Aufgabenabsolvierung wurde im Vergleich zur Kontrollgruppe ebenso höher bewertet. Die Versuchsteilnehmer, die

den Chatbot verwendeten, gaben im Durchschnitt an, während der Aktivitäten entspannter (5,67) und vitalisierter (4,47) gewesen zu sein. Im Gegensatz dazu wiesen die Teilnehmer der Kontrollgruppe mit Werten von 3,93 (Entspannung) und 3,80 (Vitalisierung) niedrigere Bewertungen auf. Des Weiteren neigten die Probanden ohne Chatbot-Nutzung dazu anzugeben, dass die Interaktion mit dem Onlineshop eher Energie raubte (3,60), im Vergleich zur Gruppe mit Chatbot-Nutzung, die angab, dass die Interaktion eher Energie gegeben hat (4,73).

6.2.2 Auswertung der GSR-Messungen

Nachfolgend werden die Ergebnisse der Hautreaktionsmessung ausgewertet. Der Hautleitwert der Versuchspersonen wurde im Laborexperiment in drei wesentlichen Phasen gemessen. Für die Messung der Hautreaktion während der Aufgabenabsolvierung sind nur die Intervalle relevant, in denen die Versuchspersonen tatsächlich im Onlineshop aktiv waren.

Tabelle 4: Hautreaktion - Messwerte

	MIT	OHNE
Basiswert VOR*	2,16	2,59
Aufgabenabsolvierung*	2,77 (+28,06%)	3,15 (+21,37%)
Basiswert NACH*	2,92 (+34,85%)	3,43 (+32,35%)

*in Mikrosiemens (μS)

Quelle: Eigene Darstellung basierend auf eigener Erhebung

Wie in Tabelle 4 zu erkennen ist, ist der Hautleitwert der Kontrollgruppe während des gesamten Laborexperiments höher im Vergleich zur Gruppe mit Chatbot-Nutzung. Außerdem sind die Werte innerhalb beider Gruppen im Verlauf des Versuchs konstant angestiegen. Jedoch ist der Anstieg des Hautleitwerts zum Basiswert vor der Aufgabenabsolvierung sowohl während als auch nach der Absolvierung in der Gruppe mit Chatbot-Interaktion höher im Vergleich zur Kontrollgruppe. Dennoch muss überprüft werden, ob die Unterschiede statistisch signifikant oder zufällig entstanden sind (vgl. Bühl 2016, S. 176). Für die Auswahl der korrekten statistischen Methode zur Überprüfung der Gruppenunterschiede müssen zunächst verschiedene Voraussetzungen geprüft werden (vgl. Bühl 2016, S. 171 ff.). Es kann angenommen werden, dass die Werte metrisch skaliert sind und aus zwei unabhängigen Stichproben stammen. Eine Überprüfung auf Normalverteilung der Werte

mittels Kolmogorov-Smirnov-Tests (K-S-T) und Shapiro-Wilk-Tests (S-W-T) ergibt, dass nicht alle Werte innerhalb der Gruppen normalverteilt sind, denn die Nullhypothese $H_0: p > 0,05$ besagt, dass die Daten normalverteilt sind. Die Alternativhypothese $H_1: p < 0,05$ besagt, dass keine Normalverteilung der Daten vorliegt. Da das Signifikanzniveau (p) beider Tests sowohl für die absoluten Werte als auch für die prozentualen Anstiege zum Großteil kleiner als 0,05 ist, muss die Nullhypothese für diese Gruppen abgelehnt und H_1 angenommen werden (vgl. Bühl 2016, S. 380 f.; Janssen/Laatz 2007, S. 249 ff.). Daher wird der Mann-Whitney-U-Test ganzheitlich angewendet, welcher sowohl für normalverteilte Daten als auch Daten ohne Normalverteilung anwendbar ist (vgl. Bühl 2016, S. 360 f.). Eine detaillierte Auflistung der Ergebnisse der Prüfung auf Normalverteilung ist im Anhang D-3 verfügbar.

Tabelle 5: Hautreaktion - Mann-Whitney-U-Test

Teststatistiken ^a					
	Basiswert VOR	Aufgaben- absolvierung	Basiswert NACH	Zuwachs Aufgaben- absolvierung	Zuwachs Basiswert NACH
Mann-Whitney-U-Test	104,000	111,000	108,000	107,000	106,000
Wilcoxon-W	224,000	231,000	228,000	227,000	226,000
Z	-0,353	-0,062	-0,187	-0,228	-0,270
Asymp. Sig. (2-seitig)	0,724	0,950	0,852	0,820	0,787
Exakte Sig. [2*(1-seitige Sig.)]	,744 ^b	,967 ^b	,870 ^b	,838 ^b	,806 ^b
a. Gruppenvariable: Nutzung Chatbot					
b. Nicht für Bindungen korrigiert.					

Quelle: Eigene Darstellung basierend auf eigener Erhebung

Zunächst muss im Szenario-Check geprüft werden, ob ein signifikanter Unterschied in den Basiswerten bereits vor der Aufgabenabsolvierung vorherrschte (vgl. Verhulst et al. 2020, S. 6). Die Nullhypothese H_0 besagt, dass es keinen signifikanten Unterschied bezüglich der zentralen Tendenzen beider Gruppen gibt, wenn $p > 0,05$. Da $n = 30$, muss die asymptotische Signifikanz (2-seitig) betrachtet werden (vgl. Janssen/Laatz 2007, S. 574). Wie aus Tabelle 5 hervorgeht, ist diese mit einem Wert von 0,724 größer als die Verwerfungsgrenze, sodass die Alternativhypothese H_1 , die einen signifikanten Unterschied annimmt, verworfen werden und H_0 angenommen werden muss. Nun wird der Unterschied der absoluten Hautleitwerte zwischen den beiden Gruppen während der Absolvierung der drei Aufgaben getestet. Die Anwendung des Mann-Whitney-U-Tests ergibt eine asymptotische Signifikanz von 0,950, welche ebenfalls größer als 0,05 ist. Somit muss H_0 erneut angenommen

werden. Die Überprüfung des Unterschieds der Basiswerte nach den Aufgaben ergibt eine asymptotische Signifikanz von $0,852 > 0,05$. Damit muss ebenfalls angenommen werden, dass kein statistisch signifikanter Unterschied zwischen den Werten beider Gruppen vorliegt. Eine Überprüfung der Anstiege von der ersten Basiswertmessung (VOR) zur Aufgabenabsolvierung als auch zur zweiten Basiswertmessung (NACH) belegt mit Werten von 0,820 und 0,787 für die asymptotische Signifikanz, dass in den Zuwächsen ebenfalls kein statistischer Unterschied nachweisbar ist. Zusammenfassend kann gesagt werden, dass Unterschiede in den absoluten Werten und relativen Veränderungen der Hautleitwerte zwischen den Gruppen erkennbar sind, diese jedoch nicht als signifikant eingestuft werden können (vgl. Bühl 2016, S. 360 f.; Janssen/Laatz 2007, S. 572 ff.).

6.2.3 Auswertung der Eye-Tracking-Aufzeichnungen

Für die Auswertung der Eye-Tracking-Aufzeichnungen wird die durchschnittliche Anzahl an Fixationen, die durchschnittliche Dauer je Fixation (Blickdauer) sowie der durchschnittliche Pupillendurchmesser pro Versuchsperson verwendet. Es werden ebenso nur die Intervalle in Betracht gezogen, in denen die Versuchspersonen die Aufgaben im Onlineshop absolviert haben. Die Werte der drei Phasen im Onlineshop werden pro Proband zu einem Wert aggregiert. Daraufhin wird das arithmetische Mittel je Kennzahl pro Proband und Gruppe gebildet. Die Ergebnisse werden in Tabelle 6 dargestellt.

Tabelle 6: Eye-Tracking - ausgewählte Metriken

	MIT	OHNE
Anzahl Fixationen	489,53	453,00
Blickdauer je Fixation*	252,87	230,33
Pupillendurchmesser**	2,99	2,92

*in Millisekunden (ms); **in Millimeter (mm)

Quelle: Eigene Darstellung basierend auf eigener Erhebung

Es ist zu erkennen, dass alle drei Schlüsselmetriken im Vergleich zu den Werten der Kontrollgruppe bei den Chatbot-Nutzern größer sind. Die Unterschiede müssen jedoch ebenfalls auf Signifikanz geprüft werden (vgl. Bühl 2016, S. 176). Es ist anzunehmen, dass die Daten aus zwei unabhängigen Stichproben stammen und metrisch skaliert sind. Die Normalverteilung muss wiederum geprüft werden. Die

Anwendung der K-S-T und S-W-Tests ergibt, dass nicht alle Daten innerhalb der Gruppen normalverteilt sind. Das Signifikanzniveau ist zum Teil kleiner als 0,05, wodurch H_0 abgelehnt werden muss und H_1 mit der Aussage, dass nicht alle Daten normalverteilt sind, angenommen wird (vgl. Janssen/Laatz 2007, S. 249 ff.). Im Anhang D-4 werden die Resultate der Überprüfung auf Normalverteilung detailliert aufgeführt.

Tabelle 7: Eye-Tracking - Mann-Whitney-U-Test

Teststatistiken ^a			
	Anzahl Fixationen	Blickdauer je Fixation	Pupillendurchmesser
Mann-Whitney-U-Test	101,500	74,000	99,000
Wilcoxon-W	221,500	194,000	219,000
Z	-0,456	-1,597	-0,560
Asymp. Sig. (2-seitig)	0,648	0,110	0,576
Exakte Sig. [2*(1-seitige Sig.)]	,653 ^b	,116 ^b	,595 ^b
a. Gruppenvariable: Nutzung Chatbot			
b. Nicht für Bindungen korrigiert.			

Quelle: Eigene Darstellung basierend auf eigener Erhebung

Zur Überprüfung der Unterschiede der Eye-Tracking-Daten wird infolgedessen ebenfalls der Mann-Whitney-U Test für jede Metrik angewendet. Zuerst wird die Anzahl der Fixation getestet. Aus Tabelle 7 ist abzulesen, dass die Überprüfung eine asymptotische Signifikanz von 0,648 ergibt, welche größer ist als der Grenzwert 0,05, sodass H_0 angenommen werden muss. Die Durchführung des Tests für die Blickdauer je Fixation führt zu einer asymptotischen Signifikanz von 0,110, welche im Vergleich zur vorherigen Metrik kleiner ist, jedoch größer als die Verwerfungsgrenze, sodass in diesem Fall angenommen werden muss, dass kein Unterschied bezüglich der zentralen Tendenzen beider Gruppe vorliegt. Die Anwendung des Tests auf die Werte des Pupillendurchmessers resultiert in einer Signifikanz von $0,576 > 0,05$. Hiermit muss ebenfalls H_0 angenommen werden. Abschließend kann festgehalten werden, dass Unterschiede in den Eye-Tracking-Metriken zwischen beiden Gruppen erkennbar sind, diese sich jedoch nicht statistisch signifikant unterscheiden (vgl. Bühl 2016, S. 360 f.; Janssen/Laatz 2007, S. 572 ff.).

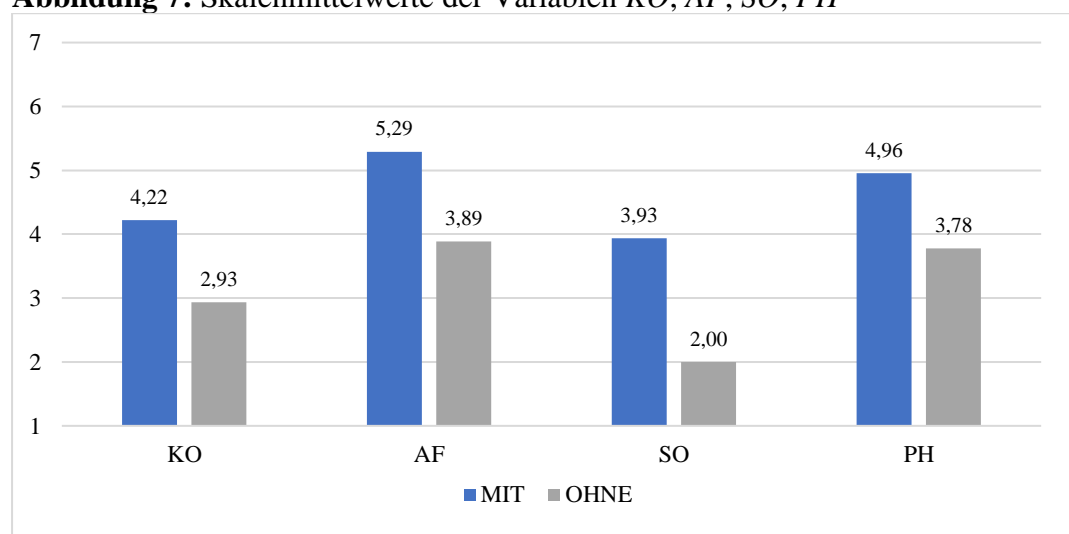
6.3 Hypothesentests

In diesem Abschnitt wird die Überprüfung der theoretischen Hypothesen auf Basis der im Fragebogen erfassten Daten durchgeführt. Die Daten der GSR- und

Eye-Tracking-Aufzeichnungen werden im späteren Verlauf der Arbeit zur Interpretation der Ergebnisse verwendet. Die Analyse der retrospektiven Befragung erfolgt mittels univariater und multivariater Verfahren. Auf eine zusätzliche Faktoren- und Reliabilitätsanalyse wird im Rahmen dieser Studie verzichtet. Die Entscheidung basiert darauf, dass die Items aus früheren Studien übernommen wurden, in denen bereits eine Prüfung auf interne Konsistenz und Validität durchgeführt wurde, sodass angenommen werden kann, dass die Items jeweils homogene Konstrukte messen (vgl. Bleier et al. 2019, S. 104; Bustamante/Rubio 2017, S. 897).

Zunächst werden die individuellen Items jeder Dimension der CX in SPSS transformiert und zu einer Variable zusammengefasst. Dieser Schritt wird unternommen, um eine verbesserte Übersichtlichkeit zu gewährleisten und die Interpretation der Ergebnisse zu erleichtern. Hierbei werden die Mittelwerte der einzelnen Items gebildet, was zur Schaffung der Variablen *AF* (Affektive Erfahrung), *KO* (Kognitive Erfahrung), *PH* (Physische Erfahrung) und *SO* (Soziale Erfahrung) führt. Es wird ein Skalenmittelwert pro Dimension und Proband ermittelt. Zusätzlich wird aus diesen Werten das arithmetische Mittel für jede Dimension und Gruppe errechnet. Der Vergleich zwischen den Gruppen wird in Abbildung 7 sichtbar gemacht (vgl. Bühl 2016, S. 213 ff.).

Abbildung 7: Skalenmittelwerte der Variablen *KO*, *AF*, *SO*, *PH*



Quelle: Eigene Darstellung basierend auf eigener Erhebung

Zur Überprüfung der in Kapitel 4 aufgestellten Hypothesen werden die zuvor gebildeten Variablen genutzt und es werden T-Tests für unabhängige Stichproben

durchgeführt. Dabei wird überprüft, ob die beobachtbaren Mittelwertunterschiede signifikant sind oder sich mit zufälligen Schwankungen erklären lassen (vgl. Bühl 2016, S. 343 ff.). Bevor ein T-Test für unabhängige Stichproben erfolgen kann, müssen im ersten Schritt bestimmte Voraussetzungen erfüllt werden. Zunächst werden zwei unabhängige Stichproben vorausgesetzt. Diese Bedingung ist erfüllt. Darüber hinaus müssen die Variablen metrisch skaliert sein. Wie bereits erläutert, kann diese Annahme ebenfalls getroffen werden. Die dritte Anforderung besagt, dass die Werte innerhalb der Gruppen normalverteilt sein müssen (vgl. Bühl 2016, S. 175). Hierzu werden die Daten pro Gruppe mittels K-S- und S-W-Tests auf eine Normalverteilung geprüft (siehe Anhang D-5). Die Überprüfung ergibt, dass alle Signifikanzniveaus (p) je Konstrukt und Gruppe sowohl für den K-S- als auch den S-W-Test größer als 0,05 sind. Somit muss die Nullhypothese angenommen werden und es kann die Annahme getroffen werden, dass alle Daten annähernd normalverteilt sind (vgl. Bühl 2016, S. 380 f.; Janssen/Laatz 2007, S. 249 ff.). Des Weiteren bestätigt eine visuelle Überprüfung der Q-Q-Diagramme diese Annahme (siehe Anhang D-5). Da alle Bedingungen erfüllt sind, können nun Hypothesentests mittels T-Test für unabhängige Stichproben durchgeführt werden.

Tabelle 8: T-Test Variable *KO*

Test bei unabhängigen Stichproben											
		Levene-Test der Varianzgleichheit		t-Test für die Mittelwertgleichheit							
		F	Sig.	T	df	Signifikanz		Mittlere Differenz	Differenz für Standardfehler	95% Konfidenzintervall der Differenz	
						p (1-seitig)	p (2-seitig)			Unterer Wert	Oberer Wert
KO	Varianzen sind gleich	0,567	0,458	2,512	28	0,009	0,018	1,28889	0,51318	0,23769	2,34008
	Varianzen sind nicht gleich			2,512	27,057	0,009	0,018	1,28889	0,51318	0,23604	2,34174

Quelle: Eigene Darstellung basierend auf eigener Erhebung

Als erstes wird die Hypothese H1 überprüft, die sich auf die kognitive Erfahrung während des Experiments bezieht. Dabei wird die Variable *KO* betrachtet. Die Gruppe mit Chatbot erreicht einen Mittelwert von 4,22, wobei sich bei der Kontrollgruppe ohne Chatbot ein Mittelwert von 2,93 beobachten lässt. Im Verlauf des Mittelwertvergleichs muss zunächst die Varianzhomogenität geprüft werden. Diese wird mit dem Levene-Test kontrolliert. Die Nullhypothese lautet hierbei, dass die Varianzen homogen sind. Demzufolge sollte die Signifikanz größer als 0,05 sein,

damit sie nicht verworfen und die homogene Varianz beider Stichproben bestätigt werden kann. Die Alternativhypothese besagt, dass die Varianzen nicht gleich sind. Wie aus Tabelle 8 hervorgeht, ergibt das Testergebnis eine Signifikanz von $0,458 > 0,05$. Demnach darf die Nullhypothese nicht abgelehnt werden und die Varianzgleichheit muss angenommen werden. Somit müssen die Werte aus der ersten Zeile in Tabelle 8 zur weiteren Überprüfung genutzt werden (vgl. Janssen/Laatz 2007, S. 347 f.). Der T-Wert ergibt 2,512, welcher als sehr gut zu interpretieren ist, da dieser größer als 0 ist und somit nochmals bestätigt wird, dass der erste Mittelwert (Gruppe mit Chatbot) größer ist als der Mittelwert der Kontrollgruppe. Da das 95%-Konfidenzintervall den Wert 0 nicht beinhaltet, ist davon auszugehen, dass der Unterschied der Differenz signifikant ist. Dies wird auch bei der Überprüfung des Signifikanzniveaus aus dem T-Test bestätigt. Beim T-Test besagt die Nullhypothese H_0 , wenn $p > 0,05$, dass es keinen signifikanten Unterschied gibt. Jedoch beträgt die Signifikanz (2-seitig) 0,018 und die einseitige Signifikanz 0,009, womit die Nullhypothese abgelehnt werden muss und die Alternativhypothese H_1 bestätigt wird. Die einseitige Signifikanz darf betrachtet werden, da eine Wirkungsvermutung vorherrscht und gerichtete Hypothesen aufgestellt worden sind. Die Hypothese H_1 kann somit auf Basis der Daten des begleitenden Fragebogens bestätigt werden (vgl. Bühl 2016, S. 344 ff.; Janssen/Laatz 2007, S. 348 ff.). Zudem kann die Effektstärke mit Cohen's d ermittelt werden, da eine Stichprobengröße $n > 20$ vorliegt. Die Punktschätzung gemäß Cohen (1988) ergibt einen Wert von 0,875 und ist somit über der Grenze von 0,8 zum starken Effekt. Der Unterschied zwischen beiden Gruppen kann somit als stark eingestuft werden (vgl. Cohen 1988, S. 25 f.). Weitere Berechnungen zur Effektgröße sind in Anhang D-5 zu finden.

Tabelle 9: T-Test Variable *AF*

		Test bei unabhängigen Stichproben									
		Levene-Test der Varianzgleichheit		t-Test für die Mittelwertgleichheit							
		F	Sig.	T	df	Signifikanz		Mittlere Differenz	Differenz für Standardfehler	95% Konfidenzintervall der Differenz	
						p (1-seitig)	p (2-seitig)			Unterer Wert	Oberer Wert
AF	Varianzen sind gleich	0,221	0,642	4,189	28	0,000	0,000	1,400	0,33418	0,71547	2,08453
	Varianzen sind nicht gleich			4,189	27,386	0,000	0,000	1,400	0,33418	0,71477	2,08523

Quelle: Eigene Darstellung basierend auf eigener Erhebung

Folgend wird die Hypothese H2 bezüglich der affektiven Erfahrung getestet. Hierbei wird das Augenmerk auf die Variable *KO* gelegt. Die Chatbot-Nutzer haben die affektive Erfahrung mit einem Mittelwert von 5,29 positiver bewertet als die Besucher des Onlineshops ohne Chatbot-Nutzung, die einen Mittelwert von 3,89 erzielt haben. Aus Tabelle 9 geht hervor, dass bei der Durchführung des T-Tests für die dazugehörige Variable der Levene-Test eine Signifikanz von $0,642 > 0,05$ ergab. Damit kann die Annahme getroffen werden, dass die Varianzen gleich sind (vgl. Janssen/Laatz 2007, S. 347 f.). Der T-Wert beträgt 4,189 und bestätigt, dass der Mittelwert der ersten Gruppe höher ist im Vergleich zur Kontrollgruppe. Beide Werte des 95%-Konfidenzintervalls sind nicht null, was darauf hindeutet, dass der Unterschied signifikant ist. Die Signifikanzniveaus des T-Tests bestätigen diese Annahme. Sowohl die Werte der zweiseitigen als auch der einseitigen Signifikanz sind mit $p < 0,001$ kleiner als 0,05. Hier darf ebenfalls die einseitige Signifikanz betrachtet werden, da im Vorhinein eine gerichtete Hypothese aufgestellt wurde. Die Nullhypothese, die besagt, dass kein signifikanter Unterschied besteht, muss demnach abgelehnt werden und H_1 wird angenommen. Somit kann die Hypothese H2 auf Basis der individuellen Wahrnehmung der Probanden ebenfalls angenommen werden (vgl. Bühl 2016, S. 344 ff.; Janssen/Laatz 2007, S. 348 ff.). Der Effekt kann zudem als stark eingestuft werden, da die Punktschätzung nach Cohen's d den Wert 1,530 ausweist (vgl. Cohen 1988, S. 25 f.).

Tabelle 10: T-Test Variable *SO*

Test bei unabhängigen Stichproben											
		Levene-Test der Varianzgleichheit		t-Test für die Mittelwertgleichheit							
		F	Sig.	T	df	Signifikanz		Mittlere Differenz	Differenz für Standardfehler	95% Konfidenzintervall der Differenz	
						p (1-seitig)	p (2-seitig)			Unterer Wert	Oberer Wert
SO	Varianzen sind gleich	9,212	0,005	3,827	28	0,000	0,001	1,9333	0,50521	0,89846	2,96821
	Varianzen sind nicht gleich			3,827	20,101	0,001	0,001	1,9333	0,50521	0,87982	2,98685

Quelle: Eigene Darstellung basierend auf eigener Erhebung

Im weiteren Verlauf wird Hypothese H3 getestet. Werden nur die Skalenmittelwerte aus der Befragung betrachtet, lässt sich diese zunächst bestätigen. Die Chatbot-Nutzer haben ihr Erlebnis mit einem Wert von durchschnittlich 3,93 bewertet, während die Besucher des Onlineshops ohne Chatbot-Nutzung ihre soziale

Erfahrung mit einem Mittelwert von 2,00 eingestuft haben. Jedoch muss dieser Unterschied auf Signifikanz geprüft werden. Zunächst werden beide Stichproben mittels Levene-Tests auf Varianzgleichheit getestet, welcher eine Signifikanz von 0,005 ergibt. Da diese kleiner als 0,05 ist, muss die Annahme der Varianzhomogenität verworfen und folglich die Alternativhypothese, die aussagt, dass die Varianzen nicht gleich sind, angenommen werden. Aufgrund der Varianzheterogenität muss der Welch-Test angewendet werden, indem die untere Zeile aus Tabelle 10 betrachtet wird. Das weitere Vorgehen der Testung bleibt gleich (vgl. Janssen/Laatz 2007, S. 348). Mit einem T-Wert von 3,827 und einer mittleren Differenz von 1,93 kann bestätigt werden, dass ein Unterschied vorliegt und der Wert der ersten Gruppe höher ist. Die Werte des 95%-Konfidenzintervalls nehmen vorweg, dass der Unterschied signifikant sein muss. Dies bestätigen ebenfalls beide Signifikanzen p mit Werten von jeweils $0,001 < 0,05$. Auch in diesem Zusammenhang kann die einseitige Signifikanz in Betracht gezogen werden, da im Voraus eine Annahme über die Wirkung und eine gerichtete Hypothese aufgestellt wurden. Die Nullhypothese H_0 , welche behauptet, dass kein signifikanter Unterschied besteht, muss demnach abgelehnt werden, während H_1 akzeptiert wird. Die Folge ist die Annahme der Hypothese H_3 (vgl. Bühl 2016, S. 344 ff.; Janssen/Laatz 2007, S. 348 ff.). Außerdem kann der signifikante Unterschied zwischen den beiden Gruppen als stark eingeordnet werden. Die Schätzung der Effektgröße gibt einen Wert für Cohen's d von 1,397 wieder (vgl. Cohen 1988, S. 25 f.).

Tabelle 11: T-Test Variable *PH*

Test bei unabhängigen Stichproben											
		Levene-Test der Varianzgleichheit		t-Test für die Mittelwertgleichheit							
		F	Sig.	T	df	Signifikanz		Mittlere Differenz	Differenz für Standardfehler	95% Konfidenzintervall der Differenz	
						p (1-seitig)	p (2-seitig)			Unterer Wert	Oberer Wert
PH	Varianzen sind gleich	0,333	0,568	3,157	28	0,002	0,004	1,17778	0,37308	0,41356	1,94200
	Varianzen sind nicht gleich			3,157	25,159	0,002	0,004	1,17778	0,37308	0,40965	1,94590

Quelle: Eigene Darstellung basierend auf eigener Erhebung

Zuletzt wird die Hypothese H_4 bezogen auf das physische Wohlbefinden überprüft. Das physische Empfinden der Teilnehmer wird in der Variable *PH* widergespiegelt. Im Durchschnitt haben die Chatbot-Nutzer ihre physische Erfahrung mit einem Wert von 4,96 positiver bewertet als die Versuchsgruppe ohne Chatbot, die

einen Mittelwert von 3,78 erzielt. Bei der Überprüfung, ob ein signifikanter Unterschied zwischen den Werten vorliegt, kann bei dem Levene-Test der Varianzgleichheit eine Signifikanz von $0,568 > 0,05$ festgestellt werden, wie in Tabelle 11 zu erkennen ist. Demnach kann die Nullhypothese, die besagt, dass die Varianzen gleich sind, angenommen und mit dem T-Test fortgefahren werden (vgl. Janssen/Laatz 2007, S. 349). Ein T-Wert von 3,157 und eine mittlere Differenz von 1,18 deuten darauf hin, dass ein Unterschied in den Mittelwerten vorliegt und der Wert der ersten Gruppe höher ist. Die Werte des 95%-Konfidenzintervalls unterstützen die Annahme, dass dieser Unterschied statistisch signifikant ist. Die zweiseitige Signifikanz des T-Tests erzielt einen Wert von 0,004 und die einseitige Signifikanz erreicht den Wert 0,002. Da beide $p < 0,05$ sind, muss H_0 abgelehnt und die Alternativhypothese angenommen werden. Gestützt auf der Selbstauskunft der Probanden, kann die Hypothese H_4 somit bestätigt werden (vgl. Bühl 2016, S. 344 ff.; Janssen/Laatz 2007, S. 348 ff.). Zusätzlich liegt der Wert für Cohen's d mit einer Punktschätzung von 1,153 über der Grenze von 0,8, sodass der Mittelwertunterschied innerhalb dieser Dimension, ebenso wie in den anderen Konstrukten, als stark eingestuft werden kann (vgl. Cohen 1988, S. 25 f.).

7 Diskussion der Ergebnisse

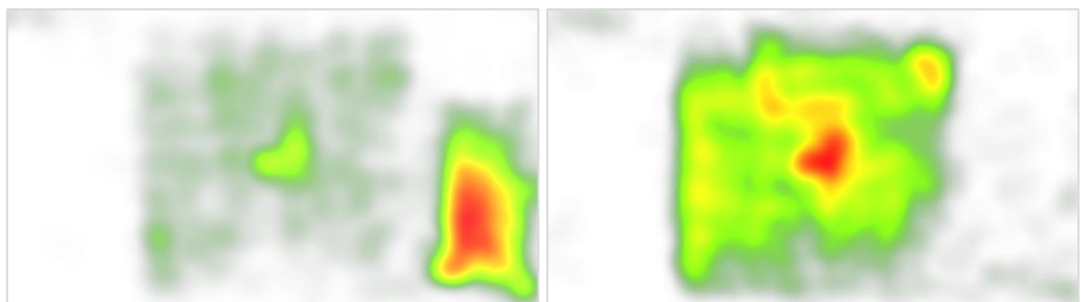
7.1 Interpretation der Ergebnisse

Die Ergebnisse des Laborexperiments bieten ein geteiltes Bild hinsichtlich der Auswirkungen von KI-gestützten Chatbots auf die CX im E-Commerce. Zwar sind Unterschiede sowohl in den Messwerten der Eye-Tracking- und GSR-Aufzeichnungen als auch in den Ergebnissen des Fragebogens beobachtbar, jedoch lassen sich statistisch relevante Unterschiede nicht flächendeckend erkennen. Insbesondere bei den ausgewählten Metriken des Eye-Trackings und der Hautwiderstandsmessung können keine signifikanten Unterschiede bezüglich der zentralen Tendenzen der Chatbot-Nutzer und der Kontrollgruppe festgestellt werden. Nichtsdestotrotz können alle aufgestellten Hypothesen aus Kapitel 4 auf Basis der subjektiven Wahrnehmung der Versuchsteilnehmer angenommen werden. In diesem Zusammenhang können gemäß Cohen (1988) in allen vier betrachteten Dimensionen der CX hohe Effektstärken nachgewiesen werden, die darauf hindeuten, dass der Chatbot einen starken Effekt auf die individuellen Einschätzungen der Teilnehmer hat.

Hergeleitet aus bisher durchgeführten Studien wird angenommen, dass Besucher eines Onlineshops, die mit einem Chatbot kommunizieren, eine aktivere Rolle im Informationsaustausch während der Customer Journey einnehmen als Kunden, die solch eine Interaktion nicht in Anspruch nehmen. Dies kann sich in der Kognition der Besucher widerspiegeln und wird in Hypothese 1 angenommen (vgl. Lou et al. 2022, S. 673 ff.; Xu et al. 2020, S. 190 f.). In der subjektiven Wahrnehmung der Versuchspersonen, kann diese Hypothese angenommen werden. Die Teilnehmer, die die Aufgaben mit einem Chatbot absolviert haben, bewerteten die kognitive Erfahrung im Vergleich zur Kontrollgruppe bedeutend höher. Demnach vermittelt die Nutzung von Chatbots nicht nur mehr interessante Informationen, sondern fördert auch stärker die Kreativität und inspiriert zu vielfältigeren Ideen (vgl. Bustamante/Rubio 2017, S. 901 f.). Da die subjektive Wahrnehmung der Probanden aus diversen Gründen verzerrt sein kann, empfiehlt es sich einen Blick auf die Hautreaktionswerte und Ergebnisse der Eye-Tracking-Aufzeichnungen zu werfen. Da in den Ergebnissen jedoch keine signifikanten Unterschiede nachgewiesen werden können, sind diese mit Vorsicht zu genießen (vgl. Verhulst et al. 2020, S. 2 ff.).

Die durchschnittlich höhere Anzahl der Fixationen pro Person in der Chatbot-Gruppe kann auf eine gesteigerte Aufmerksamkeit hindeuten. Die Nutzer richten ihre visuelle Aufmerksamkeit auf den Chatbot, wodurch sich eine gesteigerte kognitive Interaktion und eine vertiefte Auseinandersetzung mit den Inhalten abzeichnet (vgl. Holmqvist 2011, S. 560 ff.). Jedoch ist zu beachten, dass einige Teilnehmer nach dem Experiment angegeben haben, dass der Chatbot schwierig zu finden sei. Daher kann die höhere Anzahl an Fixationen mit der geringeren Sucheffizienz zusammenhängen (vgl. Bojko 2013, S. 126).

Abbildung 8: Heatmaps (links mit und rechts ohne Chatbot)



Quelle: Eigene Bildschirmaufnahmen aus Tobii Pro Lab auf Basis eigener Erhebung

Die qualitative Betrachtung der aggregierten Heatmaps beider Gruppen widerlegt diese Überlegung. Die in Abbildung 8 dargestellten Heatmaps zeigen, dass der Großteil der Aufmerksamkeit tatsächlich auf den Chatbot gerichtet ist, der sich in der unteren rechten Ecke der Webseite befindet.

Die durchschnittliche längere Blickdauer je Fixation und der durchschnittlich größere Pupillendurchmesser pro Chatbot-Nutzer kann darauf hinweisen, dass eine intensivere kognitive Verarbeitung während der Aufgabenabsolvierung stattfindet. Dennoch kann dies ebenso negativ interpretiert werden, indem eine höhere kognitive Beanspruchung ausgelöst wird, da es für die Nutzer schwieriger ist, benötigte Informationen zu extrahieren (vgl. Bojko 2013, S. 272 ff.; Holmqvist 2011, S. 542 f.). Einzelne Probanden gaben an, dass sie die Interaktion mit dem digitalen Assistenten als schwierig empfunden haben, jedoch wurde dies mit der mangelnden Affinität zu neuer Technologie begründet. Der Großteil der Versuchspersonen empfand die Bedienung des Bots hingegen als einfach, sodass der Zusammenhang zwischen KI-Technologien und positiven kognitiven Reaktionen wie in einer vorausgegangen Studie von Baek et al. (2022) bestätigt werden kann.

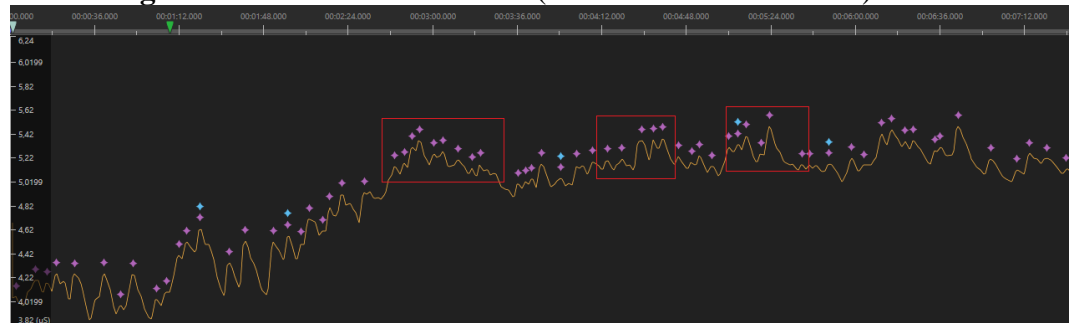
In beiden Gruppen kann über die gesamte Dauer des Versuchs ein Anstieg im Hautleitwert beobachtet werden. Dies deckt sich mit den Ergebnissen aus bisherigen Studien und kann auf die Anspannung innerhalb der Laborsituation zurückzuführen sein (vgl. Verhulst et al. 2020, S. 6 ff.). Jedoch können bei der Chatbot-Gruppe im Vergleich zur Kontrollgruppe höhere Anstiege in den Hautleitwerten sowohl während als auch unmittelbar nach der Aufgabenabsolvierung beobachtet werden. Dies kann ein Anzeichen für eine Fokussierung und Erhöhung der Aufmerksamkeit sein und kann wiederum als stärkere kognitive Erfahrung interpretiert werden (vgl. Boucsein 2013, S. 329).

Darüber hinaus wird in Hypothese 2 angenommen, dass die Nutzer eines Chatbots während der Interaktion im Onlineshop ein höheres Maß an positiver Affektivität aufweisen als Besucher ohne Chatbot-Interaktion, da KI-gestützte Chatbots in der Lage sind, bestimmte Kundenanfragen effektiv zu bearbeiten und Zufriedenheit auszulösen (vgl. Bogula 2022, S. 160; Jenneboer et al. 2022, S. 224 f.; Lou et al. 2022, S. 662). Die Nutzer des Chatbots stufen ihre Erfahrung optimistischer ein und gaben zugleich an, sowohl glücklicher als auch besser gelaunt gewesen zu sein.

Dies resultiert in einer bemerkenswerten positiveren Gesamtbewertung der erlebten Affektivität im Vergleich zur Kontrollgruppe. Die retrospektive Bewertung von Emotionen ist jedoch mit Vorbehalt zu behandeln, da Probanden befangen sein können (vgl. Verhulst et al. 2020, S. 2 ff.). Obwohl keine Signifikanz in den unterschiedlichen Messwerten der Hautreaktion beider Gruppen beobachtet werden kann, werden diese dennoch im weiteren Verlauf thematisiert, um ein umfassenderes Analysebild zu präsentieren. Trotzdem sollten die Ergebnisse mit Vorsicht betrachtet werden.

Wie bereits erwähnt, hat die Chatbot-Gruppe im Vergleich zur Kontrollgruppe stärkere Anstiege der Hautleitwerte während des Laborversuchs vorzuweisen. Dies kann darauf hindeuten, dass die Interaktion mit dem Chatbot stärkere affektive Reaktionen hervorgerufen hat. Diese können auf Freude, Aufregung oder auch Frustration hinweisen (vgl. Christopoulos et al. 2019, S. 403 ff.). Im Zusammenhang mit der subjektiven Wahrnehmung der Probanden können die höheren affektiven Reaktionen zum Großteil als eher positiv betrachtet werden. Wird die zeitliche Dynamik der GSR-Messungen im Kontext der Eye-Tracking Daten betrachtet, lässt sich beobachten, dass insbesondere Momente, in denen mit dem Chatbot kommuniziert wird und dieser die Aufgabe der Probanden löst, mit einem Anstieg des Hautleitwerts korrelieren. Diese Beobachtung wird in Abbildung 9 am Beispiel der Versuchsperson mit der Nummer 22 dargestellt. Ein ähnlicher Verlauf lässt sich ebenso bei weiteren Versuchspersonen, die den Chatbot genutzt haben, erkennen. Zudem lassen sich Parallelen in der Kontrollgruppe beobachten, allerdings sind die Ausschläge hierbei weniger prägnant. Diese Erkenntnis bestätigt die zuvor ausgewertete positivere Affektivität der Chatbot-Nutzer (vgl. Boucsein 2013, S. 310 ff.).

Abbildung 9: Verlauf des Hautleitwerts (Proband 22: mit Chatbot)



Quelle: Eigene Bildschirmaufnahme aus Tobii Pro Lab auf Basis eigener Erhebung

Hinzukommend kann der größere Pupillendurchmesser der Teilnehmer in der Chatbot-Gruppe mit einer erhöhten emotionalen Intensität in Verbindung gebracht werden (vgl. Holmqvist 2011, S. 542). Dies deckt sich mit den Ergebnissen des Fragebogens und der Hautleitwertmessung und bestätigt die Annahmen zusätzlich.

Des Weiteren können statistisch relevante Unterschiede in der Wahrnehmung der sozialen Erfahrung der Versuchspersonen mittels T-Test festgestellt werden, sodass ebenfalls Hypothese 3 nicht verworfen werden kann. Die Chatbot-Gruppe gab im Durchschnitt sowohl für das Gefühl des menschlichen Kontakts als auch der menschlichen Wärme und Sensibilität höhere Zustimmungswerte an als die Kontrollgruppe (vgl. Bleier et al. 2019, S. 99 f.; Gefen/Straub 2003, S. 13). Da der Aufbau sozialer Beziehungen stark mit kognitiven und emotionalen Prozessen korreliert, ist davon auszugehen, dass auch die soziale Erfahrung der Chatbot-Nutzer steigt (vgl. Bleier et al. 2019, S. 101). Dieses Ergebnis steht im Einklang mit der Auswertung der kognitiven und affektiven Dimension, sodass die gestellte Erwartung erfüllt wird. Gemäß den Ergebnissen lässt sich herleiten, dass der Chatbot sich stark auf dessen Nutzer und deren soziales Gefühl auswirkt. Somit wird die getroffenen Erwartungen erfüllt, da KI-gestützte Chatbots zunehmend in der Lage sind, menschliche Konversationen zu führen, wie der digitale Assistent im Laborexperiment (vgl. Bogula 2022, S. 158 f.). Darüber hinaus bekräftigen die Ergebnisse die Erkenntnisse von Liebrecht et al. (2021), indem festgestellt werden konnte, dass der Chatbot einen Effekt auf die soziale Präsenz hat.

Weiterführend wurde Hypothese 4 auf Basis der individuellen Wahrnehmung der Versuchspersonen überprüft. Das Ergebnis zeigt, dass die Versuchspersonen mit Chatbot-Nutzung die physische Erfahrung positiver bewertet haben als die Kontrollgruppe. Zudem bestätigt der statistische Test den signifikanten Unterschied der Bewertungen beider Gruppen. Demnach waren die Chatbot-Nutzer entspannter, energiereicher als auch vitalisierter (vgl. Bustamante/Rubio 2017, S. 901 f.). Infolgedessen konnten die Erwartungen erfüllt und die Hypothese 4 bestätigt werden. Die Nutzer des Chatbots weisen ein höheres Maß an physischem Wohlbefinden auf im Vergleich zu Nutzern, die keinen Chatbot genutzt haben und eher zum Unbehagen tendieren.

Da zwischen der physischen und affektiven als auch kognitiven Erfahrung ein starker Zusammenhang herrscht, empfiehlt es sich die Hautreaktions- und Eye-Tracking-Aufzeichnungen hinzuzuziehen. Die größeren Anstiege der Hautleitwerte, die höhere Anzahl an Fixationen und die längere Blickdauer der Chatbot-Nutzer können Anzeichen für intensivere emotionale und kognitive Reaktionen sein, die negativ oder positiv interpretiert werden können. Unter Beachtung der positiven Wahrnehmung der Probanden lässt sich eine negative physische Reaktion jedoch ausschließen, sodass die Ergebnisse weiterhin als positiv interpretiert werden können und zum Wohlbefinden der Teilnehmer führen (vgl. Boucsein 2013, S. 328 ff.). Hingegen kann die kürzere Blickdauer der Probanden der Kontrollgruppe als Stressreaktion interpretiert werden und auf Unbehagen deuten (vgl. Holmqvist 2011, S. 542).

Zusammenfassend belegt die Auswertung der Ergebnisse, dass sich die Nutzung eines Chatbots stark auf die subjektive Wahrnehmung der Nutzer auswirkt und unterschiedlich Reaktionen auslösen kann, die überwiegend als positiv betrachtet werden können. Übereinstimmend mit den Erkenntnissen von Hoyer et al. (2020), die ein Rahmenwerk zur Nutzung von KI-Technologien im E-Commerce entwickelt haben, können Chatbots einen hohen Wert in jeder einzelnen Dimension der CX kreieren. Infolgedessen kann eine Wertsteigerung innerhalb der Dimensionen einen starken Einfluss auf die Phasen der Customer Journey haben. Insbesondere vor und während des Kaufprozesses können sich digitale Assistenten stark auf das zukünftige Verhalten der Kunden auswirken. Darüber hinaus sollte beachtet werden, dass zahlreiche weitere Faktoren eine Rolle spielen, wie sich Chatbots auf die Kundenerfahrung auswirken. Eine wichtige Rolle dabei spielt die Wahrnehmung der Nutzer. Haben Nutzer das Gefühl, dass die Technologie autonom handelt, werden andere Reaktionen ausgelöst, als wenn der Nutzer ein Gefühl der Kontrolle hat. Zusätzlich beeinflusst der Benutzerkomfort und die Bedienung die Reaktionen stark (vgl. Hoyer et al. 2020, S. 66). Daher ist eine weiterführende Untersuchung erforderlich, die das resultierende Verhalten, wie beispielweise die Kaufentscheidung, sowie weitere Faktoren miteinbezieht und die Korrelationen mit einer gestiegenen CX analysiert (vgl. Hoyer et al. 2020, S. 64 ff.).

7.2 Kritische Diskussion der Methode und weiterer Forschungsbedarf

Im folgenden Abschnitt wird die Vorgehensweise der empirischen Studie kritisch betrachtet. Während das Experiment durch die Auswahl von drei verschiedenen Erhebungsmethoden ein breites Spektrum für die Ergebnisauswertung bietet, besteht andererseits die Gefahr einer Datenüberflutung. Daher wird in dieser Arbeit eine Fokussierung vorgenommen, indem nur ausgewählte Indikatoren in Bezug zu den Erfahrungsdimensionen analysiert werden.

Außerdem sind für eine optimale Datenerhebung und -auswertung verschiedene Gütekriterien einzuhalten, damit eine angemessene Interpretation durchgeführt werden kann. Dazu zählen Objektivität, Reliabilität und Validität (vgl. Magerhans 2016, S. 87). Um die Objektivität sicherzustellen, waren der Versuchsablauf und -aufbau sowie der Fragebogen für alle Probanden identisch. Der Versuchsleiter erteilte klare Anweisungen gemäß einem zuvor erstellten Leitfaden und es wurde darauf geachtet, dass keine persönlichen Vorannahmen in die Datenerhebung und -analyse einfließen (vgl. Magerhans 2016, S. 87).

Die Reliabilität wurde durch regelmäßige Tests der verwendeten Messinstrumente, darunter Eye-Tracking und Hautreaktionsmessung, gewährleistet. Unvollständige Datensätze wurden in der Auswertung nicht berücksichtigt. Die Messskalen im Fragebogen weisen hohe Werte für das Cronbach-Alpha auf, wodurch die Reliabilität der Konstrukte ebenfalls sichergestellt wird (vgl. Kuß et al. 2014, S. 88 f.; Magerhans 2016, S. 88 f.).

Die interne Validität wurde durch die Minimierung möglicher Störfaktoren während des Experiments gewährleistet. Es wurde auf eine konstante Raumtemperatur und -helligkeit geachtet. Beide Gruppen hatten ein identisches Forschungsdesign und denselben Fragebogen. Jedoch ist zu betonen, dass externe Faktoren nicht komplett ausgeschlossen werden können. Zum Beispiel zeigen Menschen erhebliche Unterschiede in ihrem Blickverhalten und Hautreaktionen, welche durch unterschiedliche nicht kontrollierbare Faktoren beeinflusst werden können. Daher ist es schwierig eine klare Standardisierung zu erreichen (vgl. Boucsein 2013, S. 192 ff.; Holmqvist 2011, S. 542 f. / 560 f.). Außerdem sind die Fragen aus früheren Studien entnommen, in denen sie bereits auf Konstruktvalidität geprüft wurden und zufriedenstellende Ergebnisse erzielt haben (vgl. Kuß et al. 2014, S. 106 ff.).

Die externe Validität dieser Studie ist durch die begrenzte Übertragbarkeit der Laborergebnisse auf den realen E-Commerce-Kontext eingeschränkt. Die künstliche Umgebung des Laborexperiments stellt eine wesentliche Einschränkung dar, da das Verhalten der Probanden im Labor möglicherweise nicht vollständig die natürlichen Interaktionen während des Online-Shoppings widerspiegelt. Außerdem wurden die Probanden gebeten, die Aufgaben nur mit einer Hand zu absolvieren. Dies hat zum einen Verzerrungen in der Hautreaktionsmessung minimiert, zum anderen hat es das natürliche Verhalten der Probanden beeinflusst, wodurch eine Generalisierbarkeit der Ergebnisse erschwert wird. Zudem kann die Stichprobenauswahl zu Verzerrungen in den Ergebnissen führen. Es wurde sich bemüht ähnliche Stichproben bezüglich Alter, Geschlecht und beruflichen Status zu schaffen, jedoch ist die Übertragung der Ergebnisse in die Realität aufgrund des hohen Anteils an Studierenden und des relativ jungen Durchschnittsalters kritisch zu betrachten (vgl. Kuß et al. 2014, S. 186 f.; Magerhans 2016, S. 89 f.).

Ferner ist darauf zu achten, dass die begrenzte Experiment- und Interaktionsdauer der Probanden mögliche Auswirkungen auf die CX haben kann. Langfristige Studien können dazu beitragen, umfassendere Einblicke in die nachhaltigen Auswirkungen von KI-gestützten Chatbots auf die Kundenerfahrung zu gewinnen. Außerdem können weitere alternative Forschungsmethoden, wie Feldstudien oder qualitative Ansätze mit einem größeren repräsentativen Stichprobenumfang, wertvolle Erkenntnisse in dem Bereich liefern (vgl. Kuß et al. 2014, S. 39 ff.). Jedoch ist zu beachten, dass die Technologie nicht vollständig ausgereift ist, sodass eine fortführende Forschung in den nächsten Jahren notwendig ist (vgl. Bakkouri et al. 2022, S. 435).

8 Fazit

Aufgrund der starken Weiterentwicklung von Chatbots und dem Einsatz von KI bietet die Technologie viele Potentiale in der Bereicherung der Kundenerfahrung. Jedoch stellt sich die Frage, welchen spezifischen Einfluss KI-gestützte Chatbots auf die Dimensionen der CX haben. Das Ziel der vorliegenden Arbeit ist es, die Auswirkungen der digitalen Assistenten auf die vier Dimensionen der CX aufzuzeigen und Unterschiede zur CX ohne Nutzung der Technologie darzustellen. In

diesem Zusammenhang wurde eine empirische Analyse mit zwei Gruppen durchgeführt, in dem sowohl quantitative als auch qualitative Erhebungs- und Auswertungsmethoden genutzt worden sind.

Die Analyse zeigt, dass in der subjektiven Wahrnehmung zwischen den Chatbot-Nutzern und der Kontrollgruppe wesentliche Unterschiede in allen vier Dimensionen zu beobachten sind. Sowohl die kognitive und affektive Erfahrung als auch die soziale und physische Wahrnehmung werden von der Chatbot-Gruppe höher eingestuft. Es kann nachgewiesen werden, dass der Chatbot einen starken Effekt auf seine Nutzer hat. In der objektiven Betrachtung der Hautreaktionsmesswerte und den ausgewählten Metriken der Eye-Tracking-Aufzeichnungen können ebenfalls Unterschiede zwischen den Werten festgestellt werden, dennoch können statistisch signifikante Unterschiede nicht bestätigt werden. Die qualitative Auswertung der Aufzeichnungen verdeutlicht, dass der Chatbot die Aufmerksamkeit der Nutzer auf sich zieht und eine vertiefte kognitive Verarbeitung anregt sowie eine emotionale Reaktion auslöst. Die Studie zeigt, dass der Chatbot ein wesentlicher Faktor ist, der die Reaktionen der Nutzer beeinflusst und zu einem intensiveren Erlebnis im E-Commerce führt.

Trotz einiger methodischer Einschränkungen, wie der begrenzten Stichprobe und der Laborsituation, liefern die Ergebnisse wichtige Einblicke in der Erforschung moderner Chatbots. Zukünftige Forschung kann von umfassenderen Studien in realen E-Commerce Szenarien profitieren. Zusätzlich wurde die Untersuchung ausschließlich an dem KI-gestützten Chatbot eines Anbieters durchgeführt und berücksichtigt somit andere digitale Assistenten nicht. Dies beschränkt den Grad der Verallgemeinerung der Ergebnisse.

Nichtsdestotrotz liefern die vorliegenden Ergebnisse wichtige Erkenntnisse zur bestehenden Literatur im Bereich von Chatbots im E-Commerce. Die Ergebnisse dieser Arbeit verdeutlichen die Rolle von Chatbots als Touchpoint in der Customer Journey und unterstreichen das Potential, die Kundeninteraktion und somit die Wettbewerbsfähigkeit von Unternehmen zu stärken. Die Erkenntnisse bieten wertvolle Informationen für die Entwicklung von Strategien zur Optimierung von Kundenerlebnissen, Personalisierung und Servicequalität. Eine Anreicherung der CX mit der Technologie kann zu intensiveren Kundenerlebnissen führen. Weitere Folgen können Kundenzufriedenheit und Kundenbindung sein. Unternehmen sollten

die Technologie nach umfangreicher Analyse der Customer Journey einführen und sorgfältig prüfen. Es ist wichtig zu beachten, dass ein Chatbot als ergänzender Berührungspunkt nahtlos in bestehende Verfahren implementiert wird. Zusätzlich ist es wichtig, die Grenzen der Technologie zu verstehen. Händlern sollte bewusst sein, in welchen Situationen ein digitaler Assistent in der Kundenkommunikation nicht mehr behilflich sein kann und das Gespräch daher an einen menschlichen Mitarbeiter übergeben werden muss.

Insgesamt liefert die Arbeit nicht nur einen Einblick in die aktuelle Situation, sondern betont ebenfalls die Notwendigkeit weiterer Forschung auf diesem Gebiet, um die sich rasant entwickelnde Technologie und deren Auswirkung auf die Kunden besser zu verstehen und fundierte Strategien für die Zukunft zu entwickeln.

Literaturverzeichnis

- Abu Shawar**, Bayan/**Atwell**, Eric (2007). Chatbots: Are they Really Useful?, *Journal for Language Technology and Computational Linguistics* (22), 29–49, Online: [10.21248/jlcl.22.2007.88](https://doi.org/10.21248/jlcl.22.2007.88).
- Adamopoulou**, Eleni/**Moussiades**, Lefteris (2020). An Overview of Chatbot Technology, in: Maglogiannis, I./Iliadis, L./Pimenidis, E. (Hg.) *IFIP Advances in Information and Communication Technology*, 373–383.
- Baek**, Tae Hyun et al. (2022). Smiling AI agents: How anthropomorphism and broad smiles increase charitable giving, *International Journal of Advertising* (41), 850–867, Online: [10.1080/02650487.2021.2011654](https://doi.org/10.1080/02650487.2021.2011654).
- Bagdare**, Shilpa/**Jain**, Rajnish (2013). Measuring retail customer experience, *International Journal of Retail & Distribution Management* (41), 790–804, Online: [10.1108/IJRDM-08-2012-0084](https://doi.org/10.1108/IJRDM-08-2012-0084).
- Bakkouri**, Bouchra El/**Raki**, Samira/**Belgnaoui**, Touhfa (2022). The Role of Chatbots in Enhancing Customer Experience: Literature Review, *Procedia Computer Science* (203), 432–437, Online: [10.1016/j.procs.2022.07.057](https://doi.org/10.1016/j.procs.2022.07.057).
- Basalisco**, Bruno et al. (2021). The Impact of E-commerce on the German Economy, Online: https://copenhageneconomics.com/wp-content/uploads/2021/12/2021_copenhagen-economics_the-impact-of-e-commerce-on-the-german-economy.pdf (16.09.2023).
- Bawack**, Ransome Epie et al. (2022). Artificial intelligence in E-Commerce: a bibliometric study and literature review, *Electronic Markets* (32), 297–338, Online: [10.1007/s12525-022-00537-z](https://doi.org/10.1007/s12525-022-00537-z).
- Bernhard**, Michael/**Mühling**, Thorsten (2020). Verantwortungsvolle KI Im E-Commerce. Eine kurze Einführung in Verfahren der Künstlichen Intelligenz in der Webshop-Personalisierung, Wiesbaden.
- Bianchin**, Marta/**Angrilli**, Alessandro (2012). Gender differences in emotional responses: a psychophysiological study, *Physiology & Behavior* (105), 925–932, Online: [10.1016/j.physbeh.2011.10.031](https://doi.org/10.1016/j.physbeh.2011.10.031).
- Bitkom** (2017). Jeder Vierte will Chatbots nutzen, Online: <https://www.bitkom.org/Presse/Presseinformation/Jeder-Vierte-will-Chatbots-nutzen.html> (23.12.2023).

- Bitkom** (2022). Künstliche Intelligenz – Wo steht die deutsche Wirtschaft?, Online: https://www.bitkom.org/sites/main/files/2022-09/Charts_Kuenstliche_Intelligenz_130922.pdf (02.01.2024).
- Bitner**, Mary Jo (1992). Servicescapes: The Impact of Physical Surroundings on Customers and Employees, *Journal of Marketing* (56), 57–71, Online: 10.2307/1252042.
- Bleier**, Alexander/**Harmeling**, Colleen M./**Palmatier**, Robert W. (2019). Creating Effective Online Customer Experiences, *Journal of Marketing* (83), 98–119, Online: 10.1177/0022242918809930.
- Bogula**, Werner (2022). KI in der Wirtschaft. Sprach-KI und ihre Anwendungen, in: Dahm, Markus H. (Hg.) *Wie Künstliche Intelligenz unser Leben prägt. KI verständlich erklärt*, Freiburg, 157–164.
- Bojko**, Aga (2013). *Eye tracking the user experience. A practical guide to research*, Brooklyn, New York.
- Boucsein**, Wolfram (2013). *Elektrodermale Aktivität. Grundlagen, Methoden und Anwendungen*, Berlin.
- Bühl**, Achim (2016). *SPSS 23. Einführung in die moderne Datenanalyse*, München.
- Bustamante**, Juan Carlos/**Rubio**, Natalia (2017). Measuring customer experience in physical retail environments, *Journal of Service Management* (28), 884–913, Online: 10.1108/JOSM-06-2016-0142.
- Charness**, Gary/**Gneezy**, Uri/**Kuhn**, Michael A. (2012). Experimental methods: Between-subject and within-subject design, *Journal of Economic Behavior & Organization* (81), 1–8, Online: 10.1016/j.jebo.2011.08.009.
- Cheng**, Xusen et al. (2022). Exploring consumers' response to text-based chatbots in e-commerce: the moderating role of task complexity and chatbot disclosure, *Internet Research* (32), 496–517, Online: 10.1108/INTR-08-2020-0460.
- Christopoulos**, George I./**Uy**, Marilyn A./**Yap**, Wei Jie (2019). The Body and the Brain: Measuring Skin Conductance Responses to Understand the Emotional Experience, *Organizational Research Methods* (22), 394–420, Online: 10.1177/1094428116681073.

- Clore**, Gerald L. et al. (2001). Affective Feelings as Feedback: Some Cognitive Consequences, in: Erlbaum, Lawrence (Hg.) Theories of Mood and Cognition.
- Cohen**, Jacob (1988). Statistical Power Analysis for the Behavioral Sciences, Hoboken.
- Da Silva**, Rui/**Syed**, Sharifah (2006). Cognitive, affective attributes and conative, behavioural responses in retail corporate branding, Journal of Product & Brand Management (15), 293–305, Online: 10.1108/10610420610685703.
- EOS-Gruppe** (2021). EOS Chatbot-Studie 2021. Digitale Helfer setzen sich durch, Online: https://de.eos-solutions.com/de/dam/jcr:4e1eeb45-de9f-467f-8aa8-25d49c0ef3cf/EOS_Chatbot-Studie2021.pdf (27.12.2023).
- Erevelles**, Sunil (1998). The Role of Affect in Marketing, Journal of Business Research (42), 199–215, Online: 10.1016/S0148-2963(97)00118-5.
- Gefen**, David/**Straub**, Detmar W. (2003). Managing User Trust in B2C e-Services, e-Service Journal (2), 7–24, Online: 10.2979/esj.2003.2.2.7.
- Gentsch**, Peter (2019). Künstliche Intelligenz für Sales, Marketing und Service, Wiesbaden.
- Gondorf**, Linda (2023). Wie funktioniert Produktberatung mit KI, Niklas Renner?, Online: <https://www.otto.de/unternehmen/de/kundenfokus/wie-funktioniert-produktberatung-mit-ki-niklas-renner> (05.08.2023).
- HDE** (2021). Online Monitor 2021, Online: https://einzelhandel.de/index.php?option=com_attachments&task=download&id=10572 (19.09.2023).
- HDE** (2023). Online-Monitor 2023, Online: https://einzelhandel.de/index.php?option=com_attachments&task=download&id=10735 (19.09.2023).
- Heinemann**, Gerrit (2018). Die Neuausrichtung des App- und Smartphone-Shopping, Wiesbaden.
- Hermes**, Anna/**Riedl**, René (2021). Dimensions of Retail Customer Experience and Its Outcomes: A Literature Review and Directions for Future Research, in: Nah, Fiona Fui-Hoon/Siau, Keng (Hg.) HCI in Business, Government and Organizations. 8th International Conference, 71–89.
- Holmqvist**, Kenneth (2011). Eye tracking. A comprehensive guide to methods and measures, Oxford/New York.

- Hoyer**, Wayne D. et al. (2020). Transforming the Customer Experience Through New Technologies, *Journal of Interactive Marketing* (51), 57–71, Online: 10.1016/j.intmar.2020.04.001.
- Iovine**, Andrea/**Narducci**, Fedelucio/**Semeraro**, Giovanni (2020). Conversational Recommender Systems and natural language, *Decision Support Systems* (131), 113250, Online: 10.1016/j.dss.2020.113250.
- Janssen**, Jürgen/**Laatz**, Wilfried (2007). *Statistische Datenanalyse mit SPSS für Windows. Eine anwendungsorientierte Einführung in das Basissystem und das Modul Exakte Tests*, Berlin.
- Jenneboer**, Liss/**Herrando**, Carolina/**Constantinides**, Efthymios (2022). The Impact of Chatbots on Customer Loyalty: A Systematic Literature Review, *Journal of Theoretical and Applied Electronic Commerce Research* (17), 212–229, Online: 10.3390/jtaer17010011.
- Kuß**, Alfred/**Wildner**, Raimund/**Kreis**, Henning (2014). *Marktforschung. Grundlagen der Datenerhebung und Datenanalyse*, Wiesbaden.
- Lemon**, Katherine N./**Verhoef**, Peter C. (2016). Understanding Customer Experience Throughout the Customer Journey, *Journal of Marketing* (80), 69–96, Online: 10.1509/jm.15.0420.
- Li**, Meichan/**Wang**, Rui (2023). Chatbots in e-commerce: The effect of chatbot language style on customers' continuance usage intention and attitude toward brand, *Journal of Retailing and Consumer Services* (71), 103209, Online: 10.1016/j.jretconser.2022.103209.
- Liebrecht**, Christine/**Sander**, Lena/**van Hooijdonk**, Charlotte (2021). Too Informal? How a Chatbot's Communication Style Affects Brand Attitude and Quality of Interaction, *Chatbot Research and Design* (2021), 16–31, Online: 10.1007/978-3-030-68288-0_2.
- Lo Presti**, Letizia/**Maggiore**, Giulio/**Marino**, Vittoria (2021). The role of the chatbot on customer purchase intention: towards digital relational sales, *Italian Journal of Marketing* (2021), 165–188, Online: 10.1007/s43039-021-00029-6.

- Lou, Chen/Kang, Hyunjin/Tse, Caleb H.** (2022). Bots vs. humans: how schema congruity, contingency-based interactivity, and sympathy influence consumer perceptions and patronage intentions, *International Journal of Advertising* (41), 655–684, Online: 10.1080/02650487.2021.1951510.
- Luo, Xueming et al.** (2019). Frontiers: Machines vs. Humans: The Impact of Artificial Intelligence Chatbot Disclosure on Customer Purchases, *Marketing Science*, 937–947, Online: 10.1287/mksc.2019.1192.
- Magerhans, Alexander** (2016). *Marktforschung. Eine praxisorientierte Einführung*, Wiesbaden.
- Meyer, Christopher/Schwager, Andre** (2007). Understanding customer experience, *Harvard Business Review* (2007), 116–126.
- Moriuchi, Emi et al.** (2021). Engagement with chatbots versus augmented reality interactive technology in e-commerce, *Journal of Strategic Marketing* (29), 375–389, Online: 10.1080/0965254X.2020.1740766.
- Ngai, Eric W.T. et al.** (2021). An intelligent knowledge-based chatbot for customer service, *Electronic Commerce Research and Applications* (50), 101098, Online: 10.1016/j.elerap.2021.101098.
- OpenAI** (2023). Introducing ChatGPT, Online: <https://openai.com/blog/chatgpt> (19.11.2023).
- Poulsson, Susanne H.G./Kale, Sudhir H.** (2004). The Experience Economy and Commercial Experiences, *The Marketing Review* (4), 267–277, Online: 10.1362/1469347042223445.
- Röder, Katharina-Maria** (2023). Chatbots im E-Commerce: Vorteile und Funktionsweise, Online: <https://omr.com/de/reviews/contenthub/e-commerce-chatbot> (05.08.2023).
- Roemer, Ellen et al.** (2022). Eye Tracking as a Research Method for Social Media, in: Hanlon, Annmarie/Tuten, Tracy L. (Hg.) *The SAGE Handbook of Social Media Marketing*, 161–180.
- Roy, Sanjit K./Gruner, Richard L./Guo, Jingjing** (2022). Exploring customer experience, commitment, and engagement behaviours, *Journal of Strategic Marketing* (30), 45–68, Online: 10.1080/0965254X.2019.1642937.
- Schmitt, Bernd** (1999). Experiential Marketing, *Journal of Marketing Management* (15), 53–67, Online: 10.1362/026725799784870496.

- Statista Market Insights** (2023). Umsatz im Markt für E-Commerce weltweit in den Jahren 2018 bis 2027, Online: <https://de.statista.com/prognosen/484763/prognose-der-umsaetze-im-e-commerce-markt-in-der-welt> (27.12.2023).
- Tajfel**, Henri (2010). Human groups and social categories. Studies in social psychology, Cambridge.
- Trivedi**, Jay (2019). Examining the Customer Experience of Using Banking Chatbots and Its Impact on Brand Love: The Moderating Role of Perceived Risk, *Journal of Internet Commerce* (18), 91–111, Online: 10.1080/15332861.2019.1567188.
- Verhoef**, Peter C. et al. (2009). Customer Experience Creation: Determinants, Dynamics and Management Strategies, *Journal of Retailing* (85), 31–41, Online: 10.1016/j.jretai.2008.11.001.
- Verhulst**, Nanouk et al. (2020). A neurophysiological exploration of the dynamic nature of emotions during the customer experience, *Journal of Retailing and Consumer Services* (57), 102217, Online: 10.1016/j.jretconser.2020.102217.
- Wübbenhorst**, Klaus (2018). Definition: Hautwiderstandsmessung, Online: <https://wirtschaftslexikon.gabler.de/definition/hautwiderstandsmessung-33447> (19.08.2023).
- Xu**, Yingzi et al. (2020). AI Customer Service: Task Complexity, Problem-Solving Ability, and Usage Intention, *Australasian Marketing Journal* (28), 189–199, Online: 10.1016/j.ausmj.2020.03.005.

Anhang

Anhang A: Laborversuch.....	58
Anhang A-1: Versuchsaufbau.....	58
Anhang A-2: Hautwiderstandssensoren (Shimmer)	59
Anhang A-3: Einverständniserklärung	60
Anhang A-4: Hinweise für Versuchsleiter.....	61
Anhang A-5: Dokumentation während des Versuchs.....	62
Anhang B: Ablauf und Aufgabenstellung.....	63
Anhang C: Fragebogen	68
Anhang C-1: Fragebogen in SosciSurvey.....	68
Anhang C-2: Codierung des Fragebogens	71
Anhang D: Darstellung und Auswertung der Ergebnisse.....	73
Anhang D-1: Zusammengeführter Datensatz (bereinigt)	73
Anhang D-2: Darstellung Stichprobe.....	75
Anhang D-3: Auswertung Hautreaktion	76
Anhang D-4: Auswertung Eye-Tracking.....	80
Anhang D-5: Auswertung Fragebogen	82

Anhang A: Laborversuch

Anhang A-1: Versuchsaufbau

Foto 1: Allgemeiner Versuchsaufbau



Foto 2: Perspektive Versuchsleiter

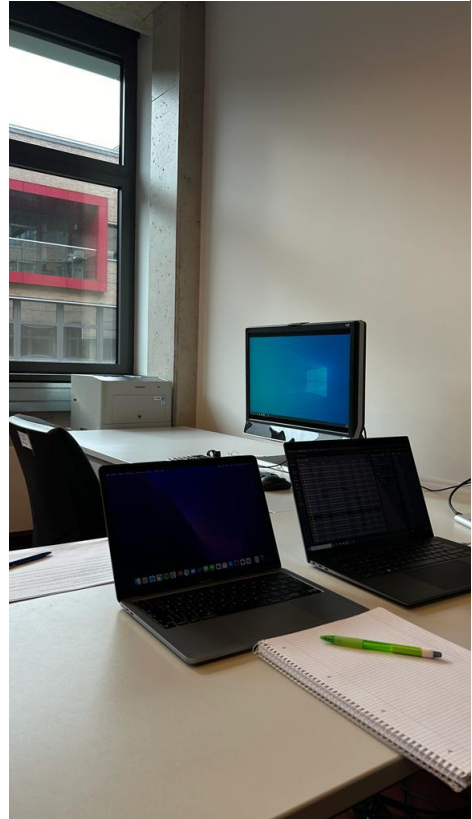


Foto 3: Perspektive Versuchsperson



Foto 4: Zweite Perspektive auf Bildschirm



Anhang A-2: Hautwiderstandssensoren (Shimmer)

Foto 5: Shimmer Sensoren mit Empfänger (1)



Foto 6: Shimmer Sensoren mit Empfänger (2)



Anhang A-3: Einverständniserklärung

Einverständniserklärung – Studie im Rahmen der Bachelorarbeit von *Mike Trynczyk* an der *Hochschule Ruhr West* 1

Einverständniserklärung zur Teilnahme an einer wissenschaftlichen Studie

Ich erkläre mich freiwillig bereit, an der wissenschaftlichen Studie teilzunehmen, die von Mike Trynczyk an der Hochschule Ruhr West im Rahmen seiner Bachelorarbeit durchgeführt wird.

Ich verstehe, dass die Teilnahme an dieser Studie einen Laborversuch mit Eye-Tracking, Hautwiderstandsmessung und einem Fragebogen umfassen wird. Ich habe ausreichend Gelegenheit gehabt, Fragen zu stellen und die Studie zu verstehen.

Ich erkläre mich damit einverstanden, dass meine Daten und meine Teilnahme an dieser Studie vertraulich behandelt werden. Mein Name und andere personenbezogene Informationen werden anonymisiert.

Mir ist bewusst, dass meine Teilnahme freiwillig ist, und ich kann meine Teilnahme jederzeit ohne Angabe von Gründen und ohne negative Konsequenzen abbrechen.

Ich verstehe, dass die Ergebnisse dieser Studie möglicherweise in der Bachelorarbeit von Mike Trynczyk veröffentlicht werden, jedoch ohne meine persönlichen Daten.

Durch meine Unterschrift bestätige ich, dass ich die Einverständniserklärung gelesen und verstanden habe und dass ich freiwillig an dieser Studie teilnehme.

Kontakt: mike.trynczyk@stud.hs-ruhrwest.de

Anhang A-4: Hinweise für Versuchsleiter

Versuchsablauf – Hinweise für Versuchsleiter

- **Begrüßung:** Studie im Rahmen meiner Bachelorarbeit: Analyse der Customer Experience im E-Commerce (Onlinehandel) mit Eye-Tracking und Hautwiderstandsmessung gefolgt von einem kurzen Fragebogen
- **Einverständniserklärung** durchlesen und unterschreiben (offene Fragen?)
- Anbringung der **Sensoren** an nicht-dominanter Hand (Links- oder Rechtshänder? Zeige- & Mittelfinger)
- Kurze Erklärung zum **Versuchsablauf:**
 - Versuch findet ausschließlich am Computer statt
 - Dauer ca. 6 – 9 Minuten
 - Alles, was erklärt wird, steht auch nochmal auf dem Bildschirm
 - Versuch startet mit einer kurzen Kalibrierung (Punkte auf Bildschirm)
 - Dann kurze Messung, die 15 Sekunden dauert, danach automatische Weiterleitung! (**Bitte nicht überspringen!**)
 - Danach folgen nacheinander 3 unterschiedliche Aufgaben die einzeln in einem Apotheken-Onlineshop zu lösen sind (**Lösungen müssen nicht aufgeschrieben oder gemerkt werden!**)
 - **Aufgaben gründlich lesen!**
 - Wenn Sie der Meinung sind, die jeweilige Aufgabe erledigt zu haben. Dann zur nächsten weitergehen.
 - Nach den Aufgaben erfolgt eine zweite 15-sekündige Messung, danach automatische Weiterleitung! (**Bitte nicht überspringen!**)
 - Zum Schluss ein kurzer Fragebogen mit 12 Fragen auf einer Skala von 1 bis 7 (**Nur Fragen zu dem Versuch – bitte intuitiv antworten und keine großen Gedanken machen!**)
- **Hand:** Bitte nutzen Sie für das Tippen und Bewegen der Maus nur die Hand ohne Sensoren. Halten Sie die Hand mit Sensoren so gut es geht still.
- **Nicht reden:** Wenn möglich bitte während des gesamten Versuches nicht reden.
- **Wartezeit beim Klicken:** Es kann 1-2 Sekunden länger dauern bis etwas geladen hat, wenn sie etwas eintippen oder klicken (Geduldig sein!)
- **Probanden-Nummer:** Wichtig für **Fragebogen** nachher. In erster Frage kurz notieren.
- **Offene Fragen?**

Anhang A-5: Dokumentation während des Versuchs

Proband	Geschlecht	Unterschrift (Ja/Nein)	Studie	Chatbot (MIT/OHNE)	Hinweise/Dokumentation
03	m	Ja	HAUPTSTUDIE	MIT	Nicht nur Chatbot für die Aufgaben genutzt. Feedback: Chatbot schwierig zu finden. Eigene Suche effizienter.
04	w	Ja	HAUPTSTUDIE	OHNE	keine
05	m	Ja	HAUPTSTUDIE	MIT	Während Aufgabe 3 gesprochen: Rückfrage zur Bedienung der Tastatur (Y und Z), trotzdem keine Auffälligkeit in GSR-Daten
06	w	Ja	HAUPTSTUDIE	OHNE	Während Aufgabe 1 gesprochen: Rückfrage zur Bedienung der Tastatur (F10), trotzdem keine Auffälligkeit in GSR-Daten Basiswert Messung 2 übersprungen (kürzer als 15 Sekunden: trotzdem Werte bekommen)
07	w	Ja	FAIL	MIT	Ausfälle bei GSR-Messung (ab Aufgabe 2): Datensatz unvollständig!
08	w	Ja	HAUPTSTUDIE	OHNE	keine
09	m	Ja	HAUPTSTUDIE	OHNE	Zum Ende der Aufgabe 1 kurz gesprochen (keine Auffälligkeit in GSR-Daten zu finden)
10	w	Ja	HAUPTSTUDIE	MIT	Während Aufgabe 1 und 3 Rückfragen gestellt (Aufgabe vergessen), keine Auffälligkeit in GSR-Daten. Basiswert-Messung 1 übersprungen (kürzer als 15 Sekunden: trotzdem Werte erhalten)
11	m	Ja	FAIL	OHNE	keine GSR-Messung (Verbindungsfehler - Akku leer)
12	m	Ja	FAIL	OHNE	keine GSR-Messung (Verbindungsfehler - Akku leer)
13	w	Ja	FAIL	OHNE	keine GSR-Messung (Verbindungsfehler - Akku leer)
14	m	Ja	HAUPTSTUDIE	MIT	keine
15	w	Ja	HAUPTSTUDIE	MIT	keine
16	m	Ja	HAUPTSTUDIE	OHNE	Falsche Szenariozuteilung in Tobii. Trotzdem Aufgabe ohne Chatbot gelöst und auch so angegeben im Fragebogen.
17	m	Ja	HAUPTSTUDIE	MIT	Aufgabe 1 ohne Chatbot gelöst (Feedback: nicht gefunden; lange Suche)
18	m	Ja	FAIL	MIT	keine GSR-Messung (Verbindungsfehler - Akku leer)
19	m	Ja	HAUPTSTUDIE	OHNE	keine
20	m	Ja	HAUPTSTUDIE	OHNE	keine
21	w	Ja	FAIL	OHNE	keine GSR-Messung (Verbindungsfehler - Akku leer)
22	w	Ja	HAUPTSTUDIE	MIT	keine
23	w	Ja	HAUPTSTUDIE	OHNE	keine
24	m	Ja	HAUPTSTUDIE	MIT	keine
25	w	Ja	HAUPTSTUDIE	MIT	keine
26	w	Ja	FAIL	OHNE	keine
27	w	Ja	HAUPTSTUDIE	MIT	keine
28	m	Ja	HAUPTSTUDIE	OHNE	keine
29	m	Ja	HAUPTSTUDIE	MIT	keine
30	w	Ja	HAUPTSTUDIE	MIT	Aufgabe 1 ohne Chatbot erledigt, Begründung: nicht gefunden
31	m	Ja	HAUPTSTUDIE	OHNE	keine
32	m	Ja	HAUPTSTUDIE	OHNE	keine
33	m	Ja	HAUPTSTUDIE	MIT	Schwierigkeiten mit Tastatur bei Aufgabe 1: kurz beide Hände genutzt (keine Auffälligkeiten bei den GSR-Daten erkennbar)
34	m	Ja	HAUPTSTUDIE	MIT	keine
35	w	Ja	HAUPTSTUDIE	OHNE	keine
36	w	Ja	HAUPTSTUDIE	OHNE	keine
37	w	Ja	HAUPTSTUDIE	OHNE	Schwierigkeiten bei Aufgabe 3: Lange Suche nach Lösung
38	w	Ja	HAUPTSTUDIE	MIT	Aufgabenauflosung ziemlich schnell. Proband hat Erfahrung mit andere Chatbots.
39	m	Ja	HAUPTSTUDIE	OHNE	keine

Anhang B: Ablauf und Aufgabenstellung

Begrüßung (beide Gruppen)

Herzlich Willkommen zur Studie im Rahmen meiner Bachelorarbeit mit dem Fokus E-Commerce.

Im Folgenden werden Sie gebeten mehrere Aufgaben in einem Onlineshop durchzuführen.

*Durch Drücken der Taste **F10** gelangen Sie zur nächsten Seite.*

Ablauf (beide Gruppen)

Ablauf

Zunächst erfolgt eine erste 15-sekündige Messung.

Im Anschluss werden die Aufgaben auf den nachfolgenden Seiten genauer erläutert. Jede Aufgabe wird separat erklärt und ist unabhängig von den anderen zu lösen. Insgesamt sind drei Aufgaben nacheinander zu absolvieren.

Bitte nutzen Sie für das Tippen und Bewegen der Maus nur Ihre dominante Hand (ohne Sensoren). Halten Sie Ihre andere Hand (mit Sensoren) während der Messungen und Absolvierung der Aufgaben so gut wie möglich ruhig.

Nach Absolvierung der Aufgaben erfolgt eine zweite 15-sekündige Messung.

Danach werden Sie zu automatisch zu einem Fragebogen weitergeleitet. Alle Daten werden vertraulich behandelt und dienen nur zu diesem Forschungszweck.

*Durch Drücken der Taste **F10** starten Sie die erste Messung.*

Basiswert-Messung 1 (VOR) (beide Gruppen)

Die erste Messung des Basiswertes startet jetzt.

Bitte halten Sie still und bewegen Ihre Hand nicht.

Nach ca. 15 Sekunden werden Sie **automatisch** zur Aufgabenstellung weitergeleitet.

Allgemeine Aufgabenstellung (mit Chatbot)

Aufgabenstellung

Bitte absolvieren Sie die folgenden Aufgaben im Apotheken-Onlineshop www.shop-apotheke.de. Sie brauchen die Lösungen nicht zu dokumentieren und sich nicht zu merken.

Die Aufgaben sind nacheinander und unabhängig voneinander zu bewältigen. Jede Aufgabe wird separat erklärt. Nach Absolvierung einer Aufgabe werden Sie zur nächsten Aufgabe weitergeleitet.

(!) Hinweis: Finden Sie den digitalen Assistenten (Chatbot) auf der Startseite im Onlineshop und nutzen Sie diesen zur Lösung aller Aufgaben.

Wenn Sie der Meinung sind, die jeweilige Aufgaben absolviert zu haben, werden Sie durch Drücken der Taste F10 aus dem Onlineshop zurückgeleitet.

Drücken Sie jetzt die Taste F10, um zur ersten Aufgabe weitergeleitet zu werden.

Alternative: Allgemeine Aufgabenstellung (ohne Chatbot)

Aufgabenstellung

Bitte absolvieren Sie die folgenden Aufgaben im Apotheken-Onlineshop www.shop-apotheke.de. Sie brauchen die Lösungen nicht zu dokumentieren und sich nicht zu merken.

Die Aufgaben sind nacheinander und unabhängig voneinander zu bewältigen. Jede Aufgabe wird separat erklärt. Nach Absolvierung einer Aufgabe werden Sie zur nächsten Aufgabe weitergeleitet.

Hinweis: Nutzen Sie nicht den digitalen Assistenten (Chatbot) im Onlineshop zur Lösung der Aufgabe.

Wenn Sie der Meinung sind, die jeweilige Aufgaben absolviert zu haben, werden Sie durch Drücken der Taste F10 aus dem Onlineshop zurückgeleitet.

Drücken Sie jetzt die Taste F10, um zur ersten Aufgabe weitergeleitet zu werden.

Aufgabe 1 (mit Chatbot)

Aufgabe 1:

Finden Sie zunächst den digitalen Assistenten (Chatbot) auf der Startseite. Informieren Sie sich dann über folgendes **Produkt**: „Paracetamol-ratiopharm 1000mg Tabletten“. Wie hoch ist der **Preis** und wie groß ist die **Packungsgröße**?

(!) Hinweis: Finden Sie den digitalen Assistenten (Chatbot) auf der Startseite im Onlineshop und nutzen Sie diesen zur Lösung der Aufgabe.

Wenn Sie der Meinung sind, die Aufgabe absolviert zu haben, werden Sie durch Drücken der Taste F10 aus dem Onlineshop zurückgeleitet.

Drücken Sie jetzt die Taste F10, um zum Onlineshop weitergeleitet zu werden.

Alternative: Aufgabe 1 (ohne Chatbot)

Aufgabe 1:

Informieren Sie sich über folgendes **Produkt**: „Paracetamol-ratiopharm 1000mg Tabletten“
Wie hoch ist der **Preis** und wie groß ist die **Packungsgröße**?

*Hinweis: Nutzen Sie **nicht** den digitalen Assistenten (Chatbot) im Onlineshop zur Lösung der Aufgabe.*

*Wenn Sie der Meinung sind, die Aufgabe absolviert zu haben, werden Sie durch Drücken der **Taste F10** aus dem Onlineshop zurückgeleitet.
Drücken Sie jetzt die **Taste F10**, um zum Onlineshop weitergeleitet zu werden.*

Aufgabe 2 (mit Chatbot)

Aufgabe 2:

Informieren Sie sich über die **Versandkosten** im Onlineshop.
Ab welchem **Mindestbestellwert** ist die Bestellung **versandkostenfrei**?

*(!) Hinweis: Finden Sie den **digitalen Assistenten (Chatbot)** auf der Startseite im Onlineshop und nutzen Sie diesen zur Lösung der Aufgabe.*

*Wenn Sie der Meinung sind, die Aufgaben absolviert zu haben, werden Sie durch Drücken der **Taste F10** aus dem Onlineshop zurückgeleitet.*

*Drücken Sie jetzt die **Taste F10**, um zum Onlineshop weitergeleitet zu werden.*

Alternative: Aufgabe 2 (ohne Chatbot)

Aufgabe 2:

Informieren Sie sich über die **Versandkosten** im Onlineshop.
Ab welchem **Mindestbestellwert** ist die Bestellung **versandkostenfrei**?

*Hinweis: Nutzen Sie **nicht** den digitalen Assistenten (Chatbot) im Onlineshop zur Lösung der Aufgabe.*

*Wenn Sie der Meinung sind, die Aufgaben absolviert zu haben, werden Sie durch Drücken der **Taste F10** aus dem Onlineshop zurückgeleitet.*

*Drücken Sie jetzt die **Taste F10**, um zum Onlineshop weitergeleitet zu werden.*

Aufgabe 3 (mit Chatbot)

Aufgabe 3:

Informieren Sie sich über die **Zahlarten**.
Welche Zahlarten gibt es?

*(!) Hinweis: Finden Sie den **digitalen Assistenten (Chatbot)** auf der Startseite im Onlineshop und nutzen Sie diesen zur Lösung der Aufgabe.*

*Wenn Sie der Meinung sind, die Aufgaben absolviert zu haben, werden Sie durch Drücken der **Taste F10** aus dem Onlineshop zurückgeleitet.*

*Drücken Sie jetzt die **Taste F10**, um zum Onlineshop weitergeleitet zu werden.*

Alternative: Aufgabe 3 (ohne Chatbot)

Aufgabe 3:

Informieren Sie sich über die **Zahlarten**.
Welche Zahlarten gibt es?

*Hinweis: Nutzen Sie **nicht** den digitalen Assistenten (Chatbot) im Onlineshop zur Lösung der Aufgabe.*

*Wenn Sie der Meinung sind, die Aufgaben absolviert zu haben, werden Sie durch Drücken der **Taste F10** aus dem Onlineshop zurückgeleitet.*

*Drücken Sie jetzt die **Taste F10**, um zum Onlineshop weitergeleitet zu werden.*

Verabschiedung (beide Gruppen)

Vielen Dank für die Absolvierung der Aufgaben.

Auf der nächsten Seite wird eine zweite Basiswert-Messung durchgeführt.
Diese dauert ebenfalls ca. 15 Sekunden.

Danach werden Sie automatisch zum Fragebogen weitergeleitet.

*Durch Drücken der **Taste F10** starten Sie die Messung.*

Basiswert-Messung 2 (NACH) und Weiterleitung Fragebogen (beide Gruppen)

Die Messung startet jetzt.

Bitte halten Sie still und bewegen Ihre Hand nicht.

Nach ca. 15 Sekunden werden Sie **automatisch** zum Fragebogen weitergeleitet.

Anhang C: Fragebogen

Anhang C-1: Fragebogen in SosciSurvey



0% ausgefüllt

Fragebogen

Im letzten Schritt bitte ich Sie, die nachfolgenden Fragen zu beantworten.

Bitte bewerten Sie ausschließlich die eigenen Erfahrungen, die Sie während des Besuchs des Onlineshops in diesem Laborversuch gemacht haben.

Bitte beantworten Sie jede Frage.

Alle Angaben werden vertraulich und anonym behandelt. Sie dienen nur zu Forschungszwecken im Rahmen der Bachelorarbeit.

Weiter



14% ausgefüllt

1. Bitte geben Sie die Probanden-Nummer ein, die Ihnen zugeordnet wurde:

2. Haben Sie zur Absolvierung der Aufgaben den digitalen Assistenten (Chatbot) im Onlineshop genutzt?

Ja, ich habe den Chatbot genutzt.

Nein, ich habe keinen Chatbot genutzt.

Zurück

Weiter



29% ausgefüllt

3. Kognitive Erfahrung – Inwiefern stimmen Sie folgenden Aussagen zu?

Bitte antworten Sie jeweils auf einer Skala von 1 bis 7, wobei 1 für „stimme gar nicht zu“ und 7 für „stimme voll und ganz zu“ steht.

Die Darstellung, der Service und die Interaktion mit dem Onlineshop...	Stimme gar nicht zu						Stimme voll und ganz zu
	1	2	3	4	5	6	7
a) lehren mich interessante Dinge.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b) wecken meine Kreativität.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c) bringen mich auf interessante Ideen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Zurück

Weiter



43% ausgefüllt

4. Soziale Erfahrung – Inwiefern stimmen Sie folgenden Aussagen zu?

Bitte antworten Sie jeweils auf einer Skala von 1 bis 7, wobei 1 für „stimme gar nicht zu“ und 7 für „stimme voll und ganz zu“ steht.

Die Darstellung, der Service und die Interaktion mit dem Onlineshop	Stimme gar nicht zu						Stimme voll und ganz zu
	1	2	3	4	5	6	7
a) geben mir ein Gefühl des menschlichen Kontakts.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b) geben mir ein Gefühl der menschlichen Wärme.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c) geben mir ein Gefühl der menschlichen Sensibilität.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Zurück

Weiter



57% ausgefüllt

5. Affektive Erfahrung – Die Darstellung, der Service und die Interaktion mit dem Onlineshop machen mich...

Bitte bewerten Sie jeweils auf einer Skala von 1 bis 7.

	1	2	3	4	5	6	7	
schlecht gelaunt	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	gut gelaunt
unglücklich	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	glücklich
pessimistisch	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	optimistisch

Zurück

Weiter



71% ausgefüllt

6. Physische Erfahrung – Die Darstellung, der Service und die Interaktion mit dem Onlineshop haben...

Bitte bewerten Sie jeweils auf einer Skala von 1 bis 7.

	1	2	3	4	5	6	7	
mich überfordert	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	mich entspannt
mir Energie geraubt	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	mir Energie gegeben
mich ermüdet	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	mich vitalisiert

Zurück

Weiter



86% ausgefüllt

7. Wie alt sind Sie?

Jahre

keine Angabe

8. Mit welchem Geschlecht identifizieren Sie sich?

männlich

weiblich

divers

keine Angabe

9. Was ist Ihr derzeitiger beruflicher Status?

Bitte wählen Sie ihre Haupttätigkeit (nur eine Auswahl möglich)

Selbstständiger

Angestellter

Beamter

Arbeitssuchender

Schüler:in

Auszubildender

Student:in

Sonstiges:

keine Angabe

Zurück

Weiter



Vielen Dank für Ihre Teilnahme!

Ich möchte mich ganz herzlich für Ihre Mithilfe bedanken. Der Laborversuch ist nun zu Ende.

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

Anhang C-2: Codierung des Fragebogens

Variable	Variable Label	Response Code	Response Label	Variable Type	Input Type
CASE	Interview-Nummer (fortlaufend)			METRIC	TXT
SERIAL	Personenkennung oder Teilnahmecode (sofern verwendet)			TEXT	TXT
REF	Referenz (sofern im Link angegeben)			TEXT	TXT
QUESTNNR	Fragebogen, der im Interview verwendet wurde			TEXT	TXT
MODE	Interview-Modus			TEXT	TXT
STARTED	Zeitpunkt zu dem das Interview begonnen hat (Europe/Berlin)			TIME	TXT
AF01_01	Affektiv: schlecht gelaunt/gut gelaunt	1	schlecht gelaunt [1]	ORDINAL	MC
AF01_01	Affektiv: schlecht gelaunt/gut gelaunt	7	gut gelaunt [7]	ORDINAL	MC
AF01_01	Affektiv: schlecht gelaunt/gut gelaunt	2	[2]	ORDINAL	MC
AF01_01	Affektiv: schlecht gelaunt/gut gelaunt	3	[3]	ORDINAL	MC
AF01_01	Affektiv: schlecht gelaunt/gut gelaunt	4	[4]	ORDINAL	MC
AF01_01	Affektiv: schlecht gelaunt/gut gelaunt	5	[5]	ORDINAL	MC
AF01_01	Affektiv: schlecht gelaunt/gut gelaunt	6	[6]	ORDINAL	MC
AF01_01	Affektiv: schlecht gelaunt/gut gelaunt	-9	nicht beantwortet	ORDINAL	MC
AF01_02	Affektiv: unglücklich/glücklich	1	unglücklich [1]	ORDINAL	MC
AF01_02	Affektiv: unglücklich/glücklich	7	glücklich [7]	ORDINAL	MC
AF01_02	Affektiv: unglücklich/glücklich	2	[2]	ORDINAL	MC
AF01_02	Affektiv: unglücklich/glücklich	3	[3]	ORDINAL	MC
AF01_02	Affektiv: unglücklich/glücklich	4	[4]	ORDINAL	MC
AF01_02	Affektiv: unglücklich/glücklich	5	[5]	ORDINAL	MC
AF01_02	Affektiv: unglücklich/glücklich	6	[6]	ORDINAL	MC
AF01_02	Affektiv: unglücklich/glücklich	-9	nicht beantwortet	ORDINAL	MC
AF01_03	Affektiv: pessimistisch/optimistisch	1	pessimistisch [1]	ORDINAL	MC
AF01_03	Affektiv: pessimistisch/optimistisch	7	optimistisch [7]	ORDINAL	MC
AF01_03	Affektiv: pessimistisch/optimistisch	2	[2]	ORDINAL	MC
AF01_03	Affektiv: pessimistisch/optimistisch	3	[3]	ORDINAL	MC
AF01_03	Affektiv: pessimistisch/optimistisch	4	[4]	ORDINAL	MC
AF01_03	Affektiv: pessimistisch/optimistisch	5	[5]	ORDINAL	MC
AF01_03	Affektiv: pessimistisch/optimistisch	6	[6]	ORDINAL	MC
AF01_03	Affektiv: pessimistisch/optimistisch	-9	nicht beantwortet	ORDINAL	MC
KO01_01	Kognitiv: a) lehren mich interessante Dinge.	1	Stimme gar nicht zu [1]	ORDINAL	MC
KO01_01	Kognitiv: a) lehren mich interessante Dinge.	2	[2]	ORDINAL	MC
KO01_01	Kognitiv: a) lehren mich interessante Dinge.	3	[3]	ORDINAL	MC
KO01_01	Kognitiv: a) lehren mich interessante Dinge.	4	[4]	ORDINAL	MC
KO01_01	Kognitiv: a) lehren mich interessante Dinge.	5	[5]	ORDINAL	MC
KO01_01	Kognitiv: a) lehren mich interessante Dinge.	6	[6]	ORDINAL	MC
KO01_01	Kognitiv: a) lehren mich interessante Dinge.	7	Stimme voll und ganz zu	ORDINAL	MC
KO01_01	Kognitiv: a) lehren mich interessante Dinge.	-9	nicht beantwortet	ORDINAL	MC
KO01_02	Kognitiv: b) wecken meine Kreativität.	1	Stimme gar nicht zu [1]	ORDINAL	MC
KO01_02	Kognitiv: b) wecken meine Kreativität.	2	[2]	ORDINAL	MC
KO01_02	Kognitiv: b) wecken meine Kreativität.	3	[3]	ORDINAL	MC
KO01_02	Kognitiv: b) wecken meine Kreativität.	4	[4]	ORDINAL	MC
KO01_02	Kognitiv: b) wecken meine Kreativität.	5	[5]	ORDINAL	MC
KO01_02	Kognitiv: b) wecken meine Kreativität.	6	[6]	ORDINAL	MC
KO01_02	Kognitiv: b) wecken meine Kreativität.	7	Stimme voll und ganz zu	ORDINAL	MC
KO01_02	Kognitiv: b) wecken meine Kreativität.	-9	nicht beantwortet	ORDINAL	MC
KO01_03	Kognitiv: c) bringen mich auf interessante Ideen	1	Stimme gar nicht zu [1]	ORDINAL	MC
KO01_03	Kognitiv: c) bringen mich auf interessante Ideen	2	[2]	ORDINAL	MC
KO01_03	Kognitiv: c) bringen mich auf interessante Ideen	3	[3]	ORDINAL	MC
KO01_03	Kognitiv: c) bringen mich auf interessante Ideen	4	[4]	ORDINAL	MC
KO01_03	Kognitiv: c) bringen mich auf interessante Ideen	5	[5]	ORDINAL	MC
KO01_03	Kognitiv: c) bringen mich auf interessante Ideen	6	[6]	ORDINAL	MC
KO01_03	Kognitiv: c) bringen mich auf interessante Ideen	7	Stimme voll und ganz zu	ORDINAL	MC
KO01_03	Kognitiv: c) bringen mich auf interessante Ideen	-9	nicht beantwortet	ORDINAL	MC
PH01_01	Physisch: mich überfordert/mich entspannt	1	mich überfordert [1]	ORDINAL	MC
PH01_01	Physisch: mich überfordert/mich entspannt	7	mich entspannt [7]	ORDINAL	MC
PH01_01	Physisch: mich überfordert/mich entspannt	2	[2]	ORDINAL	MC
PH01_01	Physisch: mich überfordert/mich entspannt	3	[3]	ORDINAL	MC
PH01_01	Physisch: mich überfordert/mich entspannt	4	[4]	ORDINAL	MC
PH01_01	Physisch: mich überfordert/mich entspannt	5	[5]	ORDINAL	MC
PH01_01	Physisch: mich überfordert/mich entspannt	6	[6]	ORDINAL	MC
PH01_01	Physisch: mich überfordert/mich entspannt	-9	nicht beantwortet	ORDINAL	MC
PH01_02	Physisch: mir Energie geraubt/mir Energie gegeben	1	mir Energie geraubt [1]	ORDINAL	MC
PH01_02	Physisch: mir Energie geraubt/mir Energie gegeben	7	mir Energie gegeben [7]	ORDINAL	MC
PH01_02	Physisch: mir Energie geraubt/mir Energie gegeben	2	[2]	ORDINAL	MC
PH01_02	Physisch: mir Energie geraubt/mir Energie gegeben	3	[3]	ORDINAL	MC
PH01_02	Physisch: mir Energie geraubt/mir Energie gegeben	4	[4]	ORDINAL	MC
PH01_02	Physisch: mir Energie geraubt/mir Energie gegeben	5	[5]	ORDINAL	MC
PH01_02	Physisch: mir Energie geraubt/mir Energie gegeben	6	[6]	ORDINAL	MC
PH01_02	Physisch: mir Energie geraubt/mir Energie gegeben	-9	nicht beantwortet	ORDINAL	MC
PH01_03	Physisch: mich ermüdet/mich vitalisiert	1	mich ermüdet [1]	ORDINAL	MC
PH01_03	Physisch: mich ermüdet/mich vitalisiert	7	mich vitalisiert [7]	ORDINAL	MC
PH01_03	Physisch: mich ermüdet/mich vitalisiert	2	[2]	ORDINAL	MC
PH01_03	Physisch: mich ermüdet/mich vitalisiert	3	[3]	ORDINAL	MC
PH01_03	Physisch: mich ermüdet/mich vitalisiert	4	[4]	ORDINAL	MC
PH01_03	Physisch: mich ermüdet/mich vitalisiert	5	[5]	ORDINAL	MC
PH01_03	Physisch: mich ermüdet/mich vitalisiert	6	[6]	ORDINAL	MC
PH01_03	Physisch: mich ermüdet/mich vitalisiert	-9	nicht beantwortet	ORDINAL	MC

Anhang

SO01_01	Sozial: a) geben mir ein Gefühl des menschlich	1 Stimme gar nicht zu [1]	ORDINAL	MC
SO01_01	Sozial: a) geben mir ein Gefühl des menschlich	2 [2]	ORDINAL	MC
SO01_01	Sozial: a) geben mir ein Gefühl des menschlich	3 [3]	ORDINAL	MC
SO01_01	Sozial: a) geben mir ein Gefühl des menschlich	4 [4]	ORDINAL	MC
SO01_01	Sozial: a) geben mir ein Gefühl des menschlich	5 [5]	ORDINAL	MC
SO01_01	Sozial: a) geben mir ein Gefühl des menschlich	6 [6]	ORDINAL	MC
SO01_01	Sozial: a) geben mir ein Gefühl des menschlich	7 Stimme voll und ganz zu	ORDINAL	MC
SO01_01	Sozial: a) geben mir ein Gefühl des menschlich	-9 nicht beantwortet	ORDINAL	MC
SO01_02	Sozial: b) geben mir ein Gefühl der menschliche	1 Stimme gar nicht zu [1]	ORDINAL	MC
SO01_02	Sozial: b) geben mir ein Gefühl der menschliche	2 [2]	ORDINAL	MC
SO01_02	Sozial: b) geben mir ein Gefühl der menschliche	3 [3]	ORDINAL	MC
SO01_02	Sozial: b) geben mir ein Gefühl der menschliche	4 [4]	ORDINAL	MC
SO01_02	Sozial: b) geben mir ein Gefühl der menschliche	5 [5]	ORDINAL	MC
SO01_02	Sozial: b) geben mir ein Gefühl der menschliche	6 [6]	ORDINAL	MC
SO01_02	Sozial: b) geben mir ein Gefühl der menschliche	7 Stimme voll und ganz zu	ORDINAL	MC
SO01_02	Sozial: b) geben mir ein Gefühl der menschliche	-9 nicht beantwortet	ORDINAL	MC
SO01_03	Sozial: c) geben mir ein Gefühl der menschliche	1 Stimme gar nicht zu [1]	ORDINAL	MC
SO01_03	Sozial: c) geben mir ein Gefühl der menschliche	2 [2]	ORDINAL	MC
SO01_03	Sozial: c) geben mir ein Gefühl der menschliche	3 [3]	ORDINAL	MC
SO01_03	Sozial: c) geben mir ein Gefühl der menschliche	4 [4]	ORDINAL	MC
SO01_03	Sozial: c) geben mir ein Gefühl der menschliche	5 [5]	ORDINAL	MC
SO01_03	Sozial: c) geben mir ein Gefühl der menschliche	6 [6]	ORDINAL	MC
SO01_03	Sozial: c) geben mir ein Gefühl der menschliche	7 Stimme voll und ganz zu	ORDINAL	MC
SO01_03	Sozial: c) geben mir ein Gefühl der menschliche	-9 nicht beantwortet	ORDINAL	MC
SZ01_01	Alter: ... Jahre		METRIC	TXT
SZ01_01a	Alter: ... Jahre: keine Angabe	1 nicht gewählt	DICHOTOMOUS	CK
SZ01_01a	Alter: ... Jahre: keine Angabe	2 ausgewählt	DICHOTOMOUS	CK
SZ02	Geschlecht	1 männlich	NOMINAL	MC
SZ02	Geschlecht	2 weiblich	NOMINAL	MC
SZ02	Geschlecht	3 divers	NOMINAL	MC
SZ02	Geschlecht	-1 keine Angabe	NOMINAL	MC
SZ02	Geschlecht	-9 nicht beantwortet	NOMINAL	MC
SZ03	Beruf	1 Selbstständige:r	NOMINAL	MC
SZ03	Beruf	2 Angestellte:r	NOMINAL	MC
SZ03	Beruf	3 Beamte:r	NOMINAL	MC
SZ03	Beruf	4 Arbeitssuchende:r	NOMINAL	MC
SZ03	Beruf	5 Schüler:in	NOMINAL	MC
SZ03	Beruf	6 Auszubildende:r	NOMINAL	MC
SZ03	Beruf	7 Student:in	NOMINAL	MC
SZ03	Beruf	8 Sonstiges:	NOMINAL	MC
SZ03	Beruf	-1 keine Angabe	NOMINAL	MC
SZ03	Beruf	-9 nicht beantwortet	NOMINAL	MC
VF01	Nutzung Chatbot	1 Ja, ich habe den Chatbot	NOMINAL	MC
VF01	Nutzung Chatbot	2 Nein, ich habe keinen Chatbot	NOMINAL	MC
VF01	Nutzung Chatbot	-9 nicht beantwortet	NOMINAL	MC
VF02_01	Probanden-Nummer: [01]		METRIC	TXT
TIME001	Verweildauer Seite 1		METRIC	TXT
TIME002	Verweildauer Seite 2		METRIC	TXT
TIME003	Verweildauer Seite 3		METRIC	TXT
TIME004	Verweildauer Seite 4		METRIC	TXT
TIME005	Verweildauer Seite 5		METRIC	TXT
TIME006	Verweildauer Seite 6		METRIC	TXT
TIME007	Verweildauer Seite 7		METRIC	TXT
TIME_SUM	Verweildauer gesamt (ohne Ausreißer)		METRIC	TXT
MAILSENT	Versandzeitpunkt der Einladungsmail (nur für nicht-anonyme Adressaten)		TIME	TXT
LASTDATA	Zeitpunkt als der Datensatz das letzte mal geändert wurde		TIME	TXT
FINISHED	Wurde die Befragung abgeschlossen (letzte Seite beantwortet)	0 abgebrochen	BOOL	TXT
FINISHED	Wurde die Befragung abgeschlossen (letzte Seite beantwortet)	1 ausgefüllt	BOOL	TXT
Q_VIEWER	Hat der Teilnehmer den Fragebogen nur angesehen	0 Teilnehmer	BOOL	TXT
Q_VIEWER	Hat der Teilnehmer den Fragebogen nur angesehen	1 Durchklicker	BOOL	TXT
LASTPAGE	Seite, die der Teilnehmer zuletzt bearbeitet hat		METRIC	TXT
MAXPAGE	Letzte Seite, die im Fragebogen bearbeitet wurde		METRIC	TXT
MISSING	Anteil fehlender Antworten in Prozent		METRIC	TXT
MISSREL	Anteil fehlender Antworten (gewichtet nach Relevanz)		METRIC	TXT
TIME_RSI	Ausfüll-Geschwindigkeit (relativ)		METRIC	TXT

Anhang D: Darstellung und Auswertung der Ergebnisse

Anhang D-1: Zusammengeführter Datensatz (bereinigt)

VF02_01 SZ02	VF01	SZ01_01	SZ03	AF01_01 AF01_02 AF01_03 KO01_01 KO01_02 KO01_03 PH01_01 PH01_02 PH01_03 SO01_01 SO01_02 SO01_03	AF01_01 AF01_02 AF01_03 KO01_01 KO01_02 KO01_03 PH01_01 PH01_02 PH01_03 SO01_01 SO01_02 SO01_03										
Proband Geschlecht	Nutzung	Alter: ...	Beruf	Affektiv: Affektiv: Affektiv: Kognitiv Kognitiv Kognitiv Physisch Physisch Physisch Sozial: a Sozial: t Sozial: c	Affektiv: Affektiv: Affektiv: Kognitiv Kognitiv Kognitiv Physisch Physisch Physisch Sozial: a Sozial: t Sozial: c										
3 m	MIT	22	Student:in	5	4	6	4	5	5	4	4	4	5	5	4
4 w	OHNE	23	Student:in	4	4	5	4	4	1	1	5	5	6	2	2
5 m	MIT	22	Student:in	5	6	5	6	4	4	4	4	4	4	4	5
6 w	OHNE	22	Student:in	4	5	5	4	4	4	4	4	4	4	4	4
8 w	OHNE	24	Student:in	3	3	2	1	2	2	2	2	2	2	2	1
9 m	OHNE	23	Student:in	6	5	6	4	5	5	4	5	4	5	2	2
10 w	MIT	21	Student:in	6	6	7	4	3	3	6	5	1	1	1	1
14 m	MIT	22	Auszubildende:r	6	5	6	2	2	5	7	5	5	4	3	3
15 w	MIT	24	Student:in	6	6	5	3	3	3	5	5	5	6	2	2
16 m	OHNE	20	Student:in	5	4	6	3	4	4	3	4	2	2	3	4
17 m	MIT	29	Student:in	6	6	6	5	6	4	6	5	5	4	4	4
19 m	OHNE	20	Student:in	5	1	2	3	2	2	4	4	4	3	2	2
20 m	OHNE	19	Student:in	4	4	4	4	2	2	2	2	4	4	2	2
22 w	MIT	22	Student:in	6	6	6	4	3	3	7	4	4	5	4	4
23 w	OHNE	22	Student:in	4	4	4	2	1	1	4	4	4	1	1	1
24 m	MIT	25	Student:in	4	4	5	4	1	1	6	4	4	2	1	1
25 w	MIT	20	Student:in	4	4	5	3	6	6	5	4	4	6	6	6
27 w	MIT	51	Beamte:r	3	4	4	6	4	4	5	2	2	2	2	1
28 m	OHNE	22	Student:in	4	1	4	1	1	1	6	4	3	1	1	1
29 m	MIT	22	Student:in	6	7	6	6	7	7	7	6	6	7	6	6
30 w	MIT	29	Student:in	4	4	5	6	5	5	5	5	5	2	2	2
31 m	OHNE	26	Student:in	3	2	2	4	5	4	3	2	3	2	1	1
32 m	OHNE	35	Student:in	4	4	4	4	2	2	4	4	4	2	1	1
33 m	MIT	23	Student:in	4	5	5	1	2	1	7	7	6	6	5	5
34 m	MIT	23	Student:in	6	6	5	7	6	6	7	6	7	7	5	6
35 w	OHNE	23	Angestellte:r	4	3	4	1	1	2	5	4	4	1	1	1
36 w	OHNE	27	Angestellte:r	3	3	3	3	2	2	3	3	4	3	2	1
37 w	OHNE	41	Angestellte:r	5	4	5	5	4	4	5	4	4	2	2	2
38 w	MIT	19	Student:in	5	6	7	5	5	6	6	4	5	6	5	4
39 m	OHNE	31	Angestellte:r	5	4	5	5	3	4	3	4	4	4	3	2

AF	KO	PH	SO	Baseline GSR			Shop GSR			Shop (avera GSR_Zuwachs_Shop	GSR_Zuwachs_BL	Eye-Tracking		
				Baseline VOR	Baseline NACH	Shop 1	Shop 2	Shop 3	Fixationen_A			Blickdauer_av Pupillendur		
5,00	4,67	4,67	4,67	1,58	2,21	1,73	1,95	2,31	2,00	0,2656	0,4016	336	279,33	3,08
4,33	3,33	5,33	2,00	2,62	3,92	3,41	3,81	3,93	3,71	0,4158	0,4938	364	248,33	3,18
5,33	4,00	4,00	4,67	2,39	2,89	2,54	2,64	2,79	2,66	0,1096	0,2061	452	295,33	2,65
4,67	4,00	4,00	4,00	8,81	12,30	10,97	11,31	12,42	11,57	0,3135	0,3971	334	255,00	3,00
2,67	1,67	2,33	1,67	1,61	1,58	1,70	1,65	1,61	1,65	0,0293	-0,0191	310	222,33	2,68
5,67	5,00	4,67	2,00	2,02	2,64	1,96	2,08	2,09	2,04	0,0130	0,3095	670	258,67	3,03
6,33	3,33	4,00	1,00	4,21	4,66	4,62	5,05	4,91	4,86	0,1560	0,1062	349	209,67	3,28
5,67	3,00	5,67	3,33	3,05	3,74	3,36	3,63	3,64	3,54	0,1609	0,2254	491	240,67	3,42
5,67	3,00	5,00	3,33	1,90	1,96	1,99	1,94	2,00	1,98	0,0427	0,0352	376	239,67	2,76
5,00	3,67	3,00	3,00	1,82	2,84	2,32	2,32	2,45	2,37	0,2975	0,5575	758	218,00	2,85
6,00	5,00	5,33	4,00	3,94	3,96	3,84	4,18	4,04	4,02	0,0193	0,0040	426	236,33	2,62
2,67	2,33	4,00	2,33	4,29	4,33	4,58	4,61	4,45	4,55	0,0592	0,0080	149	222,00	3,00
4,00	2,67	2,67	2,67	1,53	2,40	1,68	1,96	2,05	1,90	0,2358	0,5634	491	260,33	3,59
6,00	3,33	5,00	4,33	4,33	5,15	5,12	5,19	5,26	5,19	0,1989	0,1899	420	280,00	3,26
4,00	1,33	4,00	1,00	5,15	5,76	5,55	5,49	5,75	5,60	0,0864	0,1185	607	219,00	3,00
4,33	2,00	4,67	1,33	2,06	2,66	2,39	2,48	2,62	2,50	0,2134	0,2915	819	216,00	2,61
4,33	5,00	4,33	6,00	1,00	1,67	1,38	1,55	1,68	1,54	0,5342	0,6636	531	243,67	3,33
3,67	5,00	2,00	1,67	0,15	0,18	0,17	0,18	0,18	0,18	0,1756	0,2158	327	278,33	2,47
3,00	1,00	4,33	1,00	2,78	3,31	3,05	3,39	3,30	3,25	0,1699	0,1906	535	282,33	3,16
6,33	6,67	6,33	6,33	2,17	2,13	2,17	2,03	2,08	2,09	-0,0341	-0,0149	801	238,67	3,12
4,33	5,33	5,00	2,00	0,62	1,24	0,83	1,08	1,23	1,05	0,6990	1,0175	339	292,67	3,17
2,33	4,33	2,67	1,33	1,72	2,38	1,90	2,01	2,21	2,04	0,1832	0,3799	613	245,67	2,66
4,00	2,67	4,00	1,33	2,69	3,45	3,21	3,38	3,42	3,34	0,2421	0,2834	492	143,67	2,62
4,67	1,33	6,67	5,33	0,57	3,73	2,69	3,22	3,42	3,11	4,4515	5,5443	626	256,67	2,97
5,67	6,33	6,67	6,00	3,99	6,88	5,49	6,57	6,86	6,30	0,5793	0,7223	676	265,00	3,46
3,67	1,33	4,33	1,00	0,75	2,49	1,06	1,71	2,21	1,66	1,2278	2,3469	193	230,33	2,92
3,00	2,33	3,33	2,00	0,91	1,16	0,96	0,94	0,95	0,95	0,0403	0,2727	280	227,00	2,37
4,67	4,33	4,33	2,00	0,77	1,33	0,85	1,04	1,34	1,08	0,4039	0,7348	555	229,33	3,13
6,00	5,33	5,00	5,00	0,52	0,72	0,52	0,58	0,61	0,57	0,1035	0,3927	374	221,00	2,57
4,67	4,00	3,67	2,67	1,43	1,60	1,49	1,53	1,53	1,52	0,0586	0,1146	444	193,00	2,59

Anhang D-2: Darstellung Stichprobe

Alter: ... Jahre * Nutzung Chatbot Kreuztabelle				
Anzahl				
		Nutzung Chatbot		Gesamt
		MIT	OHNE	
Alter: ... Jahre	19	1	1	2
	20	1	2	3
	21	1	0	1
	22	5	3	8
	23	2	3	5
	24	1	1	2
	25	1	0	1
	26	0	1	1
	27	0	1	1
	29	2	0	2
	31	0	1	1
	35	0	1	1
41	0	1	1	
51	1	0	1	
Gesamt		15	15	30

Geschlecht * Nutzung Chatbot Kreuztabelle				
Anzahl				
		Nutzung Chatbot		Gesamt
		MIT	OHNE	
Geschlecht	männlich	8	8	16
	weiblich	7	7	14
Gesamt		15	15	30

Beruf					
		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	Angestellte:r	4	13,3	13,3	13,3
	Beamte:r	1	3,3	3,3	16,7
	Auszubildende:r	1	3,3	3,3	20,0
	Student:in	24	80,0	80,0	100,0
	Gesamt	30	100,0	100,0	
Gruppenstatistiken					
Nutzung Chatbot		N	Mittelwert	Std.- Abweichung	Standardfehler des Mittelwertes
Alter: ... Jahre	MIT	15	24,93	7,741	1,999
	OHNE	15	25,20	6,109	1,577

Test bei unabhängigen Stichproben												
		Levene-Test		t-Test für die Mittelwertgleichheit							95% Konfidenzintervall der Differenz	
		F	Sig.	T	df	Signifikanz		Mittlere Differenz	Differenz für Standardfehler	Unterer Wert	Oberer Wert	
						Einseitiges p	Zweiseitiges p					
Alter: (Jahre)	Varianzen sind gleich	0,000	0,985	-0,105	28	0,459	0,917	-0,267	2,546	-5,482	4,949	
	Varianzen sind nicht gleich			-0,105	26,564	0,459	0,917	-0,267	2,546	-5,495	4,961	

Anhang D-3: Auswertung Hautreaktion

Deskriptive Statistik				Statistik	Standard Fehler		
Nutzung Chatbot							
Baseline VOR	MIT	Mittelwert		2,1653	0,37583		
		95%	Untergrenze	1,3593			
		Konfidenzinte	Obergrenze	2,9714			
		5% getrimmtes Mittel		2,1570			
		Median		2,0600			
		Varianz		2,119			
		Standard Abweichung		1,45559			
		Minimum		0,15			
		Maximum		4,33			
		Spannweite		4,18			
		Interquartilbereich		3,32			
		Schiefe		0,237	0,580		
		Kurtosis		-1,345	1,121		
		OHNE	Mittelwert		2,5933	0,54687	
			95%	Untergrenze	1,4204		
	Konfidenzinte		Obergrenze	3,7662			
	5% getrimmtes Mittel			2,3504			
	Median			1,8200			
	Varianz			4,486			
	Standard Abweichung			2,11801			
	Minimum			0,75			
	Maximum			8,81			
	Spannweite			8,06			
	Interquartilbereich			1,35			
	Schiefe			2,079	0,580		
	Kurtosis			4,860	1,121		
	Aufgabenabsolvierung		MIT	Mittelwert		2,7727	0,45114
				95%	Untergrenze	1,8051	
		Konfidenzinte		Obergrenze	3,7403		
		5% getrimmtes Mittel			2,7207		
Median				2,5000			
Varianz				3,053			
Standard Abweichung				1,74724			
Minimum				0,18			
Maximum				6,30			
Spannweite				6,12			
Interquartilbereich				2,48			
Schiefe				0,512	0,580		
Kurtosis				-0,332	1,121		
OHNE		Mittelwert			3,1487	0,69043	
		95%		Untergrenze	1,6678		
		Konfidenzinte	Obergrenze	4,6295			
		5% getrimmtes Mittel		2,8030			
		Median		2,0400			
		Varianz		7,150			
		Standard Abweichung		2,67402			
		Minimum		0,95			
		Maximum		11,57			
		Spannweite		10,62			
		Interquartilbereich		2,06			
		Schiefe		2,497	0,580		
		Kurtosis		7,265	1,121		
		Baseline NACH	MIT	Mittelwert		2,9187	0,46262
				95%	Untergrenze	1,9265	
Konfidenzinte				Obergrenze	3,9109		
5% getrimmtes Mittel					2,8507		
Median				2,6600			
Varianz				3,210			
Standard Abweichung				1,79171			
Minimum				0,18			
Maximum				6,88			
Spannweite				6,70			
Interquartilbereich				2,29			
Schiefe				0,590	0,580		
Kurtosis				0,216	1,121		
OHNE	Mittelwert				3,4327	0,70887	
	95%			Untergrenze	1,9123		
	Konfidenzinte		Obergrenze	4,9530			
	5% getrimmtes Mittel			3,0663			
	Median			2,6400			
	Varianz			7,537			
	Standard Abweichung			2,74544			
	Minimum			1,16			
	Maximum			12,30			
	Spannweite			11,14			
	Interquartilbereich			2,32			
	Schiefe			2,682	0,580		
	Kurtosis			8,389	1,121		

Tests auf Normalverteilung							
Nutzung Chatbot		Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Statistik	df	Signifikanz	Statistik	df	Signifikanz
Zuwachs Aufgabenabsolvierung	MIT	0,366	15	0,000	0,447	15	0,000
	OHNE	0,227	15	0,037	0,695	15	0,000
Zuwachs Basiswert NACH	MIT	0,351	15	0,000	0,465	15	0,000
	OHNE	0,287	15	0,002	0,650	15	0,000

a. Signifikanzkorrektur nach Lilliefors

Ränge				
Nutzung Chatbot		N	Mittlerer Rang	Rangsumme
Zuwachs Aufgabenabsolvierung	MIT	15	15,87	238,00
	OHNE	15	15,13	227,00
	Gesamt	30		
Zuwachs Basiswert NACH	MIT	15	15,07	226,00
	OHNE	15	15,93	239,00
	Gesamt	30		

Teststatistiken ^a		
	Zuwachs Aufgabenabsolvierung	Zuwachs Basiswert NACH
Mann-Whitney-U-Test	107,000	106,000
Wilcoxon-W	227,000	226,000
Z	-0,228	-0,270
Asymp. Sig. (2-seitig)	0,820	0,787
Exakte Sig. [2*(1-seitige Sig.)]	,838 ^b	,806 ^b

a. Gruppenvariable: Nutzung Chatbot
b. Nicht für Bindungen korrigiert.

Tests auf Normalverteilung							
Nutzung Chatbot		Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Statistik	df	Signifikanz	Statistik	df	Signifikanz
Basiswert VOR	MIT	0,155	15	,200*	0,915	15	0,162
	OHNE	0,265	15	0,006	0,770	15	0,002
Aufgabenabsolvierung	MIT	0,126	15	,200*	0,967	15	0,812
	OHNE	0,217	15	0,056	0,716	15	0,000
Basiswert NACH	MIT	0,120	15	,200*	0,972	15	0,888
	OHNE	0,239	15	0,021	0,695	15	0,000

*. Dies ist eine untere Grenze der echten Signifikanz.
a. Signifikanzkorrektur nach Lilliefors

Deskriptive Statistiken					
	N	Mittelwert	Std.-Abweichung	Minimum	Maximum
Basiswert VOR	30	2,3793	1,79885	0,15	8,81
Aufgabenabsolvierung	30	2,9607	2,22762	0,18	11,57
Basiswert NACH	30	3,1757	2,29278	0,18	12,30
Nutzung Chatbot	30	1,50	0,509	1	2

Ränge				
Nutzung Chatbot		N	Mittlerer Rang	Rangsumme
Basiswert VOR	MIT	15	14,93	224,00
	OHNE	15	16,07	241,00
	Gesamt	30		
Aufgabenabsolvierung	MIT	15	15,60	234,00
	OHNE	15	15,40	231,00
	Gesamt	30		
Basiswert NACH	MIT	15	15,20	228,00
	OHNE	15	15,80	237,00
	Gesamt	30		

Teststatistiken ^a			
	Basis wert VOR	Aufgabenabsolvierung	Basis wert NACH
Mann-Whitney-U-Test	104,000	111,000	108,000
Wilcoxon-W	224,000	231,000	228,000
Z	-0,353	-0,062	-0,187
Asymp. Sig. (2-seitig)	0,724	0,950	0,852
Exakte Sig. [2*(1-seitige Sig.)]	,744 ^b	,967 ^b	,870 ^b

a. Gruppenvariable: Nutzung Chatbot
b. Nicht für Bindungen korrigiert.

Anhang D-4: Auswertung Eye-Tracking

Deskriptive Statistik				Statistik	Standard Fehler	
Nutzung Chatbot						
Anzahl Fixationen	MIT	Mittelwert		489,5333	43,07202	
		95% Konfidenzintervall des Mittelwerts	Untergrenze	397,1530		
			Obergrenze	581,9136		
		5% getrimmtes Mittel		480,2593		
		Median		426,0000		
		Varianz		27827,981		
		Standard Abweichung		166,81721		
		Minimum		327,00		
		Maximum		819,00		
		Spannweite		492,00		
		Interquartilbereich		277,00		
		Schiefe		1,018	0,580	
		Kurtosis		-0,180	1,121	
		OHNE	Mittelwert		453,0000	45,88184
	95% Konfidenzintervall des Mittelwerts		Untergrenze	354,5932		
			Obergrenze	551,4068		
	5% getrimmtes Mittel			452,9444		
	Median			491,0000		
	Varianz			31577,143		
	Standard Abweichung			177,69959		
	Minimum			149,00		
	Maximum			758,00		
	Spannweite			609,00		
	Interquartilbereich			297,00		
	Schiefe			-0,119	0,580	
	Kurtosis			-0,801	1,121	
	Blickdauer je Fixation		MIT	Mittelwert		252,8667
		95% Konfidenzintervall des Mittelwerts		Untergrenze	237,5156	
Obergrenze				268,2178		
5% getrimmtes Mittel				252,9074		
Median				243,6667		
Varianz				768,425		
Standard Abweichung				27,72049		
Minimum				209,67		
Maximum				295,33		
Spannweite				85,67		
Interquartilbereich				43,00		
Schiefe				0,082	0,580	
Kurtosis				-1,206	1,121	
OHNE		Mittelwert			230,3333	8,42483
		95% Konfidenzintervall des Mittelwerts	Untergrenze	212,2639		
			Obergrenze	248,4028		
		5% getrimmtes Mittel		232,2593		
		Median		229,3333		
		Varianz		1064,667		
		Standard Abweichung		32,62923		
		Minimum		143,67		
		Maximum		282,33		
		Spannweite		138,67		
		Interquartilbereich		36,00		
		Schiefe		-1,169	0,580	
		Kurtosis		2,794	1,121	
		Durchschnitt Pupillendurchmesser	MIT	Mittelwert		2,9858
95% Konfidenzintervall des Mittelwerts				Untergrenze	2,7967	
	Obergrenze			3,1748		
5% getrimmtes Mittel				2,9878		
Median				3,0796		
Varianz				0,117		
Standard Abweichung				0,34132		
Minimum				2,47		
Maximum				3,46		
Spannweite				0,99		
Interquartilbereich				0,66		
Schiefe				-0,143	0,580	
Kurtosis				-1,604	1,121	
OHNE	Mittelwert				2,9192	0,07797
	95% Konfidenzintervall des Mittelwerts		Untergrenze	2,7519		
			Obergrenze	3,0864		
	5% getrimmtes Mittel			2,9127		
	Median			2,9968		
	Varianz			0,091		
	Standard Abweichung			0,30196		
	Minimum			2,37		
	Maximum			3,59		
	Spannweite			1,22		
	Interquartilbereich			0,47		
	Schiefe			0,243	0,580	
	Kurtosis			0,549	1,121	

Tests auf Normalverteilung							
Nutzung Chatbot		Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Statistik	df	Signifikanz	Statistik	df	Signifikanz
Anzahl Fixationen Shop	MIT	0,189	15	0,156	0,852	15	0,018
	OHNE	0,118	15	,200 ^a	0,977	15	0,947
Durchschnittliche Blickdauer pro Fixation	MIT	0,163	15	,200 ^a	0,940	15	0,382
	OHNE	0,219	15	0,050	0,908	15	0,127
Durchschnitt Pupillendurchmesser	MIT	0,168	15	,200 ^a	0,911	15	0,139
	OHNE	0,135	15	,200 ^a	0,964	15	0,770

*. Dies ist eine untere Grenze der echten Signifikanz.
a. Signifikanzkorrektur nach Lilliefors

Deskriptive Statistiken					
	N	Mittelwert	Std.-	Minimum	Maximum
Anzahl Fixationen	30	471,2667	170,36289	149,00	819,00
Blickdauer je Fixation	30	241,6000	31,87877	143,67	295,33
Pupillendurchmesser	30	2,9525	0,31845	2,37	3,59
Nutzung Chatbot	30	1,50	0,509	1	2

Ränge				
Nutzung Chatbot		N	Mittlerer Rang	Rangsumme
Anzahl Fixationen	MIT	15	16,23	243,50
	OHNE	15	14,77	221,50
	Gesamt	30		
Blickdauer je Fixation	MIT	15	18,07	271,00
	OHNE	15	12,93	194,00
	Gesamt	30		
Pupillendurchmesser	MIT	15	16,40	246,00
	OHNE	15	14,60	219,00
	Gesamt	30		

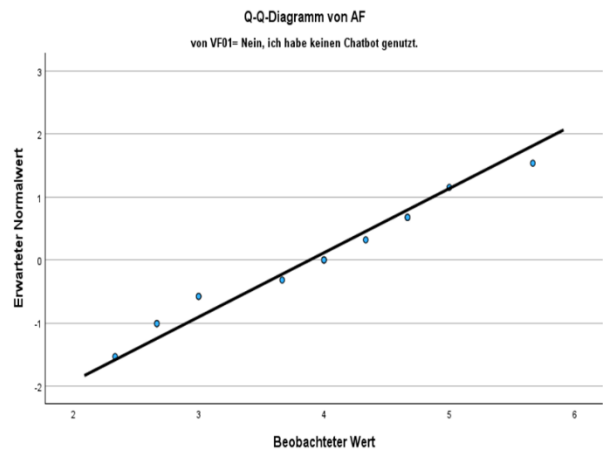
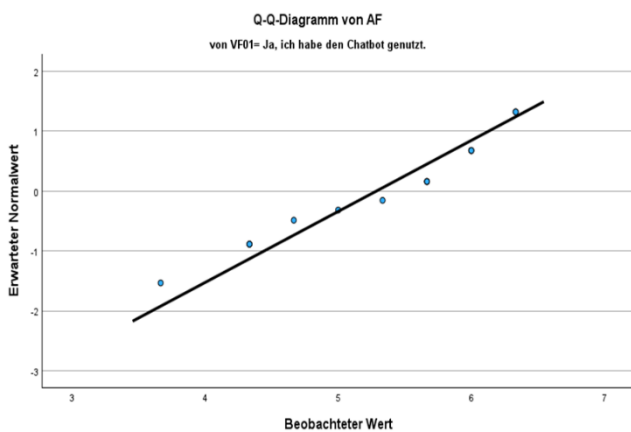
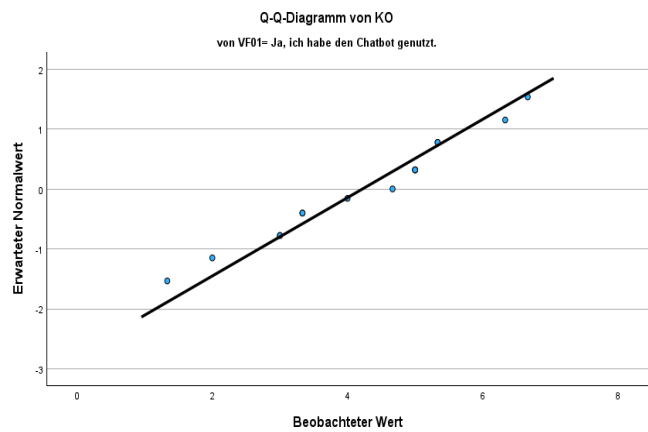
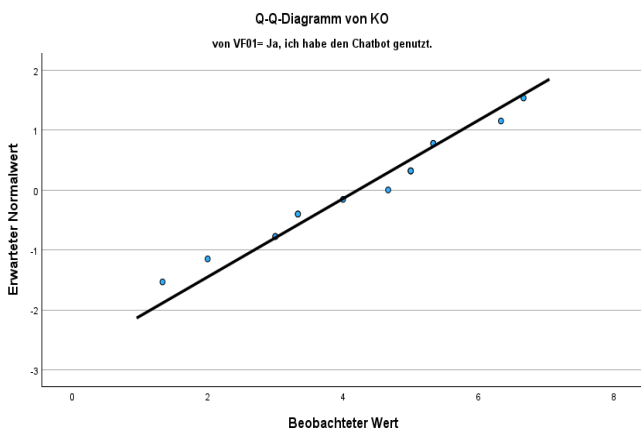
Teststatistiken ^a			
	Anzahl	Blickdauer	Pupillendurch
Mann-Whitney-U-Test	101,500	74,000	99,000
Wilcoxon-W	221,500	194,000	219,000
Z	-0,456	-1,597	-0,560
Asymp. Sig. (2-seitig)	0,648	0,110	0,576
Exakte Sig. [2*(1-seitige Sig.)]	,653 ^b	,116 ^b	,595 ^b

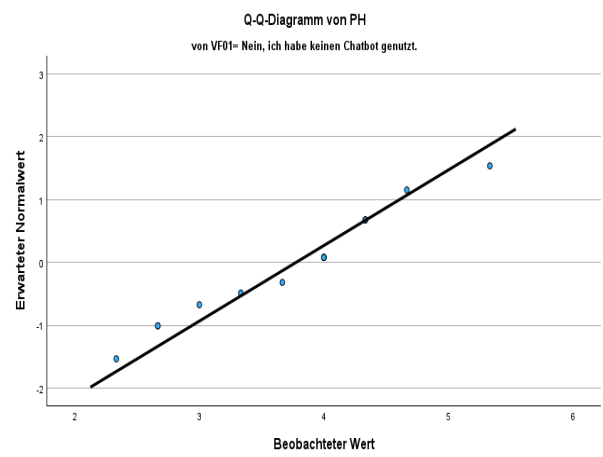
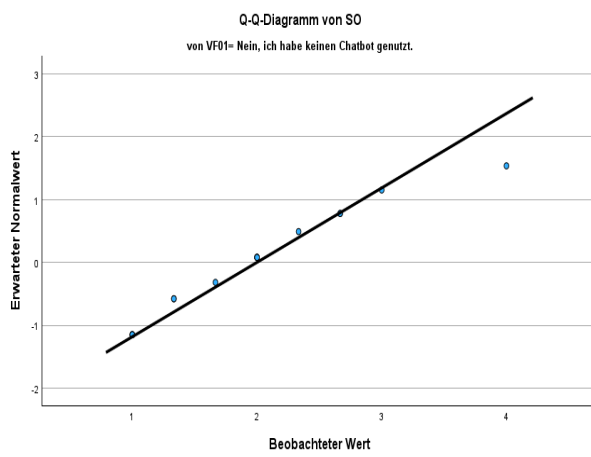
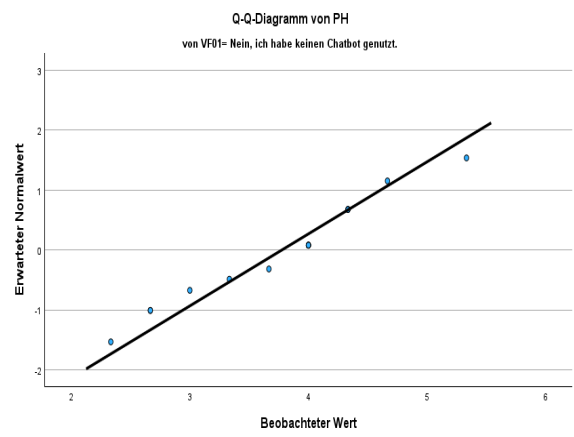
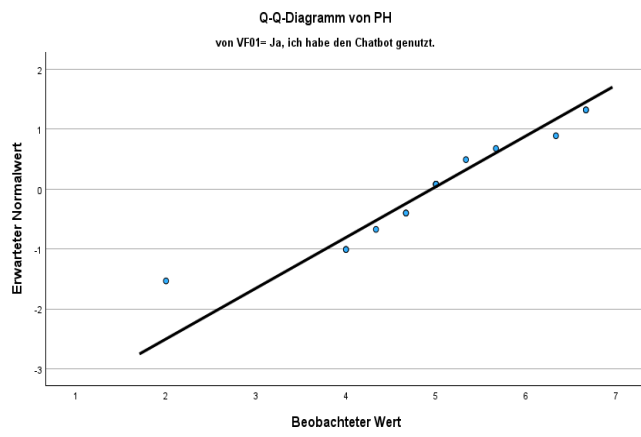
a. Gruppenvariable: Nutzung Chatbot
b. Nicht für Bindungen korrigiert.

Anhang D-5: Auswertung Fragebogen

Tests auf Normalverteilung							
Nutzung Chatbot		Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Statistik	df	Signifikanz	Statistik	df	Signifikanz
KO	MIT	0,161	15	,200*	0,961	15	0,712
	OHNE	0,133	15	,200*	0,945	15	0,451
AF	MIT	0,206	15	0,086	0,913	15	0,150
	OHNE	0,151	15	,200*	0,953	15	0,573
SO	MIT	0,130	15	,200*	0,930	15	0,274
	OHNE	0,167	15	,200*	0,922	15	0,209
PH	MIT	0,152	15	,200*	0,919	15	0,188
	OHNE	0,205	15	0,089	0,951	15	0,533

*. Dies ist eine untere Grenze der echten Signifikanz.
a. Signifikanzkorrektur nach Lilliefors





D-5.1: Kognitive Erfahrung (T-Test)

Gruppenstatistiken					
Nutzung Chatbot		N	Mittelwert	Std.- Abweichung	Standardfehler des Mittelwertes
KO	Ja, ich habe den Chatbot genutzt.	15	4,2222	1,53099	0,39530
	Nein, ich habe keinen Chatbot genutzt.	15	2,9333	1,26742	0,32725

Test bei unabhängigen Stichproben											
		Levene-Test der Varianzgleichheit		t-Test für die Mittelwertgleichheit							
		F	Sig.	T	df	Signifikanz		Mittlere Differenz	Differenz für Standardfehler	95% Konfidenzintervall der Differenz	
						P (1-seitig)	P (2-seitig)			Unterer Wert	Oberer Wert
KO	Varianzen sind gleich	0,567	0,458	2,512	28	0,009	0,018	1,28889	0,51318	0,23769	2,34008
	Varianzen sind nicht gleich			2,512	27,057	0,009	0,018	1,28889	0,51318	0,23604	2,34174

Effektgrößen bei unabhängigen Stichproben					
		Standardisierer ^a	Punktschätzung	95% Konfidenzintervall	
				Unterer Wert	Oberer Wert
KO	Cohen's d	1,40539	0,917	0,155	1,664
	Hedges' Korrektur	1,44449	0,892	0,151	1,619
	Glass' Delta	1,26742	1,017	0,196	1,810

a. Der bei der Schätzung der Effektgrößen verwendete Nenner.

D-5.2: Affektive Erfahrung (T-Test)

Gruppenstatistiken					
Nutzung Chatbot		N	Mittelwert	Std.- Abweichung	Standardfehler des Mittelwertes
AF	Ja, ich habe den Chatbot genutzt.	15	5,2889	0,84390	0,21789
	Nein, ich habe keinen Chatbot genutzt.	15	3,8889	0,98131	0,25337

Test bei unabhängigen Stichproben											
		Levene-Test der Varianzgleichheit		t-Test für die Mittelwertgleichheit							
		F	Sig.	T	df	Signifikanz		Mittlere Differenz	Differenz für Standardfehler	95% Konfidenzintervall der Differenz	
						p (1-seitig)	p (2-seitig)			Unterer Wert	Oberer Wert
AF	Varianzen sind gleich	0,221	0,642	4,189	28	0,000	0,000	1,400	0,33418	0,71547	2,08453
	Varianzen sind nicht gleich			4,189	27,386	0,000	0,000	1,400	0,33418	0,71477	2,08523

Effektgrößen bei unabhängigen Stichproben					
		Standardisierer ^a	Punktschätzung	95% Konfidenzintervall	
				Unterer Wert	Oberer Wert
AF	Cohen's d	0,91519	1,530	0,700	2,339
	Hedges' Korrektur	0,94065	1,488	0,681	2,275
	Glass' Delta	0,98131	1,427	0,524	2,297

a. Der bei der Schätzung der Effektgrößen verwendete Nenner.

D-5.3: Soziale Erfahrung (T-Test)

Gruppenstatistiken					
Nutzung Chatbot		N	Mittelwert	Std.- Abweichung	Standardfehler des Mittelwertes
SO	Ja, ich habe den Chatbot genutzt.	15	3,9333	1,76473	0,45565
	Nein, ich habe keinen Chatbot genutzt.	15	2,0000	0,84515	0,21822

Test bei unabhängigen Stichproben											
		Levene-Test der Varianzgleichheit		t-Test für die Mittelwertgleichheit							
		F	Sig.	T	df	Signifikanz		Mittlere Differenz	Differenz für Standardfehler	95% Konfidenzintervall der Differenz	
						p (1-seitig)	p (2-seitig)			Unterer Wert	Oberer Wert
SO	Varianzen sind gleich	9,212	0,005	3,827	28	0,000	0,001	1,9333	0,50521	0,89846	2,96821
	Varianzen sind nicht gleich			3,827	20,101	0,001	0,001	1,9333	0,50521	0,87982	2,98685

Effektgrößen bei unabhängigen Stichproben					
		Standardisierer ^a	Punktschätzung	95% Konfidenzintervall	
				Unterer Wert	Oberer Wert
SO	Cohen's d	1,38358	1,397	0,584	2,190
	Hedges' Korrektur	1,42207	1,360	0,569	2,131
	Glass' Delta	0,84515	2,288	1,168	3,373

a. Der bei der Schätzung der Effektgrößen verwendete Nenner.

D-5.4: Physische Erfahrung (T-Test)

Gruppenstatistiken					
Nutzung Chatbot		N	Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes
PH	Ja, ich habe den Chatbot genutzt.	15	4,9556	1,18098	0,30493
	Nein, ich habe keinen Chatbot genutzt.	15	3,7778	0,83254	0,21496

Test bei unabhängigen Stichproben											
		Levene-Test der Varianzgleichheit		t-Test für die Mittelwertgleichheit							
		F	Sig.	T	df	Signifikanz		Mittlere Differenz	Differenz für Standardfehler	95% Konfidenzintervall der Differenz	
						p (1-seitig)	p (2-seitig)			Unterer Wert	Oberer Wert
PH	Varianzen sind gleich	0,333	0,568	3,157	28	0,002	0,004	1,17778	0,37308	0,41356	1,94200
	Varianzen sind nicht gleich			3,157	25,159	0,002	0,004	1,17778	0,37308	0,40965	1,94590

Effektgrößen bei unabhängigen Stichproben					
		Standardisierter ^a	Punktschätzung	95% Konfidenzintervall	
				Unterer Wert	Oberer Wert
PH	Cohen's d	1,02172	1,153	0,368	1,920
	Hedges' Korrektur	1,05015	1,122	0,358	1,868
	Glass' Delta	0,83254	1,415	0,514	2,282

a. Der bei der Schätzung der Effektgrößen verwendete Nenner.

Eidesstattliche Erklärung bei Verwendung von IT-/KI-gestützter Software

Ich versichere an Eides statt durch meine Unterschrift, dass ich die vorliegende wissenschaftliche Arbeit durchgehend selbständig und ohne fremde Hilfe angefertigt habe und beim Einsatz IT-/KI-gestützter Schreibwerkzeuge steuernd gearbeitet habe. Es wurden nur die in der Arbeit ausdrücklich genannten Quellen benutzt. Wörtlich oder sinngemäß übernommenes Gedankengut wurde als solches kenntlich gemacht. Den konkreten Einsatz von IT-/KI-gestützten Schreibwerkzeugen habe ich im Rahmen meines Reflexionsportfolios sorgfältig dokumentiert und reflektiert.

Des Weiteren hat die Arbeit in dieser oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegen. Ich bin mir bewusst, dass die Arbeit durch eine Plagiatssoftware überprüft werden kann.

Essen, 10.01.2024

Ort, Datum



Unterschrift

Reflexionsportfolio

Name des Tools	Funktion des Tools	Begründung für die Verwendung	Reflexion der Verwendung
ChatGPT	Textproduktion	<i>Brainstorming/ Ideenfindung</i>	<p>Nutzen (Vorteile):</p> <ul style="list-style-type: none"> • Schnelle und unkomplizierte Vorschläge • (Vor)urteilsfreies und nicht emotionales Denken, d.h. Unterbreitung von rationalen Vorschlägen • Unterbreitung von unkonventionellen Ideen, die zum weiteren Nachdenken anregen <p>Begrenzungen (Nachteile):</p> <ul style="list-style-type: none"> • Meistens nur Aufgreifen allgemeiner Probleme/Ideen der Wirtschaft • Fehlende Aktualität (Wissenschaftlicher Stand bis Januar 2022) • Kein neues Denken/neue Ideen. Nur Reproduktion vorhandener Daten und Quellen
		<i>Übersetzung</i>	<p>Nutzen (Vorteile):</p> <ul style="list-style-type: none"> • Übersetzung langer Texte problemlos möglich • Übersetzung sowohl in Fachsprache als auch leichter Sprache möglich • Keine Begrenzung in Wortzahl oder Zeit • Großteils fehlerfreie Grammatik <p>Begrenzungen (Nachteile):</p> <ul style="list-style-type: none"> • Einzelne Fehler in Übersetzungen vorhanden (z.B. einzelne Wörter oder kleine Grammatikfehler) • Ab und zu werden Fachbegriffe falsch/nicht übersetzt oder fehlen in der Übersetzung • Insbesondere in langen Texten fehlt manchmal der Kontext
		<i>Zusammenfassen langer Texte</i>	<p>Nutzen (Vorteile):</p> <ul style="list-style-type: none"> • Schnelles Zusammenfassen langer Texte in Originalsprache oder andere Sprachen • Zusammenfassen sowohl in Fachsprache als auch leichter Sprache möglich • Zeitersparnis für Anwender <p>Begrenzungen (Nachteile):</p> <ul style="list-style-type: none"> • Keine Unterscheidung zwischen relevanten und unwichtigen Abschnitten, sodass manchmal für den Kontext wichtige Abschnitte fehlen • Fehlen wichtiger Definitionen oder Fachbegriffe (Verlust wissenschaftlicher Charakter, da leichter Sprachstil)

Textgenerierung

Nutzen (Vorteile):

- Vorschläge mehrere Synonyme für einzelne Wörter
- Korrektur der Grammatik und Zeichensetzung einzelner Sätze und Abschnitte
- Umformulierung von eigener Umgangssprache in Text mit wissenschaftlichen Charakter

Begrenzungen (Nachteile):

- Unterbreitung von Korrekturvorschlägen, nach denen nicht gefragt wurde
- Insbesondere bei einzelnen Sätzen falsche Grammatikkorrektur durch fehlenden Kontext im Gesamtzusammenhang

Semantic Scholar

Literaturrecherche

Finden nützlicher Quellen zu bestimmten Themen

Nutzen (Vorteile):

- Suche mit Schlüsselwörtern reicht aus
- Vorschläge vieler Quellen, die in der Regel zum Forschungsthema passen
- Automatische Sortierung der Quellen nach Relevanz (bspw. durch Anzahl Zitationen)
- Direkte Verknüpfung zu ähnlichen Quellen, die ebenfalls automatisch sortiert sind
- Gewährleistung von Aktualität durch regelmäßiges Aktualisieren

Begrenzungen (Nachteile):

- Nicht alle Quellen sind kostenlos zugänglich
- Viele Quellen können nicht heruntergeladen werden
- Manchmal werden Quellen verknüpft angezeigt, die keinen wesentlichen inhaltlichen Bezug zueinander haben

Connected Papers

Literaturrecherche

Finden nützlicher Quellen zu bestimmten Themen

Nutzen (Vorteile):

- Suche über Schlüsselwörter, DOI oder Autorennamen möglich
- Direkte Sortierung relevanter Quellen
- Kurze automatische Zusammenfassung zu jedem Paper
- Darstellung in Mindmaps, sortiert nach Relevanz und Verknüpfungen zu anderen Veröffentlichungen

Begrenzungen (Nachteile):

- Zusammenfassungen teilweise unverständlich oder mit fehlendem Kontext
 - Nicht alle Quellen sind kostenlos zugänglich
 - Manchmal werden doch Quellen verknüpft angezeigt, die keinen wesentlichen inhaltlichen Bezug zueinander haben
-

Zusammenfassung Reflexionsportfolio

Zusammenfassend kann festgestellt werden, dass IT/KI-gestützte Software bei der Anfertigung der Arbeit helfen können. Das sorgfältige wissenschaftliche Arbeiten und Denken können jedoch nicht ersetzt werden. Wesentliche Gründe sind die mangelnde Aktualität und die reine Reproduktion des bisherigen Wissens. Die Tools sind nicht in der Lage neue Erkenntnisse zu erschließen.

Jedoch kann KI-gestützte Software insbesondere in der Anfangsphase helfen, indem es zu unkonventionellem Denken anregt. Außerdem kann mithilfe der Tools Zeit eingespart werden, vor allem in der Literaturrecherche. Das manuelle Suchen wird zum Teil minimiert und Vorschläge zu relevanten Themen werden automatisch generiert. Jedoch sind auch diese mit Skepsis zu behandeln und müssen kontrolliert werden. Zudem unterstützt die Software bei der Überprüfung der Grammatik und Zeichensetzung von eigenen Texten und ist in der Lage, Synonyme vorzuschlagen, sodass Fehler minimiert werden und der Lesefluss verbessert werden kann. Nichtsdestotrotz ist eine manuelle Überprüfung aller Antworten und Vorschläge notwendig, da sowohl inhaltliche als auch sprachliche Fehler auftreten können.