

**Chatbots im Kundenservice:
Eine empirische Untersuchung von Einflussfaktoren auf
die Wiederverwendungsabsicht**

Vom Fachbereich Wirtschaftswissenschaften
der Rheinland-Pfälzischen Technischen Universität
Kaiserslautern-Landau
zur Verleihung des akademischen Grades
Doctor rerum politicarum (Dr. rer. pol.) genehmigte

DISSERTATION

vorgelegt von
Sebastian Düppre, M. Sc.

Tag der mündlichen Prüfung:	23. Juli 2025
Dekan:	Prof. Dr. Florian Sahling
Vorsitzende:	Prof. Dr. Katharina Spraul
Berichterstattende:	1. Prof. Dr. Stefan Roth 2. Prof. Dr. Oliver Wendt

D 386

(2025)

Inhalt

Inhalt.....	I
Abbildungsverzeichnis	IV
Tabellenverzeichnis	V
Abkürzungsverzeichnis	VIII
1 Einleitung.....	1
1.1 Relevanz des Themas und Zielsetzung.....	1
1.2 Aufbau der Arbeit.....	5
2 Konzeptionelle Grundlagen.....	7
2.1 Chatbots.....	7
2.1.1 Begriffsbestimmung – Chatbot.....	7
2.1.2 Kategorisierungsmöglichkeiten von Chatbots.....	12
2.1.3 Integration von Chatbots im Kundenservice	14
2.2 Bewertungen.....	18
3 Theoretische Bezugspunkte zu Mensch-Chatbot-Interaktionen.....	22
3.1 Anthropomorphismus	22
3.2 Social Response Theorie und Media Equation Theorie	24
3.3 Uncanny Valley Effekt	26
3.4 Attributionstheorie.....	28
3.5 Familiarity-Heuristik	29
3.6 Confirmation/Disconfirmation-Paradigma.....	31
4 Empirische Evidenz als Grundlage der eigenen Forschung	33
4.1 Anthropomorphe Gestaltung von Chatbots	34
4.2 Offenlegung der Chatbot-Identität	42
4.3 Übergabe von Chatbots an Mitarbeiter.....	48
4.4 Bewertungen von Chatbot-Interaktionen.....	51
5 Experimentelle Untersuchungen.....	58
5.1 Überblick der Forschungsprojekte	58
5.2 Grundlagen zu den Experimenten und deren Auswertung.....	60
5.2.1 Charakteristika von Experimenten	60

5.2.2	Gütebeurteilung der Konstrukte	63
5.2.3	Auswertung der Daten	65
5.3	Forschungsprojekt I: Anthropomorphismus	67
5.3.1	Herleitung der Forschungshypothesen	67
5.3.2	Methodik.....	68
5.3.3	Stichprobe und Gütebeurteilung.....	73
5.3.4	Ergebnisse.....	76
5.3.5	Diskussion der Ergebnisse.....	78
5.4	Forschungsprojekt II: Offenlegungszeitpunkt	80
5.4.1	Experiment 1: Offenlegungszeitpunkt eines Chatbots	82
5.4.1.1	Herleitung der Forschungshypothesen	82
5.4.1.2	Methodik.....	83
5.4.1.3	Stichprobe und Gütebeurteilung.....	88
5.4.1.4	Ergebnisse.....	90
5.4.1.5	Diskussion der Ergebnisse.....	92
5.4.2	Experiment 2: Offenlegungszeitpunkt eines Chatbots und die Kommunikationsplattform	93
5.4.2.1	Herleitung der Forschungshypothesen	94
5.4.2.2	Methodik.....	95
5.4.2.3	Stichprobe und Gütebeurteilung.....	100
5.4.2.4	Ergebnisse.....	103
5.4.2.5	Diskussion der Ergebnisse.....	106
5.4.3	Experiment 3: Offenlegungszeitpunkt eines Chatbots und Lösbarkeit der Nutzeranfrage.....	107
5.4.3.1	Herleitung der Forschungshypothesen	107
5.4.3.2	Methodik.....	109
5.4.3.3	Stichprobe und Gütebeurteilung.....	113
5.4.3.4	Ergebnisse.....	115
5.4.3.5	Diskussion der Ergebnisse.....	119
5.4.4	Abschließende Diskussion – Forschungsprojekt II	120
5.5	Forschungsprojekt III: Übergabe der Chatbot-Interaktion an einen Menschen	123
5.5.1	Herleitung der Forschungshypothesen	124
5.5.2	Methodik.....	126

5.5.3	Stichprobe und Gütebeurteilung.....	130
5.5.4	Ergebnisse.....	133
5.5.5	Diskussion der Ergebnisse.....	136
5.6	Forschungsprojekt IV: Bewertungen von Chatbots.....	138
5.6.1	Herleitung der Forschungshypothesen	139
5.6.2	Experiment 1: Rating.....	141
5.6.2.1	Methodik.....	141
5.6.2.2	Stichprobe und Gütebeurteilung.....	145
5.6.2.3	Ergebnisse.....	148
5.6.3	Experiment 2: Rating und Review.....	152
5.6.3.1	Methodik.....	152
5.6.3.2	Stichprobenbeschreibung und Gütebeurteilung.....	155
5.6.3.3	Ergebnisse.....	157
5.6.4	Diskussion der Ergebnisse – Forschungsprojekt IV	161
6	Schlussbetrachtung	165
6.1	Zusammenfassung	165
6.2	Implikationen für Forschung und Praxis	167
6.3	Forschungsausblick	170
Anhang		174
Literaturverzeichnis		176

Abbildungsverzeichnis

Abbildung 1: Aufbau der Arbeit.....	6
Abbildung 2: Uncanny Valley Effekt (Mori, 1970, S. 33).....	27
Abbildung 3: Übersicht der Forschungsprojekte.....	60
Abbildung 4: Mediationsmodell (FP I)	77
Abbildung 5: Ergebnisse der ANOVA (FP II: E 1) – Effekte des Offenlegungszeitpunkts auf die Soziale Präsenz	91
Abbildung 6: Mediationsmodell (FP II: E 1)	91
Abbildung 7: Ergebnisse der ANOVA (FP II: E 2) – Effekte des Offenlegungszeitpunkts auf die Soziale Präsenz	104
Abbildung 8: Moderiertes Mediationsmodell (FP II: E 2)	104
Abbildung 9: Ergebnisse der ANOVAs (FP II: E 3) – Effekte des Offenlegungszeitpunkts auf den wahrgenommenen Betrug, das vermutete Motiv und die Fairnesswahrnehmung...	116
Abbildung 10: Moderiertes Mediationsmodell (FP II: E 3)	117
Abbildung 11: Ergebnisse der zweifaktorielle ANOVA (FP III) – Effekte der Lösung und des Wechsels auf die Wiederverwendungsabsicht	134
Abbildung 12: Moderationsmodell (FP III).....	135
Abbildung 13: Ergebnisse der ANOVA (FP IV: E 1) – Effekte des Ratings auf die Nutzungsabsicht	149
Abbildung 14: Mediationsmodell (FP IV: E 1).....	150
Abbildung 15: Ergebnisse der zweifaktoriellen ANOVAs (FP VI: E 2) – Effekte des Ratings und Review auf die Erwartung sowie die Nutzungsabsicht	158
Abbildung 16: Mediationsmodell (FP IV: E 2).....	159

Tabellenverzeichnis

Tabelle 1: Gestaltungsmerkmale anthropomorpher Chatbots	36
Tabelle 2: Auswirkungen von Anthropomorphismus im Kontext von Chatbots	41
Tabelle 3: Auswirkungen der Offenlegung von Chatbots	47
Tabelle 4: Aktuelle Erkenntnisse zur Übergabe der Chatbot-Interaktion	51
Tabelle 5: Auswirkungen von Bewertungen	56
Tabelle 6: Szenariobeschreibung (FP I)	70
Tabelle 7: Operationalisierung der Variablen und Manipulationscheck (FP I).....	72
Tabelle 8: Ergebnisse des Manipulationschecks (FP I).....	74
Tabelle 9: Soziodemographische Merkmale (FP I).....	75
Tabelle 10: Gütekriterien (FP I)	76
Tabelle 11: Regressionskoeffizienten der Mediationsanalyse (FP I)	78
Tabelle 12: Zusammenfassung der Hypothesen (FP I)	78
Tabelle 13: Szenariobeschreibung (FP II: E 1)	84
Tabelle 14: Operationalisierung der Variablen und Manipulationscheck (FP II: E 1).....	87
Tabelle 15: Ergebnisse des Manipulationschecks (FP II: E 1).....	88
Tabelle 16: Soziodemographische Merkmale (FP II: E 1)	89
Tabelle 17: Gütekriterien (FP II: E 1)	90
Tabelle 18: Regressionskoeffizienten der Mediationsanalyse (FP II: E 1)	92
Tabelle 19: Zusammenfassung der Hypothesen (FP II: E 1).....	92
Tabelle 20: Szenariobeschreibung (FP II: E 2)	97
Tabelle 21: Operationalisierung der Variablen und Manipulationscheck (FP II: E 2).....	99
Tabelle 22: Soziodemographische Merkmale (FP II: E 2)	101
Tabelle 23: Nutzung von WhatsApp (FP II: E 2).....	102

Tabelle 24: Gütekriterien (FP II: E 2)	103
Tabelle 25: Regressionskoeffizienten der moderierten Mediationsanalyse (FP II: E 2)	105
Tabelle 26: Zusammenfassung der Hypothesen (FP II: E 2)	106
Tabelle 27: Szenariobeschreibung (FP II: E 3)	110
Tabelle 28: Operationalisierung der Variablen und Manipulationscheck (FP II: E 3)	112
Tabelle 29: Soziodemographische Merkmale (FP II: E 3)	114
Tabelle 30: Gütekriterien (FP II: E 3)	115
Tabelle 31: Regressionskoeffizienten der moderierten Mediation (FP II: E 3)	118
Tabelle 32: Zusammenfassung der Hypothesen (FP II: E 3)	119
Tabelle 33: Szenariobeschreibung (FP III)	127
Tabelle 34: Operationalisierung der Variablen und Manipulationscheck (FP III)	130
Tabelle 35: Ergebnisse des Manipulationschecks (FP IV: E 2)	131
Tabelle 36: Soziodemographische Merkmale (FP III)	132
Tabelle 37: Gütekriterien (FP III)	133
Tabelle 38: Regressionskoeffizienten der Moderationsanalyse (FP III)	135
Tabelle 39: Zusammenfassung der Hypothesen (FP III)	136
Tabelle 40: Szenariobeschreibung (FP IV: E 1)	143
Tabelle 41: Operationalisierung der Variablen und Manipulationscheck (FP IV: E 1)	144
Tabelle 42: Ergebnisse des Manipulationschecks (FP IV: E 1)	146
Tabelle 43: Soziodemographische Merkmale (FP IV: E 1)	147
Tabelle 44: Gütekriterien (FP IV: E 1)	148
Tabelle 45: Regressionskoeffizienten der Mediationsanalyse (FP IV: E 1)	151
Tabelle 46: Zusammenfassung der Hypothesen (FP IV: E 1)	151
Tabelle 47: Szenariobeschreibung (FP IV: E 2)	153
Tabelle 48: Manipulationscheck (FP IV: E 2)	155

Tabelle 49: Ergebnisse des Manipulationschecks (FP IV: E 2)	155
Tabelle 50: Soziodemographische Merkmale (FP IV: E 2)	156
Tabelle 51: Gütekriterien (FP IV: E 2).....	157
Tabelle 52: Regressionskoeffizienten der Mediationsanalyse (FP IV: E 2).....	160
Tabelle 53: Zusammenfassung der Hypothesen (FP IV: E 2)	161

Abkürzungsverzeichnis

a, b, c, c', d	Regressionskoeffizienten
ANOVA	Varianzanalyse
E	Experiment
eWoM	Electronic Word of Mouth
FAQs	Frequently Asked Questions
FP	Forschungsprojekt
KI	Konfidenzintervall
KI _o	Obere Grenze des Konfidenzintervalls
KI _u	Untere Grenze des Konfidenzintervalls
KITK	Korrigierte Item-to-Total-Korrelation
KMO	Kaiser-Meyer-Olkin-Kriterium
Koeff.	Koeffizient
M	Mittelwert
ns	nicht signifikant
p	Signifikanzniveau
SD	Standardabweichung
SE	Standardfehler
sig.	Signifikant
WoM	Word of Mouth
z. B.	zum Beispiel

1 Einleitung

1.1 Relevanz des Themas und Zielsetzung

Die fortschreitende Digitalisierung verändert die Kommunikation zwischen Menschen und Unternehmen. Insbesondere im Kundenservice, der direkten Schnittstelle beider Akteure, zeigen sich informationstechnologische Trends, welche die Art und Weise der Kommunikation verändern. Nachdem zahlreiche Jahre Kanäle wie Telefon und E-Mail breitflächig Anwendung fanden, gewinnen Chatbots an Relevanz. Diese virtuellen Schnittstellen entlasten oder ersetzen die menschlichen Servicemitarbeiter. Die zunehmende Zeit, die Kunden in digitalen Umgebungen verbringen, verstärkt die Relevanz dieser virtuellen Schnittstellen zusätzlich (Chen et al., 2024, S. 3636; Gentsch, 2018, S. 84; Sands et al., 2022, S. 2039-2040). Angetrieben von den kontinuierlichen Fortschritten im Bereich der künstlichen Intelligenz kommen immer häufiger Chatbots zum Einsatz, die eine Kommunikation zwischen Menschen und Maschinen in natürlicher Sprache ermöglichen (Chen et al., 2024, S. 3636). Durch weitere technologische Fortschritte werden automatisierte und menschenfreie Kundenserviceinteraktionen für viele Bereiche des Kundenservice zum Standard werden (Robinson et al., 2020, S. 373). So wird erwartet, dass das weltweite Marktvolumen von Chatbots bis 2032 auf 30 bis 42 Milliarden US-Dollar steigen wird. Es wird dabei von einer durchschnittlichen jährlichen Wachstumsrate von 23,0 % ausgegangen (Statista, 2025; Straits Research, 2024). Auch das Forschungsinteresse zu Chatbots ist in den letzten Jahren erheblich gestiegen (Mariani et al., 2023, S. 4). Beispielsweise wurden zahlreiche Studien zu Anthropomorphismus und Sozialer Präsenz durchgeführt (Adam et al., 2021; Blut et al., 2021; Chen et al., 2024; Jin & Youn, 2023; Liu et al., 2024; Morana et al., 2020)

Chatbots sind in die Kommunikationsinfrastruktur von Unternehmen eingebunden und ermöglichen eine effiziente und kostengünstige Interaktion mit Kunden. Dabei werden Chatbots nicht nur im Kundenservice genutzt. Sondern auch in vielen weiteren Bereichen, wie z. B. als Museumsführer, als Assistent beim Erlernen von Sprachen, im medizinischen Bereich oder einfach zum Unterhalten (Io & Lee, 2017, S. 215; Lester et al., 2004, S. 220; Shankar, 2018, S. vii; Shawar & Atwell, 2007, S. 30). Denn Chatbots sind nicht nur jederzeit verfügbar, sondern auch

in der Lage große Datenmengen zu verarbeiten und maßgeschneiderte Dienstleistungen anzubieten. Dabei helfen Chatbots, den Anforderungen vernetzter und gut informierter Kunden gerecht zu werden, indem Unternehmen schnell und kompetent auf Kundenanfragen reagieren. Dank ihrer vielfältigen Einsatzmöglichkeiten und der vielen Vorteile gewinnen Chatbots in der modernen Servicewelt zunehmend an Beliebtheit und werden beispielsweise auf Unternehmenswebseiten und in (Messenger-)Apps häufig integriert (Crollic et al., 2022, S. 132; De Cicco et al., 2020, S. 1214; Gentsch, 2018, S. 84; Kruse Brandão & Wolfram, 2018, S. 278-279). In der Literatur wird hier teilweise auch vom Conversational Commerce gesprochen, wobei durch intelligente Automatisierung Kundeninteraktionen optimiert werden (Gentsch, 2018, S. 84). Conversational Commerce beschreibt im Wesentlichen den Einsatz von Konversationsagenten wie beispielsweise Chatbots zur Interaktion mit Marken und deren Kundenservice. Dadurch können Kunden personalisierte Empfehlungen im Onlinehandel erhalten, ähnlich wie sie es von Verkäufern im Offline-Handel kennen. Conversational Commerce kann das Kundenerlebnis und die Kundenzufriedenheit beim Online-Einkauf verbessern (Lim et al., 2022, S. 1132). Dabei ist es das Ziel des Conversational Commerce, durch ein Gespräch einen Nutzer zum Kauf einer Dienstleistung oder eines Produkts zu führen. Hierbei können Chatbots eine zentrale Rolle einnehmen (Gentsch, 2018, S. 84). Chatbots sind folglich in den Kaufentscheidungsprozess von Kundinnen und Kunden eingebunden und können diesen gegebenenfalls beeinflussen.

Trotz des großen Potenzials von Chatbots im Kundenservice die Effizienz von Unternehmen zu steigern, bleiben Fragen in Bezug auf die Konzeption, den Einsatz und deren Auswirkungen offen (Crollic et al., 2022, S. 132; Lim et al., 2022, S. 1132). Bis vor einigen Jahren gab es nur wenig Forschung zu Chatbots im Bereich Marketing und viele Veröffentlichungen waren lediglich im Rahmen von Konferenzbeiträgen und wenige in wissenschaftlichen Zeitschriften. Es wurden einige relevante Forschungslücken identifiziert, die im Rahmen der Forschungsprojekte dieser Arbeit angesprochen werden. Dabei liefert diese Arbeit Erkenntnisse zu den Themengebieten Anthropomorphismus, Zeitpunkt der Offenlegung, Übergabe der Interaktion an einen menschlichen Mitarbeiter sowie Einsatz von Bewertungen im Kontext von Chatbots. Im Fokus dieser Arbeit stehen die folgenden Forschungsziele:

1. Diese Arbeit soll konzeptionelle Grundlagen zu Chatbots liefern und den aktuellen empirischen Forschungsstand im Bereich Anthropomorphismus, Offenlegungszeitpunkt, Übergabe an menschliche Mitarbeiter und Bewertung im Kontext von Chatbots aufarbeiten.

2. Ziel dieser Arbeit ist es, den wissenschaftlichen Kenntnisstand zu Chatbots zu erweitern, insbesondere in den Bereichen Anthropomorphismus, Offenlegungszeitpunkt, Übergabe an menschliche Mitarbeiter und Bewertungen.

Anthropomorphismus und Soziale Präsenz werden in Forschungen zu Chatbots mittlerweile häufig betrachtet. Dabei kann Anthropomorphismus zu einer höheren Sozialen Präsenz des Chatbots führen (Chen et al., 2024, S. 3646-3647; De Cicco et al., 2020, S. 1224; Gnewuch et al., 2018, S. 11; Han, 2021, S. 56; Jiang et al., 2023, S. 8; Jin & Youn, 2023, S. 1879; Liu et al., 2024, S. 12; Morana et al., 2020, S. 11). Diese Arbeit liefert hier einen Beitrag zur bisherigen Forschung, indem die folgende Forschungsfrage untersucht wird: Inwiefern beeinflusst eine anthropomorphe Gestaltung eines Chatbots die Soziale Präsenz, das Vertrauen sowie die Wiederverwendungsabsicht des Chatbots? Diese Forschungsfrage wird in Forschungsprojekt eins (*FP I*) untersucht. Wobei anzumerken gilt, dass *FP I* hierzu Anfang 2022 durchgeführt wurde, als einige Publikationen noch nicht erschienen waren. Darüber hinaus liefert *FP I* eine Grundlage für das anthropomorphe Chatbot Design, das für die weiteren Forschungsprojekte in angepasster Form verwendet wird.

Durch technologische Fortschritte und den Einsatz anthropomorpher Gestaltungsmöglichkeiten ist es teilweise schwierig bis unmöglich, eine Konversation mit einem Chatbot von der mit einem Menschen zu unterscheiden (Söderlund & Oikarinen, 2021, S. 95). Daher muss die nicht-menschliche Identität offengelegt werden, um den Konversationspartner als Chatbot zu identifizieren. Allerdings stehen Unternehmen hier vor einem Dilemma der Offenlegung, denn sie müssen zwischen Transparenz und Effizienz abwägen. Aus Gründen der Transparenz und Ehrlichkeit gegenüber Kunden erscheint die Offenlegung als sinnvoll. Allerdings werfen Ergebnisse der Forschung Zweifel in Bezug auf diese Annahme auf (Luo et al., 2019, S. 938; Mozafari et al., 2020, S. 3). Denn die Offenlegung von Chatbots geht häufig mit negativen Kundenreaktionen einher (Ishowo-Oloko et al., 2019; Luo et al., 2019; Mozafari et al., 2022). Hier könnte der Zeitpunkt der Offenlegung helfen, negative Kundenreaktionen zu verringern. Daher wird in dieser Arbeit in Forschungsprojekt zwei (*FP II*) der Frage nachgegangen, inwiefern der Zeitpunkt der Offenlegung die Wiederverwendungsabsicht beeinflusst. Da Chatbots auch häufig auf unterschiedlichen Kommunikationsplattformen eingesetzt werden sowie aktuell auch teilweise noch nicht zufriedenstellende Ergebnisse auf Kundenanfragen liefern, wird der Aspekt der Kommunikationsplattform sowie der Lösung der Nutzeranfrage im Rahmen unterschiedlicher Experimente untersucht.

Eine Möglichkeit zur Wiederherstellung der Interaktion nach einem Servicefehler des Chatbots besteht in der Übergabe an einen menschlichen Mitarbeiter (Meyer et al., 2022, S. 2501). Denn Fehler, Missverständnisse oder Gesprächsabbrüche treten trotz technischer Fortschritte von Chatbots immer noch auf (Zarouali et al., 2018, S. 495). Zwar nehmen solche Situationen durch den Einsatz künstlicher Intelligenz ab, sie lassen sich jedoch aktuell nicht vollständig vermeiden. Manche Anfragen können weder von einem menschlichen Mitarbeiter noch von einem Chatbot zur Zufriedenheit des Nutzers gelöst werden. In solchen Fällen stellt sich die Frage: Wer sollte die Übermittlung einer solchen Nachricht übernehmen? Mit dieser Fragestellung befasst sich Forschungsprojekt drei (*FP III*).

Bewertungen sind ein allgemein häufig untersuchtes Mittel, um Kunden in ihrer Entscheidungsfindung zu helfen und auch zu beeinflussen. Hier findet sich im Bereich von Chatbots bisher keine Forschung. Allerdings sind immer noch einige Kunden skeptisch gegenüber der Nutzung von Chatbots im Kundenservice (Donath, 2024). Hier können Bewertungen helfen, diese Skepsis zu reduzieren und Kunden zur Nutzung zu bewegen. Im Rahmen dieser Arbeit wird daher die Forschungsfrage untersucht, welchen Einfluss Bewertungen auf der Grundlage des Confirmation/Disconfirmation-Paradigmas haben. Dieser Frage wird in Forschungsprojekt vier (*FP IV*) nachgegangen und dabei die Auswirkung auf die Wiederverwendungsabsicht untersucht.

Die Forschungsfragen dieser Arbeit werden mithilfe wissenschaftlicher Experimente untersucht, die den einzelnen Forschungsprojekten zugeordnet sind. Die Ergebnisse der experimentellen Untersuchungen zeigen Faktoren, die sich auf die Wiederverwendungsabsicht eines Chatbots auswirken. Dabei werden unterschiedliche Faktoren gefunden, die als Mediatoren oder Moderatoren im Rahmen des Einsatzes von Chatbots fungieren. Die vorliegende Arbeit liefert eine gründliche Analyse des aktuellen Forschungsstands zu Chatbots in den Bereichen Anthropomorphismus, Offenlegung, Übergabe und Bewertungen. Darüber hinaus unterstützt und erweitert sie bestehende wissenschaftliche Ergebnisse und leistet so einen Beitrag zur aktuellen Forschung.

1.2 Aufbau der Arbeit

Die vorliegende Arbeit gliedert sich in sechs Hauptkapitel. Nach der Einleitung folgen in Kapitel zwei die konzeptionellen Grundlagen. Zuerst werden aktuelle Definitionsvorschläge vergleichend betrachtet und der Begriff Chatbots wird definiert. Anschließend folgt ein Unterkapitel zur Kategorisierung von Chatbots sowie ein weiteres Unterkapitel zum Einsatz von Chatbots im Kundenservice.

In Kapitel drei werden theoretische Grundlagen beschrieben, die für die empirische Evidenz sowie für die Herleitung der Hypothesen relevant sind. Dabei werden zuerst Grundlagen zu Anthropomorphismus betrachtet. Anschließend werden die Social Response Theorie und die Media Equation Theorie, der Uncanny Valley Effekt, die Attributionstheorie, die Familiarity Heuristik sowie das Confirmation/Disconfirmation Paradigma beschrieben.

In Kapitel vier wird die empirische Evidenz zu den für diese Arbeit wichtigen Teilgebieten der Chatbot-Forschung betrachtet. Hier wird zuerst der Stand der Forschung zum Design anthropomorpher Chatbots aufgezeigt. Also welche Faktoren können dazu führen, dass Chatbots eine anthropomorphe Wahrnehmung hervorrufen. Anschließend werden Forschungsergebnisse zu anthropomorphen Chatbots und deren Auswirkungen mit einem Fokus auf das Vertrauen und die Soziale Präsenz beschrieben. Im darauffolgenden Unterkapitel wird der aktuelle Forschungsstand zum Thema der Offenlegung und des Offenlegungszeitpunkts von Chatbots betrachtet. Anschließend erfolgt die Betrachtung der aktuellen Forschung in Bezug auf die Übergabe an einen menschlichen Mitarbeiter. Im letzten Unterkapitel werden Forschungsergebnisse zu Kundenbewertungen allgemein betrachtet, da im Rahmen der Literaturrecherche keine Forschung im Zusammenhang mit Chatbots gefunden wurde.

Im fünften Kapitel werden die vier durchgeführten Forschungsprojekte beschrieben und die Ergebnisse ausgewertet sowie diskutiert. Kapitel fünf beginnt mit einer kurzen Übersicht der Forschungsprojekte. Im darauffolgenden Unterkapitel werden allgemeine Grundlagen zu Experimenten und deren Auswertung beschrieben. Hier wird ein Fokus auf die in dieser Arbeit verwendeten Methoden und behandelten Inhalte gelegt. In den vier folgenden Unterkapiteln werden jeweils die Forschungsprojekte mit ihren dazugehörigen Experimenten beschrieben. Dabei ist das Vorgehen jeweils analog. Zuerst werden die Hypothesen hergeleitet und anschlie-

ßend die Methodik beschrieben. Darauf folgt die Beschreibung der Stichprobe und die Gütebeurteilung der verwendeten Skalen. Bevor die Ergebnisse jeweils diskutiert werden, werden diese beschrieben.

Kapitel sechs fasst die wichtigsten Erkenntnisse der Arbeit nochmal zusammen. Darüber hinaus werden Implikationen für die Wissenschaft und die unternehmerische Praxis aufgeführt und ein Ausblick auf möglichen weiteren Forschungsbedarf gegeben. Der Aufbau dieser Arbeit ist in Abbildung 1 übersichtlich dargestellt.

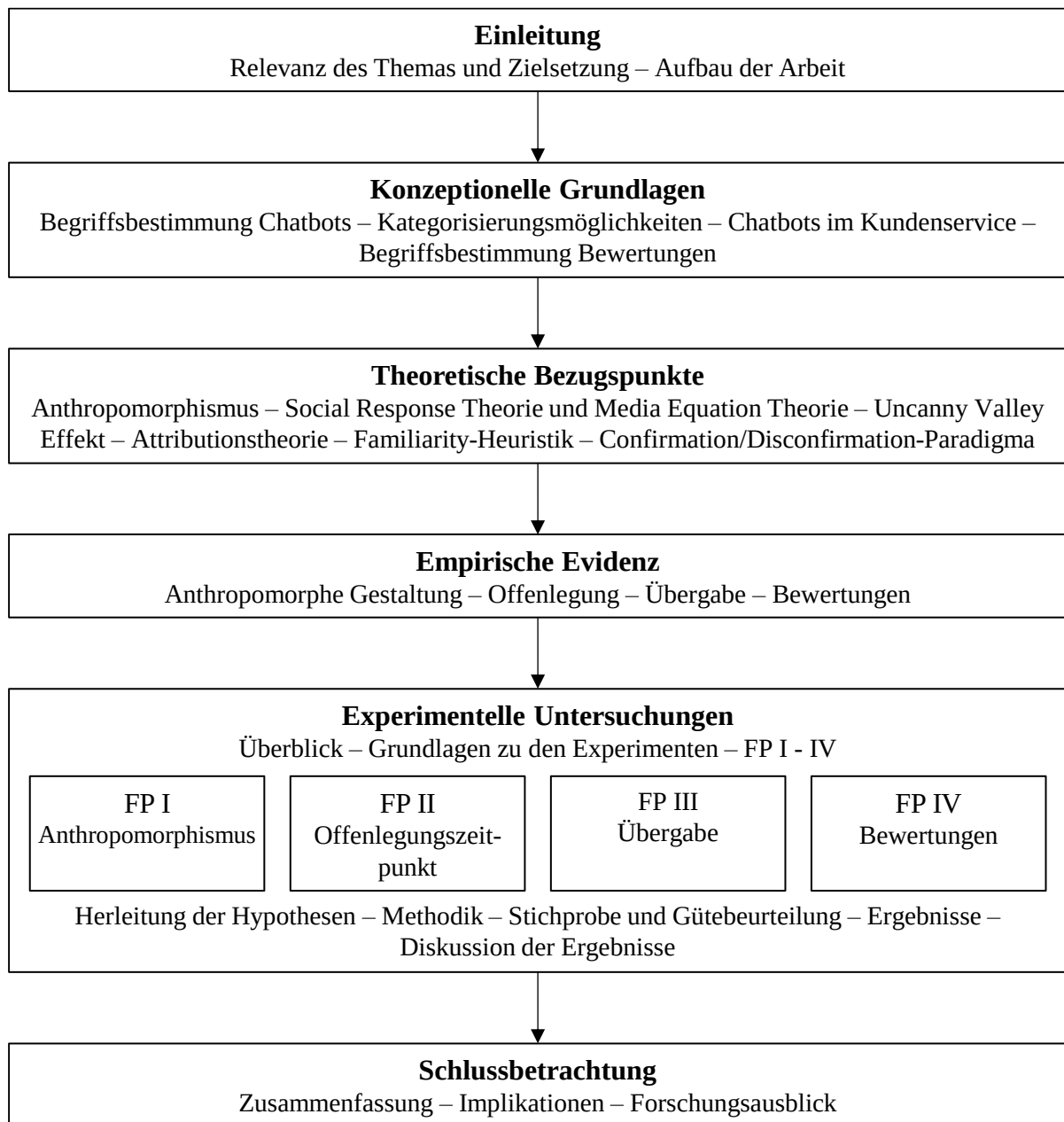


Abbildung 1: Aufbau der Arbeit

2 Konzeptionelle Grundlagen

Im zweiten Kapitel werden die konzeptionellen Grundlagen dieser Arbeit betrachtet. Dabei wird zuerst ausführlich auf das Thema Chatbots eingegangen. Anschließend werden einige Grundlagen zu Bewertungen beschrieben, die für *FP IV* von Relevanz sind.

2.1 Chatbots

In Kapitel 2.1 werden die Grundlagen zu Chatbots beschrieben. Hier wird zunächst der Begriff Chatbots für diese Arbeit definiert, indem auf Basis bisherige Definitionsansätze eine Arbeitsdefinition erarbeitet wird. Anschließend werden in der Literatur gängige Kategorisierungsmöglichkeiten von Chatbots betrachtet sowie abschließend Einsatzmöglichkeiten von Chatbots im Kundenservice erläutert.

2.1.1 Begriffsbestimmung – Chatbot

In der wirtschaftswissenschaftlichen Literatur finden sich einige Definitionsvorschläge für den Begriff Chatbot. Diese Definitionsvorschläge unterscheiden sich teilweise hinsichtlich ihrer Inhalte. Die Kernaussage ist allerdings häufig sehr ähnlich, wenn nicht sogar die Gleiche.

Allgemein stellt der Begriff Chatbot ein Kofferwort dar, das sich aus den beiden Wörtern Chat (= „plaudern“) und Bot (= „Roboter“) zusammensetzt (Kruse Brandão & Wolfram, 2018, S. 278; Rese et al., 2020, S. 2; Stucki et al., 2018, S. 728, 2020, S. 3). Weizenbaum (1966, S. 36) beschreiben ihren Chatbot ELIZA als ein Programm, das natürlichsprachliche Konversationen mit einem Computer ermöglicht, dabei die Eingaben auf der Tastatur lesen kann und auf demselben Computer antwortet. Von dieser Sichtweise haben sich aktuelle Definitionen weiterentwickelt.

Dale (2016, S. 813-814) verstehen unter einem Chatbot einen Sammelbegriff für Softwareanwendungen, die in natürlicher Sprache einen Dialog mit einem Menschen führen. Dies geschieht meist in geschriebener Form. Allerdings merken sie an, dass das Unterscheidungsmerkmal der geschriebenen Form mit Fortschritten in der Spracherkennung zunehmend schwächer wird und auch andere Kommunikationsarten möglich sind. Chatbots stellen eine Schnittstelle zwischen Kunden und Unternehmen dar (Zarouali et al., 2018, S. 495). Auch Kruse Brandão und Wolfram (2018, S. 278) beschreiben Chatbots als textbasierte Dialogsysteme, die mit Nutzern in natürlicher Sprache einen Dialog wie mit einer natürlichen Person führen.

Andere Autoren erweitern die Definitionen um weitere Kommunikationsformen. Das Verständnis von Crolic et al. (2022, S. 132) ist sehr offen und wenig technisch, allerdings auf einen Unternehmenskontext bezogen, in dem explizit von Kunden als Kommunikationspartner von Chatbots gesprochen wird. Nach Crolic et al. (2022, S. 132) kommunizieren Chatbots mit Kunden mittels Text oder Stimme, um auf eine Vielzahl von Kundenbedürfnissen einzugehen.

Hatwar et al. (2016, S. 85) nehmen die Bezeichnung Software-Agent mit in ihre Definition auf. Sie beschreiben Chatbots als Software-Agenten, die eine Entität simulieren, in der Regel einen Menschen mit vagen oder spezifisch definierten Eigenschaften, mit dem ein Benutzer in einer (schriftlichen, mündlichen oder gemischten) Konversation interagieren kann (Hatwar et al., 2016, S. 85). Ähnlich allgemein ist das Verständnis von Huang et al. (2007, S. 423) die unter einem Chatbot einen Gesprächsagenten verstehen, der mit einem Benutzer zu einem bestimmten Bereich oder einem bestimmten Thema in Sätzen in natürlicher Sprache und in Form eines Dialogs interagiert. Auch Stucki et al. (2018, S. 728, 2020, S. 3) sehen in einem Chatbot ein System, das mit einem menschlichen Nutzer einen Dialog führen und gewisse Aufgaben autonom erledigen kann. Dabei wird das System mittels natürlicher Sprache (gesprochen oder geschrieben) angesteuert und liefert die gewünschten Informationen in natürlicher Sprache zurück. Weitere Autoren zählen zu Chatbots sowohl sprachgesteuerte als auch textbasierte Systeme (Rese et al., 2020, S. 1; Sheehan et al., 2020, S. 14).

Im Gabler Wirtschaftslexikon wird ein Chatbot wie folgt definiert: „Chatbots sind Dialogsysteme mit natürlichsprachlichen Fähigkeiten textueller oder auditiver Art. Sie werden, oft in Kombination mit statischen oder animierten Avataren, auf Websites oder in Instant-Messenger-Systemen verwendet, wo sie die Produkte und Services ihrer Betreiber erklären und bewerben respektive sich um Anliegen der Interessenten und Kunden kümmern – oder einfach dem Amüsement und der Reflexion dienen. [...] Zuweilen wird der Begriff der Chatbots so weit gefasst, dass auch Sprachassistenten (Voicebots oder Voice Assistants) darunter fallen“ (Bendel, 2024).

Diese Definitionen betrachten hauptsächlich die Aspekte des natürlichen Dialogs, die Art der Kommunikation (Text, Sprache, gemischt) sowie die Aufgabenerfüllung. Die technischen Aspekte werden nicht aufgenommen. Im Folgenden werden Definitionen betrachtet, die diesen technischen Aspekt (künstliche Intelligenz, Natural Language Processing) zusätzlich beschreiben. Da die technischen Aspekte von Chatbots nicht im Mittelpunkt dieser Arbeit stehen, wird im nächsten Abschnitt kurz erläutert, was unter künstlicher Intelligenz und Natural Language

Processing zu verstehen ist, um diese Begriffe einordnen zu können und ein grundlegendes Verständnis zu schaffen.

Künstliche Intelligenz beschreibt die Fähigkeit von Computersystemen, Aufgaben eigenständig zu lösen, die aufgrund ihrer Komplexität bislang menschliche Fähigkeiten erforderten (Gethmann et al., 2022, S. 8). Anders ausgedrückt beschäftigt sich künstliche Intelligenz mit Methoden, die es Computern ermöglichen, Aufgaben zu lösen, die zuvor nur von Menschen gelöst werden konnten und dabei menschliche Intelligenz erforderten (Lackes & Siepermann, 2018; Russel & Norvig, 2022, S.20-22). Computerprogramme, die auf künstliche Intelligenz zurückgreifen, sind nicht per Definition intelligent. Sie sind lediglich in der Lage anspruchsvolle Aufgaben zu lösen, die bislang den Einsatz menschlicher Fähigkeiten erfordert haben. Zu diesen Fähigkeiten zählen beispielsweise Wahrnehmen, Lernen, Wissen, Denken, Kommunizieren und Handeln (Gethmann et al., 2022, S. 8). Damit moderne Chatbots kommunizieren können, ist Natural Language Processing notwendig. Natural Language Processing stellt dabei einen Teilbereich der künstlichen Intelligenz dar, wobei künstliche Intelligenz, Linguistik, Kognitionspsychologie und Neurowissenschaften verknüpft werden (Nadkarni et al., 2011, S. 544; Russel & Norvig, 2022, S. 874). Natural Language Processing behandelt die Verarbeitung natürlicher Sprache, z. B. die Spracherkennung, Textanalyse und die Sprachgenerierung (Gethmann et al., 2022, S. 16). Technologien nutzen Natural Language Processing, um mit Menschen zu kommunizieren und zu lernen (Russel & Norvig, 2022, S. 874). Folglich ermöglicht Natural Language Processing das Verstehen, Interpretieren und Generieren von Texten in gesprochener Sprache, so dass beispielsweise ein Chatbot, der Natural Language Processing nutzt, mit Menschen in einer Weise kommunizieren kann, die natürlich und sinnvoll erscheint. Dabei können Chatbots eigenständig die Eingabe der Nutzer erkennen, verstehen und adäquat antworten.

Die Wissensbasis eines Chatbots besteht aus einer Sammlung von Dialogmanagementregeln, die verschiedene Techniken zur Verarbeitung von Benutzereingaben verwenden. Dabei sind Chatbots eine praktische und nutzerzentrierte Form der Anwendung künstlicher Intelligenz (Ukpabi et al., 2019, S. 106). Chatbots sind eine der ersten Natural Language Processing Anwendungen (Dale, 2016, S. 814) und sind eines der bekanntesten Beispiele für intelligente Mensch-Computer-Interaktionen (Khanna et al., 2015, S. 278). Neben Natural Language Processing können weitere Anwendungen künstlicher Intelligenz, wie Bild- und Videoverarbeitung sowie Audioanalyse notwendig sein, damit Chatbots zielführend eingesetzt werden können (Bala et al., 2017, S. 2030). Chatbots, die auf künstlicher Intelligenz basieren, bieten in der

Regel eine sehr gute natürlichsprachliche Schnittstelle zwischen Software und Menschen (Rese et al., 2020, S. 2). Moderne Chatbots können Aktionen auf Grundlage von Benutzereingaben oder Veränderungen in der Umgebung reaktiv, proaktiv oder autonom ausführen. Darüber hinaus sind sie anpassungsfähig und in der Lage selbst zu lernen, um Kontextinformationen zu verarbeiten oder Benutzerpräferenzen zu berücksichtigen (Meyer von Wolff et al., 2019, S. 96). Sie helfen Nutzern individuelle Informationen zu erhalten (Wuenderlich & Paluch, 2017, S. 3). Nach diesem kurzen Exkurs werden im nächsten Abschnitt Definitionen betrachtet, die technische Begrifflichkeiten in ihre Ausführungen einbeziehen.

Shawar und Atwell (2007, S. 29-30) definieren einen Chatbot als Computerprogramm, welches mit Nutzern in natürlicher Sprache kommuniziert. Chatbots haben den Zweck, eine menschliche Konversation zu simulieren. Dabei werden ein Sprachmodell sowie Berechnungsalgorithmen in die Chatbot Architektur integriert, um die Kommunikation in natürlicher Sprache zu emulieren (Shawar & Atwell, 2007, S. 29-30). Die folgenden Definitionen nutzen künstliche Intelligenz oder Natural Language Processing explizit im Rahmen ihres Begriffsverständnisses. Mozafari et al. (2021a, S. 2916, 2022, S. 221) grenzen Chatbots auf textbasierte Benutzeroberflächen ein, die auf Natural Language Processing basieren und eine Konversation zwischen Menschen nachahmen. Hierbei verweisen sie auf Wirtz et al. (2018). Auch Hoyer et al. (2020, S. 59-60) beschreiben künstliche Intelligenz beziehungsweise Natural Language Processing als Technologie hinter Chatbots. Dabei verstehen sie unter Chatbots ein virtuelles Assistenzprogramm, das Gespräche mit Nutzern über Audio oder Text führt und somit menschliche Gespräche simuliert (Hoyer et al., 2020, S. 59-60).

Ein Chatbot kann auch als Anwendungssystem gesehen werden, das eine natürlichsprachliche kommunikative Benutzeroberfläche zur Mensch-Maschinen-Interaktion bietet. Dabei greift es auf Technologien der künstlichen Intelligenz sowie unterschiedliche (unternehmensinterne oder- externe) Datenbanken zurück. Chatbots haben dort die Aufgabe, Routinetätigkeiten zu automatisieren oder Nutzer bei Aufgaben zu unterstützen (Meyer von Wolff et al., 2019, S. 96; Meyer von Wolff & Schumann, 2018, S. 12). Allgemein versteht man unter einem Chatbot ein Computerprogramm, das mithilfe künstlicher Intelligenz eine natürliche menschliche Unterhaltung oder einen Chat simuliert, um Geschäftsprozesse zu automatisieren (Tuzovic & Paluch, 2018, S. 86). Nach Völkle und Planing (2019, S. 218) sind Chatbots eine Software, die eine Interaktion in natürlicher Sprache simuliert und teilweise auf künstlicher Intelligenz basiert.

Bala et al. (2017, S. 2030) verstehen unter einem Chatbot ein Computerprogramm, das menschliche Unterhaltungen in ihrer natürlichen Form (einschließlich geschriebenem Text oder gesprochener Sprache) nachahmt, indem es künstliche Intelligenz verwendet. Eine weitere Definition, die den technischen Aspekt mit aufgreift, liefern Khanna et al. (2015, S. 277). Sie beschreiben einen Chatbot als ein Computerprogramm, das intelligent antwortet, wenn man sich mit ihm über geschriebenen Text oder gesprochene Sprache unterhält und mindestens eine menschliche Sprache durch Natural Language Processing verstehen kann. Hierbei interpretiert das System die menschliche Spracheingabe anhand der ihm zugeführten Informationen (Khanna et al., 2015, S. 277).

Es wird deutlich, dass die Definitionen weitgehend übereinstimmen und auf denselben grundlegenden Bausteinen basieren, die jedoch unterschiedlich kombiniert und angewendet werden. Zu diesen Bausteinen der Definitionen zählen: Die natürliche Sprache, die Kommunikation in Form von Dialogen zwischen Menschen und Maschine, die Art der Kommunikation (Text, Sprache, gemischt), die Aufgabenerfüllung und die zugrundeliegende Technologie (Natural Language Processing, künstliche Intelligenz, Datenbanken).

Im Rahmen dieser Arbeit wird unter einem Chatbot eine Software verstanden, die häufig mithilfe von künstlicher Intelligenz eine natürlichsprachliche, dialogische Benutzerschnittstelle zwischen Menschen und Maschine darstellt, um Nutzeranfragen automatisiert zu bearbeiten. Dabei erfolgt die Interaktion zwischen Nutzern und Chatbots in Form von geschriebenem Text.

Es gilt anzumerken, dass neben der Bezeichnung Chatbot weitere Begriffe existieren, die teilweise synonym verwendet werden oder ähnliche Konstrukte beschreiben. Hierzu zählen beispielsweise Intelligent Virtual Assistant, Digital Assistant, Conversational Interface, Conversational Agent, Talkbot, Chatterbot, Bot, Artificial Conversational Entity, Chatter Robots, Machine Conversation System, Virtual Agent, Dialogue System, Voicebot (Bala et al., 2017, S. 2030; Bendel, 2024; Crolic et al., 2022, S. 132; Dale, 2016, S. 811; Hatwar et al., 2016, S. 85; Shawar & Atwell, 2007, S. 29; Stucki et al., 2018, S. 728, 2020, S. 3). Allen ist gemein, dass sie dasselbe Grundkonzept verfolgen, nämlich ein bestimmtes Ergebnis durch einen Dialog in natürlicher Sprache mit einer Maschine zu erzielen (Dale, 2016, S. 811). Häufig wird der Begriff Chatbot verwendet, da Nutzer Nachrichten in natürlicher, geschriebener Sprache austauschen („chatten“) können (Fischer & Lueg, 2020, S. 45).

2.1.2 Kategorisierungsmöglichkeiten von Chatbots

Es existieren unterschiedliche Kategorisierungsmöglichkeiten von Chatbots. Im Folgenden werden drei Möglichkeiten vorgestellt. Chatbots können nach ihrem Kommunikationsmedium, ihrer Funktionsweise sowie ihrem Verwendungszweck klassifiziert werden (Adamopoulou & Moussiades, 2020a, S. 5; Stucki et al., 2020, S. 6; Wernsmann, 2022, S. 359). Diese Liste der Kategorien ist nicht abschließend. In der Literatur existieren weitere Kategorisierungsmöglichkeiten (Adamopoulou & Moussiades, 2020a, S. 5, 2020b, S. 377-379; Ferner & Lienbacher, 2023, S. 162-164). Diese überschneiden sich teilweise mit den hier genannten, daher wird die zuvor genannte Kategorisierungsmöglichkeit im Rahmen dieser Arbeit gewählt.

Kategorisierung nach dem Kommunikationsmedium

Die Klassifizierung nach dem Kommunikationsmedium unterscheidet zwischen Text-, Voice- und teilweise auch Human-Bots (Ferner & Lienbacher, 2023, S. 162; Gnewuch et al., 2017, S. 3; Jain et al., 2018, S. 895; Wernsmann, 2022, S. 359-360). Bei Textbots findet die Kommunikation in Textform statt. Nutzer können in geschriebener Form mit dem Chatbot interagieren. Dabei kann die Eingabe in Form von Freitext oder mithilfe vorgegebener Dialogelemente ausgeführt werden. Textbasierte Chatbots sind heutzutage weit verbreitet und finden Anwendung auf vielen Kommunikationsplattformen, Webseiten oder Messenger-Apps (Wernsmann, 2022, S. 360-361). Textbots entsprechen somit am ehesten den in dieser Arbeit untersuchten Chatbots.

Im Gegensatz zu Textbots stellt bei Voicebots die wesentliche Kommunikationsform die gesprochene Sprache dar (Wernsmann, 2022, S. 360). Dadurch wird den Nutzern ermöglicht, mit dem Chatbot in gesprochener Sprache zu interagieren. Im Vergleich zu Textbots sind sie technisch komplexer, da sie neben der Textverarbeitung auch Sprachausgabe-Technologie benötigen. Diese finden sich allerdings heutzutage in fast jedem Smartphone oder werden in smarten Lautsprechern eingesetzt (Wernsmann, 2022, S. 360).

Human-Bots stellen ein Kommunikationssystem dar, das in Form von digital animierten Menschen auftreten und dabei neben Text und Sprache auch Mimik und Gestik für die Kommunikation nutzen. Sie nutzen eine Gesichtserkennung, um Menschen zu identifizieren und eine Mimik- beziehungsweise Gestenerkennung, um die Körpersprache der Nutzer zu interpretieren. Dabei kann der Chatbot seine Körpersprache sowie den Tonfall an die jeweilige Situation anpassen (Wernsmann, 2022, S. 360).

Kategorisierung nach der Funktionsweise

Neben der Art und Weise, wie mit dem Chatbot interagiert wird, können Chatbots auch nach ihrer Funktionsweise klassifiziert werden. Hier wird häufig zwischen regelbasierten beziehungsweise musterbasiert und selbstlernenden Chatbots unterschieden (Kruse Brandão & Wolfram, 2018, S. 282).

Regelbasierte Chatbots folgen einem programmierten Muster und vordefinierten Regeln und können nur auf bestehende Textbausteine und Antworten zurückgreifen, die im Vorfeld definiert und geskripted wurden. Dadurch sind sie nur in der Lage, Eingaben von Nutzern zu verarbeiten und Antworten zu erzeugen, die von diesen Mustern und Regeln abgedeckt werden. Häufig haben diese Chatbots eine Eingabemaske mit vorgefertigten Auswahlmöglichkeiten (Kruse Brandão & Wolfram, 2018, S. 282; Stucki et al., 2020, S. 7). Sie sind daher nur so intelligent wie die zugrundeliegende Datenbasis (Gentsch, 2019, S. 85) und in ihrer Funktionsweise auch eingeschränkt.

Intelligente, selbstlernende Chatbots weisen deutlich geringere Limitationen in ihrer Anwendung auf. Denn sie können Dialogelemente verarbeiten, die durch die Entwickler im Vorfeld nicht explizit einprogrammiert wurden. Sie sind somit flexibler, was die mögliche Eingabe und Ausgabe betrifft. Sie nutzen fortschrittliche Methoden der künstlichen Intelligenz wie maschinelles Lernen, Natural Language Processing und teilweise auch Deep Learning. Sie sind in der Lage, durch Konversationen mit Nutzern selbstständig zu lernen und sich kontinuierlich weiterzuentwickeln. Sie können auch den Kontext, die Absicht oder Emotionen von Nutzern verarbeiten (Arsenijevic & Jovic, 2019, S. 20; Izadi & Forouzanfar, 2024, S. 805-806; Kruse Brandão & Wolfram, 2018, S. 282; Stucki et al., 2020, S. 7).

Kategorisierung nach dem Verwendungszweck

Das dritte Unterscheidungskriterium stellt der Verwendungszweck des Chatbots dar. Dabei können Chatbots in zielgerichtete (aufgabenorientierte) und nicht-zielgerichtete (nicht-aufgabenorientierte) Typen unterteilt werden (Jurafsky & Martin, 2024, S. 309; Serban et al., 2015, S. 4; Stucki et al., 2020, S. 6). Zielgerichtete Chatbots lösen ein explizit definierbares Nutzerproblem, wohingegen bei nicht-zielgerichteten Chatbots das Nutzerproblem nicht eng definierbar ist. Sie dienen nicht der Lösung eines zuvor festgelegten Problems und werden eher für soziale Interaktionen genutzt. Im Unternehmenskontext sind nicht zielgerichtete Chatbots von untergeordneter Bedeutung. Zielgerichtete Chatbots können dabei beispielsweise auf Anfragen

antworten, Auskünfte zu Produkten oder Dienstleistungen liefern oder sogar definierte Tätigkeiten für Unternehmen übernehmen, wie beispielsweise im Verkauf eingesetzt werden oder Reklamationen entgegennehmen (Stucki et al., 2020, S. 6).

2.1.3 Integration von Chatbots im Kundenservice

Allgemein versteht man unter Kundenservice die Betreuung, Beratung und Unterstützung von Kunden vor, während und nach dem Kauf eines Produktes oder einer Dienstleistung durch ein Unternehmen (Haugeland et al., 2022, S. 2; Kirchgeorg, 2018, S. 1; Savage, 2022). Der Kundenservice unterstützt den Kunden während der gesamten Customer Journey. Die Customer Journey beschreibt die Interaktion von Kunden mit Unternehmen über unterschiedliche Berührungspunkte und differenziert drei Phasen, die ein Kunde im Rahmen eines Kaufprozesses durchläuft: Vorkaufphase, Kaufphase und Nachkaufphase. Aus diesem gesamten Prozess ergibt sich eine ganzheitliche Kundenerfahrung. Im Rahmen dieses Prozesses spielen auch vorherige Erfahrungen sowie externe Einflussfaktoren eine Rolle und können einen Einfluss auf die Kundenerfahrung haben (Lemon & Verhoef, 2016, S. 74-77).

Ziel des Kundenservice ist es, eine hohe Kundenzufriedenheit zu gewährleisten und Probleme oder Fragen der Kunden schnell und effizient zu lösen (Haugeland et al., 2022, S. 2; Savage, 2022). Dabei kann er eine beratende Funktion haben, bei der Lösung von Kundenanfragen und Problemen helfen oder Feedback entgegennehmen. Ein effektiver Kundenservice trägt erheblich zum Erfolg eines Unternehmens bei, da zufriedene Kunden eher zurückkehren und das Unternehmen weiterempfehlen (Kirchgeorg, 2018, S. 1).

Chatbots werden im Kundenservice zunehmend erfolgreich eingesetzt (Leitherer, 2024). 49 % der Kunden in Deutschland haben in den vergangenen zwölf Monaten bereits einen Chatbot genutzt. Ein Drittel der deutschen Kunden befürworten den Einsatz von Chatbots im Kundenservice und der Kundenkommunikation. 47 % sind bezüglich des Einsatzes unentschlossen, wobei diese dem Einsatz von Chatbots im Kundenservice und der Kundenkommunikation offen oder zwiespältig gegenüberstehen. Lediglich jeder fünfte sieht die Nutzung von Chatbots im Kundenservice und der Kundenkommunikation kritisch und lehnt deren Einsatz sogar ab. Folglich zeigt sich, dass die Mehrheit der deutschen Kunden dem Einsatz von Chatbots im Kundenservice aufgeschlossen gegenüberstehen (Donath, 2024).

Technologische Entwicklungen sowie der zunehmende Onlinehandel, die Nutzung mobiler Endgeräte, Messenger-Apps und sozialer Medien haben die Art und Weise, wie Unternehmen mit Kunden in Kontakt treten, grundlegend verändert (Savage, 2022). Menschen verbringen heutzutage viel Zeit in digitalen Umgebungen, daher gehen viele Unternehmen zu digitalen Servicelösungen wie Chatbots über. Chatbots werden immer häufiger als Schnittstellen zu Kunden im Kundenservice integriert (Sands et al., 2021, S. 1). Chatbots werden immer intelligenter und leistungsfähiger und können von Unternehmen immer besser eingesetzt werden, um Kunden effizient und effektiv anzusprechen und diese zu verstehen. Dadurch können immer mehr Kundenanliegen mithilfe von Chatbots gelöst werden (Io & Lee, 2017, S. 215; Sands et al., 2021, S. 2).

Chatbots werden beispielsweise im Onlinehandel eingesetzt, um auf unterschiedliche Bedürfnisse von Kunden einzugehen und dabei wirtschaftlicher als menschliche Arbeitskräfte zu arbeiten (Song & Shin, 2024, S. 442). Früher waren häufig der direkte Kontakt im Geschäft, ein Anruf oder eine E-Mail die einzigen Kontaktmöglichkeiten, was auch mit hohen Kosten für das Unternehmen verbunden war. Mittlerweile wird ein Großteil der Serviceanfragen durch Selbst-Service beispielsweise in Form von Frequently Asked Questions (FAQs) oder durch den Einsatz von Chatbots gelöst. Dieser Wechsel hin zu anderen Technologien wie Chatbots kann auch als „Bot Effect“ bezeichnet werden (Van Doorn & Duivestijn, 2016, S. 10). Auch soziale Medien bieten eine Anlaufstelle für Kunden, um mit Unternehmen in Kontakt zu treten (Savage, 2022). Auch auf diesen Kanälen oder in Messenger-Apps wie beispielsweise von Facebook oder WhatsApp nutzen Unternehmen zunehmend Chatbots. Diese Messenger-Apps können Unternehmen auch gezielt bei der Gestaltung von Chatbots mit der Hilfe sogenannter Bot-Builders unterstützen, z. B. Chatfuel, Botsify oder ChatterOn (Kruse Brandão & Wolfram, 2018, S. 283; Rese et al., 2020, S. 2). Dabei können Kunden einen Kommunikationskanal mit dem Unternehmen nutzen, den sie häufig aus ihrer alltäglichen Kommunikation mit Freunden kennen und müssen keine zusätzlichen Apps auf ihrem Handy installieren oder im Browser schreiben (Koumaras et al., 2018, S. 1). Häufig werden im Kundenservice textbasierte Chatbots eingesetzt. Dabei werden Avatare oder menschliche Profilbilder als Gestaltungsmerkmale genutzt, um einem Chatbot eine menschenähnliche Persönlichkeit zu verleihen (McTear, 2017, S. 43-44; Song & Shin, 2024, S. 442). Neben den textbasierten Chatbots wurden durch Smartphones auch Voicebots in Form von Sprachassistenten wie Siri, Alexa oder Google Assistant

für jeden zugänglich und können daher für den Kontakt mit Unternehmen auch verwendet werden.

Die Wurzeln von textbasierten Chatbots gehen in die 1960er Jahre in den medizinischen Bereich als Anwendungsgebiet zurück (Rese et al., 2020, S. 1). Weizenbaum (1966) entwickelte 1966 den Chatbot ELIZA, der einen Psychotherapeuten nachahmte und mit einem Menschen über geschriebene Sprache kommunizieren konnte. Chatbots konnten zu Beginn lediglich auf einfache, sich wiederholende Anfragen antworten, welche simplen Regeln folgten. Sie konnten beispielsweise die Frage nach dem heutigen Wetter beantworten (Gentsch, 2018, S. 87). Waren in ihren Funktionen allerdings stark eingeschränkt. Das erneute Interesse in den letzten Jahren auch im Bereich des Kundenservice ist auf Entwicklungen und Fortschritte von Technologien, im maschinellen Lernen sowie in künstlicher Intelligenz, in der Entwicklungen von Sprach- und Textverarbeitung sowie auf die Verfügbarkeit von Messenger-Apps für Smartphones (z. B. WhatsApp, Facebook Messenger) zurückzuführen (Gentsch, 2018, S. 87, 2019, S. 81; Rese et al., 2020, S. 1). Denn technologische Fortschritte im Rahmen der künstlichen Intelligenz haben den Weg für menschenähnliche Interaktionen geebnet (Adam et al., 2021, S. 1). Entwicklungen wie ChatGPT haben textbasierte Chatbots nochmals für jeden nutzbar gemacht. Dabei werden durch solche Chatbots sehr gute Ergebnisse geliefert und die Interaktion sowie die Antworten sind nur schwer bis hin zu unmöglich von denen eines Menschen zu unterscheiden. Chatbots sind mittlerweile auch von weniger technikaffinen Menschen sehr gut nutzbar (Gentsch, 2018, S. 44). Durch moderne Chatbots ist es möglich, Interaktionen so durchzuführen, dass diese einer Interaktion mit einem Menschen ähneln und Nutzer ohne explizite Offenlegung der nicht-menschlichen Identität glauben, mit einem Menschen zu schreiben (Hatwar et al., 2016, S. 85).

Anfänglich ging es hauptsächlich um das Chatten zwischen Menschen und Chatbot (Mauldin, 1994, S. 17). Mittlerweile werden Chatbots häufig im Kundenservice eingesetzt (Hoyer et al., 2020, S. 60). Dabei kommen sie auch in den unterschiedlichsten Branchen zum Einsatz, wie z. B. im Finanzwesen, im Gesundheitswesen, im Handel oder in der Telekommunikation. Dabei ist der Einsatz von Chatbots nicht auf den Business-to-Consumer-Bereich beschränkt. Auch im Business-to-Business-Bereich werden Chatbots eingesetzt (Chong et al., 2021, S. 3). Im Fokus dieser Arbeit steht allerdings der Business-to-Consumer Bereich. Dabei können Chatbots an vielen Stellen der Customer Journey genutzt werden und Aufgaben erledigen, die zuvor von Menschen durchgeführt wurden.

Chatbots können im Kundenservice eingesetzt werden, um Termine zu koordinieren, Informationen zu Produkten und Dienstleistungen bereitzustellen, Kundendaten zu ändern, Kundensupport anzubieten, Bestellungen und Buchungen entgegenzunehmen, Beratungen durchzuführen, Feedback und Beschwerden zu erfassen sowie einfache Verträge abzuschließen (Donath, 2024; Stucki et al., 2020, S. 6). Sie können somit entlang der gesamten Customer Journey einen Kontaktpunkt zum Unternehmen darstellen (Rese et al., 2020, S. 2). In der Vorkaufphase, in der potenzielle Käufer Informationen über ein Produkt oder eine Dienstleistung suchen, können Chatbots eine entscheidende Unterstützung bieten. Chatbots können als Verkaufsassistenten fungieren und Dienstleistungen oder Produkte empfehlen, wodurch sie die Entscheidungsfindung von Kunden unterstützen. In der Kaufphase kann der Chatbot ebenfalls unterstützen, oder sogar den gesamten Verkaufsprozess übernehmen. In der Nachkaufphase kann der Chatbot dann bei Kundenrückfragen unterstützen oder Reklamationen bearbeiten. Gerade zur Beantwortung von einfachen Routineanfragen wie beispielsweise Versand- oder Rückgabeformationen, der Paketverfolgung oder dem Zahlungsvorgang können Chatbots gut eingesetzt werden (Song & Shin, 2024, S. 442). Fortgeschrittene Chatbots können sogar personalisierte Erlebnisse bieten, indem sie sich an frühere Interaktionen und Präferenzen der Nutzer erinnern. Darüber hinaus können Chatbots auch proaktiv eingesetzt werden, um Gespräche zu initiieren, Upselling zu betreiben oder Produkte und Dienstleistungen anzubieten, mit denen der Nutzer noch nicht vertraut ist (Mucci, 2024).

Allgemein werden Chatbots somit im Kundenservice genutzt, um Kundenanfragen zu bearbeiten. Künstliche Intelligenz basierte Chatbots können dabei von Kundeninteraktionen lernen und sich verbessern. An dieser Stelle gilt es natürlich zu vermeiden, dass sich der Chatbot in eine Weise weiterentwickelt, die vom Unternehmen nicht gewünscht ist, wie es beim Chatbot Tay von Microsoft der Fall war. Dieser fing innerhalb weniger Stunden an, anzügliche, beleidigende oder rassistische Antworten zu verfassen, die aus der selbstlernenden Funktion entstanden sind (Gentsch, 2018, S. 105; Graff, 2013).

Es zeigt sich eine Vielzahl möglicher Einsatzmöglichkeiten im Kundenservice. Dabei kann der Einsatz von Chatbots im Kundenservice mit einigen Vorteilen für Unternehmen als auch für die Kunden verbunden sein. Kunden haben durch den Einsatz von Chatbots eine 24/7 Verfügbarkeit und können dadurch jederzeit den Kundenservice erreichen. In der Regel erreichen Chatbots schnelle Antwortzeiten mit (wenn überhaupt) kurzen Wartezeiten. Chatbots können auch individuell auf die Anfrage der Kunden eingehen und es müssen beispielsweise keine

FAQs durchsucht werden, um die richtige Antwort zu finden (Ferner & Lienbacher, 2023, S. 161; Sands et al., 2021, S. 1). Für Unternehmen haben Chatbots im Kundenservice eine Vielzahl an Vorteilen: Z. B. die Entlastung von Mitarbeitern, eine hohe Skalierbarkeit ohne Kapazitätsgrenzen, die Nutzung von Nutzerdaten für Erkenntnisse und Analysen des Kundenverhaltens sowie Imageeffekte (Wahrnehmung als fortschrittliches Unternehmen) (Ferner & Lienbacher, 2023, S. 161; Gentsch, 2018, S. 109). Gerade in Zeiten, in denen Unternehmen Kosten und Arbeitskräfte einsparen müssen oder Arbeitskräftemangel herrscht, können Chatbots gute Alternativen sein, da weniger Personal im Kundenservice benötigt wird und dies ergänzend zu Chatbots in reduzierten Zeiten eingesetzt werden kann. Chatbots können somit Mitarbeiter unterstützen und entlasten. Allerdings hat dies auch zur Folge, dass menschliche Servicemitarbeiter auf Webseiten, in sozialen Netzwerken oder Messenger-Apps zunehmend durch Chatbots ersetzt werden (Crollic et al., 2022, S. 132). Gerade der letzte Punkt wird von manchen Kunden kritisch gesehen, denn sie vermuten hinter dem Einsatz von Chatbots anstelle der menschlichen Mitarbeiter teilweise Kostenmotive. Dies kann in der Folge zu negative Kundenreaktionen führen, wie beispielsweise negativen Bewertungen (Castelo et al., 2023, S. 853). Kunden können auch Datenschutzbedenken haben und durch den Einsatz von Chatbots im Kundenservice ihre Privatsphäre gefährdet sehen (Gentsch, 2018, S. 109). Beim Einsatz von Chatbots im Kundenservice sollten Unternehmen sorgfältig vorgehen und die Funktionalität sowie den Datenschutz gewährleisten. Der Chatbot sollte auf den spezifischen Zweck der Anwendung im Kundenservice abgestimmt sein und einen echten Mehrwert für den Kunden bieten. Im Idealfall sollte zudem die Möglichkeit bestehen, jederzeit zu einem menschlichen Servicemitarbeiter wechseln zu können (Mucci, 2024).

2.2 Bewertungen

Durch die Abgabe von Bewertungen haben Kunden die Möglichkeit, ihre subjektiven Erfahrungen über Produkte oder Dienstleistungen zu teilen und dadurch anderen (potenziellen) Kunden durch ihre Erfahrungen zu helfen (Homburg, 2020, S. 1040). Solche Bewertungen basieren auf persönlichen Erfahrungen und Eindrücken der Kunden und können sowohl positiv als auch negativ sein. Bewertungen können zur Entscheidungsfindung anderer Kunden beitragen. Beispielsweise beim Kauf eines Produktes oder der Auswahl eines Dienstleisters können Bewertungen bei der Entscheidungsfindung einen Teil der Entscheidungsgrundlage darstellen. Gerade im Online-Kontext sind Bewertungen von besonderer Bedeutung, da (potenzielle) Kunden sich

nicht selbst von der Qualität der Dienstleistung oder der Produkte im Vorfeld überzeugen können. Sie haben keine Möglichkeit, die Produkte im Vorfeld zu betrachten und anzufassen. Daher verlassen sich Kunden in Online-Umgebungen häufig auf Bewertungen vorheriger Kunden. Auch für Unternehmen können positive Bewertungen sehr nützlich sein. Wenn Kunden die Bewertungen in ihre Entscheidungsfindung miteinbeziehen und diese positiv sind, wird die Kaufabsicht des Produktes oder der Dienstleistung höher sein. Bei negativen Bewertungen kann sich dieser positive Effekt allerdings schnell umdrehen. Daher sollte es eine hohe Priorität für Unternehmen haben, die Kunden zufriedenzustellen und zur Abgabe einer Bewertung zu motivieren. Generell sollten verärgerte Kunden und negative Kundenreaktionen vermieden werden. Unternehmen können Kunden aktiv zur Abgabe einer Bewertung auffordern und dadurch die Anzahl an Bewertungen erhöhen und somit eine bessere Entscheidungshilfe für neue Kunden liefern. Unternehmen sollten die Bewertungen auch reflektieren und gegebenenfalls ihre Dienstleistung oder ihr Produkt anpassen.

Bewertungen können in textbasierter oder numerischer Form auftreten. Hier wird häufig von „Review“ und „Rating“ gesprochen. Ein Review ist eine textbasierte Beurteilung eines Kunden, welche die Eigenschaften eines Produkts oder einer Dienstleistung beschreibt. Dabei können Kunden auf Vor- oder Nachteile eingehen und positive sowie negative Erfahrungen teilen. Sie sind bei der Formulierung eines Reviews frei und können dieses so gestalten, wie sie möchten. Reviews stellen somit offene von Nutzern erstellte Textnachrichten dar. Bei einem Rating verwenden Kunden eine numerische Skala, um ihre Bewertungen abzugeben (Floh et al., 2013, S. 647; Lackermair et al., 2013, S. 1; Mudambi & Schuff, 2010, S. 187). Ratings können dabei Skalen mit Zahlen oder Sternen verwenden. Aus den Ratings resultieren häufig Durchschnittsangaben, die den Kunden zur Verfügung gestellt werden. Oftmals werden zur Visualisierung Sternbewertungen von eins bis fünf Sterne verwendet, wobei eine fünf Sternbewertung sehr gut ist. Zusätzlich zur durchschnittlichen Bewertung geben Histogramme Informationen über die Streuung der Ratings (LaCour & Serra, 2022, S. 231; Mudambi & Schuff, 2010, S. 187). Ratings und Review können getrennt voneinander genutzt werden oder in Kombination auftreten (Floh et al., 2013, S. 647). Die Begrifflichkeiten Rating und Review werden manchmal nicht trennscharf verwendet und teilweise auch synonym genutzt (Lackermair et al., 2013, S. 1). Im Rahmen dieser Arbeit wird die angesprochene Differenzierung allerdings genutzt und die Begriffe „Rating“ und „Review“ verwendet.

Der Begriff WoM oder expliziter „electronic Word of Mouth“ (eWoM) wird im Kontext von Bewertungen häufig genannt und betrachtet. Dabei stellen Bewertungen einen Teil von eWoM dar (Floh et al., 2013, S. 646; Moore, 2015, S. 31; Nath et al., 2018, S. 123; Park et al., 2007, S. 127; Wei & Lu, 2013, S. 194; Wu et al., 2015, S. 739). WoM wird auf Deutsch als „Mundpropaganda“ bezeichnet und betrachtet die Weitergabe von Informationen, Meinungen oder Empfehlungen durch persönliche Gespräche zwischen Menschen. Diese Art der Kommunikation erfolgt meist auf informelle Weise und basiert auf persönlichen Erfahrungen oder Empfehlungen. In der Regel geschieht WoM von Privatperson zu Privatperson, wenn jemand eine positive, negative oder neutrale Erfahrung mit einem Produkt oder einer Dienstleistung teilt (Anderson, 1998, S. 6; Lis & Korchmar, 2013, S. 5-7). Positives WoM kann sich beispielsweise durch Empfehlungen oder das Erzählen angenehmer und neuer Erfahrungen mit einem Produkt oder einer Dienstleistung äußern. Negatives WoM hingegen kann in Form von Berichten über unangenehme Erfahrungen, private Beschwerden oder einer Verunglimpfung eines Produkts oder einer Dienstleistung bis hin zum Streuen von Gerüchten geschehen. WoM geschieht dabei häufig auf eine anschauliche Art und Weise (Anderson, 1998, S. 6).

Der Nutzen von WoM ergibt sich aus dessen Wirkung auf aktuelle und potenzielle Kunden (Ennew et al., 2000, S. 76). Dabei kann WoM sehr einflussreich sein, weil die Menschen oft eher auf die Meinungen von Freunden, Familienmitgliedern oder Bekannten hören als auf Werbung oder kommerzielle Angebote. Der Einfluss von WoM ist allerdings auch abhängig vom Produkt, dem Markt und dem organisationalen Kontext. So ist der Einfluss von WoM hoch für Produkte und Dienstleistungen, mit deren Kauf ein hohes Risiko verbunden ist und bei denen Erfahrungen und Glaubwürdigkeit besonders wichtig sind (Ennew et al., 2000, S. 75). Für Unternehmen ist WoM besonders wichtig, da sie das Vertrauen in Marken stärkt und eine Form der „echten“ Empfehlung darstellt, die oft als glaubwürdiger angesehen wird. Positive Mundpropaganda kann die Bekanntheit und den Erfolg eines Unternehmens erheblich steigern (See-To & Ho, 2014, S. 188)

Während WoM in der Regel mündlich und persönlich stattfindet, erfolgt eWoM über das Internet. Dabei werden digitale Kanäle wie beispielsweise Soziale Medien, online Bewertungen, Foren oder Blogbeiträge verwendet. eWoM kann dabei beispielsweise in Form der Veröffentlichung von Produkttests auf YouTube oder durch die Abgabe von Online-Bewertungen geschehen. Andere potenzielle Kunden können sich dann über das Produkt oder die Dienstleistungen informieren und in ihrer Entscheidung beeinflusst werden. Darüber hinaus können sie

auch selbst zu einem späteren Zeitpunkt potenzielle Sender von eWoM werden (Homburg, 2020, S. 92; Reichelt, 2013, S. 13). eWoM kann als Erweiterung des klassischen WoM angesehen werden. eWoM kann durch die verwendete Onlineumgebung eine viel größere Reichweite erreichen und ist auch zeitlich gesehen dauerhaft verfügbar (Cascio et al., 2015, S. 559; Lis & Korchmar, 2013, S. 1-2). Darüber hinaus wird eWoM häufig anonym abgegeben. Die Glaubwürdigkeit von eWoM kann im Vergleich zu WoM allerdings geringer sein. WoM wird häufig durch persönliche Kontakte ausgeführt, die in der Regel als sehr vertrauenswürdig gelten. Wohingegen im Rahmen von eWoM Skepsis herrschen kann, da die dahinterstehenden Bewertungen teilweise auch gekauft oder gefälscht sein können und kein persönlicher Kontakt zu den Verfassern von eWoM existiert.

3 Theoretische Bezugspunkte zu Mensch-Chatbot-Interaktionen

In der stark gewachsenen Chatbot-Literatur der letzten Jahre wurden eine Vielzahl theoretischer Erklärungsansätze genutzt. Mariani et al. (2023, S. 5) haben im Rahmen ihrer Literaturrecherche 95 unterschiedliche verwendete Theorien und Modelle beschrieben. Zu den häufig verwendeten Theorien zählen beispielsweise Technologieakzeptanzmodelle, die Social Response Theorie, Anthropomorphismus, der Uncanny Valley Effekt, die Media Equation Theorie und die Use and Gratification Theorie.

In diesem Kapitel werden die theoretische Erklärungsansätze betrachtet, die für das Thema dieser Arbeit wichtig sind und Grundlagen für die Forschungshypothesen aus Kapitel 5 bilden. Dies sind der Anthropomorphismus, die Social Response Theorie, der Uncanny Valley Effekt, die Attributionstheorie, die Familiarity-Heuristik sowie das Confirmation/Disconfirmation-Paradigma.

3.1 Anthropomorphismus

Anthropomorphismus beschreibt im Kern die Zuschreibung menschlicher Eigenschaften, Charakteristika, Verhaltensweisen oder mentaler Zustände auf nicht-menschliche Akteure. Der Begriff Anthropomorphismus setzt sich aus den griechischen Wörtern „anthropos“ (= Mensch) und „morphe“ (= Form oder Gestalt) zusammen (Epley et al., 2008b, S. 144; Epley et al., 2007, S. 864-865). Anthropomorphismus bedeutet, dass nicht-menschlichen Objekten ein Verstand, Absichten, die Fähigkeit zu denken, emotionale Zustände oder Verhaltensmerkmale zugeschrieben werden (Puzakova et al., 2013, S. 82). Die Wahrnehmung von menschlichen Merkmalen bei nicht-menschlichen Akteuren ist das Wesen des Anthropomorphismus. Die nicht-menschlichen Akteure können real oder imaginär sein. Sie müssen allerdings scheinbar unabhängig handeln. Beispiele dafür sind Tiere, Naturgewalten, Götter sowie technische und mechanische Dinge. Solche anthropomorphe Zuschreibungen beeinflussen die Verhaltensweise des Gegenübers (Epley et al., 2008b, S. 144). Anthropomorphismus von Chatbots wird als Ausmaß, in dem der Chatbot mit menschenähnlichen Eigenschaften wie beispielsweise einem Namen oder einem Avatar ausgestattet ist, beschrieben (Crollic et al., 2022, S. 132).

So ist es ein Grundbedürfnis von Menschen soziale Bindungen zu anderen Menschen aufzubauen und aufrechtzuerhalten. Mangelt es an solchen Bindungen, können nicht-menschliche Akteure zur Kompensation genutzt werden, indem sie ein Gefühl der menschlichen Verbundenheit hervorrufen. Anthropomorphismus kann Grundbedürfnisse von Menschen befriedigen, etwa das Bedürfnis nach sozialen Kontakten (Epley et al., 2008b, S. 144). Darüber hinaus kann Anthropomorphismus auch das Bedürfnis nach Kompetenzerfahrung, Verständnis, Kontrolle und Vorhersehbarkeit erfüllen (Epley et al., 2008a, S. 114; Epley et al., 2008b, S. 149; Epley et al., 2007, S. 866). Wenn nicht-menschlichen Akteuren menschliche Eigenschaften zugeschrieben werden, können Handlungen des Akteurs besser nachvollzogen, Unsicherheiten verringert und das Vertrauen erhöht werden (Epley et al., 2007, S. 866).

Das Ausmaß, in dem eine Person einen nicht-menschlichen Akteur anthropomorphisiert, hängt von drei Determinanten ab: Sociality, Effectance und Elicited Agent Knowledge (Epley et al., 2007, S. 866). Sociality drückt das Bedürfnis und den Wunsch aus, soziale Beziehungen mit anderen Menschen zu führen und aufrechtzuerhalten. Dabei kann Anthropomorphismus die Befriedigung dieses Bedürfnisses ermöglichen, indem die Beziehung zu einem nicht-menschlichen Akteur als menschlich wahrgenommen wird. In Bezug auf Anthropomorphismus beschreibt Effectance das Bedürfnis und die Motivation effektiv mit der eigenen Umwelt und insbesondere dem nicht-menschlichen Akteur zu interagieren. Die dritte Determinante ist Elicited Agent Knowledge und stellt einen kognitiven Faktor des nicht-menschlichen Akteurs dar. Dabei wird betrachtet, welches Wissen dem nicht-menschlichen Akteur zugeschrieben wird und inwieweit dieses gut zugänglich ist (Epley et al., 2007, S. 866). So kann Anthropomorphismus beispielsweise zunehmen, wenn Menschen das Gefühl haben, keine sozialen Bindungen zu anderen zu besitzen und abnehmen, wenn sie ein starkes Gefühl der sozialen Bindung erleben (Epley et al., 2007, S. 866).

In der Literatur wird zwischen Anthropomorphismus und dem Konzept der Sozialen Präsenz unterschieden (Ischen et al., 2020a, S. 619). Anthropomorphismus kann zu einer Wahrnehmung als sozial präsent führen (Adam et al., 2021, S. 438; Gnewuch et al., 2018, S. 11; Thomaz et al., 2020, S. 51). Dabei können anthropomorphe Hinweise von Technologien (z. B. Chatbots) wie ein Name, ein Profilbild, Emojis, die persönliche Anrede oder die informelle Sprache, zur Wahrnehmung als sozial präsent führen. Anthropomorphismus von Chatbots wird häufig mit der menschlichen Gestaltung dieses gleichgesetzt. Soziale Präsenz hingegen beschreibt die sub-

jektive Wahrnehmung der Nutzer (Konya-Baumbach et al., 2023, S. 3). Sie ist in der Kommunikationstheorie verankert und beschreibt den Grad der Bedeutung anderer Personen in der Interaktion und die daraus folgende Bedeutung der zwischenmenschlichen Beziehung. Sie wird auch als Ausmaß betrachtet, in dem ein Kommunikationsmedium es ermöglicht, den Gesprächspartner während der Interaktion als psychologisch anwesend wahrzunehmen (Short et al., 1976, S. 65). Das Kommunikationsmedium wird als „realer“ sozialer Akteur wahrgenommen (Ischen et al., 2020a, S. 619).

Soziale Präsenz beschreibt dabei das Ausmaß, in dem ein Mensch glaubt, dass eine Person tatsächlich anwesend ist. Anders ausgedrückt beschreibt Soziale Präsenz das Ausmaß, in dem Maschinen (beispielsweise Roboter oder Chatbots) Konsumenten das Gefühl geben, dass sie sich in Gesellschaft eines anderen sozialen Wesens befinden. Dabei haben sie das Gefühl der Gegenwart oder Nähe, dass in sozialen Interaktionen entsteht, auch wenn die Beteiligten physisch getrennt sind. Soziale Präsenz kann verwendet werden, um die Wahrnehmung von anderen Menschen oder Technologien und deren Persönlichkeit in einem digitalen Kontext zu beschreiben (Blut et al., 2021, S. 644; Heerink et al., 2008, S. 35, 2010, S. 79; Van Doorn & Duivestijn, 2016, S. 44).

3.2 Social Response Theorie und Media Equation Theorie

Die Social Response Theorie wird auch als Computers are Social Actors Paradigma (CASA-Paradigma) bezeichnet (Adam et al., 2021, S. 429; Moon, 2000, S. 325). Die Social Response Theorie stellt ein Konzept der Medienpsychologie dar. Es ist eng mit der Media-Equation-Theorie verbunden und beschreibt Wechselwirkungen zwischen Menschen und Medien, insbesondere Computern. Beide Theorien vertreten die Annahme, dass Menschen mit Medien, die menschlichen Eigenschaften aufweisen, ähnlich interagieren wie mit anderen Menschen (Nass & Moon, 2000; Reeves & Nass, 1996).

Die Media Equation Theorie (Reeves & Nass, 1996) betrachtet die Kommunikationsschnittstelle zwischen Menschen und einer Technologie. Menschen reagieren auf mediale Darstellungen natürlich, spontan und unreflektiert in Bezug auf ihre soziale Interaktion. Reeves und Nass (1996, S. 251-253) argumentieren, dass Nutzer emotionale und soziale Reaktionen auf Medien mit sozialen und menschlichen Merkmalen zeigen, ohne kognitive Anstrengungen zu unternehmen, um zu bestimmen, wie sie reagieren sollten. Die Media Equation Theorie betrachtet eher

die Reaktionen auf physische Merkmale traditioneller Medien, während sich die Social Response Theorie mit den Reaktionen von Nutzern auf soziale Merkmale von Computern und Software-Agenten beschäftigt (Lee, 2008, S. 1). Daher liegt der Fokus in dieser Arbeit auf der Social Response Theorie und die Media Equation Theorie wurde im Rahmen dieses Unterkapitels nur kurz beschrieben.

Menschen zeigen soziale und emotionale Reaktionen auf Technologien mit menschlichen Eigenschaften auf ähnliche Weise wie auf Menschen. Es wird davon ausgegangen, dass Menschen den Technologien (beispielsweise Computern, Robotern oder Chatbots) soziale Eigenschaften zuschreiben. Allgemein beschreibt die Social Response Theorie, dass Menschen ihre sozialen Regeln auch gegenüber Technologien anwenden (Nass & Moon, 2000, S. 83; Nass et al., 1994b, S. 77).

Die soziale Orientierung von Menschen ist stärker ausgeprägt, wenn sie mit einer Technologie interagieren, die menschenähnliche Eigenschaften aufweist, also anthropomorph gestaltet ist. In Situationen mit Technologien, die menschenähnliche Eigenschaften aufweisen, reagieren Menschen sozial. Das bedeutet, dass Nutzer soziale Regeln auf Chatbots anwenden und ähnliche Erwartungen an deren Verhalten haben wie gegenüber menschlichen Interaktionspartnern. Diese Reaktion ist reflexartig und erfolgt ohne großes Nachdenken oder Überlegen. Daher neigen Menschen dazu, Computer als soziale Akteure zu behandeln (Huang & Lin, 2011, S. 1965-1966; Moon, 2000, S. 325; Nass & Moon, 2000, S. 83).

Menschen wenden soziale Skripte der Mensch-Mensch-Interaktion gedankenlos auf Mensch-Computer-Interaktionen an (Nass & Moon, 2000, S. 83). Soziale Reaktionen sind unbewusst und automatisch. Sie stellen eine natürliche Reaktion in sozialen Situationen dar (Nass et al., 1994b, S. 77). Dabei können beispielsweise Interaktionen mit Technologien wie Chatbots solche Situationen repräsentieren.

Gedankenlosigkeit ist ein zentrales Konstrukt für menschliche Reaktionen auf Technologien mit menschlichen Eigenschaften. Gedankenloses Verhalten ist ein Resultat bewusster Aufmerksamkeit auf eine Untergruppe von kontextuellen Hinweisen. Diese Hinweise können soziale Skripte hinsichtlich des Verhaltens von Menschen auslösen und Aufmerksamkeit verschieben und folglich von anderen Informationen ablenken. Demnach greift der Mensch an-

stelle aktiver Unterscheidung anhand von relevanten Merkmalen der jeweiligen Situation gedankenlos auf übermäßig vereinfachte soziale Regeln zurück, die in der Vergangenheit durch Erfahrungen in ähnlichen Situationen entstanden sind (Nass & Moon, 2000, S. 82-83).

Es gilt allerdings anzumerken, dass die Social Response Theorie nicht auf alle Maschinen oder Technologien angewendet werden kann. Technologien müssen zwei Voraussetzungen erfüllen: Soziale Hinweise und Art beziehungsweise Herkunft der Technologie (Gambino et al., 2020, S. 72). Es müssen genügend soziale Hinweise vorliegen, damit eine Person dazu veranlasst wird, Verhaltensweisen der Mensch-Mensch-Interaktion auf die Technologie zu übertragen und anzuwenden (Gambino et al., 2020, S. 73; Nass & Moon, 2000, S. 83). Dieses Verhalten unterscheidet sich von Mensch zu Mensch und von Situation zu Situation stark (Gambino et al., 2020, S. 73). Die zweite Voraussetzung stellt die Art oder Herkunft der Technologie dar. Dabei ist die Wahrnehmung der Technologie als autonome Quelle entscheidend. Technologien, die im Kontext der Social Response Theorie als soziale Akteure wahrgenommen werden, dürfen nicht ausschließlich als Kanal der Kommunikation dienen. Sie müssen das Potenzial aufweisen als Quelle für soziale Interaktion wahrgenommen zu werden (Gambino et al., 2020, S. 73).

Nass et al. (1995) zeigen, dass Menschen auf die Persönlichkeiten von Computern ähnlich reagieren wie auf Menschen. Nass et al. (1994a) zeigen, dass Menschen soziale Regeln anwenden, wenn sie das Verhalten von Maschinen bewerten. Die Social-Response-Theorie wurde in verschiedenen Studien als theoretische Grundlage verwendet, um zu zeigen, wie Menschen soziale Regeln auf anthropomorph gestaltete Computer anwenden (Adam et al., 2021, S. 429).

3.3 Uncanny Valley Effekt

Menschen sind mit Technologien wie Robotern, Chatbots oder künstlichen Figuren vertrauter, je menschlicher diese gestaltet sind. Es gibt jedoch einen bestimmten Bereich, in dem die positive Einstellung umschlägt und Menschen die Technologie als seltsam, gruselig oder unheimlich empfinden (Mori, 1970, S. 33, Rajaobelina et al., 2021, S. 2340-2341). In diesem Bereich nehmen die Vertrautheit und die Akzeptanz abrupt ab. Unnatürliche Bewegungen oder Gesichtszüge bei Robotern können beispielsweise zu einem solchen Einbruch führen. Die Technologien sehen in diesem Bereich einem Menschen sehr ähnlich, sind allerdings noch nicht perfekt menschlich. Dieser Bereich wird als „Uncanny Valley“ bezeichnet. Sobald die Gestaltung der Technologie noch menschlicher wird, steigt die Vertrautheit sowie die Akzeptanz

nochmals an (Mori, 1970, S. 33-34; Mori et al., 2012, S. 98-99). Das Uncanny Valley beschreibt ein Phänomen, bei dem Objekte oder Technologien, die dem menschlichen Erscheinungsbild sehr nahekommen, jedoch nicht vollkommen menschlich wirken, bei Nutzern Unbehagen oder ablehnende Reaktionen hervorrufen. Dabei kann ein Gefühl des Unheimlichen oder des Unbehagens auftreten. In der Folge kann dies zur Ablehnung der Technologie führen (Blut et al., 2021, S. 635). Der Uncanny Valley Effekt ist in Abbildung 2 dargestellt.

Ursprünglich wurde der Effekt im Rahmen von Robotern beobachtet und beschrieben (Mori, 1970, S. 33). Der Uncanny Valley Effekt wird auch für andere Technologien wie beispielsweise Chatbots beschrieben (Ciechanowski et al., 2019; Skjuve et al., 2019; Song & Shin, 2024). Mori (1970, S. 33) unterscheidet zwischen bewegten und unbewegten Robotern. Folgt man Abbildung 2, ist der Einbruch bei unbewegten Dingen wie beispielsweise Chatbots nicht so extrem wie bei Robotern, die menschlich aussehen und sich auch menschenähnlich bewegen.

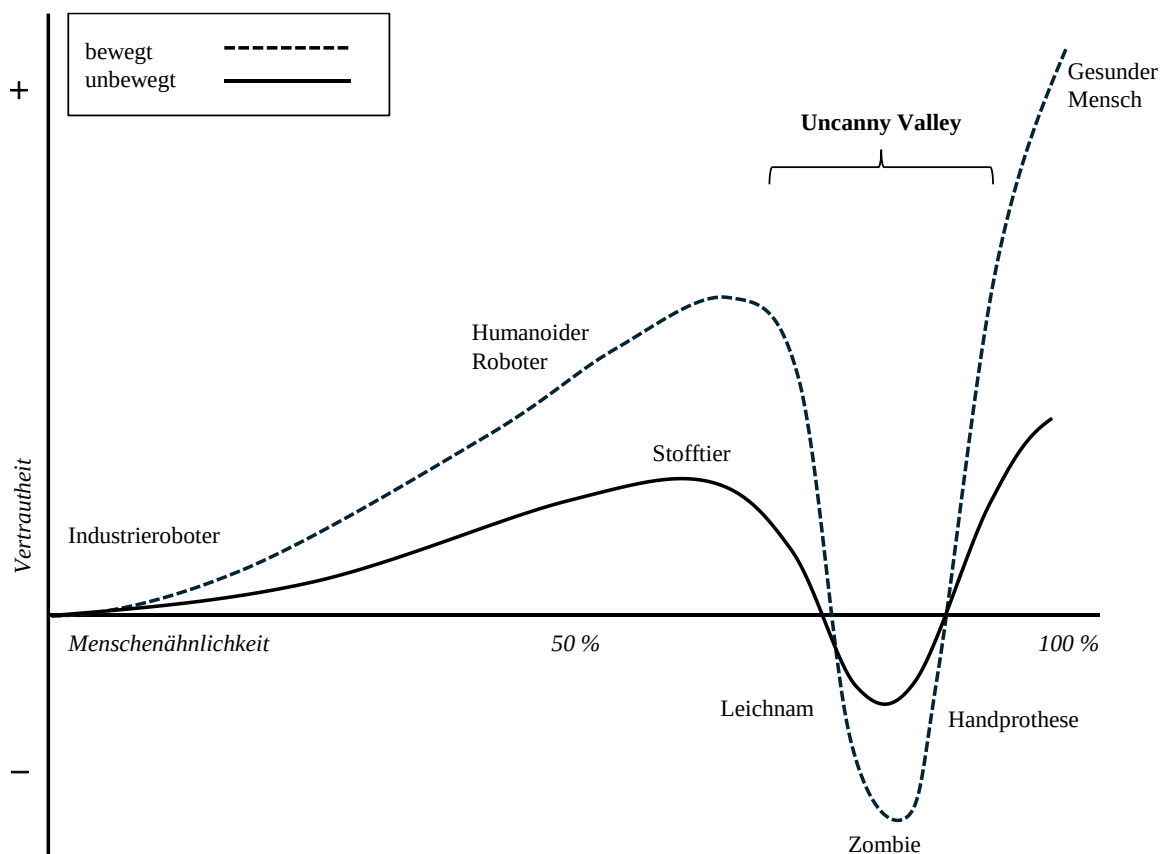


Abbildung 2: Uncanny Valley Effekt (Mori, 1970, S. 33)

Der Uncanny Valley Effekt wird in der Forschung verwendet, um zu erklären, warum es bei Mensch-Maschinen-Interaktionen zu negativen Reaktionen kommt, wenn der Roboter oder Chatbot zu menschlich ist (Mende et al., 2019; Song & Shin, 2024).

3.4 Attributionstheorie

Nach der Attributionstheorie sind Menschen rationale Informationsverarbeiter und ihre kausalen Schlussfolgerungen beeinflussen ihr Handeln. Sie beschreibt, wie Menschen Informationen nutzen, um Verhaltensweisen kausal zu erklären (Folkes, 1984, S. 398). Dabei versucht der Mensch, Handlungen oder Ergebnisse zu erklären und Rückschlüsse auf zugrundeliegende Motive zu ziehen (Folkes, 1988, S. 548; Homburg & Koschate, 2005, S. 403; Kelley, 1973, S. 107). Die Attributionstheorie geht auf Heider (1958) zurück und besteht aus mehreren Theorien mit gemeinsamen Grundannahmen (Folkes, 1988, S. 548; Kelley & Michela, 1980, S. 458; Weiner, 1994, S. 220). Attributionstheorien beschäftigen sich mit der Wahrnehmung von Kausalität, also den wahrgenommenen Ursachen für das Eintreten eines bestimmten Ereignisses (Weiner, 1994, S. 220). Heider (1958, S. 82) stellte fest, dass das Ergebnis einer Handlung sowohl von individuellen Faktoren einer Person als auch von Faktoren der Umwelt abhängt. Die kognitiven Prozesse sind der Kern der Attributionstheorien. Der Mensch schreibt seinem eigenen Verhalten oder dem Verhalten anderer Menschen Ursachen zu. Die Attributionstheorie beschreibt, wie Menschen dafür kausale Erklärungen abgeben. Vereinfacht ausgedrückt beschäftigt sie sich damit, wie Menschen Fragen beantworten, die mit „warum“ beginnen (Kelley, 1973, S. 107). Dabei ist das Urteil von Menschen über Verhaltensweisen unterschiedlich, je nachdem welche Ursachen sie dem Verhalten zuschreiben (Homburg & Koschate, 2005, S. 404). Negative oder unerwartete Ereignisse können das Attributionsdenken fördern und Konsumenten eher dazu bewegen kausale Erklärungsansätze zu finden (Kanazawa, 1992, S. 659; Weiner, 1985a, S. 77).

In der Attributionstheorie haben sich die folgenden drei Forschungsrichtungen entwickelt. (1) Die Spezifikation der wahrgenommenen Verhaltensursachen, wobei hier zwischen internaler (in der Person lokalisiert) und externaler (in der Umwelt lokalisiert) Kausalität unterschieden wird. (2) Die Entwicklung allgemeiner Gesetzmäßigkeiten zu Kausalinferenzen, die sich aus der Beziehung der Ursacheninformation und dem kognitiven Verhaltens ergeben. (3) Die Erforschung der Auswirkungen von Kausalattributionen auf unterschiedliche Hinweise durch beobachtbares Verhalten (Weiner, 1994, S. 220-221).

Die Art der Ursachenzuschreibung (internal oder external) führt dabei zu unterschiedlichen Verhaltenskonsequenzen. Sie können unterschiedliche Gefühle, Erwartungen oder Handlungen zur Folge haben (Weiner, 1994, S. 223). Dies wurde in einem Teil der Attributionsforschung untersucht. Eine weitere Frage, mit der sich Attributionsforscher beschäftigt haben, ist die Frage: „Wie kommen Kausalattributionen zustanden?“. Hier geht es um die Beantwortung von Fragen nach der Informationsverwendung. Also welche Informationen werden herangezogen oder welche Prozesse und Strukturen werden aktiviert, um eine Erklärung für das Auftreten eines bestimmten Ereignisses zu erlangen (Weiner, 1994, S. 226). Dabei wird die Bedingung für die Wirkung verantwortlich gemacht, die vorhanden ist, wenn die Wirkung auftritt, aber fehlt, wenn die Wirkung ausbleibt (Heider, 1958, S. 152).

In der Literatur wird zwischen drei Kausaldimensionen unterschieden: Lokation, Stabilität und Kontrollierbarkeit (Folkes, 1984, S. 399; Weiner, 1985b, S. 549). Die Lokation beschreibt die Ursache des Ereignisses genauer. Dabei wird betrachtet, ob der Grund für ein eingetretene Ereignis in der Person selbst liegt oder äußeren Umständen geschuldet ist (Weiner, 1985b, S. 551). Also der Ort der Ursache ist hier ausschlaggebend für die Attribution. Bei fehlerhaften Produkten stellt sich beispielsweise die Frage, ob die Fehlerursache etwas mit dem Kunden, dem Vertrieb oder der Produktion zu tun hat. So kann das Versagen eines Bücherregals z. B. durch eine fehlerhafte Montage des Käufers oder eine mangelhafte Herstellung verursacht worden sein (Folkes, 1984, S. 399). Die (zeitliche) Stabilität betrachtet die Zeitkomponente der Ursache, ob diese vorübergehend, schwankend oder dauerhaft ist (Folkes, 1988, S. 556). Kontrollierbarkeit beschreibt die mögliche Kontrolle einer Ursache. Hier wird unterschieden, ob die Ursache willentlich oder unwillentlich herbeigeführt wurde. Also ob die Ursache willentlich durch eine Person hervorgerufen wurde oder ob ein bestimmtes Ereignis vermeidbar gewesen wäre (Folkes, 1984, S. 399; Weiner, 1985b, S. 551).

3.5 Familiarity-Heuristik

Menschen nutzen eine begrenzte Anzahl an heuristischen Prinzipien. Heuristische Prinzipien werden in komplexen Situationen oder bei komplexen Aufgaben verwendet, um ein einfaches Urteil fällen zu können. Häufig sind Heuristiken hilfreich, können aber auch zu Fehlern führen. Entscheider müssen nicht alle verfügbaren Informationen, Werte oder Wahrscheinlichkeiten

berücksichtigen, um eine Entscheidung zu treffen. Mithilfe von Heuristiken können sie komplexe Entscheidungssituationen vereinfachen und so leichter zu einem Entschluss gelangen (Tversky & Kahneman, 1974, S. 1124).

Vertrautheit kann beispielsweise für Heuristiken eine Grundlage bilden. Die Familiarity-Heuristik ist ein psychologisches Konzept, das beschreibt, wie Individuen Entscheidungen treffen (beispielsweise Dinge als sicherer, wahrscheinlicher oder besser bewerten), basierend auf der Vertrautheit mit einer spezifischen Situation oder einem Objekt (Chira et al., 2008, S. 13; Giacalone & Jaeger, 2016, S. 121; Johnson & Russo, 1984, S. 542). Vertrautheit ist dabei das Wissen über ein Produkt oder einen Service, das auf der Anzahl an Erfahrungen und vorherigem Kontakt beruht (Alba & Hutchinson, 1987, S. 411; Chen et al., 2011, S. 2143). Nach der Familiarity-Heuristik werden Schlussfolgerungen von Entscheidungsträgern auf Grundlage der Vertrautheit getroffen. So hat die Vertrautheit beispielsweise einen Einfluss auf die Informationsgewinnung, Reaktionen auf Werbung oder die Wahl der Entscheidungsregel, die in einer Situation angewendet wird (Chira et al., 2008, S. 13; Giacalone & Jaeger, 2016, S. 121; Johnson & Russo, 1984, S. 542).

In einer Entscheidungssituation, in der ein Kunde beispielsweise wählen muss, welchen Kommunikationskanal (z. B. E-Mail, Telefon, persönlicher Kontakt, Online-Chat oder Chatbot) er nutzt, kann die Wahl auf Basis früherer Erfahrungen und der damit verbundenen Vertrautheit getroffen werden. So wird er voraussichtlich eher den vertrauten Kanal wählen, wenn er mit diesem schon im Vorfeld gute Erfahrungen gesammelt hat. Auf Grundlage früherer Erfahrungen und vorhandenen Wissens sind Menschen in der Lage, zeitnah fundierte Entscheidungen zu treffen. Hierbei können auch Fehlentscheidungen getroffen werden, da nicht objektiv beurteilt wird und wichtige Faktoren gegebenenfalls keine ausreichende Berücksichtigung finden (Giacalone & Jaeger, 2016, S. 121).

Die Anwendung der Heuristiken basiert auf der Annahme, dass unser Verhalten dann als angemessen gilt, wenn es sich in der Vergangenheit bereits als erfolgreich erwiesen hat. Die Familiarity-Heuristik betrachtet eine Art mentale Gewohnheit, bei der frühere Handlungen ein Beweis dafür sind, dass ein bestimmtes Verhalten angemessen ist (McCammon, 2002, S. 246). Menschen hinterfragen ihre Gewohnheiten oft nicht und verhalten sich stattdessen automatisch gemäß diesen eingeübten Handlungsmustern. Vertraute Kunden verfügen über ein sehr gutes Wissen über die ihnen bekannten Produkte, daher werden sie Alternativen weniger Beachtung

schenken. „Zudem ermöglicht es ihnen dieser Prozess, ihre Aufmerksamkeit auf relevante Informationen zu lenken und gleichzeitig irrelevante oder redundante Informationen zu ignorieren. Dies erleichtert häufig die Verarbeitung neuer Informationen (Johnson & Russo, 1984, S. 543).

3.6 Confirmation/Disconfirmation-Paradigma

Das Confirmation/Disconfirmation-Paradigma ist ein Konzept aus der Psychologie und Marketing-Forschung, das im Zusammenhang mit der Entstehung von Kundenzufriedenheit häufig genutzt wird (Homburg, 2020, S. 46). Das Confirmation/Disconfirmation-Paradigma beschreibt, wie Menschen auf die Übereinstimmung mit oder Abweichung von ihren Erwartungen reagieren. Es wird vor allem verwendet, um die Zufriedenheit von Konsumenten mit Produkten oder Dienstleistungen zu erklären. Dabei beschreibt es im Kern, wie sich Zufriedenheit aus dem Vergleich der tatsächlichen Erfahrung bei Nutzung einer Leistung (Ist-Leistung) mit einem bestimmten Vergleichsstandard (Soll-Leistung) ergibt. Dieser Vergleichsstandard kann beispielsweise durch vorherige eigene Erfahrungen oder Erfahrungen anderer Nutzer und deren Bewertungen beeinflusst werden. Zufriedenheit hängt also von der Diskrepanz zwischen den Erwartungen und der tatsächlichen Erfahrung ab (Homburg, 2020, S. 46).

Dieser Vergleichsprozess kann unterschiedlich ausfallen und es werden drei mögliche Ergebnisse unterschieden. Wenn die wahrgenommene Ist-Leistung der Soll-Leistung entspricht, stellt dies eine Bestätigung dar und man bezeichnet diesen Zustand als Konfirmation. Hier entsteht Zufriedenheit auf Konfirmationsniveau. Übersteigt die Ist-Leistung die Soll-Leistung führt dies zu einer hohen Zufriedenheit, die über dem Konfirmationsniveau liegt. Dies bezeichnet man als positive Diskonfirmation. Ist hingegen die Ist-Leistung schlechter als die Soll-Leistung entsteht Unzufriedenheit. Die Zufriedenheit liegt dabei unter dem Konfirmationsniveau und man bezeichnet diesen Zustand als negative Diskonfirmation (Homburg, 2020, S. 46). Werden die Erwartungen folglich nicht erfüllt, ist der Kunde unzufrieden. Entspricht das Ergebnis den Erwartungen, zeigt sich der Kunde zufrieden. Übertrifft das Ergebnis die Erwartungen, führt dies zu hoher Zufriedenheit (Kotler et al., 2017, S. 169). Das Modell hilft zu verstehen, warum Menschen mit bestimmten Produkten oder Dienstleistungen zufrieden oder unzufrieden sind. Ein wichtiger Aspekt dabei ist, dass die Erwartungen und die Wahrnehmung der Konsumenten eine entscheidende Rolle bei der Bewertung ihrer Erlebnisse spielen. Wenn die Erwartungen zu hoch

oder zu niedrig sind, wird dies die Zufriedenheit stark beeinflussen. Das Confirmation/Disconfirmation-Paradigma stellt also ein Rahmenwerk dar, dass die Entstehung von Zufriedenheit beschreibt.

4 Empirische Evidenz als Grundlage der eigenen Forschung

Die wissenschaftliche Literatur zu anthropomorphen Chatbots hat sich in den letzten fünf Jahren erheblich vergrößert. Unter den Stichwörtern „anthropomorphism, chatbots“ liefert die Suche bei GoogleScholar 2025 siebenmal mehr Treffer als 2019. Diese Zunahme verdeutlicht auch die hohe Dynamik in diesem Forschungsbereich sowie die Wichtigkeit von anthropomorphen Chatbots. Die Literaturrecherche zeigte jedoch, dass viele der gefundenen Beiträge nur bedingt Relevanz für das Thema der vorliegenden Arbeit aufweisen. Die relevanten Ergebnisse werden im Folgenden vorgestellt. In der Literatur wurden unterschiedliche theoretische Erklärungsansätze herangezogen und eine Vielzahl abhängiger und unabhängiger Variablen sowie Mediator- und Moderatorvariablen analysiert.

Diese Arbeit trägt zum aktuellen Wissenstand bei, indem Erkenntnisse zu anthropomorphen Chatbots, zum Offenlegungszeitpunkt von Chatbots, der Übergabe einer Chatbot-Mensch-Interaktion an einen Menschen und Bewertungen von Chatbots untersucht werden. Nachfolgend erfolgt eine Zusammenstellung der bisherigen empirischen Befunde zu den genannten Themenbereichen. Die Forschung zum Anthropomorphismus von Chatbots befasst sich einerseits mit den Möglichkeiten der anthropomorphen Gestaltung und andererseits mit den Auswirkungen solcher Gestaltungsformen auf Nutzer. Zuerst werden in Kapitel 4.1 die anthropomorphen Gestaltungsmöglichkeiten betrachtet. Auf deren Grundlage wurden die anthropomorphen Chatbot Designs für die in dieser Arbeit durchgeführten Experimente erstellt. Anschließend werden Auswirkungen anthropomorpher Chatbots in Kapitel 4.1 betrachtet. In Kapitel 4.2 wird der Forschungsstand zur Offenlegung beziehungsweise dem Offenlegungszeitpunkt von Chatbots dargelegt. Dieses Kapitel liefert damit eine Grundlage für *FP II*. Kapitel 4.3 betrachtet bisherige Forschung zur Übergabe einer Chatinteraktion an einen Menschen. Dieses Unterkapitel dient als Grundlage für *FP III*. Kapitel 4.4 setzt sich mit Forschung zu Nutzerbewertungen und deren Auswirkungen auseinander. Hier wurden im Rahmen der Literaturrecherche keine Forschungsarbeiten gefunden, die sich explizit mit Bewertungen im Kontext von Chatbots auseinandersetzen. Daher werden in diesem Unterkapitel bisherige Ergebnisse der allgemeinen Forschung zu Bewertungen dargelegt. Dieses Unterkapitel dient als Grundlage für *FP IV*.

4.1 Anthropomorphe Gestaltung von Chatbots

Gestaltungsmerkmale anthropomorpher Chatbots

Technologien wie Chatbots sind häufig vermenschlicht (Schmitt, 2019, S. 828). Dabei werden unterschiedliche Gestaltungselemente verwendet, um Chatbots menschenähnlich beziehungsweise anthropomorph darzustellen. Blut et al. (2021, S. 637) unterscheiden zwischen physischen und nicht physischen Eigenschaften. Physische Eigenschaften sind dabei beispielsweise das Erscheinungsbild oder die „Verkörperung“ des Chatbots. Hier ist es ziemlich offensichtlich, dass diese Eigenschaften das Ausmaß an Anthropomorphismus beeinflussen. Nicht physische Eigenschaften beziehen sich auf die Verhaltensweisen des Chatbots, wie beispielsweise die Art und Weise der Kommunikation, die Art der Ansprache oder die verwendete Sprache (Blut et al., 2021, S. 637). Neben den beiden genannten Chatbot-Eigenschaften können auch individuelle Nutzerfaktoren wie Bildungsniveau, Alter, Einkommen oder Geschlecht einen Einfluss auf die anthropomorphe Wahrnehmung haben. Auch die Kompetenz von Nutzern, deren vorherige Erfahrungen, das Bedürfnisse nach Interaktionen, Angst vor Computern oder eine allgemeine negative Einstellung zu Technologien wie Chatbots oder Robotern können einen Einfluss auf Anthropomorphismus haben (Blut et al., 2021, S. 635-637).

Menschliche Merkmale von Robotern oder Chatbots dienen als Anhaltspunkte für Menschenähnlichkeit beziehungsweise Anthropomorphismus. Je mehr menschliche, also anthropomorphe Merkmale vorhanden sind, desto stärker wird der Chatbot auch anthropomorphisiert. Bei Robotern erhöht die Gestaltung des Kopfes, des Gesichts sowie des Körpers die wahrgenommene Menschenähnlichkeit und den Anthropomorphismus (Blut et al., 2021, S. 637). Übertragen auf Chatbots kann die Gestaltung des Profilbilds als parallele gesehen werden. Dabei können Chatbots unterschiedlich gestaltet werden und unterschiedlich hohe Anthropomorphismus-Level haben. Ein menschliches Profilbild im Vergleich zu einem roboterähnlichen Profilbild hat ein höheres Level an Anthropomorphismus. Neben dem Profilbild kann auch der Name des Chatbots vergeben werden, um anthropomorphe Wahrnehmung zu erhöhen. Ein menschlicher Name stellt ein hohes Level an Anthropomorphismus dar (Araujo, 2018, S. 187-188; Go & Sundar, 2019, S. 304-305; Janson, 2023, S. 10; Morana et al., 2020, S. 7; Park et al., 2023, S. 358; Song & Shin, 2024, S. 443; Sun et al., 2024, S. 3; Yuan et al., 2025, S. 3). Bei der Verwendung eines Namens kann es erforderlich sein, diesen besonders sorgfältig auszuwählen.

So kann es sinnvoll sein, einen geschlechtsneutralen Namen zu verwenden, da ansonsten stereotypische Vorurteile auftreten können (Adam et al., 2021, S. 434; McDonnell & Baxter, 2019, S. 119-120; Nass et al., 1997, S. 874).

Neben dem Namen kann auch die Verwendung von Pronomen in der ersten Person Singular durch den Chatbot zu einer anthropomorphen Wahrnehmung des Chatbots führen. Denn ein solches Verhalten wird ansonsten nur von Menschen gezeigt (Adam et al., 2021, S. 433). Weitere anthropomorphe Faktoren sind die Herzlichkeit und die Kompetenz des Chatbots (Cheng et al., 2022b, S. 5; Davis et al., 2023, S. 49; Li et al., 2023a, S. 1377; Li et al., 2023c, S. 2; Roy & Naidoo, 2021, S. 24). So kann ein Chatbot, der als sehr herzlich und kompetent wahrgenommen wird, anthropomorpher wirken. Darüber hinaus sollte ein anthropomorpher Chatbot den Nutzer begrüßen und verabschieden (Al Farisi et al., 2022, S. 371-372; Cai et al., 2022, S. 231; Morana et al., 2020, S. 7; Seeger et al., 2021, S. 937). Ein allgemein höfliches Auftreten sowie das Erinnern an den Namen des Nutzers haben einen positiven Einfluss auf Anthropomorphismus (Cai et al., 2022, S. 231; Morana et al., 2020, S. 7).

Allgemein hat die Verwendung einer menschenähnlichen Sprache einen Einfluss auf Anthropomorphismus (Al Farisi et al., 2022, S. 371-372; Araujo, 2018, S. 187-188; Diederich et al., 2020, S. 201). Daneben kann die Verwendung von Smalltalk und Empathie zu einem höheren Level an Anthropomorphismus führen, da dies als menschlicher und lebensechter wahrgenommen wird (Adam et al., 2021, S. 433; Hu & Sun, 2023, S. 3). Auch die Verwendung von Emojis wird zu den anthropomorphen Gestaltungsmerkmalen von Chatbots gezählt (Al Farisi et al., 2022, S. 371-372; Nguyen et al., 2023, S. 2202).

Darüber hinaus kann die Antwortzeit einen Einfluss haben. So wird eine direkte Antwort weniger anthropomorph wahrgenommen als eine eher dynamische. Dabei beschreibt dynamisch, dass sich der Chatbot bei der Antwort kurz Zeit lässt und damit eine kurze Verzögerung auftritt. Dadurch wirkt es, als würde ein Mensch die Anfrage lesen und anschließend erst antworten (Cheng et al., 2022b, S. 7; Gnewuch et al., 2018, S. 3; Toader et al., 2020, S. 10). Bei der Verwendung der dynamischen Antwort kann auch gekennzeichnet werden, dass der Chatbot gerade schreibt. Dies ist durch Nutzung von drei Punkten oder der Anzeige „schreibt gerade“ möglich (Al Farisi et al., 2022, S. 371-372).

Häufig wird eine Mischung dieser Einflussfaktoren verwendet, um einen Chatbot anthropomorph zu gestalten. Es ist nicht abschätzbar, welcher Faktor welchen Einfluss auf Anthropomorphismus hat und welche genau Wechselwirkungen haben. Es wird angenommen, je mehr der zuvor beschriebenen Faktoren genutzt werden, desto anthropomorpher wird der Chatbot wahrgenommen. Tabelle 1 zeigt übersichtlich die im Rahmen dieser Arbeit aufgezeigten anthropomorphen Gestaltungsmöglichkeiten von Chatbots.

Tabelle 1: Gestaltungsmerkmale anthropomorpher Chatbots

Gestaltungsmerkmal		Art	Quelle
Physisch	Profilbild	kein Bild, Avatar, menschliches Bild	Blut et al., 2021; Crolic et al., 2022; Diederich et al., 2020; Go & Sundar, 2019; Gong, 2008; Janson, 2023; Morana et al., 2020; Park et al., 2023; Song & Shin, 2024; Sun et al., 2024; Yuan et al., 2025
	Demographie	Name	Araujo, 2018; Crolic et al., 2022; Diederich et al., 2020
		Geschlecht	Adam et al., 2021; Diederich et al., 2020; McDonnell & Baxter, 2019; Nass et al., 1997
Nicht Physisch	Kommunikation/ Sprache	Empathie	Adam et al., 2021; Diederich et al., 2020; Hu & Sun, 2023
		Personalpronomen	Adam et al., 2021
		Begrüßen und Verabschieden	Al Farisi et al., 2022; Cai et al., 2022; Morana et al., 2020; Seeger et al., 2021
		Smalltalk	Adam et al., 2021
		Verzögerte Antwortzeit	Al Farisi et al., 2022; Cheng et al., 2022b; Gnewuch et al., 2018; Toader et al., 2020
		Höfliches Auftreten	Diederich et al., 2020; Morana et al., 2020
		Emojis	Al Farisi et al., 2022; Nguyen et al., 2023
		Herzlichkeit	Cheng et al., 2022b; Davis et al., 2023; Li et al., 2023a; Li et al., 2023c; Roy & Naidoo, 2021
		Kompetenz	Cheng et al., 2022b; Li et al., 2023a; Li et al., 2023c; Roy & Naidoo, 2021
		Erinnern des Nutzernamens	Cai et al., 2022; Morana et al., 2020

Auswirkungen anthropomorpher Chatbots

Neben Studien, die sich mit Einflussfaktoren auf Anthropomorphismus beschäftigen, gibt es auch Studien zu Auswirkungen von Anthropomorphismus oder anthropomorphen Gestaltungsmerkmalen. Deren Ergebnisse werden im Folgenden beschrieben.

Anthropomorphismus kann direkt oder indirekt über Mediatoren zu positiven Nutzerreaktionen führen. Hier findet sich in der Literatur Forschung zu Anthropomorphismus allgemein sowie einzelnen anthropomorphen Gestaltungsmerkmalen. Im Folgenden werden sowohl die allgemeinen Untersuchungen als auch die Untersuchungen zu den spezifischen Gestaltungsmerkmalen betrachtet.

Es wird ein besonderer Fokus auf die für diese Arbeit relevanten Konstrukte (Soziale Präsenz, Vertrauen und Wiederverwendungsabsicht) gelegt. Weitere interessante Ergebnisse werden auch kurz betrachtet. Dazu zählen beispielsweise die wahrgenommene Kompetenz sowie die Authentizität, die die Beziehung zwischen Anthropomorphismus und Kundenbindung medieren (Nguyen et al., 2023, S. 2213). Des Weiteren hat Anthropomorphismus einen positiven Einfluss auf die wahrgenommene Freude bei der Nutzung des Chatbots (Han, 2021, S. 56; Klein & Martinez, 2023, S. 2809). Dadurch kann eine höhere Kaufabsicht erreicht werden (Han, 2021, S. 56). Weiterhin kann Anthropomorphismus die Nützlichkeit eines Chatbots erhöhen (Epley et al., 2007, S. 879). Anthropomorphismus kann die Beziehung zu dem Chatbot verbessern und dadurch zu einem besseren WoM-Verhalten der Nutzer führen (Kayeser Fatima et al., 2024, S. 6). Anthropomorphismus kann das WoM-Verhalten auch direkt beeinflussen und verbessern (Konya-Baumbach et al., 2023, S. 8). Privatsphärebedenken gegenüber Chatbots können durch Anthropomorphismus reduziert werden (Ischen et al., 2020b, S. 45).

Je stärker die anthropomorphen Gestaltungsmerkmale des Profilbilds von Chatbots zunehmen, desto eher reagieren Menschen mit sozialen Reaktionen auf den Chatbot. Nutzer schreiben ihnen ein positiveres, soziales Urteilsvermögen und eine größere Homophilie zu. Sie bewerten die Kompetenz sowie die Vertrauenswürdigkeit höher. Anthropomorphismus führt zu einem größeren sozialen Einfluss (Gong, 2008, S. 1505). Des Weiteren kann Empathie von Chatbots die wahrgenommene Kompetenz sowie die Herzlichkeit beeinflussen, was sich beides auf die Servicequalität und die Zufriedenheit auswirkt (Han et al., 2022, S. 5-6). Liu et al. (2023b, S. 7) sowie Hu & Sun (2023, S. 11) beschreiben einen direkten positiven Effekt der Empathie auf die Zufriedenheit. Neben der Zufriedenheit kann durch Herzlichkeit und Kompetenz auch die

Wiederverwendungsabsicht von Chatbots erhöht werden. Dabei hat Empathie einen Einfluss auf die Herzlichkeit (Haupt et al., 2023, S. 16). Li et al. (2023a, S. 1386) bestätigen einen positiven direkten Effekt der Kompetenz sowie der Herzlichkeit auf die weitere Nutzungsabsicht. Die wahrgenommene Nützlichkeit mediiert diese Beziehung (Li et al., 2023a, S. 1386). Eine anthropomorphe Sprache, wie die Verwendung von Emojis, erhöht die wahrgenommene Kompetenz des Chatbots (Nguyen et al., 2023, S. 2214). Die Zufriedenheit der Nutzer kann von Anthropomorphismus positiv beeinflusst werden (Chang-Chien et al., 2022, S. 7; Chen et al., 2024, S. 3646-3647; Hsiao & Chen, 2022, S. 940; Konya-Baumbach et al., 2023, S. 8). Das Geschlecht des Chatbots kann ebenfalls einen Einfluss auf die Zufriedenheit mit dem Chatbot haben und Geschlechtsstereotypen werden angewandt (McDonnell & Baxter, 2019, S. 119).

Vertrauen ist ein wichtiger Faktor für den Erfolg von Technologien wie Chatbots und einer Kommunikation (Yen & Chiang, 2021, S. 1177). Ein höheres Vertrauen in den Chatbot kann zu einer größeren Wiederverwendungsabsicht führen (Lei et al., 2021, S. 3987; Liu et al., 2024, S. 11). Silva et al. (2023, S. 467) zeigen hingegen keinen signifikanten Einfluss des Vertrauens auf die Wiederverwendungsabsicht. Die anthropomorphe Gestaltung mit einem menschlichen Profilbild kann einen positiven Einfluss auf die Wiederverwendungsabsicht haben (Park et al., 2023, S. 361). Auch Anthropomorphismus allgemein kann das Vertrauen in den Chatbot erhöhen (Chen et al., 2024, S. 3646-3647; Cheng et al., 2022b, S. 10; Hsiao & Chen, 2022, S. 939-940; Konya-Baumbach et al., 2023, S. 8; Li et al., 2023b, S. 13; Liu et al., 2024, S. 10; Morana et al., 2020, S. 11). So können anthropomorphe Gestaltungselemente das Vertrauen erhöhen (Klein & Martinez, 2023, S. 2809). In Summe führt ein größeres Vertrauen in den Chatbot zu einem größeren Vertrauen in den Anbieter und dadurch kann die Kaufabsicht erhöht werden (Waytz et al., 2014, S. 113; Yen & Chiang, 2021, S. 1188). Auch die Absicht, einen Chatbot zu nutzen, ist vom Vertrauen abhängig (Fu et al., 2023, S. 6). Studien zeigen, dass das Vertrauen in eine Technologie wie Chatbots durch Soziale Präsenz erhöht werden kann (De Cicco et al., 2020, S. 1224; Hassanein & Head, 2007, S. 702; Janson, 2023, S. 10; Qiu & Benbasat, 2009, S. 165; Yen & Chiang, 2021, S. 1188). Neben dem Vertrauen kann die Soziale Präsenz auch die wahrgenommene Freude erhöhen. Vertrauen und wahrgenommene Freude führen zu einer besseren Einstellung zum Chatbot (De Cicco et al., 2020, S. 1224). Die anthropomorphen Gestaltungsmerkmale Herzlichkeit und Kompetenz können das Vertrauen in den Chatbot ebenfalls verbessern (Li et al., 2023c, S. 13).

Nutzer wenden soziale Regeln der Mensch-Mensch-Kommunikation auf Computer und Chatbots an, die anthropomorphe Gestaltungsmerkmale aufweisen (Kim et al., 2013, S. 1800; Nass et al., 1995, S. 229). Anthropomorphismus kann ein Gefühl der Sozialen Präsenz erzeugen (Adam et al., 2021, S. 438; Gnewuch et al., 2018, S. 11). Dabei ist die Soziale Präsenz anthropomorpher Chatbots höher als die nicht anthropomorpher Chatbots (Chen et al., 2024, S. 3646-3647; De Cicco et al., 2020, S. 1224; Gnewuch et al., 2018, S. 11; Han, 2021, S. 56; Jiang et al., 2023, S. 8; Jin & Youn, 2023, S. 1879; Liu et al., 2024, S. 12; Morana et al., 2020, S. 11). Des Weiteren haben Adam et al. (2021, S. 438) gezeigt, dass die Soziale Präsenz sowie die anthropomorphen Gestaltungselemente einen Einfluss auf die Benutzerfreundlichkeit von Chatbots haben. Auch in diesem Bereich gibt es Forschung, die die bisher beschriebenen Ergebnisse nicht bestätigen kann. So haben Sun et al. (2024, S. 7) keine Bestätigung gefunden, dass ein menschenähnliches Profilbild als Gestaltungsmerkmal von Anthropomorphismus die Soziale Präsenz erhöht. Die Soziale Präsenz anthropomorpher Chatbots kann die Zufriedenheit der Nutzer erhöhen (Janson, 2023, S. 9). Darüber hinaus kann neben der Zufriedenheit auch das Vertrauen in den Chatbot durch Soziale Präsenz erhöht werden (Klein & Martinez, 2023, S. 2809). Die Soziale Präsenz kann über die soziale Attraktivität des Chatbots einen Einfluss auf das Vertrauen haben (Lei et al., 2021, S. 3986).

Neben diesen positiven Nutzerreaktionen auf Chatbots gibt es auch Ergebnisse, die zeigen, dass eine menschenähnliche Gestaltung nicht immer optimal für die Nutzerwahrnehmung des Chatbots ist (Klein et al., 2024, S. 13). Forschung zu Anthropomorphismus bei Chatbots und dessen Auswirkungen finden auch negative Ergebnisse. So können negative Nutzerreaktionen durch anthropomorphe Chatbots hervorgerufen werden. Anthropomorphismus kann Unbehagen bei Nutzern auslösen, wie beispielsweise ein Gefühl von Unheimlichkeit oder dass sich Nutzer in ihrer menschlichen Identität bedroht fühlen. Dieses Ergebnis ist in Übereinstimmung mit dem Uncanny Valley Effekt (Blut et al., 2021, S. 635; Mende et al., 2019, S. 549; Song & Shin, 2024, S. 449). Die menschenähnliche Gestaltung des Profilbilds, welches zu den anthropomorphen Gestaltungsmerkmalen zählt, erhöht das Gefühl der Unheimlichkeit, wodurch das Vertrauen und die Wiederverwendungsabsicht reduziert werden können (Song & Shin, 2024, S. 450). Eine verzögerte Antwortzeit kann zu einem geringeren Vertrauen in den Chatbot führen (Cheng et al., 2022b, S. 10). Allgemein kann Anthropomorphismus das Vertrauen in den Chatbot auch verringern (Fu et al., 2023, S. 6).

Tabelle 2 fasst die im Rahmen dieser Arbeit gefundenen Ergebnisse übersichtlich zusammen. Dabei wird hinsichtlich der abhängigen Variablen unterschieden, ob in den Studien allgemeine Untersuchungen zu Anthropomorphismus oder zu einzelnen anthropomorphen Gestaltungsmerkmale durchgeführt wurden. Es finden sich teilweise unterschiedliche Ergebnisse in der Literatur. Anthropomorphismus im Bereich von Chatbots hat in den letzten Jahren zu einer Reihe wissenschaftlicher Veröffentlichungen geführt. Wichtige Faktoren sind das Vertrauen sowie die Soziale Präsenz. Daher werden diese in *FP I* genutzt und ein Beitrag zur bisherigen Forschung geleistet. Die Wiederverwendungsabsicht als abhängige Variable fand bisher wenig Beachtung in der Literatur, daher wird diese in den Fokus dieser Arbeit gestellt. Im Rahmen dieser Arbeit wird ein anthropomorphes Design verwendet und dabei auf Teile der anthropomorphen Gestaltungsmerkmale dieses Kapitels zurückgegriffen.

Tabelle 2: Auswirkungen von Anthropomorphismus im Kontext von Chatbots

Unabhängige Variable	Mediatoren	Abhängige Variable	Quelle
Anthropomorphismus: Allgemein	Kompetenz, Authentizität	Kundenbindung	Nguyen et al., 2023
		Nutzungsfreude	Klein & Martinez, 2023
		Nützlichkeit	Epley et al., 2007
		Benutzerfreundlichkeit	Adam et al., 2021
	Nutzungsfreude	Kaufabsicht	Han, 2021
	Beziehung zu Chatbot	WoM	Kayeser Fatima et al., 2024
			Konya-Baumbach et al., 2023
		Zufriedenheit	Chang-Chien et al., 2022; Chen et al., 2024; Hsiao & Chen, 2022; Konya-Baumbach et al., 2023
		Privatsphärebedenken	Ischen et al., 2020b
		Vertrauen	Chen et al., 2024; Cheng et al., 2022b; Fu et al., 2023; Hsiao & Chen, 2022; Konya-Baumbach et al., 2023; Li et al., 2023b; Liu et al., 2024; Morana et al., 2020
		Unbehagen (Unheimlichkeit, Bedrohung)	Blut et al., 2021; Mende et al., 2019; Song & Shin, 2024
		Soziale Präsenz	Adam et al., 2021; Chen et al., 2024; De Cicco et al., 2020; Gnewuch et al., 2018; Han, 2021; Jin & Youn, 2023; Liu et al., 2024; Morana et al., 2020
Empathie	Kompetenz, Herzlichkeit	Servicequalität	Han et al., 2022
			Han et al., 2022
		Zufriedenheit	Hu & Sun, 2023; Liu et al., 2023b
	Kompetenz, Herzlichkeit	Wiederverwendungsabsicht	Haupt et al., 2023
Kompetenz, Herzlichkeit	Nützlichkeit	Wiederverwendungsabsicht, Vertrauen	Li et al., 2023a
Emojis		Wahrgenommene Kompetenz	Nguyen et al., 2023
Geschlecht		Zufriedenheit	McDonnell & Baxter, 2019
Profilbild		Wiederverwendungsabsicht	Park et al., 2023
	Unheimlichkeit	Wiederverwendungsabsicht, Vertrauen	Song & Shin, 2024
		Soziale Präsenz	Sun et al., 2024
Antwortzeit		Vertrauen	Cheng et al., 2022b

4.2 Offenlegung der Chatbot-Identität

Chatbots werden häufig anthropomorph dargestellt und haben durch technologische Fortschritte einen Kommunikationsstil, der dem menschlichen sehr ähnelt. Sie werden häufig menschenähnlich gestaltet. Dadurch ist das Vorhandensein eines Chatbots für Nutzer nur schwer oder gar nicht erkennbar, und seine Identität muss explizit offengelegt werden, damit klar wird, dass es sich nicht um einen Menschen handelt (Belanche et al., 2024, S. 151). Dies kann von Unternehmen aktiv gezeigt werden, indem beispielsweise vor der Interaktion ein Hinweis mit Informationen zum Chatbot gegeben wird oder der Chatbot sich selbst als nicht-menschlich vorstellt. Die nicht-menschliche Identität kann allerdings auch erst nach der Konversation beschrieben werden. Dies kann rein informell sein oder als eine mögliche Wiederherstellungsstrategie einer gescheiterten Interaktion genutzt werden. Kommt es zu Gesprächsabbrüchen mit Chatbots oder kann dieser die Anfrage nicht lösen, kann mit der Offenlegung versucht werden, den Kontakt mit dem Kunden nochmals herzustellen. Hierbei können entweder der Chatbot selbst, die Schwächen und/oder die Kompetenzen des Chatbots beschrieben werden (Benner et al., 2021, S. 8-9).

Es sollten im Rahmen der Offenlegung nicht nur die positiven Eigenschaften sowie die Fähigkeiten des Chatbots hervorgehoben werden, sondern zusätzlich auch mögliche Einschränkungen genannt werden. Dies ist in Form einer zweiseitigen Nachricht möglich, die sowohl positive als auch negative Aspekte umfasst. Denn dadurch kann die Service-Authentizität und in der Folge auch die Nutzungsbereitschaft erhöht werden (Meng et al., 2023, S. 8). Das Vertrauen in den Chatbot kann steigen, wenn die Expertise oder die Schwächen des Chatbots im Rahmen der Offenlegung kommuniziert werden. Dadurch kann ein ähnliches Vertrauensniveau erzeugt werden, wie es nicht offengelegte Chatbots erreichen (Mozafari et al., 2021a, S. 2920-2921). Offenlegung ist ein wichtiges Instrument, denn problematisch wird es, wenn sich Nutzer durch das Verschweigen der nicht-menschlichen Identität durch den Anbieter getäuscht fühlen. Dadurch können negative Kundenreaktionen hervorgerufen werden. Offenlegung kann dabei helfen, diese negative Kundenreaktionen zu reduzieren oder zu vermeiden (Castillo et al., 2021, S. 919).

Viele Forschungsergebnisse zeigen negative Auswirkungen der Offenlegung von Chatbots (Is-howo-Oloko et al., 2019; Luo et al., 2019; Mozafari et al., 2022). Allgemein kann Transparenz, die durch Offenlegung erreicht wird, dem Erfolg einer Technologie wie beispielsweise eines

Chatbots schaden und mögliche Vorteile der Technologie verringern (Ishowo-Oloko et al., 2019, S. 519). Denn das Beschreiben der nicht-menschlichen Identität kann die Einstellung zum Onlinehändler negativ beeinflussen, indem die Soziale Präsenz und das Vertrauen diesen Effekt seriell mediierten. Offenlegung führt dabei zu geringerer Sozialer Präsenz, geringerem Vertrauen und zu einer schlechteren Einstellung zum Onlinehändler (De Cicco et al., 2021, S. 11). Chatbots, die ihre nicht-menschliche Identität offenlegen, werden als weniger menschlich wahrgenommen und ihnen wird eine geringere Soziale Präsenz zugeschrieben (Hendriks et al., 2020, S. 276; Mozafari et al., 2021b, S. 9). Van der Goot et al. (2024, S. 10) haben hingegen keinen signifikanten Einfluss der Offenlegung auf die Wahrnehmung des Chatbots als anthropomorph und sozial präsent gefunden. Die aktuelle Forschungslage zeigt hierbei also unterschiedliche Befunde.

Nutzer von offengelegten Chatbots reagieren häufig schroff und tätigen weniger Käufe, da ein solcher Chatbot als weniger sachkundig und einfühlsam wahrgenommen wird (Luo et al., 2019, S. 944). Wird ein Chatbot beispielsweise im Rahmen der Sammlung von Spenden für ein soziales Projekt eingesetzt, kann eine Information über den Chatbot auch zu negativen Reaktionen führen und die Wahrnehmung des Zwischenmenschlichen bei der Interaktion ändern. So kann die Offenlegung die Wahrscheinlichkeit, eine Spende zu tätigen, verringern (Shi et al., 2020, S. 10). Allgemein haben Nutzer eher negative Vorurteile gegenüber den Antworten eines offengelegten Chatbots (Murgia et al., 2016, S. 1277).

Die Offenlegung der Identität des Chatbots führt zu einem geringeren Vertrauen, gerade wenn das Anliegen des Nutzers kritisch und die richtige Bearbeitung sehr wichtig ist. Darüber hinaus findet sich ein indirekter negativer Effekt auf die Kundenbindung, der durch Vertrauen mediert wird (Mozafari et al., 2022, S. 230-231). Auf der anderen Seite kann das Vertrauen auch höher sein, wenn die Identität preisgegeben wird. Dies gilt allerdings nur, wenn die Chatinteraktion scheitert, also das Problem des Kunden nicht im Online-Chat gelöst werden kann. In dem Fall eines Fehlers ist auch der indirekte Effekt der Offenlegung auf die Kundenbindung über das Vertrauen positiv (Mozafari et al., 2022, S. 234). So kann auf der einen Seite durch Offenlegung das Vertrauen und indirekt die Kundenbindung erhöht werden und auf der anderen Seite können negative Auswirkungen die Folge sein. Das Aufdecken der nicht-menschlichen Identität des Chatbots kann den negativen Effekt des Scheiterns abschwächen (Mozafari et al., 2022, S. 234-235; Mozafari et al., 2020, S. 15). Funke et al. (2023, S. 14) haben keinen signifikanten

Einfluss der Offenlegung auf die Kundenbindung festgestellt. Durch das Preisgeben der Chatbot-Identität sinken die Erwartungen an das Gesprächsengagement. Unter Gesprächsengagement wird dabei verstanden, dass der Chatbot gut kommunizieren kann, indem er überlegt agiert und auf die Benutzereingaben gut antwortet (Grimes et al., 2021, S. 7). Negative Auswirkungen können durch eine späte Offenlegungsstrategie sowie vorherige Chatbot Erfahrung der Nutzer verringert werden (Luo et al., 2019, S. 938).

Offenlegung der Chatbot-Identität wurde auch als Moderator untersucht (Cheng et al., 2022a; Fan et al., 2024; Xie et al., 2024). Dabei moderiert die Preisgabe der Chatbot-Identität die positive Beziehung zwischen Empathie und Vertrauen der Kunden negativ, während sie die positive Beziehung zwischen Freundlichkeit und Vertrauen der Kunden verstärkt (Chen et al., 2022a, S. 508). Auch Fan et al. (2024) untersuchten die moderierende Rolle von Offenlegung in Bezug auf Empathie, genauer gesagt zwischen der Empathie und der Kundenbindung. Ihre Ergebnisse zeigen, dass die negativen Auswirkungen der Inkongruenz von empathischen Antworten auf die Kundenbindung schwächer sind, wenn die Identität des Chatbots aufgedeckt wird. Darüber hinaus sind die positiven Auswirkungen der Kongruenz empathischer Antworten auf die Kundenbindung geringer, wenn die Identität preisgegeben wird (Fan et al., 2024, S. 1881-1882). Neben Empathie wurde auch der Humor von Chatbots und dessen Auswirkungen auf die Kompetenz, die Soziale Präsenz, das Entertainment sowie die Zufriedenheit mit dem Service betrachtete und hierbei auf die moderierende Rolle der Offenlegung geschaut (Xie et al., 2024). Dabei haben Xie et al. (2024, S. 9-10) gezeigt, dass der von Chatbots ausgedrückte Humor einen positiven Einfluss auf die Servicezufriedenheit und -kompetenz hat, wenn die Identität des Chatbots nicht dargelegt wird. Wird die Identität des Chatbots allerdings preisgegeben, wird dieser positive Einfluss verringert oder verschwindet.

Neben möglichen negativen Auswirkungen kann die Offenlegung von Chatbots auch mit positiven Reaktionen verbunden sein. So reagieren Kunden auf nicht anthropomorphe Chatbots positiver als auf menschliche Agenten, wenn sie wissentlich mit einem Chatbot kommunizieren, gerade wenn es sich um ein unangenehmes Thema oder Produkt handelt. Dieser Einfluss ist allerdings geringer oder kann ganz verschwinden, wenn der Chatbot anthropomorph gestaltet ist (Jin et al., 2024, S. 8). Die Offenlegung von Chatbots kann auch zu einem respektvollen und ehrlichen Umgang führen (Murtarelli et al., 2021, S. 931). Durch Preisgabe der Chatbot-Identität kann die Organisation hinter dem Chatbot ethischer und vertrauenswürdiger wirken (McGuire et al., 2023, S. 5).

Forschung im Bereich der Offenlegung von Chatbots beschäftigt sich mit Auswirkungen dieser und differenziert häufig zwischen nicht offengelegten und offengelegten Chatbots (Gnewuch et al., 2024; Mozafari et al., 2022). Dem Zeitpunkt der Offenlegung wurde bisher wenig Aufmerksamkeit geschenkt. Luo et al. (2019) haben diesen beispielsweise untersucht. Hier wird gezeigt, dass eine Offenlegung vor der Interaktion die Kaufabsicht um mehr als 78 % reduziert (Luo et al., 2019, S. 938). Handelt es sich bei der Chatbot Anwendung um eine Aufgabe, die lediglich einen geringen Dienstleistungsanteil benötigen, führt die Offenlegung nach der Interaktion zu einer höheren Zufriedenheit mit dem Service. Ist allerdings ein hoher Dienstleistungsanteil notwendig, dreht sich die Wirkung um und eine Offenlegung vor der Interaktion führt zu einer höheren Zufriedenheit (Yu et al., 2024, S. 1098). Interaktionen mit einem hohen Dienstleistungsanteil im Vergleich zu einem geringen sind dadurch gekennzeichnet, dass die Kommunikationszeit länger, die Vertrautheit höher und der Informationsaustausch intensiver ist (Yu et al., 2024, S. 1091). Darüber hinaus verringert sich die Interaktionsdauer, wenn der Chatbot vor der Interaktion offengelegt wird (Luo et al., 2019, S. 944).

Die vorherigen Ausführungen zeigen, dass die Offenlegung von Chatbots häufig zu einer Vielzahl negativer Reaktionen führt. In manchen Fällen können allerdings auch positive Reaktionen mit der Offenlegung verbunden sein. Selbst bei gleicher Leistung oder gleichem Ergebnis werden Chatbots häufig schlechter beurteilt als menschliche Mitarbeiter und negative Reaktionen überwiegen bei offengelegten Chatbots. Teilweise wird hier von einem Offenlegungsdilemma gesprochen (Mozafari et al., 2020, S. 3), gleiche Leistung wird nicht gleich bewertet. In der Folge können für Unternehmen Schäden und hohe Kosten entstehen.

Auf der einen Seite scheint es sinnvoll zu sein, die nicht-menschliche Identität zu verschweigen, um diese häufig negativen Reaktionen zu vermeiden. Gerade wenn eine nahtlose Übergabe an einen Mitarbeiter gewährleistet ist und irrtümliche Erwartungen an die Fähigkeiten des Chatbots nicht zu unangemessenen oder frustrierten Serviceergebnissen führen, kann es ebenfalls sinnvoll sein, den Nutzer im Unklaren über die nicht-menschliche Identität zu lassen (Skjuve et al., 2019, S. 48). Andererseits könnte im Hinblick auf offensichtliche gesellschaftliche und ethische Bedenken bezüglich Chatbots gefordert werden, die Offenlegung von Chatbots zu regulieren oder als Standardpraxis zu etablieren (Huseynov, 2023, S. 63; Robinson et al., 2020, S. 371; Schanke et al., 2021, S. 746). Das Verschweigen widerspricht etischen Grundsätzen sowie rechtlichen Anforderungen. In Kalifornien gibt es beispielsweise seit 2019 ein Gesetz, das den Dienstleister verpflichtet, die nicht-menschliche Identität von Chatbots offenzulegen

(Gershgorn, 2018). Auch das Europäische Parlament hat ein Gesetz auf den Weg gebracht, das Transparenzanforderungen an Chatbots wie ChatGPT stellt. Dabei muss offengelegt werden, dass der Inhalt durch künstliche Intelligenz generiert wurde (Europäisches Parlament, 2023).

Darüber hinaus kann das Verschweigen der Chatbot-Identität auch als Täuschung gesehen werden (McGuire et al., 2023, S. 3). Chatbots bieten eine Möglichkeit, überzeugende Antworten zu liefern. Daher ist es weder notwendig noch wünschenswert, den Nutzern vorzuspiegeln, dass sie mit einem Menschen interagieren (Adam et al., 2021, S. 438). Sollte der Chatbot keine zufriedenstellenden Kommunikationsfähigkeiten haben, fällt häufig auf, dass das Gegenüber kein Mensch sein kann. Daher kann eine offene Kommunikation der Identität sinnvoll sein. Weitere Forschung im Bereich der Art und Weise der Offenlegung von Chatbots sollte vorangetrieben werden (Meng et al., 2023, S. 2). Es ist auch unklar, ob eine frühe Offenlegung tatsächlich mit negativen Konsequenzen einhergeht und ob die Ergebnisse von Luo et al. (2019) bestätigt werden können. Hier ist der Zeitpunkt der Offenlegung interessant, um negative Kundenreaktionen zu vermeiden oder zu reduzieren. Die im Rahmen der Literaturrecherche dieser Arbeit gefundenen wissenschaftlichen Ergebnisse werden in Tabelle 3 übersichtlich zusammengefasst.

Untersucht wurde dabei nicht nur, ob die Identität des Chatbots offengelegt wird, sondern auch, ob die Möglichkeit besteht, zu einem menschlichen Ansprechpartner zu wechseln. Hier spricht man von hybriden Serviceagenten, Mensch und Chatbot werden ergänzend eingesetzt. Durch die Kommunikation dieser Möglichkeit wird klar, dass die ursprüngliche Interaktion nicht mit einem Menschen stattgefunden hat. Dabei kann die Offenlegung dieser Möglichkeit einen Einfluss auf den Kommunikationsstil der Nutzer haben. Nutzer verwenden folglich einen eher menschlichen Kommunikationsstil, wenn sie zu Beginn oder während der Interaktion über die menschliche Beteiligung informiert werden (Gnewuch et al., 2024, S. 944). Allerdings kann Offenlegung auch dazu führen, dass die Arbeitsbelastung für die menschlichen Mitarbeiter höher wird (Gnewuch et al., 2024, S. 951) und die Nutzer eher wechseln möchten und direkt mit einem Menschen interagieren. Die wissenschaftlichen Ergebnisse zur Übergabe eines Chatbots an einen Menschen werden in Kapitel 4.3 genauer betrachtet sowie in Kapitel 5.5 ergänzt.

Tabelle 3: Auswirkungen der Offenlegung von Chatbots

Unabhängige Variable	Mediatoren	Moderatoren	Abhängige Variable	Quelle
Offenlegung			Negative Kundenreaktionen	Castillo et al., 2021; Ishowo-Oloko et al., 2019; Murgia et al., 2016
		Zeitpunkt, Chatboterfahrung		Luo et al., 2019
	Sachkunde, Einfühlungsvermögen			Luo et al., 2019
		Anthropomorphismus, (unangenehmes) Gesprächsthema		Jin et al., 2024
	Soziale Präsenz, Vertrauen		Einstellung zum Händler	De Cicco et al., 2021
			Menschlichkeit	Hendriks et al., 2020; Mozafari et al., 2021b
			Soziale Präsenz	Hendriks et al., 2020; Mozafari et al., 2021b; Van der Goot et al., 2024
			Anthropomorphismus	Van der Goot et al., 2024
	Vertrauen	Wichtigkeit	Kundenbindung	Mozafari et al., 2022
				Funke et al., 2023
			Effekt des Scheiterns	Mozafari et al., 2022; Mozafari et al., 2020
			Gesprächsengagement	Grimes et al., 2021
			Spendenabsicht	Shi et al., 2020
			Respektvoller Umgang	Murtarelli et al., 2021
Offenlegungszeitpunkt			Kaufabsicht	Luo et al., 2019
			Interaktionsdauer	Luo et al., 2019
		Dienstleistungsanteil	Zufriedenheit	Yu et al., 2024
Offenlegung (zweiseitige Nachricht)	Service-Authentizität		Nutzungsbereitschaft	Meng et al., 2023
			Vertrauen	Funke et al., 2023; Mozafari et al., 2021a, 2022

Tabelle 3: Auswirkungen der Offenlegung von Chatbots (Fortsetzung)

Unabhängige Variable	Mediatoren	Moderatoren	Abhängige Variable	Quelle
Offenlegung als Moderator				
Empathie		Offenlegung	Vertrauen	Cheng et al., 2022b
			Kundenbindung	Fan et al., 2024
Freundlichkeit			Vertrauen	Cheng et al., 2022b
Humor			Kompetenz, Soziale Präsenz, Entertainment, Zufriedenheit	Xie et al., 2024

4.3 Übergabe von Chatbots an Mitarbeiter

Übergaben von Chatbots an Menschen werden oft im Zusammenhang mit Servicefehlern untersucht (Meyer et al., 2022; Poser et al., 2021, 2022, 2023). Fehler, Error, Missverständnisse oder Gesprächsabbrüche kommen trotz technischer Fortschritte im Rahmen von Konversationen mit Chatbots noch häufig vor (Zarouali et al., 2018, S. 495). Diese werden zwar weniger, gerade durch den Einsatz von künstlicher Intelligenz, können aber nicht gänzlich vermieden werden. Menschen haben eine geringe Toleranz gegenüber Fehlern von Chatbots (Murgia et al., 2016, S. 1277). Daher ist es im Falle eines solchen Fehlers wichtig, dass die Anfrage des Kunden nicht endet, sondern im besten Fall ohne Unterbrechung an einen menschlichen Mitarbeiter übergeben wird und ein Wiederherstellungsprozess gestartet wird (Meyer et al., 2022, S. 2496). In der Praxis werden solche Übergaben an einen menschlichen Mitarbeiter teilweise genutzt (Braggaar et al., 2024, S. 28). Weitere gängige Wiederherstellungsstrategien, die nach einem Fehler eingesetzt werden, sind Entschädigungen oder ein Entgegenkommen bei Anfragen, eine Entschuldigung oder ein Neustart des Prozesses mit dem Ziel, den Fehler zu korrigieren und den Nutzer um Nachsicht zu bitten und dadurch zufrieden zu stellen (Liu et al., 2023a, S. 9). Positive sowie negative Erfahrungen mit dem Chatbot wirken sich auf die Einstellung zum Unternehmen aus. Denn wenn Technologien wie Chatbots Fehler machen, geben Nutzer nicht der Technologie selbst die Schuld, sondern teilweise dem Unternehmen, das sie einsetzt (Meyer et al., 2022, S. 2501; Zarouali et al., 2018, S. 495). Daher ist es auch wichtig, den Kunden nicht allein zu lassen und gegebenenfalls die Möglichkeit zu eröffnen mit einem menschlichen Mitarbeiter zu kommunizieren. Die Übergabe nach einem Fehler sollte schnell und reibungslos erfolgen, damit keine weiteren Unterbrechungen für den Kunden wahrnehmbar

sind (Ozuem et al., 2024, S. 2270). Der Umgang mit Fehlern ist für Unternehmen daher sehr wichtig.

Je nachdem muss es allerdings gar nicht erst zum Fehler kommen. Der Chatbot kann während des Gesprächs merken, dass das Gespräch in eine nicht gewollte Richtung läuft und einen möglichen Fehler vermeiden, indem die Interaktion an einen Menschen übergeben wird, automatische Wiederherstellungsprozesse gestartet werden oder sich der Chatbot entschuldigt (Feine et al., 2019, S. 13).

Hybride Serviceagenten sollen die jeweiligen Stärken von menschlicher und künstlicher Intelligenz kombinieren, um deren Schwächen auszugleichen und herausragende Serviceergebnisse zu erzielen. Einfache, allgemeine Anfragen übernimmt der Chatbot, während alle übrigen an einen Menschen weitergeleitet werden. Oder im Falle eines Fehlers kann der Chatbot an einen Menschen übergeben (Gnewuch et al., 2024, S. 944; Keyser et al., 2019, S. 161). Diese Wiederherstellungsstrategie kann effektiv sein, um vermeintlich verlorene Nutzer zurückzugewinnen, da sie eine schnelle Lösung bietet und so die Nutzererfahrung verbessert (Meyer et al., 2022, S. 2494).

Im Rahmen der Möglichkeit einer Übergabe an einen Menschen haben sich Huang und Dootson (2022) mit Bewältigungsstrategien beschäftigt. Dabei hat eine späte Offenlegung der Möglichkeit, mit einem Menschen zu sprechen – also nach dem Fehler – die emotionsorientierte Bewältigungsstrategie erhöht. Dies kann sich darin äußern, dass negative Emotionen zum Ausdruck gebracht werden, man nach emotionaler Unterstützung sucht oder mit Verleumdung reagiert. Eine mögliche Reaktion kann sich dabei in Form von Aggression äußern. Dieser Einfluss wird durch den Grad der Kundenbeteiligung moderiert (Huang & Dootson, 2022, S. 6).

Durch die Übergabe an einen Menschen und die damit einhergehende Wiederherstellung der Interaktion kann die externe Schuldzuschreibung des Fehlers abgemildert werden. Denn die Nutzer übernehmen selbst nochmals eher einen Anteil der Verantwortung für den Fehler und geben sich somit eher selbst eine Mitschuld am Entstehen des Fehlers (Meyer et al., 2022, S. 2492-2493). Das Vertrauen in einen Chatbot kann die Absicht von Nutzern verringern, von Chatbots zu einem menschlichen Chatpartner wechseln zu wollen (Cheng et al., 2022b, S. 9).

Durch die Übergabe des Gesprächs an einen Menschen kann die Kundeninteraktion mit geringem Aufwand zielgerichtet fortgesetzt werden. Aufgrund einer guten Übergabe benötigen die

Servicemitarbeiter nur wenig Zeit, um das Gespräch mit dem Kunden zu beginnen, denn relevante Informationen liegen teilweise schon vor. Darüber hinaus reduziert sich die Zahl der wiederholenden Anfragen. Dadurch kann die Arbeitsbelastung der Mitarbeiter reduziert werden und der wahrgenommene Zeitdruck sinkt, was eine kundenorientierte Interaktion ermöglicht (Poser et al., 2021, S. 1188, 2022, S. 541, 2023, S. 12). Neben den Kundenbedürfnissen sollte bei der Gestaltung von Chatbots und der Übergabe auch die Anforderungen der Mitarbeiter betrachtet werden. Denn nur so kann eine erfolgreiche Weiterführung der Nutzeranfrage nach der Übergabe sichergestellt werden (Poser et al., 2023, S. 12). Perspektivisch sollte untersucht werden, wie und ob Warnsysteme für bevorstehende Übergaben eingeführt werden sollen (Poser et al., 2021, S. 1188).

Sollte die Interaktion nicht direkt von einem Menschen übernommen werden können und Kunden eine Wartezeit in Kauf nehmen müssen, können Informationen über die Wartezeit sowie die Position in der Warteschlange oder Zeitfüller helfen, die Wartezeit zu überbrücken. Zeitfüller lenken den Kunden ab, indem er von dem Chatbot aktiv in ein Gespräch verwickelt wird (Wintersberger et al., 2020, S. 2). Informationen über die Wartezeit sowie die Position in der Warteschlange oder Zeitfüller reduzieren das Zeitempfinden und führen zu einer als kürzer empfundenen Wartezeit. Darüber hinaus wird die Servicequalität als höher bewertet, wenn Zeitfüller oder die Informationen über Wartezeit und -position eingesetzt werden. Die Nutzererfahrung kann höher sein, wenn Zeitfüller von Chatbots genutzt werden (Wintersberger et al., 2020, S. 5). Die im Rahmen dieser Arbeit betrachteten bisherigen wissenschaftlichen Ergebnisse sind in Tabelle 4 dargestellt.

Wissenschaftliche Erkenntnisse zur Übergabe an einen Menschen als Wiederherstellungsstrategie gibt es bisher wenig. Bei Servicefehlern wurden eher Entschuldigungen als mögliche Strategie untersucht (Meyer et al., 2022, S. 2501). Daher wird in *FP III* der Einfluss der Übergabe an einen menschlichen Mitarbeiter untersucht.

Tabelle 4: Aktuelle Erkenntnisse zur Übergabe der Chatbot-Interaktion

Zentrales Ergebnis	Quelle
Durch die Übergabe an einen Mitarbeiter können verloren geglaubte Kunden zurückgewonnen werden, da ihnen eine schnelle Lösung geliefert wird. Die Nutzererfahrung kann durch diese Wiederherstellungsstrategie gesteigert werden.	Meyer et al., 2022
Übergaben an einen Menschen sollten schnell und reibungslos erfolgen.	Ozuem et al., 2024
Eine späte Kommunikation der Möglichkeit mit einem Menschen zu schreiben, beispielsweise nach einem Fehler kann die emotionsorientierte Bewältigungsstrategie der Nutzer erhöhen.	Huang & Dootson, 2022
Vertrauen Nutzer einem Chatbot ist deren Wechselabsicht zu einem Menschen geringer.	Chang-Chien et al., 2022
Durch die Übergabe an einen Menschen wird die externe Schuldzuschreibung eines Fehlers im Rahmen der Interaktion verringert und der Nutzer gibt sich eher eine Mitschuld an der Entstehung des Fehlers.	Meyer et al., 2022
Sollte die Übergabe an einen Menschen mit einer Wartezeit verbunden sein, helfen Informationen über die Wartezeit, die Position in der Warteschlange oder allgemeine Zeitfüller. Dadurch wird das Zeitempfinden verringert. Darüber hinaus hat dies positive Auswirkungen auf die Beurteilung der Servicequalität und die Nutzererfahrung kann besser sein.	Wintersberger et al., 2020

4.4 Bewertungen von Chatbot-Interaktionen

Kundenbewertungen stellen allgemein ein wichtiges Instrument bei Kauf- oder Nutzungsentscheidungen dar. Dabei kann durch positive Kundenbewertungen das Vertrauen erhöht (Gavilan et al., 2018, S. 58) oder Unsicherheiten und Risiken beim Onlinekauf verringert werden (Bae & Lee, 2011, S. 262; Hong et al., 2017, S. 1). Allerdings können schlechte Kundenbewertungen auch zu Zweifeln an der Qualität des Produktes oder des Service führen. Für potenzielle Käufer oder Nutzer stellen Bewertungen eine Entscheidungsgrundlage für oder gegen den Kauf oder die Nutzung dar. Wichtig ist, dass diese Bewertungen von anderen Kunden oder Nutzern stammen und nicht vom Unternehmen selbst.

Im Folgenden werden sowohl Ergebnisse zu Bewertungen im Allgemeinen und nicht explizit im Kontext von Chatbots betrachtet. Denn Auswirkungen von Bewertungen von Chatbots wurden in der Literatur bisher nicht untersucht. Zu Bewertungen in anderen Kontexten wurde eine Vielzahl unterschiedlicher Faktoren untersucht, die einen Einfluss auf das Verfassen einer Bewertung haben oder die von Bewertungen beeinflusst werden. Bewertungen wurden dabei auf unterschiedlichen Bewertungsplattformen aus verschiedenen Branchen, von unterschiedlichen Unternehmen wie beispielsweise Amazon, TripAdvisor, Yelp.com oder Ebay untersucht (Zheng, 2021, S. 235). Im Folgenden werden Faktoren der aktuellen Literatur betrachtet. Dabei

wird der Differenzierung aus Kapitel 2.2 gefolgt und zwischen Rating, Review sowie Bewertung (Rating und Review) unterschieden. Bisher wurden Auswirkungen von Bewertungen auf die Wahrnehmung von Kunden und Nutzern untersucht. Dabei werden Faktoren wie beispielsweise wahrgenommene Glaubwürdigkeit, Nützlichkeit, Vertrauen oder Akzeptanz von Bewertungen betrachtet (Zheng, 2021, S. 235).

Ratings können für Kunden den Entscheidungsprozess vereinfachen. Gerade wenn eine hohe Anzahl sehr guter oder sehr schlechter Ratings vorhanden ist (LaCour & Serra, 2022, S. 242). Darüber hinaus können Ratings das Vertrauen der Kunden erhöhen (Kostyk et al., 2017, S. 137). Auch die Nützlichkeit einer Bewertung kann durch ein Rating beeinflusst werden (Guha Majumder et al., 2022, S. 161). Höhere Ratings müssen allerdings nicht zwingend zu stärkeren Verkäufen von Produkten führen (Duan et al., 2008, S. 1014). Erwartungen können durch Bewertungen beeinflusst werden. Positive Bewertungen führen zu höheren Erwartungen als negative Bewertungen (Nath et al., 2018, S. 130).

Gavilan et al. (2018) haben sich den Einfluss der Anzahl der Reviews angeschaut. Dabei gehen sie von einer Interaktion zwischen Ratings und Anzahl an Reviews aus. Wenn ein Rating gut ist, kann die Vertrauenswürdigkeit dieses von der Anzahl an Reviews abhängen. Wenn das Rating schon schlecht ist, hat die Anzahl der Reviews keinen Einfluss auf die Vertrauenswürdigkeit der Bewertung (Gavilan et al., 2018, S. 58). Auch Duan et al. (2008, S. 1014) beschreiben, dass die Anzahl der Reviews wichtig ist und den Verkauf von Produkten beeinflussen kann. Die Absicht, ein Produkt zu kaufen, wird durch Reviews beeinflusst (Park et al., 2007, S. 140). Allgemein haben Reviews einen Einfluss auf die Kaufabsicht, die Nutzungsabsicht oder das Buchungsverhalten von Unterkünften (Arief et al., 2023, S. 91; Ye et al., 2009, S. 182).

Neben der Review Anzahl können gerade negative Reviews einen Einfluss auf die Kaufabsicht haben (Cui et al., 2012, S. 53; Zhao et al., 2015, S. 1357). Auch die Qualität der Reviews kann die Kaufabsicht der Kunden beeinflussen (Park et al., 2007, S. 140). Die Einstellung von Kunden wird schlechter, wenn der Anteil negativer Reviews zunimmt (Lee et al., 2008, S. 349). Positive Reviews haben einen Einfluss auf die Vertrauenswürdigkeit eines Onlinehändlers (Utz et al., 2012, S. 55).

Die Dienstleistungsqualität kann höher bewertet werden, wenn die Bewertungen überwiegend positiv sind (Browning et al., 2013, S. 35). Des Weiteren müssen gelesene Reviews auch nütz-

lich für den Kunden sein. Denn die wahrgenommene Nützlichkeit eines Reviews hat einen Einfluss auf das Vertrauen sowie die Kaufabsicht (Elwalda et al., 2016, S. 314; Ventre & Kolbe, 2020, S. 294; Zhao et al., 2015, S. 1358).

Es gibt aber auch Zweifel an Bewertungen. So ist fraglich, ob die Bewertungsschreiber repräsentativ für die Nutzer sind. Oftmals sind es insbesondere Nutzer mit sehr positiven oder sehr negativen Erfahrungen, die Bewertungen abgeben (Langhe et al., 2016, S. 819). Die objektiven Qualitätsinformationen, die in Nutzerbewertungen enthalten sind, sind viel schwächer als Kunden glauben. Kunden nutzen die Bewertungen trotzdem häufig, um die Qualität von Produkten oder Dienstleistungen zu beurteilen. Obwohl es auch andere, bessere Prädiktoren gibt, wie z. B. den Preis (Langhe et al., 2016, S. 828).

Des Weiteren ist die Wirkung eines Reviews von unterschiedlichen Faktoren abhängig. Dabei wird unterschieden zwischen Charakteristika des Rezensenten, Charakteristika des Empfängers, Kontextfaktoren sowie Charakteristika des Reviews selbst. Diese werden teilweise als Moderatoren in der Literatur untersucht.

Die Offenlegung von Informationen über den Rezensenten, dessen Reputation sowie dessen Expertise, als Faktoren des Rezensenten, haben einen Einfluss auf die Nützlichkeit des Reviews (Baek et al., 2012, S. 112; Hong et al., 2017, S. 8; Racherla & Friske, 2012, S. 556; Zhao et al., 2015, S. 1354; Zheng, 2021, S. 235). Nicht nur Merkmale des Rezensenten haben einen Einfluss auf die Wirkung des Reviews, sondern auch die Charakteristika des Empfängers. So können die Vertrautheit und die Identifikation mit dem Produkt als empfängerbezogene Merkmale einen Einfluss auf den Erfolg einer Bewertung haben (Zheng, 2021, S. 235). Kulturelle Unterschiede der Empfänger können den Einfluss von Bewertungen moderieren (Nath et al., 2018, S. 130). Auch das Involvement der Kunden sowie deren bisherige Erfahrung mit Reviews können eine moderierende Rolle einnehmen und dabei den Einfluss auf die Vertrauenswürdigkeit des Reviews beeinflussen (Bae & Lee, 2011, S. 262; Filieri, 2016, S. 58). Hier gilt anzumerken, dass Bae und Lee (2011) die moderierende Rolle nur beschrieben und nicht explizit untersucht haben. Mudambi & Schuff (2010, S. 195) zeigten das die bisherige Erfahrung der Nutzer einen Einfluss auf die Nützlichkeit von Bewertungen haben kann. So werden extreme Bewertungen als weniger hilfreich angesehen als moderate Bewertungen, wenn eine hohe Erfahrung vorliegt. Der Einfluss von Reviews kann auch von Marken oder den Produkten selbst abhängen. Sowohl positive als auch negative Reviews haben einen größeren positiven beziehungsweise negativen

Einfluss bei schwachen Marken, als es bei starken Marken der Fall wäre (Ho-Dac et al., 2013, S. 46). Neben der Marke kann auch das Produkt selbst, also die Art des Produkts einen moderierenden Einfluss auf die Nützlichkeitsbewertung oder Glaubwürdigkeitsbewertung eines Reviews haben (Bae & Lee, 2011, S. 262; Cui et al., 2012, S. 53; Guha Majumder et al., 2022, S. 161). Auch die Bekanntheit eines Produkts beeinflusst die Auswirkung einer Bewertung. So haben Zhu und Zhang (2010, S. 144) für Videospiele herausgefunden, dass der Einfluss von Reviews bei unbekannten Spielen größer ist. Neben der Marke oder dem Produkt beziehungsweise der Dienstleistung selbst kann der Kommunikationskanal, über den die Bewertung zugänglich ist, einen Einfluss auf die Wahrnehmung von Bewertungen haben (Zheng, 2021, S. 235). Der Kommunikationskanal, auf dem die Bewertung veröffentlicht ist (z. B. Unternehmensseite, unabhängige Seite oder kommerzielle Drittanbieterseite), kann dabei eine moderierende Rolle einnehmen und die Vertrauenswürdigkeit indirekt beeinflussen (Bae & Lee, 2011, S. 262; Filieri, 2016, S. 58).

Ein weiterer Teil der Forschung beschäftigt sich mit Gestaltungsmerkmalen von Reviews und den damit verbundenen Auswirkungen. Durch die Gestaltung kann gegebenenfalls die Vertrauenswürdigkeit erhöht werden und Zweifel an den Reviews verringert werden. Dabei können Gestaltungsmerkmale die wahrgenommene Vertrauenswürdigkeit beeinflussen. Diese können direkt von der Quelle der Bewertung abhängen, wie beispielsweise der Anzahl an Beiträgen oder der Gestaltung des Nutzerprofils. Die wahrgenommene Vertrauenswürdigkeit kann auch von der Nachricht selbst abhängen. Hier können die Länge, die Art der Informationen, der Detailgrad oder der Schreibstil Beispiele sein. Auch der Inhalt kann Auswirkungen haben, ob dieser negativ oder positiv dargestellt ist oder gegebenenfalls auch beide Seiten im Review betrachtet werden. Erkennbare Muster können sich ebenfalls auswirken, wie beispielsweise ein übermäßig positives Review nach einer Vielzahl negativer Reviews oder Dopplungen von Inhalten oder ganzer Reviews (Filieri, 2016, S. 58). Die Länge eines Reviews wirkt sich positiv auf den wahrgenommenen Wert aus, allerdings nur bis zu einem gewissen Umfang. Durch zu viele Informationen in einem Review kann es schwer zu verarbeiten sein (Schindler & Bickart, 2012, S. 239). Auch die Nützlichkeit eines Reviews kann von der Tiefe und Länge beeinflusst werden. Darüber hinaus kann auch die Wertigkeit des Reviews, die Lesbarkeit sowie das Alter des Reviews einen Einfluss auf die Nützlichkeit haben (Baek et al., 2012, S. 112; Guha Majumder et al., 2022, S. 161; Hong et al., 2017, S. 8; Mudambi & Schuff, 2010, S. 195). Des Weiteren kann die Qualität des Reviews einen Einfluss auf die Kundenbeeinflussung haben. So

beeinflussen qualitativ hochwertige negative Reviews die Einstellung von Kunden stärker als negative Reviews von geringer Qualität (Lee et al., 2008, S. 349). Allgemein hat der Inhalt einen positiven Einfluss auf die wahrgenommene Nützlichkeit. Im Rahmen von Reviews ist es wichtig, dass die Kernelemente der Dienstleistung in der Kundenbewertung berücksichtigt wurden (Browning et al., 2013, S. 35). Allerdings kann auch der Inhalt Kunden misstrauisch machen. Denn durch eine zu positive Formulierung können Kunden die Motive des Rezensenten in Frage stellen und misstrauisch werden (Schindler & Bickart, 2012, S. 240).

Generell lässt sich sagen, dass gerade unzufriedene Kunden anfangs eher Reviews schreiben oder Ratings abgeben als zufriedene Kunden. Daher kann es passieren, dass mit steigender Anzahl auch die Bewertungen positiver ausfallen (Melián-González et al., 2013, S. 279). Dies passt auch zu den Ergebnissen, dass die Anzahl an Bewertungen ein entscheidendes Kriterium für den Erfolg von Bewertungen ist. Allerdings kann es häufig passieren, dass die Anzahl an Reviews sehr groß ist und es daher für Kunden sehr schwierig bis unmöglich ist, alle Reviews zu lesen und darauf ihre Kaufentscheidung aufzubauen (Singh et al., 2017, S. 346). Bewertungen sollten aus Rating und Review bestehen. Diese Kombination hat einen größeren Einfluss auf das Kundenverhalten, als es Ratings oder Reviews allein hätten (Nazlan et al., 2018, S. 457). Aufgrund der zunehmenden Nutzung von Bewertungen im Bereich von Chatbots ist es wichtig, deren Einfluss in diesem Kontext zu untersuchen.

Tabelle 5: Auswirkungen von Bewertungen

Unabhängige Variable	Mediatoren	Moderatoren	Abhängige Variable	Quelle
Bewertung			Vertrauen	Gavilan et al., 2018
			Unsicherheiten, Risiken	Bae & Lee, 2011; Hong et al., 2017
			Erwartungen	Nath et al., 2018
			Qualität	Browning et al., 2013
		Erfahrung	Nützlichkeit	Mudambi & Schuff, 2010
		Marke oder Produkt	Wirksamkeit der Bewertung	Ho-Dac et al., 2013
		Bekanntheit des Produkts		Zhu & Zhang, 2010
		Produktart	Nützlichkeit, Glaubwürdigkeit	Bae & Lee, 2011; Cui et al., 2012; Guha Majumder et al., 2022
		Kommunikationskanal	Vertrauenswürdigkeit	Filieri, 2016
Rating		Anzahl positiver oder negativer Ratings	Entscheidungsprozess	LaCour & Serra, 2022
			Vertrauen	Kostyk et al., 2017
			Nützlichkeit Bewertung	Guha Majumder et al., 2022
Review			Kauf-, Nutzungsabsicht, Buchungsverhalten	Arief et al., 2023; Ye et al., 2009
			Produktkauf	Park et al., 2007
		Involvement, Erfahrung	Vertrauenswürdigkeit	Utz et al., 2012
				Filieri, 2016
	Nützlichkeit		Vertrauen, Kaufabsicht	Elwalda et al., 2016; Ventre & Kolbe, 2020; Zhao et al., 2015
	Qualität Review		Einstellung	Lee et al., 2008
Anzahl Reviews		Rating	Vertrauenswürdigkeit	Gavilan et al., 2018
			Produktkauf	Duan et al., 2008
			Kaufabsicht	Cui et al., 2012; Zhao et al., 2015
			Einstellung	Lee et al., 2008
Qualität Review			Kaufabsicht	Park et al., 2007

Tabelle 5: Auswirkungen von Bewertungen (Fortsetzung)

Unabhängige Variable	Mediatoren	Moderatoren	Abhängige Variable	Quelle
Gestaltung Review			Vertrauenswürdigkeit	Filieri, 2016
			Wahrgenommener Wert des Reviews	Schindler & Bickart, 2012
			Nützlichkeit	Baek et al., 2012; Guha Majumder et al., 2022; Hong et al., 2017; Mudambi & Schuff, 2010; Racherla & Friske, 2012; Zhao et al., 2015; Zheng, 2021

5 Experimentelle Untersuchungen

In Kapitel 5 werden vier Forschungsprojekte vorgestellt, welche Einflussfaktoren auf die Wiederverwendungsabsicht von Chatbots untersuchen. Bevor die Forschungsprojekte in Kapitel 5.3 bis Kapitel 5.6 beschrieben und deren Ergebnisse dargestellt und diskutiert werden, gibt das folgende Unterkapitel (Kapitel 5.1) zunächst einen Überblick über die vier Forschungsprojekte. Anschließend wird in Kapitel 5.2 auf Grundlagen zu experimenteller Forschung sowie der in dieser Arbeit verwendeten Methoden eingegangen.

Die Kapitel 5.3 bis 5.6 sind strukturell ähnlich aufgebaut. Zunächst wird der Hintergrund der jeweiligen Forschungsprojekte erläutert und die Forschungsfrage des Projekts kurz vorgestellt. Dabei umfassen *FP II* und *FP IV* jeweils mehreren Experimenten. *FP I* und *FP III* bestehen jeweils aus nur einem Experiment. Zuerst werden die zu untersuchenden Hypothesen hergeleitet. Im Folgenden wird die Methodik erläutert, einschließlich des Aufbaus der durchgeführten Experimente. Danach werden die Merkmale der Stichprobe analysiert und die verwendeten Konstrukte einer Güteprüfung unterzogen. Bevor die Ergebnisse der statistischen Analyse diskutiert werden, werden diese zunächst in übersichtlicher Form dargelegt.

5.1 Überblick der Forschungsprojekte

Ein entscheidendes Kriterium für den Erfolg von Produkten und Dienstleistungen ist die Absicht der Nutzer, diese erneut zu verwenden. Oftmals ist es das Ziel, Kunden langfristig an ein Unternehmen zu binden und aus einmaligen Kunden wiederkehrende Kunden zu generieren. Hierbei stellt die Wiederverwendungsabsicht eine entscheidende Rolle dar. Kunden, die nicht gewillt sind, den Service wiederzuverwenden, werden ohne zusätzliche Bemühungen des Unternehmens dies auch nur in Ausnahmefällen tun, beispielsweise bei fehlenden Alternativen. Gerade auch für Technologien wie Chatbots, die erst in den letzten Jahren vermehrt im Kundenservice eingesetzt werden, ist die Wiederverwendungsabsicht ein entscheidendes Kriterium für den Erfolg. Der Wiederverwendungsabsicht von Chatbots wurde jedoch in der bisherigen Literatur nur wenig Beachtung geschenkt. Die Forschungsprojekte dieser Arbeit widmen sich in einzelnen Studien den Einflussfaktoren auf die Absicht der Wiederverwendung eines Chatbots.

FP I beschäftigt sich mit Anthropomorphismus von Chatbots und dessen Auswirkung auf die Wiederverwendungsabsicht. Hier steht die Frage im Mittelpunkt, wie Anthropomorphismus auf die Wiederverwendungsabsicht eines Chatbots wirkt. Hierbei wird die mediierende Funktion der Sozialen Präsenz sowie des Vertrauens überprüft. Darüber hinaus dient dieses Forschungsprojekt als Grundlage für die weiteren Forschungsprojekte. Die Gestaltung des Chatbots wird für die folgenden Forschungsprojekte am anthropomorphen Design dieses Forschungsprojekts angelegt.

FP II beschäftigt sich mit Fragen zur Offenlegung der nicht-menschlichen Chatbot-Identität. Genauer mit dem expliziten Zeitpunkt der Offenlegung. Hier wird zwischen einer Offenlegung vor und einer Offenlegung nach der Interaktion unterschieden. Fehlende Offenlegung wird teilweise als Referenzkategorie genutzt. Das erste Experiment untersucht den Einfluss des Offenlegungszeitpunkts auf die Wiederverwendungsabsicht unter Beachtung der Sozialen Präsenz als Mediator. Im zweiten Experiment wird die Kommunikationsplattform (Internet oder WhatsApp), auf der die Interaktion mit dem Chatbot stattfindet, als Moderator betrachtet. Dabei wird ein moderiertes Mediationsmodell mit der Sozialen Präsenz als Mediator und der Kommunikationsplattform als Moderator untersucht. Experiment drei beschäftigt sich mit weiteren Mediatorvariablen. Dort werden der wahrgenommene Betrug, die wahrgenommene Fairness und das vermutete Motiv betrachtet. Hier wird ein moderiertes Mediationsmodell untersucht, das die Lösung der Anfrage als Moderator betrachtet. Ziel ist es den Einfluss des Offenlegungszeitpunkts auf die Wiederverwendungsabsicht unter Beachtung der Lösung der Anfrage (keine Lösung, Lösung) und die Rolle der drei Mediatoren zu untersuchen.

FP III betrachtet die moderierende Rolle des Wechsels von einem Chatbot zu einem Menschen, also den Einsatz eines hybriden Serviceagenten. Dabei wird der moderierende Einfluss des Wechsels auf die Wirkung des Offenlegungszeitpunkts auf die Wiederverwendungsabsicht untersucht. So stellt sich die Frage, wer der Überbringer von schlechten Nachrichten sein sollte.

FP IV fokussiert auf Bewertungen von Chatbots und deren Auswirkungen auf die Wiederverwendungsabsicht. Dabei wird in zwei Experimenten ein seriell Mediationsmodell untersucht, das die Erwartungen, die Nutzungsabsicht sowie die Zufriedenheit als Mediatoren betrachtet. Dabei unterscheiden sich die beiden Experimente hinsichtlich der Manipulation der Bewertung. In Experiment eins wird lediglich ein Rating betrachtet. Unterschieden wird dabei zwischen einem guten, mittleren und schlechten Rating. Eine Kontrollgruppe, die kein Rating gesehen

hat, dient als Referenz. In Experiment zwei wird die Manipulation der Bewertung noch durch ein zusätzliches Review ergänzt. Hierbei wird in Bezug auf das Rating und Review nur noch zwischen gut und schlecht differenziert. Abbildung 3 gibt einen Überblick der vier Forschungsprojekte mit den insgesamt sieben Experimenten. Die Wiederverwendungsabsicht wurde in allen Forschungsprojekten als abhängige Variable betrachtet.

FP I: Anthropomorphismus	FP III: Lösung
Mediatoren: Soziale Präsenz und Vertrauen	Moderator: Übergabe an Mensch
FP II: Offenlegungszeitpunkt	FP IV: Bewertung
Experiment 1: Mediator: Soziale Präsenz Experiment 2: Mediator: Soziale Präsenz Moderator: Kommunikationsplattform Experiment 3: Mediatoren: Wahrgenommener Betrug, Wahrgenommene Fairness, Vermutetes Motiv Moderator: Lösung	Mediatoren: Erwartung, Nutzungsabsicht, Zufriedenheit Experiment 1: Rating Experiment 2: Rating und Review

Abbildung 3: Übersicht der Forschungsprojekte

5.2 Grundlagen zu den Experimenten und deren Auswertung

Folgend werden allgemeinen Charakteristika von Experimenten, die Gütebeurteilung in Bezug auf Reliabilität und Validität sowie Verfahren zur Auswertung der experimentellen Daten beschrieben. Dabei wird ein Fokus auf die für diese Arbeit relevanten Themenbereiche gelegt. In diesem Kapitel werden die in den einzelnen Experimenten verwendeten Methoden übersichtlich beschrieben.

5.2.1 Charakteristika von Experimenten

Wissenschaftliche Experimente stellen in den Wirtschaftswissenschaften, insbesondere im Marketing, eine etablierte Forschungsmethode dar (Koschate-Fischer & Schandelmeier, 2014, S. 794). Wissenschaftliche Experimente gehören zu den quantitativen Forschungsmethoden. Sie testen Hypothesen und statistische Beziehungen auf Grundlage der gemessenen und quantifizierten Beziehungen zwischen Variablen. Dabei unterscheiden sich wissenschaftliche Experimente im Hinblick auf das Forschungsvorhaben, der Anzahl der abhängigen und unabhängigen Variablen und der Struktur des betrachteten Forschungskontexts. Häufig werden neben den Haupteffekten auch Interaktionseffekte zwischen Variablen untersucht (Koschate-Fischer &

Schandelmeier, 2014, S. 796). Der Zusammenhang zwischen einer unabhängigen und einer abhängigen Variablen kann von Mediator- oder Moderatorvariablen beeinflusst werden. Eine Moderatorvariable kann die Richtung oder die Stärke der Beziehung beeinflussen (Baron & Kenny, 1986, S. 1174), wohingegen eine Mediatorvariable vermittelnd zwischen einer abhängigen und einer unabhängigen Variable wirkt. Moderatorvariablen geben an, wann bestimmte Effekte auftreten. Mediatorvariablen hingegen beschreiben, warum solche Effekte auftreten (Baron & Kenny, 1986, S. 1176).

Wissenschaftliche Experimente sind durch die systematische Variation einer oder mehrerer unabhängiger Variablen, die auch als Faktoren bezeichnet werden und der Bestimmung ihrer Auswirkungen auf eine oder mehrere abhängige Variablen gekennzeichnet. Dabei werden Störgrößen, also Einflüsse von Fremdvariablen, kontrolliert und gegebenenfalls eliminiert. In wissenschaftlichen Experimenten werden Kontroll- und Experimentalgruppen eingesetzt, um die hypothetisierten Beziehungen zwischen unabhängigen und abhängigen Variablen zu testen (Kumar et al., 2019, S. 321). Die Teilnehmer werden einer der Experimental- oder der Kontrollgruppe zufällig zugeordnet. Dadurch können externe Variablen kontrolliert und die Reliabilität der Experimente erhöht werden (Kumar et al., 2019, S. 321). In wissenschaftlichen Experimenten wird den Teilnehmern eine Situation strukturiert beschrieben, in die sie sich gedanklich hineinversetzen sollen. Dabei wird die unabhängige Variable systematisch verändert. Diese Manipulation geschieht auf Grundlage vorheriger theoretischer oder sachlogischer Überlegungen des Anwenders. Dieses Vorgehen wird als Szenario-Technik beschrieben. Eine weitere Möglichkeit, die unabhängige Variable zu variieren, stellt die Event-Technik dar. Hier werden die Teilnehmenden direkt mit einem Ereignis konfrontiert (Koschate-Fischer & Schandelmeier, 2014, S. 809). Alle experimentellen Untersuchungen dieser Arbeit wurden mithilfe der Szenario-Technik umgesetzt. Die Teilnehmer sollten sich in das beschriebene Szenario hineinversetzen und anschließend Fragen dazu beantworten. Die Szenario-Technik ist in der Marketing-Forschung und explizit in der Chatbot-Forschung ein häufig genutztes Instrument, um Kausalzusammenhänge zu überprüfen (Crollic et al., 2022; Meyer et al., 2022; Mozafari et al., 2022).

Abhängige Variablen werden im Rahmen wissenschaftlicher Experimente operationalisiert, dies geschieht in der Regel anhand von Indikatoren, die in der Regel direkt beobachtbar sind und zu einem gesamten Konstrukt zusammengefasst werden können (Koschate, 2008, S. 112). Zur Operationalisierung der abhängigen Variablen wurde in allen Experimenten dieser Arbeit

auf Konstrukte aus wissenschaftlichen Publikationen zurückgegriffen. Diese mussten für die Experimente ins Deutsche übersetzt werden. Hierfür wurde von mehreren Forschern unabhängige Übersetzungsvorschläge gemacht. Bei Unterschieden wurde in einer Forschergruppe die geeignetste Übersetzung gefunden. Die verwendeten Indikatoren wurden in den Experimenten je Konstrukt randomisiert. Die Verwendung von Konstrukten aus wissenschaftlichen Publikationen hat den Vorteil, dass die verwendeten Skalen bereits auf Reliabilität (Zuverlässigkeit) und Validität (Gültigkeit) getestet wurden (Koschate-Fischer & Schandelmeier, 2014, S. 811). Trotzdem müssen die Konstrukte mit ihren Indikatoren noch einmal auf diese beiden Gütekriterien geprüft werden, um zu beurteilen, inwiefern die Messung durch die Indikatoren ein theoretisches Konstrukt erfassen kann (Koschate, 2008, S. 112) und ob die erhobenen Daten verwendet werden können, um aussagekräftige Ergebnisse zu erhalten.

Wissenschaftliche Experimente werden mithilfe von Versuchsplänen (Designs) durchgeführt, dabei lassen sich vier Gruppen von Versuchsplänen unterscheiden: Versuchspläne mit Zufallsgruppenbildung (Between-Subjects-Design), Versuchspläne mit Wiederholungsmessung (Within-Subjects-Design), Versuchspläne mit parallelisierten Gruppen (Blockversuchspläne) und Mischversuchspläne (Koschate, 2008, S. 115-116). Im Rahmen der Experimente dieser Arbeit werden Versuchspläne mit Zufallsgruppenbildung verwendet. Dabei werden die Teilnehmer randomisiert auf zwei oder mehr Untergruppen aufgeteilt, wobei jeder Teilnehmer genau eine Gruppe durchläuft (Koschate, 2008, S. 115-117).

Wissenschaftliche Experimente werden allgemein zwischen Labor-, Feld- und Onlineexperiment unterschieden. Im Rahmen von Laborexperimenten wird eine speziell geschaffene Umgebung genutzt, wohingegen Feldexperimente in einer realen, natürlichen Umgebung stattfinden. Nehmen Teilnehmer über das Internet an Experimenten teil, spricht man von Onlineexperimenten (Koschate-Fischer & Schandelmeier, 2014, S. 798). Eine kleine Einschränkung von Onlineexperimenten ist, dass die Umgebung im Vergleich zu einem Laborexperiment nicht kontrolliert werden kann. So können beispielsweise Ablenkungen durch Umweltreize nicht ausgeschlossen oder kontrolliert werden. Die Experimente dieser Arbeit wurden in Form von Onlineexperimenten durchgeführt.

Alle Experimente dieser Arbeit wurden über die Online-Software Unipark (EFS Survey) der Tivian XI GmbH, eine professionellen Online-Befragungssoftware, umgesetzt. Die Teilnehmer

wurden an der RPTU Kaiserslautern-Landau akquiriert. Dabei wurden auch Plattformen wie SurveyCircle und Soziale Medien verwendet, um Teilnehmer für die Experimente zu gewinnen.

5.2.2 Gütebeurteilung der Konstrukte

Die im Rahmen der Experimente erfassten Daten wurden mithilfe von SPSS ausgewertet. Nachdem die Rohdaten in ein bearbeitbares Format gebracht und bereinigt wurden, wurde die Güte der Messung durch Reliabilität und Validität beurteilt. Reliabilität ist ein Maß, das beschreibt, ob wiederholte Messungen mit einem Messinstrument ein konsistentes Ergebnis liefern und somit frei von Zufallsfehlern sind. Die Reliabilitätsprüfung schätzt somit, inwiefern das Messinstrument frei von Zufallsfehlern ist. Systematische Fehler können allerdings weiter vorliegen (Peter, 1979, S. 6; Weiber & Sarstedt, 2021, S. 147). Ist die Messung frei von Zufallsfehlern sowie systematischen Fehlern spricht man von Validität. Das Messinstrument ist somit gültig und misst genau das, was es theoretisch messen soll (Weiber & Sarstedt, 2021, S. 167). Reliabilität ist eine notwendige, aber nicht hinreichende Bedingung für Validität (Peter, 1979, S. 6).

In dieser Arbeit wurde zur Prüfung der Reliabilität und Validität auf etablierte Gütekriterien der ersten und der zweiten Generation zurückgegriffen. Eine wichtige Voraussetzung, um mit den Indikatoren weiterzurechnen und daraus ein Konstrukt (Abhängige Variable, Moderations- oder Mediationsvariable) zu berechnen, bildet die Eindimensionalität. Also, dass alle Indikatoren einen Faktor beschreiben (Gerbing & Anderson, 1988, S. 186). Die Prüfung der Eindimensionalität geschieht mithilfe einer explorativen Faktorenanalyse. Die Anwendung einer explorativen Faktorenanalyse ist nur dann sinnvoll, wenn zwischen den Indikatoren eine hinreichend große Korrelation besteht (Weiber & Sarstedt, 2021, S. 142). Daher wird zunächst das Kaiser-Meyer-Olkin-Kriterium (KMO) betrachtet, das Auskunft über die Zusammengehörigkeit der Variablen gibt (Weiber & Sarstedt, 2021, S. 143). Sind die KMO-Werte kleiner 0,5 werden diese nicht akzeptiert. Sie sollten sogar größer als 0,6 sein (Kaiser & Rice, 1974, S. 112). Das Ergebnis der explorativen Faktorenanalyse liefert Informationen über die zugrundeliegende Faktorstruktur, also darüber, ob die Indikatoren auf einen oder mehrere Faktoren laden. Dadurch kann die Eindimensionalität nachgewiesen werden. Dabei müssen die Faktorladungen den geforderten Grenzwert von 0,4 überschreiten (Homburg & Giering, 1996, S. 8).

Zur Prüfung der Reliabilität eines Messinstruments wird häufig Cronbachs Alpha verwendet, welches die interne Korrelationsstärke misst (Peter, 1979, S. 8). Je näher der Wert von Cronbachs Alpha an eins liegt, desto stärker ist die Korrelation zwischen den Indikatoren, also die interne Konsistenz (Weiber & Sarstedt, 2021, S. 148). Als Grenzwert für eine gute Reliabilität wird in der Literatur ein Cronbachs Alpha von 0,7 beschrieben (Nunnally & Bernstein, 1994, S. 252). Darüber hinaus stellt die korrigierte Item-to-Total-Korrelation (KITK) eine weitere Möglichkeit der Reliabilitätsprüfung dar und beschreibt die Trennschärfe (Weiber & Sarstedt, 2021, S. 148-150). Hiermit können einzelnen Indikatoren gegebenenfalls ausgeschlossen werden, um die Reliabilität zu erhöhen und einen zu niedrigen Cronbachs Alpha Wert zu verbessern. Dabei wird die Indikatorvariable mit dem niedrigsten KITK-Wert eliminiert (Homburg & Giering, 1996, S. 8-9). Die KITK-Werte können zwischen minus eins und plus eins liegen. Ein Indikator sollte ausgeschlossen werden, wenn der KITK-Wert kleiner als 0,5 ist (Bearden et al., 1989, S. 475; Weiber & Sarstedt, 2021, S. 150).

Explorative Faktorenanalyse, KITK und Cronbachs Alpha zählen zu den Gütekriterien der ersten Generation. Sie erlauben keine Schätzung der Messfehler und Modellparameter (Weiber & Sarstedt, 2021, S. 140). Der Messfehler wird bei den Gütekriterien zweiter Ordnung betrachtet, indem ein Vergleich zwischen der Varianz eines Indikators und der Varianz des Messfehlers vorgenommen wird. Je größer die erklärte Varianz ist, desto besser ist das Kriterium (Weiber & Sarstedt, 2021, S. 157). Für die erklärte Varianz wird in der Literatur ein Schwellenwert von 50 % beschrieben (Ebert & Raithel, 2009, S. 127).

Die einzelnen Indikatoren eines Messinstruments müssen zu Konstrukten mit entsprechenden Faktorwerten zusammengefasst werden, um sie in der weiteren Analyse als Variablen verwenden zu können. Faktorwerte mit den zugrundeliegenden Skalen wurden im Rahmen dieser Arbeit aus den Indikatoren mithilfe der Faktorenanalyse berechnet. Die Nutzung der Faktorenanalyse hat den Vorteil, dass dort die Faktorladungen berücksichtigt werden. Hierzu wurde in SPSS bei der Faktorenanalyse unter Factorscore ausgewählt, dass dieser als Variable, also als neues Konstrukt gespeichert werden soll. Als Methode wurde hier eine Regression gewählt. Die Faktorwerte ergeben sich aus der Faktorenanalyse als gewichtete Mittelwerte der standardisierten Werte der in die Hauptkomponentenanalyse eingehenden Indikatoren. Dabei sind die Faktorwerte zentriert und standardisiert (Lüken & Schimmelpfennig, 2020, S. 34). Die so berechneten Skalen haben dann einen Mittelwert von null. Folglich zeigen Werte größer als null, dass die Befragten hinsichtlich der Dimension überdurchschnittlich geantwortet haben und bei Werten

kleiner als null, dass sie unterdurchschnittlich geantwortet haben. Bei einer Zufriedenheitsskala würden beispielsweise Werte unter null auf eine geringere und Werte über null auf eine höhere Zufriedenheit hinweisen (Lüken & Schimmelpfennig, 2020, S. 34).

Eine weitere Möglichkeit zur Bestimmung des Konstrukts mit den entsprechenden Faktorwerten ist die Berechnung über den Mittelwert. Hierbei wird der Mittelwert mithilfe der Indikatoren berechnet, die auf den gleichen Faktor laden. Ein Vorteil dieser Methode ist, dass die Daten und die Ergebnisse weiterer Analysen unmittelbar mit der ursprünglichen Skala interpretiert werden können (Lüken & Schimmelpfennig, 2020, S. 34). Sollte an einer Stelle dieser Arbeit aus Interpretationsgründen der Mittelwert zur Berechnung der Konstrukte verwendet worden sein, wird dies angegeben.

5.2.3 Auswertung der Daten

Die Daten der einzelnen Forschungsprojekte und respektive der Experimente werden mithilfe unterschiedlicher statistischer Verfahren ausgewertet. Zum einfachen Mittelwertvergleich zweier Gruppen dient der t-Test (Backhaus et al., 2021, S. 181; Kreis et al., 2021, S. 275). Er hilft herauszufinden, ob Mittelwertunterschiede systematisch erklärbar sind oder durch zufällige Gruppenunterschiede zustande kommen (Kreis et al., 2021, S. 275). Zur Einschätzung der Effektstärke bei Mittelwertunterschieden können Effektgrößen wie Cohen's d oder Hedges' g herangezogen werden (Kreis et al., 2021, S. 277). Cohen's d Werte um 0,2 beschreiben eine geringe Effektstärke. Die Ergebnisse des t-Test sollten dann nicht interpretiert werden, denn meist mangelt es an experimenteller und/oder messtechnischer Kontrolle. Werte ab 0,5 zeigen eine mittlere Effektstärke und ab 0,8 wird von einer hohen Effektstärke ausgegangen (Cohen, 1988, S. 25-26).

Hat eine unabhängige Variable mehr als zwei Ausprägungen oder werden mehrere unabhängige Variablen betrachtet, reichen einfache Mittelwertvergleiche nicht aus. In solchen Fällen können die Mittelwerte mithilfe einer Varianzanalyse (ANOVA) betrachtet werden. Eine ANOVA wird verwendet, um die Zusammenhänge zwischen einer metrischen abhängigen Variable und einer oder mehrerer nominal skalierten unabhängigen Variablen zu überprüfen (Backhaus et al., 2021, S. 162; Böhler et al., 2021, S. 199). Die Bezeichnung einer ANOVA richtet sich nach der Anzahl der einbezogenen Faktoren, sodass zwischen ein-, zwei-, dreifaktoriellen und weiteren Va-

rianten unterschieden wird (Backhaus et al., 2021, S. 165). Die Ergebnisse einer ANOVA beschreiben, ob der betrachtete Faktor einen signifikanten Einfluss auf die abhängige Variable besitzt, also Mittelwertunterschiede vorliegen und wie groß er Erklärungsbeitrag des untersuchten Modells ist (Backhaus et al., 2021, S. 180). Hieraus kann allerdings noch nicht geschlossen werden, wie genau sich die Gruppen voneinander unterscheiden. Hierfür können Post-hoc-Tests verwendet werden. Wurden Mittelwertunterschiede festgestellt, kann mithilfe der Post-hoc-Tests festgestellt werden, welche Faktorstufen Mittelwertunterschiede begründen. In der Praxis findet man häufig den Bonferroni-, Scheffe- oder den Tukey Test (Backhaus et al., 2021, S. 181-182).

Liegen zwei oder mehr Faktoren vor, kann die ANOVA auch Aussagen über mögliche Interaktionseffekte liefern. Liegen bei mehrfaktoriellen Versuchsplänen Interaktionseffekte vor, können diese nicht direkt interpretiert werden. Eine Lösungsmöglichkeit bietet die Simple-Effects-Analyse. Dabei kann der Haupteffekt für alle Faktorstufen einer anderen unabhängigen Variablen betrachtet werden und dadurch die Interaktionseffekte verdeutlicht werden (Field, 2018, S. 635). Eine Moderation beschreibt die Randbedingungen eines Zusammenhangs zwischen zwei Variablen. Dabei hängt die Größe oder das Vorzeichen des Zusammenhangs zwischen zwei Variablen von dem Moderator oder der Moderatoren ab (Hayes, 2022, S. 8). Statistisch gesehen ist eine Moderation eine Interaktion (Field, 2018, S. 483).

Neben den direkten Effekten und der moderierenden Funktion können Variablen auf medierende Art wirken. Dabei ist ein Mediator eine intervenierende Variable zwischen der unabhängigen und der abhängigen Variablen, die einen Teil oder den gesamten Effekt der unabhängigen Variable auf die abhängige Variable erklärt (Hayes, 2022, S. 7; Hayes et al., 2017, S. 76). Mediations- und Moderationsanalysen können mithilfe des PROCESS-Makros für SPSS gerechnet werden und liefern dabei gut interpretierbare Ergebnisse (Hayes et al., 2017, S. 77). Dabei existieren bei PROCESS unterschiedliche Modelle, um von einfachen Moderationen über einfache Mediationen bis hin zu seriellen Mediationen oder moderierten Mediationen statistische Auswertungen in SPSS auszuführen. Im Rahmen der Analysen verwendet PROCESS zur Schätzung der Parameter regressionsbasierte Pfadanalysen (Hayes et al., 2017, S. 77). Um Konfidenzintervalle zu berechnen und dadurch Aussagen über die Signifikanz der untersuchten Auswirkungen treffen zu können, wird in PROCESS Bootstrapping durchgeführt. Bootstrapping ist in der statistischen Analyse weit verbreitet und hat sich gegenüber dem Sobel Test als überlegen

erwiesen (Hayes, 2015, S. 8). Dabei sollten Bootstrap Sample zwischen 1.000 und 10.000 gewählt werden (Hayes, 2015, S. 8, 2022, S. 100). Im Rahmen dieser Arbeit wurde ein Bootstrap Sample von 5.000 gewählt, womit sich an dem Vorgehen von Hayes (2022, S. 100) orientiert wurde.

5.3 Forschungsprojekt I: Anthropomorphismus

Anthropomorphe Chatbots nehmen eine immer größer werdende Rolle im Kundenservice ein. Wie in Kapitel 4.1 gezeigt, sind die Soziale Präsenz sowie das Vertrauen Faktoren, die durch Anthropomorphismus beeinflusst werden können. Auch die Wiederverwendungsabsicht von Chatbots kann durch die Soziale Präsenz sowie das Vertrauen erhöht werden. Daher beschäftigt sich dieses Forschungsprojekt mit der Frage: Welchen Einfluss hat die anthropomorphe Gestaltung von Chatbots auf die Wiederverwendungsabsicht und welche mediiierende Funktion haben die Soziale Präsenz sowie das Vertrauen? Ziel des ersten Forschungsprojekts ist es den Einfluss von Anthropomorphismus auf die Soziale Präsenz und das Vertrauen in einen Chatbot sowie die Wiederverwendungsabsicht zu untersuchen. Hiermit wird zum aktuellen wissenschaftlichen Diskurs beigetragen, indem bestehende Forschungsergebnisse unterstützt werden und durch die serielle Mediation ergänzt werden. Darüber hinaus soll ein anthropomorphes Chatbot Design gefunden werden, das in den weiteren Forschungsprojekten in überarbeiteten Versionen verwendet wird. Dabei werden in der Literatur identifizierte Faktoren, wie die Redensart des Chatbots sowie die Gestaltung des Avatars genutzt, um den Grad der anthropomorphen Gestaltung des Chatbots zu variieren.

5.3.1 Herleitung der Forschungshypothesen

Nach der Media Equation Theorie (Reeves & Nass, 1996) sowie der Social Response Theorie (Nass & Moon, 2000; Nass et al., 1994b) reagieren Menschen auf soziale Weise, wenn sie mit Computern oder neuen Kommunikationstechnologien interagieren. Menschen wenden emotionale Reaktionen und soziale Regel auf Kommunikationstechnologien mit menschlichen Eigenschaften an. Kommunikationstechnologien wie beispielsweise Roboter oder Chatbots werden soziale Eigenschaften zugeschrieben (Nass & Moon, 2000, S. 83; Nass et al., 1994b, S. 77; Reeves & Nass, 1996, S. 5). Sozialen Regeln werden gedankenlos in Interaktionen mit solchen Kommunikationstechnologien angewandt, wenn sie anthropomorphe Hinweise enthalten (Kim

et al., 2013, S. 1800; Nass & Moon, 2000, S. 83). Die Zuschreibung menschlicher Eigenschaften erhöht die Motivation, die Handlungen der Technologie zu verstehen und verringert die Unsicherheit, die mit der Nutzung einer Technologie einhergeht. Darüber hinaus kann das Vertrauen höher sein, wenn der Technologie menschliche Eigenschaften zugeschrieben werden. Anthropomorphismus stellt eine Möglichkeit dar, um Nutzerreaktionen auf Technologien zu beeinflussen. Auch das zukünftige Nutzungsverhalten kann damit beeinflusst werden (Epley et al., 2007, S. 866). Die soziale Orientierung von Menschen ist stärker ausgeprägt, wenn die Technologie menschenähnliche Eigenschaften aufweist (Huang & Lin, 2011, S. 1965-1966; Moon, 2000, S. 325; Nass & Moon, 2000, S. 83). Anthropomorph gestaltete Chatbots können daher als soziale Akteure wahrgenommen werden (Kim et al., 2013, S. 1800). Deshalb kann die Soziale Präsenz des Chatbots höher sein, wenn der Chatbot anthropomorph gestaltet ist (Blut et al., 2021, S. 638).

Daraus folgend kann das Bedürfnis von Menschen nach sozialer Verbindung (Epley et al., 2008b, S. 144) durch anthropomorphe Chatbots befriedigt werden. Nutzer werden Chatbots, die sie als reale soziale Akteure wahrnehmen, eher vertrauen und auch eher gewillt sein, diese wiederzuverwenden. Aus diesen Überlegungen ergeben sich die folgenden drei Hypothesen:

- H₁: Anthropomorphismus erhöht die Wiederverwendungsabsicht von Chatbots.
- H₂: Die Soziale Präsenz mediiert den positiven Einfluss des Anthropomorphismus auf die Wiederverwendungsabsicht des Chatbots.
- H₃: Die Soziale Präsenz mediiert den positiven Einfluss des Anthropomorphismus auf das Vertrauen in den Chatbot. Das Vertrauen in den Chatbot mediiert den positiven Einfluss der Sozialen Präsenz auf die Wiederverwendungsabsicht des Chatbots.

5.3.2 Methodik

In *FP I* wird die Szenario-Technik genutzt, um den Einfluss des Anthropomorphismus auf die Soziale Präsenz, das Vertrauen sowie die Wiederverwendungsabsicht zu untersuchen. In dem folgenden Experiment sollen sich die Teilnehmer in die hypothetische Situation hineinversetzen, einen Kurzurlaub bei einem Online-Reiseportal buchen zu wollen. Dabei bietet das Online-Reiseportal die Möglichkeit der Buchung mithilfe eines Chatbots an, der als Serviceagent agiert und den Kunden durch den Buchungsprozess begleitet. Hier fragt der Chatbot ähnlich wie in einem Verkaufsgespräch im Reisebüro sukzessive nach den Eckdaten der Reise und schlägt den

Teilnehmern am Ende der Konversation eine Vorauswahl an Unterkünften vor. Nach der Beschreibung des Szenarios wird den Teilnehmern ein Screenshot der Chatinteraktion gezeigt, den sie aufmerksam ansehen und durchlesen sollen. Am Ende der Chatinteraktion schlägt der Chatbot eine Trefferliste vor, die unter einem angegebenen Link abgerufen werden kann. Anschließend sollten sich die Teilnehmer vorstellen, dem Link zu folgen und dort eine entsprechende Unterkunft auszuwählen, sodass von einer erfolgreichen Buchung ausgegangen werden kann. Hierbei wird ein einfaktorielles Design mit Zufallsgruppenbildung verwendet, bei dem die anthropomorphen Gestaltungsmerkmale des Chatbots manipuliert werden, so dass sich zwei Gruppen ergeben: *Anthropomorphismus_gering* und *Anthropomorphismus_hoch*. Der Chatbot stellt sich im Begrüßungstext als virtueller Assistent vor: „Hallo, ich bin Ihr virtueller Assistent.“ oder „Hallo, ich bin Tina, deine virtuelle Assistentin.“ Die Chatinteraktionen und das Profilbild des Chatbots der beiden Gruppen wurden mithilfe der in Kapitel 4.1 gefundenen Einflussfaktoren auf Anthropomorphismus entworfen. Dabei wurde die Redensart (mechanisch/künstlich, menschlich), Name sowie das Profilbild (Avatar, Mensch) als zu manipulierende Variablen ausgewählt. Die Teilnehmer wurden zufällig einer der zwei Experimentalgruppen zugeordnet. Das Szenario ist in Tabelle 6 übersichtlich dargestellt.

Tabelle 6: Szenariobeschreibung (FP I)

Einführung	
Stellen Sie sich bitte vor, Sie heißen Kim und planen einen Kurzurlaub zu zweit in Hamburg. Daher suchen Sie bei einem Online Reiseportal nach einer Unterkunft, die Ihren Wünschen entsprechen soll. Das Online Reiseportal bietet den Service an, sich von einem Chatbot, also einem virtuellen Assistenten, helfen zu lassen. Sie entscheiden sich dafür, diesen Service zu nutzen, wobei sich folgende Konversation zwischen Ihnen und dem Chatbot entwickelt:	
Gruppen	
Anthropomorphismus_gering	Anthropomorphismus_hoch
Chatinteraktion	
• Hallo ich bin ihr virtueller Assistent. Wie kann ich Ihnen helfen? • Ich suche eine Unterkunft in Hamburg. • In welchem Zeitraum? • Vom 18.07.22 bis 22.07.22 • Für wie viele Personen? • 2 Personen • Geben Sie zusätzliche Wünsche an, wie z. B. Frühstück inbegriffen, Parkplatz vorhanden, Innenstadtnähe • Frühstück inbegriffen, Parkplatz vorhanden • Wie viel darf die Unterkunft pro Nacht kosten? • Maximal 100 Euro • Ich habe 20 Treffer gefunden. Hier gelangen Sie zur Trefferliste: https://www.die-beste-unterkunft.de/hamburg/2345	• Hallo ich bin Tina, deine virtuelle Assistentin. Ich werde dir dabei helfen, deine Lieblingsunterkunft zu finden! Wie kann ich dir helfen? • Ich bin Kim und suche eine eine Unterkunft in Hamburg. • Hamburg, eine sehr schöne Stadt! In welchem Zeitraum möchtest du denn dort sein? • Vom 18.07.22 bis 22.07.22 • Vielen Dank für die Information! Wirst du eine Begleitung haben? Wie viele Personen möchten dort übernachten? • 2 Erwachsene • Hast du sonstige Wünsche zu deiner Unterkunft? Schreibe mir einfach Dinge, wie z. B.: Frühstück inbegriffen, Parkplatz vorhanden, Innenstadtnähe • Frühstück inbegriffen, Parkplatz vorhanden • Super, Danke für die Information. Wie viel darf die Unterkunft pro Nacht kosten? • Maximal 100 Euro • Alles klar Kim! Dann wollen wir uns mal die Ergebnisse anschauen. Ich habe 20 Treffer gefunden, die zu deinen Wünschen passen. Wenn du dem folgenden Link folgst, kommst du zu deiner persönlichen Trefferliste: https://www.die-beste-unterkunft.de/hamburg/2345 Viel Spaß beim Stöbern. Wenn du noch Fragen hast, kannst du mich jederzeit kontaktieren. Ich wünsche dir einen schönen Tag!

Tabelle 7: Operationalisierung der Variablen und Manipulationscheck (FP I)

Kürzel	Indikator	Skala
Manipulationscheck		
MC_1	Der Chatbot hat ein menschliches Profilbild.	7-Punkt-Likert-Skala: 1 = Stimme überhaupt nicht zu, 7 = Stimme vollkommen zu
MC_2	Der Chatbot hat mich mit „Sie“ angeredet.	
MC_3	Der Chatbot hat mich mit „Du“ angeredet.	
Anthropomorphismus (Crolic et al., 2022, S. 139)		
Bitte bewerten Sie, in welchem Ausmaß der Chatbot...		
Anth_1	...in Ihrer Vorstellung (wie eine Person) lebendig wurde.	7-Punkt-Likert-Skala: 1 = Stimme überhaupt nicht zu, 7 = Stimme vollkommen zu
Anth_2	...menschliche Eigenschaften besitzt.	
Anth_3	...eine Person zu sein scheint.	
Anth_4	...sich menschlich anfühlte.	
Anth_5	...eine Persöhnlichkeit zu haben schien.	
Anth_6	...einen eigenen Willen zu haben schien.	
Anth_7	...seine eigenen Absichten zu haben schien.	
Anth_8	...einen freien Willen zu haben schien.	
Anth_9	...ein Bewusstsein zu haben schien.	
Soziale Präsenz (Toader et al., 2020, S. 14)		
Inwiefern stimmen Sie den folgenden Aussagen zu?		
SoPr_1	Bei der Nutzung des Chatbots hatte ich das Gefühl, mit einem Menschen zu interagieren.	7-Punkt-Likert-Skala: 1 = Stimme überhaupt nicht zu, 7 = Stimme vollkommen zu
SoPr_2	Obwohl ich den Chatbot nicht im realen Leben sehen konnte, war ein Gefühl der menschlichen Wärme vorhanden.	
SoPr_3	Während der Interaktion mit dem Chatbot hatte ich ein Gefühl der Geselligkeit.	
SoPr_4	Ich hatte das Gefühl, dass es eine Person gab, die mir wirklich Sicherheit gegeben hat.	
SoPr_5	Ich hatte das Gefühl, dass es jemanden gibt, der für mich da ist, wenn ich ein Problem habe.	
Vertrauen (Bansal et al., 2005, S. 106)		
Inwiefern stimmen Sie den folgenden Aussagen zu?		
Tru_1	Ich habe das Gefühl, dass ich dem Chatbot voll und ganz vertrauen kann.	7-Punkt-Likert-Skala: 1 = Stimme überhaupt nicht zu, 7 = Stimme vollkommen zu
Tru_2	Der Chatbot ist voll und ganz aufrichtig in seinen Versprechen.	
Tru_3	Der Chatbot ist ehrlich und aufrichtig zu mir.	
Tru_4	Der Chatbot behandelt mich fair und gerecht.	
Tru_5	Ich habe das Gefühl, dass ich mich darauf verlassen kann, dass der Chatbot mir hilft, wenn ich ihn brauche.	
Tru_6 ^R	Ich habe das Gefühl, dass der Chatbot nicht ausreichend auf mich eingeht.	

Tabelle 2: Operationalisierung der Variablen und Manipulationscheck (FP I) (Fortsetzung)

Kürzel	Indikator	Skala
Wiederverwendung (Moriuchi, 2021, S. 27-28)		
Inwiefern stimmen Sie den folgenden Aussagen zu?		
Reu_1	Wenn ich Zugang zu einem Chatbot hätte, würde ich ihn ständig benutzen.	7-Punkt-Likert-Skala: 1 = Stimme überhaupt nicht zu, 7 = Stimme vollkommen zu
Reu_2	Ich denke, es wird sich für mich lohnen, Chatbots zu verwenden, wenn sie verfügbar sind.	
Reu_3	Ich würde meinen Freunden und meiner Familie die Nutzung eines Chatbots empfehlen.	
Reu_4	Wenn ich die Möglichkeit hätte mehr über Chatbots zu kaufen, würde ich dies tun.	
Reu_5	Ich nutze lieber einen Chatbot, um meine Fragen zu beantworten, als eine andere Person zu fragen.	
Wie Realistisch fanden Sie die Situation, die im Szenario dargestellt wurde?		1 = Überhaupt nicht realistisch, 7 = Sehr realistisch
Bitte beurteilen Sie, wie gut Sie sich in die beschriebene Situation hineinversetzen konnten.		1 = Es fiel mir schwer mich in die Situation hineinzuversetzen, 7 = Es fiel mir sehr leicht mich in die Situation hineinzuversetzen
R = Invers formulierter Indikator		

5.3.3 Stichprobe und Gütebeurteilung

111 Personen haben an dem Experiment teilgenommen. Diese haben sich wie folgt auf die beiden Gruppen aufgeteilt: *Anthropomorphismus_gering* 55 Teilnehmer und *Anthropomorphismus_hoch* 56 Teilnehmer. Der Manipulationscheck zeigt, dass die Teilnehmer die Manipulation richtig wahrgenommen haben. So unterscheiden sich die beiden Gruppen für die drei Manipulationscheckfragen signifikant voneinander ($p < 0,001$). Tabelle 8 zeigt die Mittelwerte der Antworten der Manipulationscheckfragen der beiden Gruppen. Hier wurden die Mittelwerte zur Berechnung des Konstrukts verwendet.

Tabelle 8: Ergebnisse des Manipulationschecks (FP I)

Gruppe	MC_1	MC_2	MC_3
Anthropomorphismus_gering	2,35	5,36	2,38
Anthropomorphismus_hoch	5,91	2,68	5,34

53,2 % der Teilnehmer sind weiblich, 45,0 % männlich und 1,8 % machen keine Angabe zu ihrem Geschlecht. Das Durchschnittsalter der Stichprobe ist 33,17 Jahre ($SD = 12,56$). Die Teilnehmer beurteilten die Situation, die im Szenario dargestellt wurde, als realistisch ($M = 5,19$, $SD = 1,56$) und es fiel ihnen leicht, sich in die Situation hineinzusetzen ($M = 5,74$, $SD = 1,43$). 74,8 % der Befragten haben mindestens einmal einen Chatbot in den letzten zwölf Monaten verwendet. 19,8 % geben sogar an, einen Chatbot häufiger als sechsmal in den letzten zwölf Monaten genutzt zu haben. Detailliertere Informationen zur Stichprobe sind in Tabelle 9 zusammengefasst. Die Verteilung der Nutzungshäufigkeit von Chatbots in den letzten zwölf Monaten der Teilnehmer ist in Anhang 1 zusammengefasst.

Tabelle 9: Soziodemographische Merkmale (FP I)

Merkmal	Ausprägung	Prozentualer Anteil
Geschlecht	Weiblich	53,2 %
	Männlich	45,0 %
	Divers	0,0 %
	Keine Angabe	1,8 %
Alter	M = 33,17 (SD = 12,56), 5 kein Angabe	
Haushaltsgröße	1 Person	18,0 %
	2 Personen	47,7 %
	3-5 Personen	32,4 %
	Mehr als 5 Personen	1,8 %
	Keine Angabe	0 %
Höchster Bildungsabschluss	Noch in schulischer Ausbildung	8,1 %
	Volks-/Hauptschulabschluss	0,9 %
	Realschulabschluss	12,6 %
	Fachschulabschluss	5,4 %
	Abitur	12,6 %
	Fachhoch-/Hochschulabschluss	60,4 %
	Keine Angabe	0 %
Monatliches Haushaltsnettoeinkommen	Bis 1000 Euro	14,4 %
	1001 - 2000 Euro	15,3 %
	2001 - 3000 Euro	27,0 %
	3001 - 4000 Euro	13,5 %
	4001 - 5000 Euro	9,0 %
	Mehr als 5000 Euro	13,5 %
	Keine Angabe	7,2 %

Ein t-Test lieferte einen signifikanten Unterschied zwischen den beiden Gruppen bezüglich der Chatbot Wahrnehmung als anthropomorph und nach Cohen (1988, S. 26) liegt eine mittlere bis große Effektstärke vor (95 % KI [-0,996, -0,271], $t(102,26) = -3,46$, $p < 0,001$, $d = 0,66$). Dabei bewerten die Teilnehmer der Gruppe *Anthropomorphismus_hoch* ($M = 0,3$, $SD = 1,09$) den Chatbot als signifikant anthropomorpher als die Teilnehmer der Gruppe *Anthropomorphismus_gering* ($M = -0,34$, $SD = 0,82$).

Die Gütebeurteilung der gemessenen Konstrukte erfolgt auf Basis der in Kapitel 5.2 beschriebenen Gütekriterien und Grenzwerte. Die Indikatoren Anth_7 und Tru_6 wurden entfernt und für die weitere Analyse nicht verwendet, da die Grenzwerte der Faktorladung (Anth_7 = 0,39,

Tru_6 = 0,30) sowie KITK (Anth_7 = 0,22, Tru_6 = 0,02) unterschritten werden. In Tabelle 10 sind die Ergebnisse der Gütebewertung nach Ausschluss der Indikatoren zusammengefasst. Hier zeigt sich, dass die Grenzwerte der Gütekriterien für Reliabilität und Validität alle überschritten sind. Daher werden die Konstrukte für die folgende Auswertung verwendet.

Tabelle 10: Gütekriterien (FP I)

Konstrukt	Indikator	Faktorladung	KITK	Chronbachs Alpha	KMO	Erklärte Varianz
		> 0,4	> 0,5	> 0,7	> 0,6	> 50 %
Anthropomorphismus	Anth_1	0,829	0,765	0,914	0,925	62,786
	Anth_2	0,804	0,734			
	Anth_3	0,837	0,775			
	Anth_4	0,780	0,705			
	Anth_5	0,844	0,782			
	Anth_6	0,599	0,510			
	Anth_8	0,777	0,704			
	Anth_9	0,840	0,776			
Soziale Präsenz	SoPr_1	0,887	0,813	0,916	0,885	75,234
	SoPr_2	0,857	0,770			
	SoPr_3	0,855	0,772			
	SoPr_4	0,889	0,820			
	SoPr_5	0,848	0,760			
Vertrauen	Tru_1	0,853	0,762	0,894	0,860	70,521
	Tru_2	0,880	0,795			
	Tru_3	0,880	0,794			
	Tru_4	0,790	0,673			
	Tru_5	0,792	0,680			
Wiederverwendung	Reu_1	0,887	0,808	0,889	0,879	70,229
	Reu_2	0,862	0,770			
	Reu_3	0,847	0,748			
	Reu_4	0,861	0,769			
	Reu_5	0,723	0,597			

5.3.4 Ergebnisse

Zur Überprüfung der Forschungshypothesen aus *FP I* wird ein serielles Mediationsmodell in SPSS gerechnet. Hierfür wird Modell 6 des PROCESS-Makros mit einem Bootstrap Sample von 5.000 verwendet (Hayes, 2022, S. 623). Das Mediationsmodell mit den entsprechenden

Regressionskoeffizienten ist in Abbildung 4 dargestellt. Als Referenzkategorie wird in diesem Experiment die Gruppe *Anthropomorphismus_gering* betrachtet.

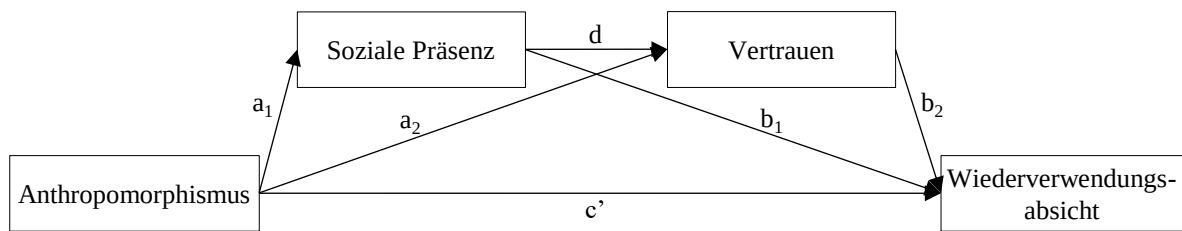


Abbildung 4: Mediationsmodell (FP I)

Der direkte Effekt von Anthropomorphismus auf die Wiederverwendungsabsicht wurde im Experiment nicht signifikant ($c' = 0,011$, $p > 0,05$). Hypothese H_1 kann folglich nicht bestätigt werden. Auch der Gesamteffekt des Anthropomorphismus auf die Wiederverwendungsabsicht des Chatbots ist nicht signifikant ($c = 0,447$, $p > 0,05$).

Der indirekte Effekt des Anthropomorphismus auf die Wiederverwendungsabsicht über die Soziale Präsenz und das Vertrauen in den Chatbot in Serie ist signifikant ($a_1 \times d \times b_2 = 0,034$, 95 % KI [0,004, 0,076]). Somit wird Hypothese H_3 durch die Mediationsanalyse bestätigt. Die beiden anderen indirekten Effekte sind ebenfalls signifikant. Hypothese H_2 kann in *FP I* ebenfalls bestätigt werden. Die Soziale Präsenz mediert die Beziehung zwischen Anthropomorphismus und Wiederverwendungsabsicht ($a_1 \times b_1 = 0,075$, 95 % KI [0,010, 0,158]). Ebenso mediert das Vertrauen die Beziehung zwischen Anthropomorphismus und der Wiederverwendungsabsicht ($a_2 \times b_2 = -0,075$, 95 % KI [-0,143, -0,022]). Die Ergebnisse der Mediationsanalyse sind in Tabelle 11 zusammengefasst. Sie gibt einen umfassenden Überblick über die Regressionskoeffizienten der Effekte auf die Wiederverwendungsabsicht. Tabelle 12 zeigt eine übersichtliche Zusammenfassung der Hypothesen.

Tabelle 11: Regressionskoeffizienten der Mediationsanalyse (FP I)

	Symbol	Koeff.	KI _u	KI _o
Effekte auf die Wiederverwendungsabsicht				
Anthropomorphismus	c'	0,011 ^{ns}	-0,211	-0,266
Soziale Präsenz	b ₁	0,348 ^{***}	0,146	0,551
Vertrauen	b ₂	0,282 ^{**}	0,075	0,490
Effekt auf die Soziale Präsenz				
Anthropomorphismus	a ₁	0,215 [*]	0,028	0,402
Effekte auf das Vertrauen				
Anthropomorphismus	a ₂	-0,264 [*]	-0,418	-0,110
Soziale Präsenz	d	0,555 [*]	0,402	0,708
	Symbol	Effekt	KI_u	KI_o
Anthropomorphismus x Soziale Präsenz	a ₁ x b ₁	0,075	0,010	0,158
Anthropomorphismus x Vertrauen	a ₂ x b ₂	-0,075	-0,143	-0,022
Anthropomorphismus x Soziale Präsenz x Vertrauen	a ₁ x d x b ₂	0,034	0,004	0,076
***: p < 0,001; **: p < 0,01; *: p < 0,05; ns: p > 0,05				

Tabelle 12: Zusammenfassung der Hypothesen (FP I)

Hypothese H _i	Ergebnis
H ₁ Anthropomorphismus erhöht die Wiederverwendungsabsicht von Chatbots.	✗
H ₂ Die Soziale Präsenz mediiert den positiven Einfluss des Anthropomorphismus auf die Wiederverwendungsabsicht des Chatbots.	✓
H ₃ Die Soziale Präsenz mediiert den positiven Einfluss des Anthropomorphismus auf das Vertrauen in den Chatbot. Das Vertrauen in den Chatbot mediiert den positiven Einfluss der Sozialen Präsenz auf die Wiederverwendungsabsicht des Chatbots.	✓
✓ = bestätigt; ✗ = nicht bestätigt	

5.3.5 Diskussion der Ergebnisse

Im Rahmen dieses Forschungsprojekts konnte kein direkter Einfluss der anthropomorphen Gestaltung auf die Wiederverwendungsabsicht gefunden werden. Allerdings konnte gezeigt werden, dass eine anthropomorphe Gestaltung eines Chatbots einen indirekten Einfluss auf die Wiederverwendungsabsicht von Chatbots hat. Dabei wirkt sich eine anthropomorphe Gestaltung positiv auf die Soziale Präsenz aus. Die Wahrnehmung eines anthropomorphen Chatbot als sozial präsent ist höher als die eines nicht anthropomorphen. Diese Ergebnisse stehen im

Einklang mit der Media Equation Theorie sowie der Social Response Theorie. Bisherige Forschungsergebnisse können im Bereich von Chatbots bestätigt werden (Gnewuch et al., 2018, S. 11; Jiang et al., 2023, S. 8; Liu et al., 2024, S. 12; Morana et al., 2020, S. 11). Die Interaktion mit einem anthropomorphen Chatbot vermittelt Nutzern eher den Eindruck eines echten Gesprächs mit einem Menschen als die Interaktion mit einem nicht-anthropomorphen. Es kann daher davon ausgegangen werden, dass vermehrt Regeln der Mensch-Mensch-Interaktion angewendet werden, wenn der Chatbot anthropomorph gestaltet ist und der Chatbot als sozialer Akteur gesehen wird. Darüber hinaus hat die Soziale Präsenz einen positiven Einfluss auf das Vertrauen. Dieses Ergebnis steht in Einklang mit bisherigen Ergebnissen (De Cicco et al., 2020, S. 1224; Janson, 2023, S. 10; Yen & Chiang, 2021, S. 1188). Ein höheres Vertrauen hat einen positiven Einfluss auf die Wiederverwendungsabsicht. Hier können die Ergebnisse von Lei et al. (2021, S. 3987) und Liu et al. (2024, S. 11) bestätigt werden. Bisherige Forschungsergebnisse werden durch die serielle Mediation erweitert. Darüber hinaus können durch *FP I* bisherige Forschungsergebnisse unterstützt werden. Bezüglich der Forschungsfrage kann kein direkter Einfluss des anthropomorphen Designs auf die Wiederverwendungsabsicht nachgewiesen werden. Allerdings mediiieren die Soziale Präsenz sowie das Vertrauen die Beziehung zwischen Anthropomorphismus und Wiederverwendungsabsicht seriell.

Implikationen für die Praxis

Chatbots sollten anthropomorph gestaltet werden. Gerade in Bereichen, in denen Nutzer gerne mit einem Menschen interagieren würden, kann es helfen, einen anthropomorph gestalteten Chatbot einzusetzen. Denn dadurch kann die Soziale Präsenz im Rahmen der Konversation erhöht werden und die Nutzer haben eher das Gefühl, mit einem Menschen zu interagieren, wodurch das Vertrauen und in der Folge auch die Wiederverwendungsabsicht steigen können.

Weitergedacht könnten auf Nutzerprofile im Vorfeld der Interaktion zurückgegriffen werden, die durch vorherige Internetnutzung angelegt wurden. Dabei könnte darauf geachtet werden, ob der Nutzer Wert auf Soziale Präsenz legt. Darauf aufbauend ließe sich ein entsprechend gestalteter Chatbot einsetzen, ähnlich wie bei Cookies, die zur Personalisierung von Werbung verwendet werden. Dort wird dem Nutzer Werbung gezeigt, die auf den Interessen beziehungsweise dem vorherigen Verhalten des Nutzers basiert. So könnte es in Zukunft möglich sein, auch im Bereich von Chatbots verstärkt einen personalisierten Chatbot einzusetzen, der auf

Vorlieben der Nutzer angepasst wird, beispielsweise in Bezug auf den Grad der anthropomorphen Gestaltung.

Limitationen und Forschungsausblick

FP I ist nicht frei von Limitationen. Die verwendeten anthropomorphen Gestaltungsmerkmale haben einen Einfluss auf die Wahrnehmung als anthropomorph. Die Teilnehmer empfanden den anthropomorph gestalteten Chatbot als anthropomorpher als den roboterähnlichen. Allerdings gilt hier anzumerken, dass auch die Wahrnehmung des anthropomorphen Designs lediglich als mittelmäßig bewertet wurde. Für die weiteren Designs der in dieser Arbeit folgenden Experimente wurde das anthropomorphe Design weiterverwendet. Dort liegt allerdings auch kein expliziter Fokus auf Anthropomorphismus. Forschungen, die diesen Fokus setzen, sollten darauf achten, dass durch anthropomorphe Gestaltungsmerkmale auch ein hohes Maß an Anthropomorphismus in der Wahrnehmung der Nutzer erreicht wird. Hierfür könnten mehr der in Kapitel 4.1 beschriebenen Faktoren oder eine andere Kombination dieser genutzt werden.

Des Weiteren wurde in dem Experiment ein hypothetisches Szenario genutzt, in das sich die Teilnehmer nach einer kurzen Szenariobeschreibung hineinversetzen sollten. Die Szenariotechnik ist eine in der Marketing-Forschung und auch explizit in der Forschung zu Chatbots verbreitete Methode, um Hypothesen zu prüfen. Allerdings könnten Feldexperimente dazu beitragen, die Effekte in realen Umgebungen zu untersuchen und zu bestätigen.

Außerdem wurden die Experimente in Deutschland durchgeführt und die Stichproben sind für die Grundgesamtheit der Deutschen etwas zu jung und haben ein sehr hohes Bildungsniveau. Das Durchschnittsalter in Deutschland lag im Jahr 2023 bei 44,6 Jahren (Statistisches Bundesamt, 2025). An dieser Stelle könnte bei der Teilnehmerakquise auf das Alter sowie den Bildungsgrad geachtet werden, um eine Stichprobe mit Werten zu erreichen, die näher an der Grundgesamtheit der Deutschen liegen.

5.4 Forschungsprojekt II: Offenlegungszeitpunkt

Chatbots werden immer häufiger im Kundenservice eingesetzt. Aufgrund des anthropomorphen Designs sowie deren technologischen Fortschritten sind Chatbots immer schwieriger von menschlichen Chat-Partnern zu unterscheiden. Häufig ist dies nicht möglich. In Deutschland sind lediglich 39 % der Menschen zuversichtlich im Kontext eines Online-Kundenservice einen

menschlichen Kundenberater von einem Chatbot, der künstliche Intelligenz nutzt, unterscheiden zu können (Statista, 2023). Daher ist es wichtig, die nicht-menschliche Identität offenzulegen, insbesondere bei aktuellen Chatbots, die künstliche Intelligenz nutzen.

Die Offenlegung der Identität des Chatbots wird häufig mit negativen Kundenreaktionen verbunden (Ishowo-Oloko et al., 2019; Luo et al., 2019; Mozafari et al., 2022). Kunden neigen trotz teilweise funktionaler Überlegenheit der Chatbots dazu, der Technologie weniger zu vertrauen als Menschen (De Cicco et al., 2021, S. 11; Mozafari et al., 2022, S. 234), weshalb die Identität manchmal auch verschwiegen wird. Anbieter können allerdings, wie in Kapitel 4.2 beschrieben, zur Offenlegung gezwungen werden. Der Zeitpunkt der Offenlegung, also ob dies vor oder nach der Interaktion geschieht, wird häufig nicht vorgegeben. Daher ist dieser Zeitpunkt ein interessantes Themengebiet, um gegebenenfalls negative Auswirkungen zu analysieren.

Der Einfluss der Offenlegung von Chatbots und explizit dessen Zeitpunkt auf die Wiederverwendungsabsicht sowie die Soziale Präsenz fand in der Literatur bisher nur wenig Beachtung. Chatbots werden mittlerweile auf unterschiedlichen Kommunikationsplattformen eingesetzt, beispielsweise Internetseiten, Sozialen Medien oder Messengern wie beispielsweise WhatsApp. Gerade WhatsApp wird häufig und hauptsächlich von Nutzern für Kommunikationen im privaten Kontext verwendet. Nutzer sind daher häufig sehr vertraut und vermuten voraussichtlich weniger den Einsatz von Chatbot im Vergleich zu Online-Chats, die vermehrt für die nicht private Kommunikation also mit Unternehmen verwendet werden. In der bestehenden Literatur findet die Kombination von Offenlegungszeitpunkt und Fehlerkultur bislang keine Berücksichtigung. Dabei kann sich der Offenlegungszeitpunkt abhängig des Ergebnisses der Chatinteraktion unterschiedlich auf den Nutzer auswirken, explizit auf das vermutete Motiv, die wahrgenommene Fairness sowie den wahrgenommenen Betrug. Hieraus ergeben sich interessante Fragestellungen in Bezug auf den Offenlegungszeitpunkt sowie die Art der Kommunikationsplattform und die Fehlerkultur, die im Rahmen dieses Forschungsprojekts untersucht werden. *FP II* beschäftigt sich mit folgenden Fragen, die in den drei Experimenten untersucht werden:

1. Welchen Einfluss hat der Offenlegungszeitpunkt auf die Soziale Präsenz sowie die Wiederverwendungsabsicht?

2. Welchen Einfluss hat der Offenlegungszeitpunkt bei unterschiedlichen Kommunikationsplattformen (Internet oder WhatsApp) auf denen der Chatbot eingesetzt wird, auf die Soziale Präsenz sowie die Wiederverwendungsabsicht und welche moderierende Rolle spielt hierbei die Lösung der Anfrage?
3. Welchen Einfluss hat der Offenlegungszeitpunkt bei Erfolg oder Misserfolg der Anfrage auf den wahrgenommenen Betrug, die wahrgenommene Fairness, das Vermutete Motiv sowie die Wiederverwendungsabsicht?

5.4.1 Experiment 1: Offenlegungszeitpunkt eines Chatbots

Das erste Experiment aus Forschungsprojekt II (*FP II: E 1*) betrachtet die Auswirkung des Offenlegungszeitpunkts eines anthropomorph gestalteten Chatbots auf die Wahrnehmung als sozial präsent sowie die Wiederverwendungsabsicht. Die Soziale Präsenz wird dabei als Mediatorvariable in das Untersuchungsdesign aufgenommen.

5.4.1.1 Herleitung der Forschungshypothesen

Laut der Media Equation Theorie (Reeves & Nass, 1996) und der Social Response Theorie (Nass & Moon, 2000; Nass et al., 1994b) verhalten sich Menschen sozial, wenn sie mit Computern oder modernen Kommunikationstechnologien interagieren. Dabei übertragen sie Verhaltensmuster und soziale Regeln, die sie aus der zwischenmenschlichen Kommunikation kennen, auf den Umgang mit technischen Systemen (Nass et al., 1994b, S. 77; Reeves & Nass, 1996, S. 5). Auch auf Chatbots, welche eine aktuelle Kommunikationstechnologie darstellen, werden diese sozialen Regeln häufig unbewusst angewendet, besonders wenn diese anthropomorph gestaltet sind (Kim et al., 2013, S. 1800; Nass & Moon, 2000, S. 93). Chatbots können daher als echte soziale Akteure wahrgenommen werden. Auch in *FP I* zeigte sich, dass die wahrgenommene Soziale Präsenz bei anthropomorph gestalteten Chatbots höher ausfällt als bei nicht-anthropomorphen.

Betrachtet man nun den Offenlegungszeitpunkt, könnte die Offenlegung der Chatbots-Identität nach dem Gespräch die Soziale Präsenz des Chatbots erhöhen. Denn während der Interaktion mit dem Chatbot ist es unklar, ob die Kunden mit einem Menschen oder einem Chatbot interagieren. Wenn die Nutzer jedoch im Voraus explizit wissen, dass sie mit einem Chatbot schreiben, kann dies zu einer geringeren Sozialen Präsenz führen. Wenn die Chatbot-Identität vor dem Gespräch offengelegt wird, wissen sie, dass sie mit einem Chatbot und nicht mit einer

echten Person schreiben. Daher kann eine Benachrichtigung unmittelbar vor der Interaktion die Wahrnehmung von Sozialer Präsenz verringern. Folglich kann Hypothese 1 formuliert werden:

H₁: Die Offenlegung der Identität eines Chatbots vor der Interaktion (im Vergleich zur Offenlegung nach der Interaktion) verringert die Soziale Präsenz.

Das Bedürfnis nach sozialen Verbindungen ist ein Grundbedürfnis von Menschen (Epley, et al., 2008b, S. 144). Soziale Präsenz beschreibt das Ausmaß, in dem das Bedürfnis nach zwischenmenschlicher Verbindung befriedigt werden kann (Blut et al., 2021, S. 638). Eine hohe Soziale Präsenz kann zu einer größeren Wiederverwendungsabsicht führen oder einen positiven Einfluss auf die Zufriedenheit der Nutzer haben (Cheng & Jiang, 2020, S. 606; Verhagen et al., 2014, S. 540). Zufriedenheit wirkt sich positiv auf die weitere Nutzung von Chatbots aus (Cheng & Jiang, 2020, S. 606). Die Soziale Präsenz beeinflusst das Vertrauen und die wahrgenommene Freude, wohingegen diese beiden Variablen die Einstellung gegenüber dem Chatbot auf positive Weise beeinflussen (De Cicco et al., 2020, S. 1224). Adam et al. (2021, S. 437) zeigen einen Mediationseffekt von Sozialer Präsenz zwischen Anthropomorphismus und Benutzerfreundlichkeit. Im Kontext eines Serviceausfalls können Chatbots die negativen Auswirkungen stärker abmildern als Menschen (Sands et al., 2022, S. 2047). Diese Erkenntnisse zeigen, dass Soziale Präsenz Beziehungen zwischen unabhängigen und abhängigen Variablen medieren kann.

Ein wichtiger Indikator für den Erfolg eines anthropomorphen Chatbots ist die Absicht, den Chatbot wiederzuverwenden. Kunden, die zufrieden sind und eine gute Erfahrung gemacht haben, sind in der Regel eher bereit, den Chatbot wiederzuverwenden (Zhang et al., 2024, S. 6). Darüber hinaus wird durch die Interaktion mit dem anthropomorph gestalteten Chatbot das Bedürfnis nach sozialer Nähe erhöht. Das kann sich positiv auf die Wiederverwendungsabsicht auswirken. Daher wird folgende Hypothese aufgestellt:

H₂: Die Soziale Präsenz mediiert den Effekt des Offenlegungszeitpunkts der Identität eines Chatbots auf die Wiederverwendungsabsicht.

5.4.1.2 Methodik

In *FP II: E 1* wird ein hypothetisches Kundenservice-Szenario verwendet. Dabei sollen sich die Teilnehmer in die Situation hineinversetzen, dass ihr neu gekauftes Tablet ein defektes Display hat und sie dies bei dem Hersteller reklamieren möchten. Nach einem kurzen Chat, in dem das

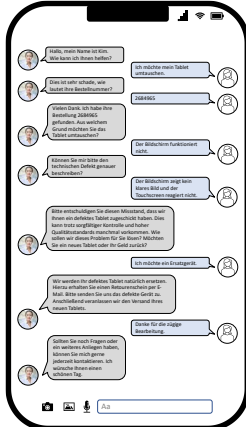
Anliegen geklärt wird, kommt es zu einer erfolgreichen Reklamation des Tablets mit Versand eines neuen Geräts. Die gesamte Chatinteraktion wurde den Teilnehmern in Form eines Screenshots gezeigt, der aufmerksam durchgelesen werden sollte. Dieser war für alle Gruppen gleich (siehe Tabelle 13).

Im Rahmen dieses Experiments wird ein einfaktorielles Design mit Zufallsgruppenbildung verwendet, bei dem der Offenlegungszeitpunkt manipuliert wird. Der Zwischensubjektfaktor Offenlegungszeitpunkt umfasst drei Ausprägungen. In der ersten Experimentalgruppe wurde die Chatbot-Identität vor der Interaktion offengelegt (*Offenlegung_vor*), in der zweiten Experimentalgruppe nach der Interaktion (*Offenlegung_nach*) und in der dritten Experimentalgruppe gab es keine Offenlegung (*Offenlegung_nein*). Den Teilnehmern der beiden Gruppen *Offenlegung_vor* und *Offenlegung_nach* wurde in einer Hinweisbox direkt vor beziehungsweise nach der Interaktion der folgende Inhalt gezeigt: „Hinweis: Die folgende/vorherige Konversation wird/hat mit einem Chatbot, einem virtuellen Assistenten, stattfinden/stattgefunden.“ Die dritte Gruppe wurde als Kontrollgruppe genutzt, welcher die Identität des Chatbots nicht offengelegt wird (*Offenlegung_nein*). Diese Gruppe sollte in Bezug auf die Soziale Präsenz einen ähnlichen Effekt wie die Gruppe *Offenlegung_nach* aufweisen, da in beiden Fällen während der Interaktion nicht klar ist, ob die Teilnehmer mit einem Chatbot oder einem Menschen interagieren. Um ein realistisches Szenario zu gewährleisten und für die Teilnehmer die Identifikation des Chatbots ohne Hinweis zu erschweren, wurde das anthropomorphe Design aus *FP I* in angepasster Form verwendet. Das gesamte Szenario ist in Tabelle 13 übersichtlich visualisiert.

Tabelle 13: Szenariobeschreibung (FP II: E 1)

Einführung		
Stellen Sie sich bitte vor, dass Sie ein Tablet bei einem Onlinehändler gekauft haben. Nach dem Erhalt des Tablets wollen Sie dieses natürlich direkt nutzen. Sie stellen fest, dass der Bildschirm kein klares Bild zeigt und der Touchscreen nicht reagiert. Daher wollen Sie das defekte Tablet reklamieren. Hierfür steht ihnen eine Chatfunktion zur Verfügung, die die Serviceanfragen an den Onlinehändler bearbeitet. Sie entscheiden sich diesen Service zu nutzen, wobei sich die folgende Chatinteraktion entwickelt:		
Gruppen		
Offenlegung_vor	Offenlegung_nach	Offenlegung_nein
Hinweis: Die folgende Konversation wird mit einem Chatbot, einem virtuellen Assistenten, stattfinden.		
 Hinweis: Die folgende Konversation wird mit einem Chatbot, einem virtuellen Assistenten, stattfinden.		

Tabelle 13: Szenariobeschreibung (FP II: E 1) (Fortsetzung)

Chatinteraktion		
- Hallo, mein Name ist Kim. Wie kann ich Ihnen helfen? - Ich möchte mein Tablet umtauschen. - Dies ist sehr schade, wie lautet ihre Bestellnummer? 2684965 - Vielen Dank. Ich habe ihre Bestellung 2684965 gefunden. Aus welchem Grund möchten Sie das Tablet umtauschen? - Der Bildschirm funktioniert nicht. - Können Sie mir bitte den technischen Defekt genauer beschreiben? - Der Bildschirm zeigt kein klares Bild und der Touchscreen reagiert nicht. - Bitte entschuldigen Sie diesen Missstand, dass wir Ihnen ein defektes Tablet zugeschickt haben. Dies kann trotz sorgfältiger Kontrolle und hoher Qualitätsstandards manchmal vorkommen. Wie sollen wir dieses Problem für Sie lösen? Möchten Sie ein neues Tablet oder Ihr Geld zurück? - Ich möchte ein Ersatzgerät. - Wir werden Ihr defektes Tablet natürlich ersetzen. Hierzu erhalten Sie einen Retourenschein per E-Mail. Bitte senden Sie uns das defekte Gerät zu. Anschließend veranlassen wir den Versand Ihres neuen Tablets. - Danke für die zügige Bearbeitung. - Sollten Sie noch Fragen oder ein weiteres Anliegen haben, können Sie mich gerne jederzeit kontaktieren. Ich wünsche Ihnen einen schönen Tag.		
		
Gruppen		
Offenlegung_vor	Offenlegung_nach	Offenlegung_nein
	Hinweis: Die vorherige Konversation hat mit einem Chatbot, einem virtuellen Assistenten, stattgefunden.	
	 Hinweis: Die vorherige Konversation hat mit einem Chatbot, einem virtuellen Assistenten, stattgefunden.	

Im Anschluss an das Szenario wurden die Teilnehmer zur Sozialen Präsenz (Toader et al., 2020, S. 14) sowie zur Wiederverwendungsabsicht des Chatbots (Moriuchi, 2021, S. 27-28) befragt. Tabelle 14 zeigt einen Überblick der verwendeten Skalen mit den übersetzten Indikatoren. Darüber hinaus wurde die Manipulation mithilfe von drei Manipulationscheckfragen überprüft. Zur

Beschreibung der Stichprobe wurden abschließend Fragen zu soziodemographischen Merkmalen der Teilnehmer, zur Nutzungshäufigkeit von Chatbots und zu ihrer Expertise in Bezug auf Chatbots (Ohanian, 1990, S. 50) gestellt. Auch in diesem Experiment sollten die Teilnehmer beurteilen, wie realitätsnah das Szenario ist und wie gut sie sich in die Situation hineinversetzen konnten. Hierfür wurden die gleichen Fragen wie aus *FP I* verwendet.

Tabelle 14: Operationalisierung der Variablen und Manipulationscheck (FP II: E 1)

Kürzel	Indikator	Skala
Manipulationscheck		
War ein Hinweis über die Konversation vorhanden?		
MC_1	Der Onlinehändler hat mir keinen Hinweis über die Konversation gegeben.	Einfachauswahl
MC_2	Der Onlinehändler hat mir einen Hinweis über die Konversation vor der Interaktion gegeben.	
MC_3	Der Onlinehändler hat mir einen Hinweis über die Konversation nach der Interaktion gegeben.	
Soziale Präsenz (Toader et al., 2020, S. 14)		
Inwiefern stimmen Sie den folgenden Aussagen zu?		
SoPr_1	Bei der Nutzung der Chatfunktion hatte ich das Gefühl, mit einem Menschen zu interagieren.	7-Punkt-Likert-Skala: 1 = Stimme überhaupt nicht zu, 7 = Stimme vollkommen zu
SoPr_2	Obwohl ich den Chatpartner nicht im realen Leben sehen konnte, war ein Gefühl der menschlichen Wärme vorhanden.	
SoPr_3	Während der Chatinteraktion hatte ich ein Gefühl der Geselligkeit.	
SoPr_4	Ich hatte das Gefühl, dass es eine Person gab, die mir Sicherheit gegeben hat.	
SoPr_5	Ich hatte das Gefühl, dass es jemanden gibt, der für mich da ist, wenn ich ein Problem habe.	
Wiederverwendung (Moriuchi, 2021, S. 27-28)		
Inwiefern stimmen Sie den folgenden Aussagen zu?		
Reu_1	Wenn ich Zugang zu einer Chatfunktion hätte, würde ich sie immer nutzen.	7-Punkt-Likert-Skala: 1 = Stimme überhaupt nicht zu, 7 = Stimme vollkommen zu
Reu_2	Ich denke, es wird sich für mich lohnen, Chatfunktionen zu verwenden, wenn sie verfügbar sind.	
Reu_3	Ich würde meinen Freunden und meiner Familie die Nutzung einer Chatfunktion empfehlen.	
Reu_4	Wenn ich die Möglichkeit hätte mehr über Chatfunktionen mit Unternehmen zu kommunizieren, würde ich dies tun.	
Reu_5	Ich nutze lieber eine Chatfunktion, um mein Anliegen zu klären, als eine andere Person zu fragen.	
Expertise (Ohanian, 1990, S. 50)		
Inwiefern stimmen Sie den folgenden Aussagen zu?		
Exp_1	Ich bin Experte im Umgang mit Chatbots.	7-Punkt-Likert-Skala: 1 = Stimme überhaupt nicht zu, 7 = Stimme vollkommen zu
Exp_2	Ich bin erfahren im Umgang mit Chatbots.	
Exp_3	Ich bin wissend im Umgang mit Chatbots.	
Exp_4	Ich bin qualifiziert im Umgang mit Chatbots.	
Exp_5	Ich bin kompetent im Umgang mit Chatbots.	

5.4.1.3 Stichprobe und Gütebeurteilung

Insgesamt nahmen 197 Teilnehmer an dem Experiment teil. Viele der Teilnehmer haben die Manipulationscheckfragen allerdings falsch beantwortet und mussten aus der Stichprobe ausgeschlossen werden (Tabelle 15). Die grün markierten Teilnehmer haben die Manipulationscheckfragen richtig beantwortet. Dadurch hat sich die Teilnehmerzahl auf 101 reduziert.

Tabelle 15: Ergebnisse des Manipulationschecks (FP II: E 1)

Gruppe	MC_1	MC_2	MC_3	Summe
Offenlegung_nein	33	24	8	65
Offenlegung_vor	21	48	2	71
Offenlegung_nach	31	10	20	61

Von der finalen Stichprobe sind 65,3 % der Teilnehmer weiblich, 33,7 % männlich und 1,0 % machten keine Angabe zu ihrem Geschlecht. Das Durchschnittsalter der Teilnehmer liegt bei 30,67 Jahre (SD 11,95). 89,1 % der Befragten haben mindestens einmal einen Chatbot in den letzten zwölf Monaten verwendet. Die Situation des Szenarios wurde von den Teilnehmern als realistisch bewertet ($M = 5,12$, $SD = 1,66$) und es fiel ihnen leicht, sich in die Situation hineinzuversetzen ($M = 5,71$, $SD = 1,51$). Die Teilnehmer schätzten ihre Expertise im Umgang mit Chatbots im Mittel mit $M = 4,19$ ($SD = 1,66$) ein. Dieses Konstrukt wurde in SPSS aus den Indikatoren mithilfe der Mean-Funktion berechnet. Eine detaillierte Stichprobenbeschreibung zeigt Tabelle 16. Die Verteilung der Nutzungshäufigkeit von Chatbots in den letzten zwölf Monaten der Teilnehmer ist in Anhang 1 zusammengefasst.

Tabelle 16: Soziodemographische Merkmale (FP II: E 1)

Merkmal	Ausprägung	Prozentualer Anteil
Geschlecht	Weiblich	65,3 %
	Männlich	33,7 %
	Divers	0,0 %
	Keine Angabe	1,0 %
Alter	M = 30,67 (SD = 11,95), 3 kein Angabe	
Haushaltsgröße	1 Person	20,8 %
	2 Personen	50,5 %
	3-5 Personen	26,7 %
	Mehr als 5 Personen	1,0 %
	Keine Angabe	1,0 %
Höchster Bildungsabschluss	Noch in schulischer Ausbildung	0,0 %
	Volks-/Hauptschulabschluss	2,0 %
	Realschulabschluss	7,9 %
	Fachschulabschluss	5,9 %
	Abitur	25,7 %
	Fachhoch-/Hochschulabschluss	58,4 %
	Keine Angabe	0,0 %
Monatliches Haushaltsnettoeinkommen	Bis 1000 Euro	19,8 %
	1001 - 2000 Euro	25,7 %
	2001 - 3000 Euro	24,8 %
	3001 - 4000 Euro	8,9 %
	4001 - 5000 Euro	8,9 %
	Mehr als 5000 Euro	11,9 %
	Keine Angabe	0,0 %

Die Gütebeurteilung der gemessenen Konstrukte erfolgt auf Basis der in Kapitel 4.2 beschriebene Gütekriterien und Grenzwerte. In Tabelle 17 ist das Ergebnis der Gütebeurteilung der Indikatoren und Konstrukte zusammengefasst. Hier zeigt sich, dass die Grenzwerte der Gütekriterien für Reliabilität und Validität alle überschritten wurden. Daher werden die Konstrukte für die folgende Auswertung verwendet.

Tabelle 17: Gütekriterien (FP II: E 1)

Konstrukt	Indikator	Faktorladung > 0,4	KITK > 0,5	Chronbachs Alpha > 0,7	KMO > 0,6	Erklärte Varianz > 50 %
Soziale Präsenz	SoPr_1	0,896	0,861	0,904	0,817	72,294
	SoPr_2	0,830	0,695			
	SoPr_3	0,797	0,850			
	SoPr_4	0,874	0,861			
	SoPr_5	0,850	0,814			
Wiederverwendung	Reu_1	0,916	0,808	0,930	0,899	78,243
	Reu_2	0,792	0,770			
	Reu_3	0,908	0,748			
	Reu_4	0,916	0,769			
	Reu_5	0,884	0,597			
Expertise	Exp_1	0,854	0,777	0,939	0,841	80,318
	Exp_2	0,909	0,851			
	Exp_3	0,894	0,831			
	Exp_4	0,913	0,861			
	Exp_5	0,911	0,854			

5.4.1.4 Ergebnisse

Zur Überprüfung der Auswirkung des Offenlegungszeitpunkts auf die Soziale Präsenz wurde in einem ersten Schritt eine einfaktorielle ANOVA mit Post-hoc-Test nach Bonferroni ausgeführt. An dieser Stelle soll auch der Unterschied zu Kontrollgruppe überprüft werden. Der Zwischensubjektfaktor Offenlegungszeitpunkt wird in der ANOVA als unabhängige Variable und die Soziale Präsenz als abhängige Variable verwendet. Die Gruppe *Offenlegung_nach* hat die höchste Wahrnehmung der Sozialen Präsenz ($M = 0,58$, $SD = 1,18$), die beiden anderen Gruppen sind sehr ähnlich (*Offenlegung_vor*: $M = -0,15$, $SD = 0,92$; *Offenlegung_nein*: $M = -0,13$, $SD = 0,89$). Der Post-hoc-Test nach Bonferroni zeigt, dass sich die Gruppen *Offenlegung_vor* und *Offenlegung_nach* signifikant unterscheiden ($p < 0,05$). Folglich kann Hypothese H_1 bestätigt werden. Eine Offenlegung vor der Interaktion verringert die Soziale Präsenz. Die Kontrollgruppe hat eine ähnliche Wahrnehmung, wie die Gruppe *Offenlegung_vor*. Die Gruppe *Offenlegung_nach* unterscheidet sich von den beiden anderen Gruppen signifikant ($p < 0,05$). Der Offenlegungszeitpunkt wirkt sich somit auf die Soziale Präsenz aus. Die Ergebnisse der ANOVA sind in Abbildung 5 grafisch dargestellt.

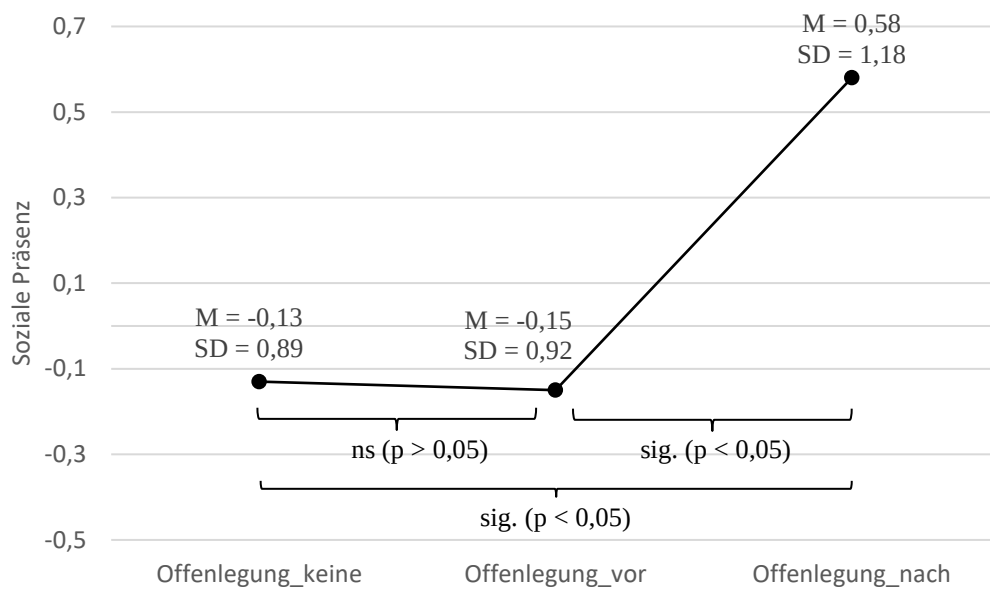


Abbildung 5: Ergebnisse der ANOVA (FP II: E 1) – Effekte des Offenlegungszeitpunkts auf die Soziale Präsenz

Zur Überprüfung der Forschungshypothese H₂ aus FP II: E 1 wird das Mediationsmodell 4 des PROCESS-Makros mit einem Bootstrap Sample von 5.000 gerechnet (Hayes, 2022, S. 622). Als unabhängige Variable wird der Offenlegungszeitpunkt (vorher vs. nachher) verwendet. Die Soziale Präsenz wird als Mediatorvariable und die Wiederverwendungsabsicht als abhängige Variable eingesetzt. Das Mediationsmodell mit den entsprechenden Regressionskoeffizienten ist in Abbildung 6 dargestellt. Als Referenzkategorie dient hier die Experimentalgruppe *Offenlegung_vor*.

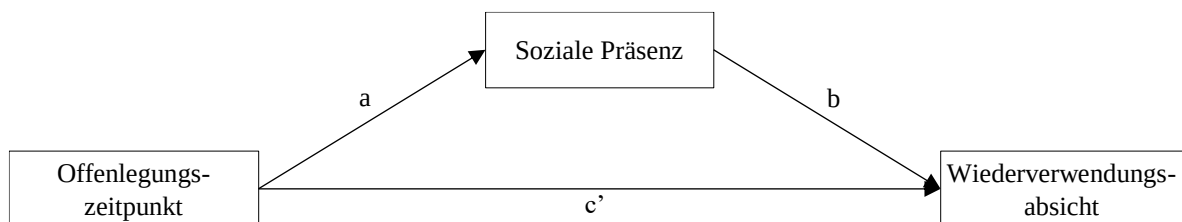


Abbildung 6: Mediationsmodell (FP II: E 1)

Die Ergebnisse zeigen einen signifikanten indirekten Effekt des Offenlegungszeitpunkts auf die Wiederverwendungsabsicht über die Soziale Präsenz ($a \times b = 0,299$, 95 % KI [0,055, 0,626]). Auch in Bezug auf den Effekt des Offenlegungszeitpunkts auf die Wiederverwendungsabsicht des Chatbots konnte ein signifikanter Effekt bestätigt werden ($c' = -0,509$, $p < 0,05$). Diese Ergebnisse bestätigen Hypothese H₂. Die Hypothesen sind in Tabelle 19 zusammengefasst. An-

zumerken gilt, dass der direkte Einfluss des Offenlegungszeitpunkts auf die Wiederverwendungsabsicht in dem Mediationsmodell entgegengerichtet ist. Eine Offenlegung nach der Interaktion führt zu einer geringeren Wiederverwendungsabsicht. Ein Überblick der Ergebnisse der Analyse wird in Tabelle 18 dargestellt.

Tabelle 18: Regressionskoeffizienten der Mediationsanalyse (FP II: E 1)

	Symbol	Koeff.	KI _u	KI _o
Effekte auf die Wiederverwendungsabsicht				
Offenlegungszeitpunkt	c'	-0,509*	-0,992	-0,026
Soziale Präsenz	b	0,406***	0,195	0,618
Effekt auf die Soziale Präsenz				
Offenlegungszeitpunkt	a	0,735**	0,204	1,266
	Symbol	Effekt	KI_u	KI_o
Offenlegungszeitpunkt x Soziale Präsenz	a x b	0,299	0,055	0,626
***: p < 0,001; **: p < 0,01; *: p < 0,05; ns: p > 0,05				

Tabelle 19: Zusammenfassung der Hypothesen (FP II: E 1)

Hypothese H _i		Ergebnis
H ₁	Die Offenlegung der Identität eines Chatbots vor der Interaktion (im Vergleich zur Offenlegung nach der Interaktion) verringert die Soziale Präsenz.	✓
H ₂	Die Soziale Präsenz mediert den Effekt des Offenlegungszeitpunkts der Identität eines Chatbots auf die Wiederverwendungsabsicht.	✓
✓ = bestätigt; ✗ = nicht bestätigt		

5.4.1.5 Diskussion der Ergebnisse

In *FP II: E 1* konnte gezeigt werden, dass eine späte Offenlegung die Soziale Präsenz während der Interaktion erhöht und die Soziale Präsenz als Mediator zwischen Offenlegungszeitpunkt und der Wiederverwendungsabsicht wirkt. Eine höhere Soziale Präsenz führt zu einer höheren Wiederverwendungsabsicht, wohingegen der direkte Effekt des Offenlegungszeitpunkts auf die Wiederverwendungsabsicht im Mediationsmodell umgekehrt gerichtet ist. Eine späte Offenlegung führt zu einer niedrigeren Wiederverwendungsabsicht. Dies könnte möglicherweise darauf zurückzuführen sein, dass der Kundenservice und insbesondere der Chatbot von Anfang an nicht ehrlich gegenüber dem Kunden war und ihn während der Interaktion im Glauben ließ, mit einem Menschen zu kommunizieren, oder zumindest keine Eindeutigkeit darüber bestand. In diesem Fall könnte sich der Nutzer getäuscht fühlen oder die Art der Kommunikation, dass

es sich um einen Chatbot handelt, als unfair erachten. In *FP II: E 3* wird daher der Einfluss des Offenlegungszeitpunkts auf den wahrgenommenen Betrug, die Fairnesswahrnehmung sowie das vermutete Motiv betrachtet.

Die Wirkung des Offenlegungszeitpunkts könnte von der Wichtigkeit der Sozialen Präsenz des Nutzers abhängen (Gambino et al., 2020, S. 73). Hier könnte möglicherweise der Mediations-effekt überwiegen und dadurch der negative direkte Effekt reduziert werden. Personen, für die soziale Interaktionen und Soziale Präsenz wichtig sind, könnten daher durch die späte Offenlegung eine größere Wiederverwendungsabsicht haben. Für Personen, denen soziale Interaktionen und Soziale Präsenz eher unwichtig sind, könnten eine späte Offenlegung zu einer niedrigeren Wiederverwendungsabsicht führen.

Diese Ergebnisse unterstützen und erweitern bisherige Erkenntnisse, beispielsweise von Hendriks et al. (2020), Mozafari et al. (2021b) und Van der Goot et al. (2024) in Bezug auf die Offenlegung von Chatbots. Die Gruppe *Offenlegung_nein* hat die Soziale Präsenz ähnlich bewertet wie die Gruppe *Offenlegung_vor*. Hier wurde in den vorherigen Ausführungen allerdings davon ausgegangen, dass die Bewertung der Sozialen Präsenz ähnlich zur Gruppe *Offenlegung_nach* sein sollte. Diese Annahme konnte im Rahmen dieses Experiments nicht bestätigt werden. Leider mussten in diesem Experiment viele Teilnehmer aus der Stichprobe entfernt werden, da sie den Manipulationcheck nicht richtig beantworteten. Dies führte zur Bildung kleiner sowie unterschiedlich großer Gruppen, welche anschließend in die Analyse einfließen.

5.4.2 Experiment 2: Offenlegungszeitpunkt eines Chatbots und die Kommunikationsplattform

Da Chatbots auch immer häufiger von Unternehmen in Messenger-Apps eingesetzt werden, ist es wichtig zu untersuchen, wie Kunden darauf reagieren (Rese et al., 2020, S. 2). So können endlose Warteschlange und verspätete Antworten vermieden werden und Kunden können in einer sehr gewohnten Umgebung interagieren, die sie in der Regel für ihre alltägliche Kommunikation nutzen. Hierfür bietet WhatsApp beispielsweise WhatsApp Business an (Pisu, 2024). Derzeit werden solche Messenger-Apps hauptsächlich für die private Kommunikation mit Freunden, Bekannten oder Familie genutzt und weniger für die Kommunikation mit dem Kun-

denservice eines Unternehmens. Daher sollten Nutzer im Rahmen solcher Messenger-Apps weniger den Einsatz von Chatbots erwarten und Chatbots sollten auf jeden Fall offengelegt werden.

Das zweite Experiment aus Forschungsprojekt II (*FP II: E 2*) betrachtet daher die Auswirkung des Offenlegungszeitpunkts eines anthropomorph gestalteten Chatbots auf die Wiederverwendungsabsicht bei Nutzung unterschiedlicher Kommunikationsplattformen. In diesem Experiment wird zwischen WhatsApp als Beispiel einer Messenger-App und einem Online-Chat unterschieden. Das Offenlegungsdesign der Studie wurde im Vergleich zu *FP II: E 1* dahingehend angepasst, dass die Offenlegungsinformation besser hervorgehoben und in den Screenshot der Chatinteraktion integriert wurde. Dies soll die Ausschlussrate verringern.

Die Soziale Präsenz wird wie in *FP II: E 1* als Mediatorvariable in dem Untersuchungsdesign aufgenommen. Neben dem Offenlegungszeitpunkt wurde die Kommunikationsplattform des Chatbots manipuliert, da angenommen wird, dass diese als Moderatorvariable zwischen Offenlegungszeitpunkt und Sozialer Präsenz interagiert. Dabei wurde zwischen einem Online-Chat auf der Internetseite des Anbieters und einem WhatsApp-Chat unterschieden.

5.4.2.1 Herleitung der Forschungshypothesen

Da sich Messenger-Apps durch die häufig eher private Nutzung und Kommunikation mit Freunden und Bekannten von Online-Chats auf Unternehmenswebseiten unterscheiden, soll in *FP II: E 2* untersucht werden, welchen Einfluss ein unterschiedlicher Offenlegungszeitpunkt im Rahmen von Messenger-Apps und explizit am Beispiel von WhatsApp im Vergleich zu Online-Chats auf die Soziale Präsenz sowie die Wiederverwendungsabsicht hat.

In diesem Experiment wird neben der meditierenden Funktion der Sozialen Präsenz die moderierende Funktion der Vertrautheit mit der Kommunikationsplattform untersucht. Dabei wird unter Kommunikationsplattformvertrautheit das Wissen über die Kommunikationsplattform und dessen Service, das auf der Anzahl an Erfahrungen und der vorherigen Nutzung beruht, verstanden. In einem ersten Schritt werden dieselben Hypothesen wie in *F II: E 1* betrachtet:

- H₁: Die Offenlegung der Identität eines Chatbots vor der Interaktion (im Vergleich zur Offenlegung nach der Interaktion) verringert die Soziale Präsenz.
- H₂: Soziale Präsenz mediiert den Effekt des Offenlegungszeitpunkts der Identität eines Chatbots auf die Wiederverwendungsabsicht.

Nutzer werden auf WhatsApp eher weniger mit Chatbots rechnen, da WhatsApp hauptsächlich für die private Kommunikation genutzt wird. Umso wichtiger ist es, die Identität eines Chatbots direkt zu Beginn der Konversation offenzulegen, damit die Nutzer von Beginn an wissen, mit wem sie kommunizieren. Hier wird davon ausgegangen, dass Nutzer eine größere Soziale Präsenz verspüren, wenn sie mit einem Chatbot in einem eher „privaten“ Kommunikationsmedium interagieren. Da sie dort eher weniger mit Chatbots rechnen, kann eine späte Offenlegung zu einer höheren Sozialen Präsenz führen. In der Folge kann die Wiederverwendungsabsicht höher sein. Dies steht im Einklang mit der Familiarity-Heuristik, der zufolge Kunden eher Marken oder Anbieter (erneut) wählen, mit denen sie bereits vertraut sind, denn ihnen ist bekannt, was sie erwarten könnten. Das Vertrauen mit einer Kommunikationsplattform kann sich auf die Freundlichkeit oder die Geselligkeit auswirken. So sind Nutzer gegebenenfalls mit einer Website vertrauter als mit einem Chatbot im Internet und dadurch wird ihnen ein Gefühl der Sicherheit vermittelt (Ischen et al., 2020b, S. 45). Kunden sind mit WhatsApp als Kommunikationsplattform vertraut und nutzen diese meist für die private und somit häufig für die soziale Interaktion. Daher kann die Soziale Präsenz dort höher sein, wenn die Nutzer während der Interaktion nicht wissen, dass sie mit einem Chatbot schreiben und eher davon ausgegangen wird, dass das Gegenüber ein Mensch und kein Chatbot ist. Daraus ergibt sich die folgende moderierte Mediationshypothese:

H₃: Die Art der Kommunikationsplattform (Online-Chat vs. WhatsApp) moderiert den Mediationseffekt des Offenlegungszeitpunkts auf die Wiederverwendungsabsicht über die Soziale Präsenz, sodass der Mediationseffekt größer ist, wenn die Kommunikation in WhatsApp stattfindet.

5.4.2.2 Methodik

Zur Überprüfung der Hypothesen in *FP II: E 2* sollten sich die Teilnehmer in einem hypothetischen Kundenservice-Szenario in die Situation versetzen, dass sich die Bestellung einer Lampe verzögert. Dabei möchte der Kunde weitere Informationen zum Lieferstatus erhalten und die Bestellung gegebenenfalls in Abhängigkeit der Lieferverzögerung stornieren. Zur Kontaktaufnahme besucht der Kunde die Webseite des Anbieters mit seinem Handy. Dort wird am rechten Bildschirmrand die Möglichkeit angezeigt mit dem Kundenservice zu interagieren. In diesem Experiment wurde den Teilnehmern Screenshots der Chatinteraktion sowie der Webseite ge-

zeigt. Das Szenario ist etwas interaktiver als die beiden vorherigen, da die einzelnen Gesprächsbestandteile erst schrittweise in den Screenshots ersichtlich werden und die Teilnehmer durch Klicken auf weiter sukzessive die Konversation angezeigt bekommen. Dies ähnelt dem Prozess, bei dem Nutzer eine eigene Nachricht eingeben und daraufhin eine entsprechende Antwort erhalten.

Es wurde ein zweifaktorielles Design mit Zufallsgruppenbildung genutzt, bei dem der Offenlegungszeitpunkt des Chatbots sowie die Kommunikationsplattform manipuliert wurden. Der Zwischensubjektfaktor Offenlegungszeitpunkt umfasst zwei Ausprägungen. In der ersten Experimentalgruppe wurde die Chatbot-Identität vor der Interaktion offengelegt (*Offenlegung_vor*), in der zweiten Experimentalgruppe nach der Interaktion (*Offenlegung_nach*). Dabei wurde den Teilnehmern im Chat eine Hinweisbox direkt vor oder nach der Interaktion mit dem folgenden Inhalt gezeigt. „Hinweis: Die folgende/vorherige Konversation wird/hat mit einem Chatbot, einem virtuellen Assistenten, stattfinden/stattgefunden.“ Die Offenlegungsinformation bezüglich des Chatpartners wurde in den Screenshots implementiert und nicht wie in *FP II: E 1* auf einer eigenen Seite vor oder nach der Interaktion angezeigt. Dadurch sollte ein realistischeres Szenario erreicht und der Hinweis eher gelesen werden. Der Zwischensubjektfaktor Kommunikationsplattform wurde dahingehend angepasst, dass das Layout der Handydisplays geändert wurde. So sahen die Teilnehmer der *WhatsApp-Chat*-Gruppen die Konversation in einem WhatsApp-Design. Die Teilnehmer der *Online-Chat*-Gruppe sahen die Konversation in einem Online-Design, dass dem Design der Homepage angepasst war. Die Stimuli der vier Gruppen sind in Tabelle 20 dargestellt. Auch in diesem Experiment wurde durch das Design eine realitätsnahe Situation dargestellt und für den Chatbot ein ähnliches Design verwendet wie in den vorherigen Experimenten.

Tabelle 20: Szenariobeschreibung (FP II: E 2)

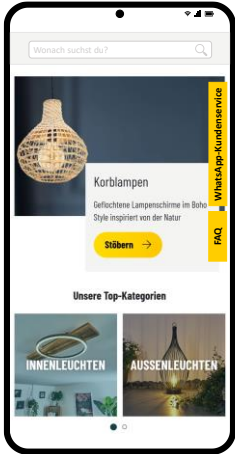
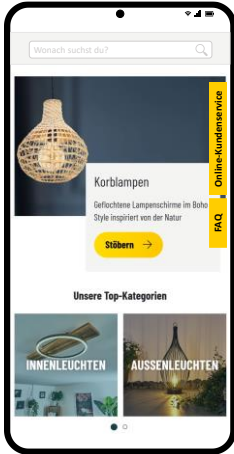
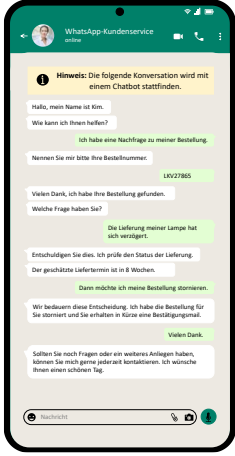
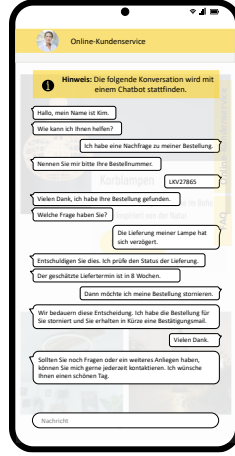
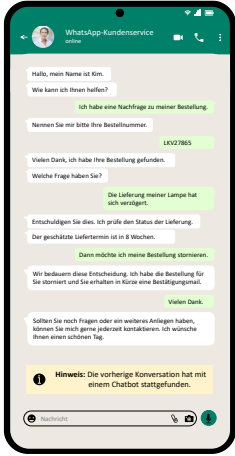
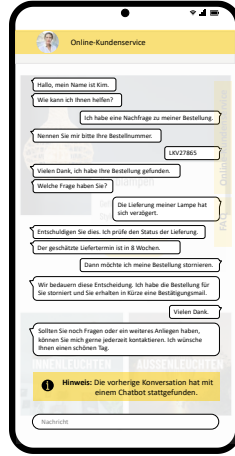
Einführung	
Bitte versetzen Sie sich in die Situation, dass Sie eine neue Lampe bestellt haben. Diese wurde nicht wie ursprünglich angegeben, innerhalb einer Woche geliefert. Sie möchten Informationen zum Lieferstatus Ihrer Bestellung erhalten. Sollte sich die Lieferung weiter verzögern, möchten Sie Ihre Bestellung stornieren. Hierfür besuchen Sie die Website des Händlers und suchen nach einer Kontaktmöglichkeit. Auf der folgenden Seite ist die Website des Händlers abgebildet.	
Gruppe	
WhatsApp-Chat	Online-Chat
	
Sie sehen am rechten Bildschirmrand die Möglichkeit eines WhatsApp-Kundenservice. Sie entscheiden sich, diesen WhatsApp-Kundenservice zu nutzen. Auf den nächsten Seiten wird Ihnen der Gesprächsverlauf mit dem Kundenservice-Chat gezeigt.	Sie sehen am rechten Bildschirmrand die Möglichkeit eines Online-Kundenservice. Sie entscheiden sich, diesen Online-Kundenservice zu nutzen. Auf den nächsten Seiten wird Ihnen der Gesprächsverlauf mit dem Kundenservice-Chat gezeigt.
Chatinteraktion	
• Hallo, mein Name ist Kim. Wie kann ich Ihnen helfen? • Ich habe eine Nachfrage zu meiner Bestellung. • Nennen Sie mir bitte Ihre Bestellnummer. • LKV27865 • Vielen Dank, ich habe Ihre Bestellung gefunden. Welche Frage haben Sie? • Die Lieferung meiner Lampe hat sich verzögert. • Entschuldigen Sie dies. Ich prüfe den Status der Lieferung. Der geschätzte Liefertermin ist in 8 Wochen. • Dann möchte ich meine Bestellung stornieren. • Wir bedauern diese Entscheidung. Ich habe die Bestellung für Sie storniert und Sie erhalten in Kürze eine Bestätigungsmail. • Vielen Dank. • Sollten Sie noch Fragen oder ein weiteres Anliegen haben, können Sie mich gerne jederzeit kontaktieren. Ich wünsche Ihnen einen schönen Tag.	

Tabelle 20: Szenariobeschreibung (FP II: E 2) (Fortsetzung)

WhatsApp-Chat / Offenlegung_vor	Online-Chat / Offenlegung_vor
Hinweis: Die folgende Konversation wird mit einem Chatbot stattfinden.	
	
WhatsApp-Chat / Offenlegung_nach	Online-Chat / Offenlegung_nach
	
Hinweis: Die vorherige Konversation hat mit einem Chatbot stattgefunden.	

Im Anschluss an das Szenario wurden die Teilnehmer zur Sozialen Präsenz des Chatbots (Toader et al., 2020, S. 14) und der Wiederverwendungsabsicht des Chatbots (Moriuchi, 2021, S. 27-28) befragt. Zur Messung dieser Variablen wurde auf Skalen aus der bisherigen Forschung zurückgegriffen, die auf Basis des in Kapitel 5.2 vorgestellten Prozess ins Deutsche übersetzt wurden. Diese Skalen wurden im Vergleich zu den vorherigen Experimenten leicht angepasst. Einen Überblick der verwendeten Skalen mit den übersetzten Indikatoren ist in Tabelle 21 zu sehen. Darüber hinaus wurde die Manipulation mithilfe von zwei Manipulationscheckfragen jeweils unter Verwendung einer Einfachauswahl überprüft. Abschließend wurden Fragen zu soziodemographischen Merkmalen der Teilnehmer, Nutzungshäufigkeit von Chat-

bots, ihrer Expertise in Bezug auf Chatbots (Ohanian, 1990, S. 50) sowie ihrem Nutzungsverhalten von WhatsApp gestellt, um die Stichprobe zu beschreiben. Auch in diesem Experiment sollten die Teilnehmer beurteilen, wie realitätsnah das Szenario ist und wie gut sie sich in die Situation hineinversetzen konnten. Hierzu wurden die gleichen Fragen wie in *FP II: E 1* genutzt.

Tabelle 21: Operationalisierung der Variablen und Manipulationscheck (FP II: E 2)

Kürzel	Indikator	Skala
Manipulationscheck		
Über welche Chat-Funktion haben Sie mit dem Kundenservice kommuniziert?		
MC_11	WhatsApp	Einfachauswahl
MC_12	Online-Chatfunktion des Händlers	
Haben Sie von der Chat-Funktion explizit einen Hinweis über den Kommunikationspartner erhalten?		
MC_21	Ich habe vor der Interaktion einen Hinweis erhalten.	Einfachauswahl
MC_22	Ich habe nach der Interaktion einen Hinweis erhalten.	
MC_23	Ich habe keinen Hinweis erhalten.	
Soziale Präsenz (Toader et al., 2020, S. 14)		
Inwiefern stimmen Sie den folgenden Aussagen zu?		
SoPr_1	Bei der Nutzung des Kundenservice-Chats hatte ich das Gefühl, mit einem Menschen zu interagieren.	7-Punkt-Likert-Skala: 1 = Stimme überhaupt nicht zu, 7 = Stimme vollkommen zu
SoPr_2	Obwohl ich den Kundenservice-Chat nicht im realen Leben sehen konnte, war ein Gefühl der menschlichen Wärme vorhanden.	
SoPr_3	Während der Interaktion mit dem Kundenservice-Chat hatte ich ein Gefühl der Geselligkeit.	
SoPr_4	Ich hatte das Gefühl, dass es eine Person gab, die mir Sicherheit gegeben hat.	
SoPr_5	Ich hatte das Gefühl, dass es jemanden gibt, der für mich da ist, wenn ich ein Problem habe.	
Wiederverwendung (Moriuchi, 2021, S. 27-28)		
Inwiefern stimmen Sie den folgenden Aussagen zu?		
Reu_1	Wenn ich Zugang zu einem Kundenservice-Chat hätte, würde ich ihn immer nutzen.	7-Punkt-Likert-Skala: 1 = Stimme überhaupt nicht zu, 7 = Stimme vollkommen zu
Reu_2	Ich denke, es wird sich für mich lohnen, Kundenservice-Chats zu verwenden, wenn sie verfügbar sind.	
Reu_3	Ich würde meinen Freunden und meiner Familie die Nutzung eines Kundenservice-Chats empfehlen.	
Reu_4	Wenn ich die Möglichkeit hätte mehr über Kundenservice-Chats zu kommunizieren, würde ich dies tun.	
Reu_5	Ich nutze lieber einen Kundenservice-Chat, um meine Anliegen zu klären, als eine andere Person zu fragen.	

5.4.2.3 Stichprobe und Gütebeurteilung

Insgesamt nahmen 128 Teilnehmer an dem Experiment teil. Diese wurden in Bezug auf ihre Antworten der beiden Manipulationscheckfragen geprüft. Dabei haben lediglich 5 Teilnehmer die Frage zur Kommunikationsplattform falsch beantwortet und wurden aus der Stichprobe ausgeschlossen. Die anschließende Frage nach dem Offenlegungszeitpunkt hat zu weiteren 23 Ausschlüssen geführt. Folglich ergibt sich eine Stichprobe von 100. 58,0 % der Teilnehmer sind weiblich, 37,0 % männlich, 3,0 % machen keine Angabe zu ihrem Geschlecht und 2,0 % geben Divers an. Das Durchschnittsalter der Teilnehmer liegt bei 29,86 Jahren (SD 9,88). 87,0 % der Befragten haben mindestens einmal einen Chatbot in den letzten zwölf Monaten verwendet. 22,0 % haben Chatbots sogar häufiger als sechsmal in den letzten zwölf Monaten genutzt. Die Teilnehmer bewerten die Situation des Szenarios als sehr realistisch ($M = 5,87$, $SD = 1,37$) und es fällt ihnen leicht, sich in die Situation hineinzusetzen ($M = 6,11$, $SD = 1,21$). Hier gibt es auch keine signifikanten Unterschiede zwischen den Gruppen. Die Teilnehmer bewerten ihre Expertise zu Chatbots im Mittel mit $M = 4,35$ ($SD = 1,54$). Die Expertise wurde in SPSS aus den Indikatoren mithilfe der Mean-Funktion berechnet. Eine detaillierte Stichprobenbeschreibung zeigt Tabelle 22. Die Verteilung der Nutzungshäufigkeit von Chatbots in den letzten zwölf Monaten der Teilnehmer ist in Anhang 1 zusammengefasst.

Tabelle 22: Soziodemographische Merkmale (FP II: E 2)

Merkmal	Ausprägung	Prozentualer Anteil
Geschlecht	Weiblich	58,0 %
	Männlich	37,0 %
	Divers	3,0 %
	Keine Angabe	2,0 %
Alter	M = 29,86 (SD = 9,883), 3 kein Angabe	
Haushaltsgröße	1 Person	19,0 %
	2 Personen	38,0 %
	3-5 Personen	40,0 %
	Mehr als 5 Personen	2,0 %
	Keine Angabe	1,0 %
Höchster Bildungsabschluss	Noch in schulischer Ausbildung	3,0 %
	Volks-/Hauptschulabschluss	2,0 %
	Realschulabschluss	5,0 %
	Fachschulabschluss	4,0 %
	Abitur	22,0 %
	Fachhoch-/Hochschulabschluss	64,0 %
	Keine Angabe	0,0 %
Monatliches Haushaltsnettoeinkommen	Bis 1000 Euro	25,0 %
	1001 - 2000 Euro	19,0 %
	2001 - 3000 Euro	18,0 %
	3001 - 4000 Euro	10,0 %
	4001 - 5000 Euro	6,0 %
	Mehr als 5000 Euro	11,0 %
	Keine Angabe	11,0 %

Die Teilnehmer des Experiments nutzen WhatsApp eher für die private Kommunikation. Dabei verwenden 87,0 % WhatsApp für die tägliche private Kommunikation. Für die Kommunikation mit Unternehmen haben 70,0 % der Teilnehmer hingegen noch nie WhatsApp genutzt. Dadurch haben sie voraussichtlich noch seltener mit einem Chatbot über WhatsApp kommuniziert. Tabelle 23 gibt einen Überblick über den Nutzungsgrund von WhatsApp.

Tabelle 23: Nutzung von WhatsApp (FP II: E 2)

Nutzungsgrund:	Private Kommunikation	Kommunikation mit Unternehmen
Nie	3,0 %	70,0 %
Einmal im Monat	3,0 %	20,0 %
Wöchentlich	7,0 %	9,0 %
Täglich	87,0 %	1,0 %

Die Gütebeurteilung der gemessenen Konstrukte erfolgt auch in diesem Experiment auf Basis der in Kapitel 4.2 beschriebene Gütekriterien und Grenzwerte. Die geforderten Grenzwerte wurden von allen Indikatoren und Konstrukten eingehalten. In Tabelle 24 ist die Gütebeurteilung zusammengefasst. Hier zeigt sich, dass die Grenzwerte der Gütekriterien für Reliabilität und Validität alle überschritten sind. Daher werden die Skalen für die folgende Auswertung verwendet.

Tabelle 24: Gütekriterien (FP II: E 2)

Konstrukt	Indikator	Faktorladung	KITK	Chronbachs Alpha	KMO	Erklärte Varianz
		> 0,4	> 0,5	> 0,7	> 0,6	> 50 %
Soziale Präsenz	SoPr_1	0,852	0,744	0,873	0,765	66,854
	SoPr_2	0,841	0,726			
	SoPr_3	0,803	0,683			
	SoPr_4	0,860	0,772			
	SoPr_5	0,725	0,594			
Wiederverwendung	Reu_1	0,854	0,755	0,895	0,849	71,438
	Reu_2	0,841	0,737			
	Reu_3	0,884	0,807			
	Reu_4	0,886	0,806			
	Reu_5	0,754	0,635			
Expertise	Exp_1	0,866	0,794	0,943	0,897	81,536
	Exp_2	0,927	0,881			
	Exp_3	0,867	0,796			
	Exp_4	0,910	0,855			
	Exp_5	0,941	0,903			

5.4.2.4 Ergebnisse

Zur Überprüfung der Hypothesen wird zuerst eine zweifaktoriellen ANOVA durchgeführt, um den Einfluss des Offenlegungszeitpunkts sowie der Kommunikationsplattform zu überprüfen und dabei auch den Interaktionseffekt zu berücksichtigen. In einem zweiten Schritt wird mit einer moderierten Mediationsanalyse untersucht, ob der Offenlegungszeitpunkt die Wiederverwendungsabsicht über die Soziale Präsenz beeinflusst und welchen Einfluss die moderierende Rolle der Kommunikationsplattform hat.

Hypothesen H_1 und H_3 wurden mithilfe einer zweifaktoriellen ANOVA in SPSS überprüft. Die Ergebnisse zeigen weder einen direkten Effekt des Offenlegungszeitpunkts auf die Soziale Präsenz ($F_{\text{Offenlegungszeitpunkt}}(1, 99) = 2,370, p > 0,05$) noch einen Interaktionseffekt zwischen dem Offenlegungszeitpunkt und der Kommunikationsplattform, also der Vertrautheit mit der dieser ($F_{\text{Offenlegungszeitpunkt} \times \text{Plattform}}(1, 99) = 0,413, p > 0,05$). Aufgrund des Fehlens der Interaktion kann kein Moderationseffekt der Kommunikationsplattform erreicht werden. Folglich können die Hypothese H_1 und H_3 nicht bestätigt werden. Die geschätzten Mittelwerte der Sozialen Präsenz der vier Gruppen sind in Abbildung 7 visualisiert.

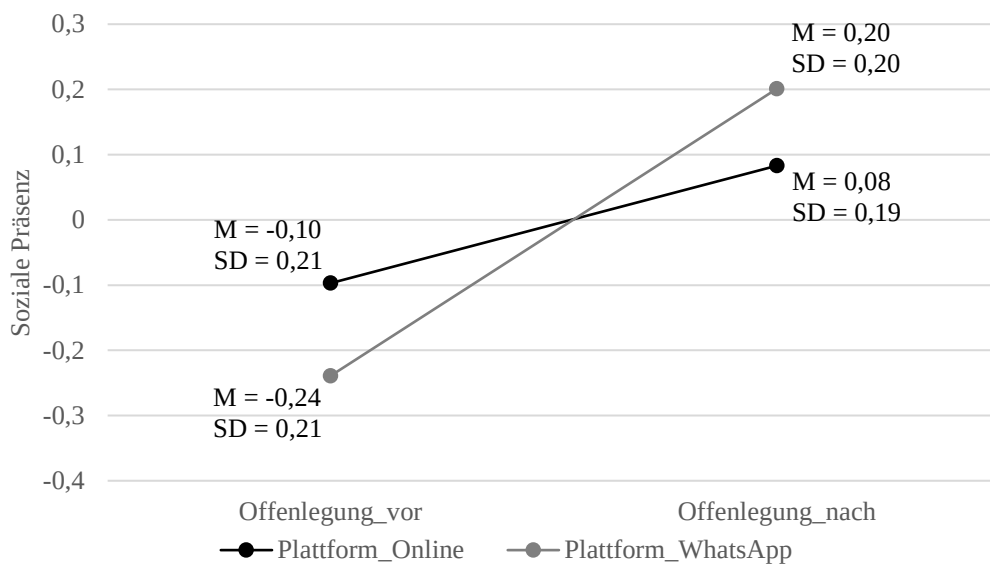


Abbildung 7: Ergebnisse der ANOVA (FP II: E 2) – Effekte des Offenlegungszeitpunkts auf die Soziale Präsenz

Zur Prüfung der Hypothesen H₂ und H₃ wird in einem weiteren Schritt die moderierte Mediation mithilfe des Modells 7 des PROCESS-Makros mit einem Bootstrap Sample von 5.000 durchgeführt (Hayes, 2022, S. 625). Hier wird nach den nicht signifikanten Ergebnissen der ANOVA auch kein Effekt erklärbar sein. Zur Vollständigkeit wird die Analyse allerdings noch durchgeführt. Als unabhängige Variable wird der Offenlegungszeitpunkt verwendet. Die Soziale Präsenz wird als Mediatorvariable und die Wiederverwendungsabsicht als abhängige Variable im PROCESS Modell verwendet. Die Kommunikationsplattform wird hier als Moderator genutzt. Das untersuchte moderierte Mediationsmodell ist in Abbildung 8 visualisiert. Als Referenzkategorie dient die Gruppe *Online-Chat* und *Offenlegung_vor*.

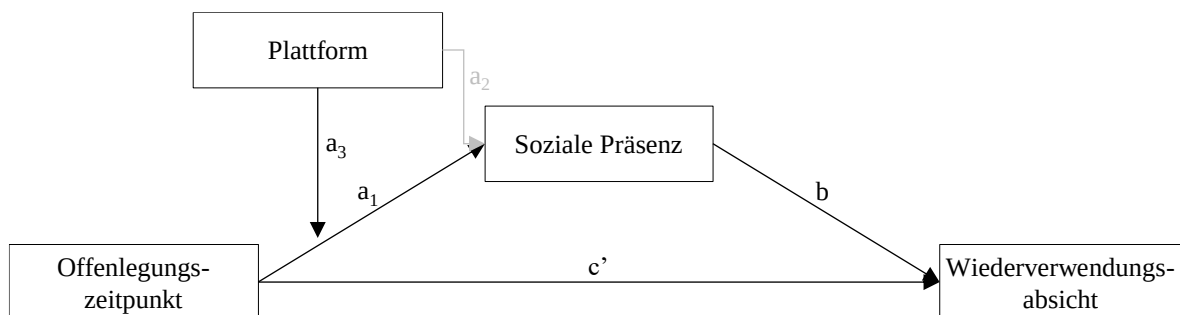


Abbildung 8: Moderiertes Mediationsmodell (FP II: E 2)

Die Ergebnisse zeigen einen signifikanten Effekt der Sozialen Präsenz auf die Wiederverwendungsabsicht ($b = 0,398$, $p < 0,001$). Der direkte Einfluss des Offenlegungszeitpunkts auf die Wiederverwendungsabsicht ist knapp nicht signifikant ($c' = -0,360$, $p = 0,058$). Der Einfluss des

Offenlegungszeitpunkts auf die Soziale Präsenz konnte auch nicht nachgewiesen werden ($a_1 = 0,181$, $p > 0,05$). Der Mediationseffekt ist sowohl für Online als auch WhatsApp nicht signifikant (Online: $a_1 \times b = 0,078$, 95 % KI [-0,146, 0,339]; WhatsApp: $a_1 \times b = 0,175$, 95 % KI [-0,042, 0,420]). Der Moderationseffekt der Kommunikationsplattform konnte nicht nachgewiesen werden ($a_3 = 0,259$, $p > 0,05$). Daher können H_2 und H_3 im Rahmen von *FP II: E 3* nicht bestätigt werden. Die Ergebnisse der Analyse sind in Tabelle 25 übersichtlich dargestellt. Tabelle 26 zeigt einen Überblick der Hypothesen und dem Ergebnis der Analysen in Bezug auf die Hypothesen.

Tabelle 25: Regressionskoeffizienten der moderierten Mediationsanalyse (FP II: E 2)

	Symbol	Koeff.	KI _u	KI _o
Effekte auf die Wiederverwendungsabsicht				
Offenlegungszeitpunkt	c'	-0,360 ^{ns}	-0,732	0,012
Soziale Präsenz	b	0,398 ^{***}	0,212	0,584
Effekt auf die Soziale Präsenz				
Offenlegungszeitpunkt	a	0,181 ^{ns}	-0,381	0,743
Moderationseffekt				
Offenlegungszeitpunkt x Kommunikationsplattform	a ₃	0,259 ^{ns}	-0,541	1,059
	Symbol	Effekt	KI _u	KI _o
Offenlegungszeitpunkt x Soziale Präsenz (Online)	a ₁₁ x b	0,072	-0,146	0,339
Offenlegungszeitpunkt x Soziale Präsenz (WhatsApp)	a ₁₂ x b	0,175	-0,042	0,420
***: $p < 0,001$; **: $p < 0,01$; *: $p < 0,05$; ns: $p > 0,05$				

Tabelle 26: Zusammenfassung der Hypothesen (FP II: E 2)

Hypothese H _i		Ergebnis
H ₁	Die Offenlegung der Identität eines Chatbots vor der Interaktion (im Vergleich zur Offenlegung nach der Interaktion) verringert die Soziale Präsenz.	✗
H ₂	Soziale Präsenz mediiert den Effekt des Offenlegungszeitpunkts der Identität eines Chatbots auf die Absicht, den Chatbot wiederzuverwenden.	✗
H ₃	Die Art der Kommunikationsplattform (Online-Chat vs. WhatsApp) moderiert den Mediationseffekt des Offenlegungszeitpunkts auf die Wiederverwendungsabsicht über die Soziale Präsenz, sodass der Mediationseffekt größer ist, wenn die Kommunikation in WhatsApp stattfindet.	✗
✓ = bestätigt; ✗ = nicht bestätigt		

5.4.2.5 Diskussion der Ergebnisse

In *FP II: E 2* konnten kein Einfluss des Offenlegungszeitpunkts auf die Soziale Präsenz gefunden werden. Dieses Ergebnis bestätigt das Ergebnis von Van der Goot et al. (2024, S. 10), die keinen signifikanten Einfluss der Offenlegung auf die Soziale Präsenz nachweisen konnten. Die in *FP II: E 1* gezeigten Erkenntnisse in Bezug auf die Soziale Präsenz konnten nicht bestätigt werden.

Auch die Art der Kommunikationsplattform beziehungsweise die Vertrautheit mit der Kommunikationsplattform kann folglich keinen moderierenden Effekt auf den Einfluss des Offenlegungszeitpunkt auf die Soziale Präsenz haben. Bei Betrachtung der Mittelwerte in Abbildung 7 fällt allerdings auf, dass sich diese für die Gruppen voneinander unterscheiden. So ist die Soziale Präsenz beim *WhatsApp-Chat* geringer als beim *Online-Chat*, wenn der Chatbot vor der Interaktion offengelegt wird. Bei der Offenlegung nach der Interaktion ist die Soziale Präsenz höher für die Gruppe *WhatsApp-Chat*. Eine mögliche Erklärung könnte in der Vertrautheit mit dem Kommunikationskanal liegen, ebenso wie darin, dass WhatsApp vorwiegend für die private Kommunikation verwendet wird. Allerdings sind die Gruppenunterschiede nicht signifikant. Daher dürfen keine allgemeingültigen Aussagen davon abgeleitet werden.

Durch die fehlerhafte Beantwortung der Manipulationscheckfragen wurde die Stichprobe klein. Für eine größere Stichprobe oder eine reale Chatinteraktion könnten bessere Ergebnisse erreicht werden. Soziale Präsenz scheint ein wichtiger Einflussfaktor auf die Wiederverwendungsabsicht von Chatbots zu sein. Hier konnte gezeigt werden, dass eine höhere Soziale Präsenz die Wiederverwendungsabsicht erhöht. In der Literatur gibt es bisher nur wenig Forschung zum

Thema des Offenlegungszeitpunkts und der Rolle der Kommunikationsplattform. Dieses Forschungsprojekt stellt einen ersten Ansatz dar, um dieses Themenfeld zu bearbeiten.

5.4.3 Experiment 3: Offenlegungszeitpunkt eines Chatbots und Lösbarkeit der Nutzeranfrage

Das dritte Experiment des Forschungsprojekts II (*FP II: E 3*) untersucht wie sich der Zeitpunkt der Offenlegung eines anthropomorph gestalteten Chatbots auf die Wiederverwendungsabsicht auswirkt. Dabei werden das vermutete Motiv, der wahrgenommene Betrug sowie die Fairnesswahrnehmung als Mediatoren betrachtet. Darüber hinaus wurde in dem Experiment eine zweite unabhängige Variable betrachtet, nämlich, ob der Chatbot eine Lösung für die Anfrage liefern konnte (*Lösung_ja*) oder nicht (*Lösung_nein*). Auf eine Weiterführung des Themas Kommunikationsplattform wurde an dieser Stelle verzichtet, da in *FP II: E 2* kein moderierender Einfluss festgestellt werden konnte.

5.4.3.1 Herleitung der Forschungshypothesen

Gerade ältere Kundenservice-Chatbots sind in ihren Funktionen noch stark eingeschränkt und nicht mit modernen, auf künstlicher Intelligenz basierenden Chatbots wie beispielsweise ChatGPT vergleichbar. Daher kommt es bei Chatbots im Kundenservice noch häufig zu Fehlern und die Anfrage des Kunden werden nicht gelöst. Die Offenlegung der nicht-menschlichen Identität eines Chatbots kann das Vertrauen und damit die Kundenbindung erhöhen, wenn der Chatbot nicht die erwartete Leistung erbringt (Mozafari et al., 2022, S. 236). Die Offenlegung stellt eine Art der Erklärung für das negative Ergebnis dar und kann wie eine „Entschuldigung“ wahrgenommen werden (Mozafari et al., 2022, S. 236). Kunden können sich von Beginn an darauf einstellen und dadurch gegebenenfalls die Erwartungen verringern, da sie explizit wissen, dass sie nicht mit einem Menschen interagieren. Der Einfluss der Offenlegung und noch detaillierter der Einfluss des Offenlegungszeitpunkts im Rahmen von fehlerhaften Chatinteraktionen im Vergleich zu erfolgreichen Chatinteraktionen wurde in der Literatur bisher wenig Beachtung geschenkt (Mozafari et al., 2022, S. 226).

Im Falle einer späten Offenlegung ist das Unternehmen nicht von Beginn der Konversation ehrlich und offen zum Kunden. Der Kunde kann sich nicht von Anfang an darauf einstellen, dass er nicht mit einem Menschen redet und erfährt dies erst im Nachgang der Konversation.

Diese Tatsache kann beim Kunden negative Reaktionen hervorrufen. Dabei kann sich der Kunde betrogen fühlen oder hinter der Offenlegung ein negatives Motiv sehen. Er könnte vermuten, dass das Unternehmen die Offenlegung nach der Interaktion aus eigennützigen Motiven vornimmt, den Kunden täuschen möchte oder ihn durch das Verschweigen der nicht-menschlichen Identität dazu bewegen will, den Chatbot zu Nutzen. Darüber hinaus kann ein solches Vorgehen auch als unfair wahrgenommen werden. Wohingegen bei einer Offenlegung vor der Interaktion das Unternehmen von Beginn der Kommunikation ehrlich zu dem Kunden ist. Dadurch können solche negativen Reaktionen verringert oder gänzlich vermieden werden. Daher werden folgende Hypothesen aufgestellt:

- H_{1a}: Kunden fühlen sich eher betrogen, wenn die Identität nach der Interaktion offengelegt wird, als wenn dies vor der Interaktion geschieht.
- H_{1b}: Die Art der Kommunikation, dass es sich um einen Chatbot handelt, wird als weniger fair wahrgenommen, wenn die Identität nach der Interaktion offengelegt wird, als wenn die Identität vor der Interaktion offengelegt wird.
- H_{1c}: Das vermutete Motiv hinter der Offenlegung ist schlechter, wenn die Identität nach der Interaktion offengelegt wird, als wenn dies vor der Interaktion geschieht.

Die Reaktion von Kunden auf Chatbots kann von der auf einen Menschen abweichen. Gerade beim Auftreten eines Fehlers im Verlauf der Chatinteraktion kann sich die Kundenreaktion unterscheiden (Belanche et al., 2020, S. 281). In Übereinstimmung mit der Attributionstheorie können negative Ereignisse das Bestreben verstärken, den kausalen Zusammenhang zu verstehen (Kanazawa, 1992, S. 659; Weiner, 1985a, S. 77). Folglich kann die Reaktion von Kunden auf den Offenlegungszeitpunkt von Chatbots je nach Ergebnis der Chatinteraktion unterschiedlich ausfallen. Denn bei Fehlern des Chatbots kann der Wille der Zuschreibung der Ursache und der sich daraus ergebenden negativen Konsequenzen stärker sein. Der von Mozafari et al. (2022) beschriebene positive Einfluss der Offenlegung kann durch eine späte Offenlegung nochmal zu Nichte gemacht werden, da negative Kundenreaktionen überwiegen. Durch eine späte Offenlegung der Chatbot-Identität kann der Kunde negative Beweggründe hinter der Offenlegung vermuten. In einer Situation, in der der Chatbot einen Fehler macht, kann eine späte Offenlegung dazu führen, dass der Kunde sich betrogen fühlt. Er könnte vermuten, dass hinter der Offenlegung ein negatives Unternehmensmotiv steckt und die Art und Weise der Kommunikation, welche die Offenlegung des Chatbots betrifft, als unfair wahrnehmen, insbesondere

im Vergleich zu einer Situation, in der der Chatbot die Kundenanfrage korrekt löst. Daher werden die folgenden Hypothesen aufgestellt:

H_{2a}: Die Lösungsmöglichkeit (Lösung vs. keine Lösung) moderiert die Beziehung zwischen Offenlegungszeitpunkt und wahrgenommenem Betrug, sodass der wahrgenommene Betrug größer ist, wenn der Chatbot keine Lösung liefert.

H_{2b}: Die Lösungsmöglichkeit (Lösung vs. keine Lösung) moderiert die Beziehung zwischen Offenlegungszeitpunkt und Fairnesswahrnehmung, sodass die Fairnesswahrnehmung geringer ist, wenn der Chatbot keine Lösung liefert.

H_{2c}: Die Lösungsmöglichkeit (Lösung vs. keine Lösung) moderiert die Beziehung zwischen Offenlegungszeitpunkt und vermutetem Motiv, sodass das vermutete negative Motiv hinter der Offenlegung höher ist, wenn der Chatbot keine Lösung liefert.

Die negativen Reaktionen werden sich auch auf die Wiederverwendungsabsicht auswirken. Betrug, das vermutete Motiv und unfaire Behandlung können zu einer geringeren Wiederverwendungsabsicht führen. Daraus ergeben sich die folgenden drei Hypothesen:

H_{3a}: Eine hohe Wahrnehmung von Betrug führt zu einer verringerten Absicht der Nutzer, den Chatbot erneut zu verwenden.

H_{3b}: Eine geringere Fairnesswahrnehmung verringert die Absicht der Nutzer, den Chatbot erneut zu verwenden.

H_{3c}: Ein schlechtes vermutetes Motiv hinter der Offenlegung verringert die Absicht der Nutzer, den Chatbot erneut zu verwenden.

5.4.3.2 Methodik

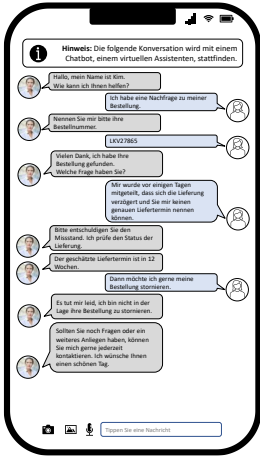
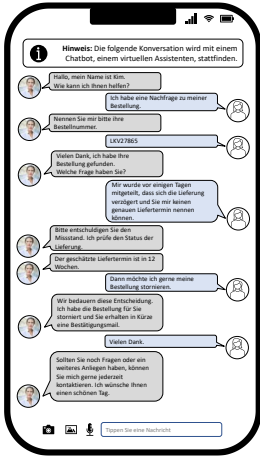
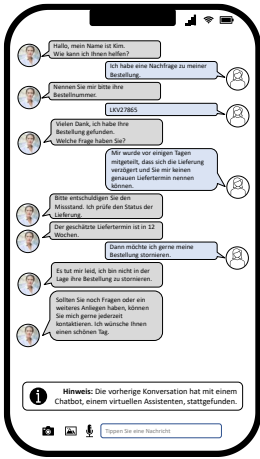
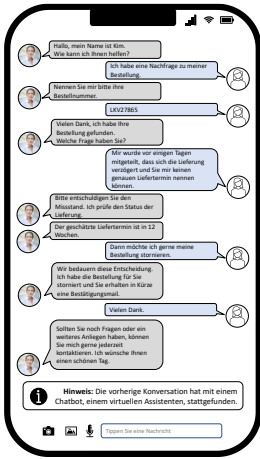
In *FP II: E 3* wird auch ein hypothetisches Kundenservice-Szenario genutzt. Hierbei sollen sich die Teilnehmer in die Situation versetzen, dass sie Informationen zum Lieferstatus ihrer Bestellung erhalten möchten, da sich die Lieferung verzögert. Sollte sich die Lieferung weiter verzögern, soll die Bestellung storniert werden. Hierfür sollen die Teilnehmer die Chatfunktion des Händlers nutzen. Auch hier werden den Teilnehmern die Teile der Screenshots der Chatinteraktion sukzessive gezeigt. Es wird ein zweifaktorielles Design mit Zufallsgruppenbildung verwendet. Dabei wird der Offenlegungszeitpunkt (*Offenlegung_vor* und *Offenlegung_nach*) sowie der Erfolg Kundenanfrage also ob diese gelöst werden kann oder nicht (*Lösung_nein* und

Lösung_ja) manipuliert. Die Gruppe *Lösung_ja* konnte die Bestellung über den Chatbot stornieren. Wohingegen in der Gruppe *Lösung_nein* die Bestellung nicht storniert werden konnte. Wie in *FP II: E 2* wurde den Teilnehmern im Chat in einer Hinweisbox direkt vor oder nach der Interaktion mit dem folgenden Inhalt gezeigt: „Hinweis: Die folgende/vorherige Konversation wird/hat mit einem Chatbot stattfinden/stattgefunden.“ Das Szenario mit den Stimuli der vier Gruppen ist in Tabelle 27 dargestellt.

Tabelle 27: Szenariobeschreibung (FP II: E 3)

Einführung	
Bitte versetzen Sie sich in die Situation, dass Sie Informationen zum Lieferstatus Ihrer Bestellung erhalten möchten. Das bestellte Produkt ist nicht wie ursprünglich angegeben lieferbar. Sollte sich die Lieferung weiter verzögern, möchten Sie Ihre Bestellung stornieren. Sie sehen auf der Webseite des Händlers nur die Möglichkeit über eine Chatfunktion mit dem Unternehmen zu kommunizieren. Sie entscheiden sich, diese Funktion zu nutzen.	
Gruppen	
Lösung_nein	Lösung_ja
Chatinteraktion	
• Hallo, mein Name ist Kim. Wie kann ich Ihnen helfen? • Ich habe eine Nachfrage zu meiner Bestellung. • Nennen Sie mir bitte ihre Bestellnummer. • LKV27865 • Vielen Dank, ich habe Ihre Bestellung gefunden. Welche Frage haben Sie? • Mir wurde vor einigen Tagen mitgeteilt, dass sich die Lieferung verzögert und Sie mir keinen genauen Liefertermin nennen können. • Bitte entschuldigen Sie den Missstand. Ich prüfe den Status der Lieferung. • Der geschätzte Liefertermin ist in 12 Wochen. • Dann möchte ich gerne meine Bestellung stornieren. • Es tut mir leid, ich bin nicht in der Lage ihre Bestellung zu stornieren. Sollten Sie noch Fragen oder ein weiteres Anliegen haben, können Sie mich gerne jederzeit kontaktieren. Ich wünsche Ihnen einen schönen Tag.	• Hallo, mein Name ist Kim. Wie kann ich Ihnen helfen? • Ich habe eine Nachfrage zu meiner Bestellung. • Nennen Sie mir bitte ihre Bestellnummer. • LKV27865 • Vielen Dank, ich habe Ihre Bestellung gefunden. Welche Frage haben Sie? • Mir wurde vor einigen Tagen mitgeteilt, dass sich die Lieferung verzögert und Sie mir keinen genauen Liefertermin nennen können. • Bitte entschuldigen Sie den Missstand. Ich prüfe den Status der Lieferung. • Der geschätzte Liefertermin ist in 12 Wochen. • Dann möchte ich gerne meine Bestellung stornieren. • Wir bedauern diese Entscheidung. Ich habe die Bestellung für Sie storniert und Sie erhalten in Kürze eine Bestätigungsmail. • Vielen Dank. • Sollten Sie noch Fragen oder ein weiteres Anliegen haben, können Sie mich gerne jederzeit kontaktieren. Ich wünsche Ihnen einen schönen Tag.

Tabelle 27: Szenariobeschreibung (FP II: E 3) (Fortsetzung)

Lösung_nein / Offenlegung_vor	Lösung_ja / Offenlegung_vor
Hinweis: Die folgende Konversation wird mit einem Chatbot, einem virtuellen Assistenten, stattfinden.	
	
Lösung_nein / Offenlegung_nach	Lösung_ja / Offenlegung_nach
	
Hinweis: Die vorherige Konversation hat mit einem Chatbot, einem virtuellen Assistenten, stattgefunden.	

Nach dem das Szenario durchlaufen wurde, werden die Teilnehmer zur Wiederverwendungsabsicht (Moriuchi, 2021, S. 27-28), dem vermuteten Motiv hinter der Offenlegung (Kukar-Kinney et al., 2007, S. 329), dem wahrgenommenen Betrug (Grégoire & Fisher, 2008, S. 252) sowie der Fairnesswahrnehmung (Garbarino & Maxwell, 2010, S. 1071) befragt. Auch in diesem Experiment wurde zur Messung der Konstrukte auf Skalen aus der bisherigen Forschung zurückgegriffen, die auf Basis des in Kapitel 5.2 vorgestellten Prozess ins Deutsche übersetzt wurden. Mithilfe zweier Manipulationscheckfragen wurde die Manipulation überprüft. Anschließend wurden die soziodemographischen Merkmale in der Umfrage erfasst sowie die Nutzungshäufigkeit von Chatbots und die Expertise der Teilnehmer in Bezug auf Chatbots (Ohanian, 1990, S. 50) abgefragt. Auch in diesem Experiment sollten die Teilnehmer beurteilen, wie

realitätsnah das Szenario ist und wie gut sie sich in die Situation hineinversetzen konnten. Tabelle 28 zeigt einen Überblick der verwendeten Skalen mit den übersetzten Indikatoren. Die nicht abgebildeten Skalen sind identisch mit denen aus vorherigen Experimenten.

Tabelle 28: Operationalisierung der Variablen und Manipulationscheck (FP II: E 3)

Kürzel	Indikator	Skala
Manipulationscheck		
Inwiefern stimmen sie den folgenden Aussagen zu?		
MC_11	Ich wurde nach der Interaktion informiert, dass ich mit einem Chatbot kommuniziere.	7-Punkt-Likert-Skala: 1 = Stimme überhaupt nicht zu, 7 = Stimme vollkommen zu
MC_12	Ich wurde vor der Interaktion informiert, dass ich mit einem Chatbot kommuniziere.	
MC_21	Ich konnte meine Bestellung stornieren .	
MC_22	Ich konnte meine Bestellung nicht stornieren .	
Wahrgenommene Fairness (Garbarino & Maxwell, 2010, S. 1071)		
Bitte beantworten Sie die folgenden Fragen.		
Fair_1	Wie würden Sie die Fairness der Kommunikation, dass es sich um einen Chatbot handelt, in dieser Situation bewerten?	7-Punkt-Likert-Skala: 1 = Überhaupt nicht fair, 7 = Sehr fair
Fair_2	Wie fair bewerten Sie die Art der Kommunikation, dass es sich um einen Chatbot handelt in dieser Situation?	
Inwiefern stimmen Sie den folgenden Aussagen zu?		
Fair_3 ^R	Ich finde diese Art der Kommunikation, dass es sich um einen Chatbot handelt, unfair.	7-Punkt-Likert-Skala: 1 = Stimme überhaupt nicht zu, 7 = Stimme vollkommen zu
Fair_4	Diese Art der Kommunikation, dass es sich um einen Chatbot handelt, ist fair für den Kunden.	
Fair_5	Ein objektiver Dritter würde sagen, die Art der Kommunikation, dass es sich um einen Chatbot handelt, ist fair.	
Wahrgenommener Betrug (Grégoire & Fisher, 2008, S. 252)		
Inwiefern stimmen Sie den folgenden Aussagen zu?		
PerBe_1	Ich fühle mich hintergangen.	7-Punkt-Likert-Skala: 1 = Stimme überhaupt nicht zu, 7 = Stimme vollkommen zu
PerBe_2	Ich fühle mich betrogen.	
PerBe_3	Ich fühle mich belogen.	
PerBe_4	Das Unternehmen hatte die Absicht, mich auszunutzen.	
PerBe_5	Das Unternehmen hat versucht, mich über den Tisch zu ziehen.	

Tabelle28: Operationalisierung der Variablen (FP II: E 3) (Fortsetzung)

Kürzel	Indikator	Skala
Vermutetes Motiv (Kukar-Kinney et al., 2007, S. 329)		
Inwiefern stimmen Sie den folgenden Aussagen zu?		
InfMo_1	Das Unternehmen nutzt die Art der Kommunikation, dass es sich um einen Chatbot handelt, um seine eigenen Interessen zu verfolgen.	7-Punkt-Likert-Skala: 1 = Stimme überhaupt nicht zu, 7 = Stimme vollkommen zu
InfMo_2	Das Unternehmen nutzt die Art der Kommunikation, dass es sich um einen Chatbot handelt, um seinen Gewinn auf Kosten der Kunden zu steigern.	
InfMo_3	Das Unternehmen will seine Kunden durch die Art der Kommunikation, dass es sich um einen Chatbot handelt, ausnutzen.	
Wiederverwendung (Moriuchi, 2021, S. 27-28)		
Inwiefern stimmen Sie den folgenden Aussagen zu?		
Reu_1	Wenn ich Zugang zu einem Chatbot hätte, würde ich ihn immer nutzen.	7-Punkt-Likert-Skala: 1 = Stimme überhaupt nicht zu, 7 = Stimme vollkommen zu
Reu_2	Ich denke, es wird sich für mich lohnen, Chatbots zu verwenden, wenn sie verfügbar sind.	
Reu_3	Ich würde meinen Freunden und meiner Familie die Nutzung eines Chatbots empfehlen.	
Reu_4	Wenn ich die Möglichkeit hätte mehr über Chatbots mit Unternehmen zu kommunizieren, würde ich dies tun.	
Reu_5	Ich nutze lieber einen Chatbot, um mein Anliegen zu klären, als eine andere Person zu fragen.	
R = Invers formulierter Indikator		

5.4.3.3 Stichprobe und Gütebeurteilung

Insgesamt nahmen 210 Teilnehmer an dem Experiment teil. Nach Prüfung der Manipulationscheckfragen wurden 60 Personen aus der Stichprobe entfernt. Diese Personen haben mindestens eine der Manipulationscheckfragen mit einem Wert von vier oder niedriger auf der 7-Punkte-Likert-Skala beantwortet. 48,0 % der Teilnehmer sind weiblich, 50,0 % männlich, 1,3 % machen keine Angabe zu ihrem Geschlecht und 0,7 % geben Divers an. Das Durchschnittsalter der Teilnehmer liegt bei 38,22 Jahren (SD 14,84). 84,0 % der Befragten haben mindestens einmal einen Chatbot in den letzten zwölf Monaten verwendet. Die Teilnehmer bewerteten die Situation des Szenarios als sehr realistisch ($M = 5,99$, $SD = 1,32$) und sie konnten sich einfach in die Situation hineinversetzen ($M = 6,26$, $SD = 1,07$). Die Teilnehmer bewerteten sich selbst als mittelmäßig erfahren im Umgang mit Chatbots ($M = 3,71$, $SD = 1,55$). Hier wurde die Skala mithilfe der Mean-Funktion berechnet. Eine detaillierte Stichprobenbeschreibung zeigt Tabelle

29. Die Verteilung der Nutzungshäufigkeit von Chatbots in den letzten zwölf Monaten der Teilnehmer ist in Anhang 1 zusammengefasst.

Tabelle 29: Soziodemographische Merkmale (FP II: E 3)

Merkmal	Ausprägung	Prozentualer Anteil
Geschlecht	Weiblich	48,0 %
	Männlich	50,0 %
	Divers	0,7 %
	Keine Angabe	1,3 %
Alter	M = 38,22 (SD = 14,84), 2 kein Angabe	
Haushaltsgröße	1 Person	16,0 %
	2 Personen	43,3 %
	3-5 Personen	37,3 %
	Mehr als 5 Personen	0,7 %
	Keine Angabe	2,7 %
Höchster Bildungsabschluss	Noch in schulischer Ausbildung	2,0 %
	Volks-/Hauptschulabschluss	4,0 %
	Realschulabschluss	16,7 %
	Fachschulabschluss	8,7 %
	Abitur	19,3 %
	Fachhoch-/Hochschulabschluss	48,0 %
	Keine Angabe	1,3 %
Monatliches Haushaltsnettoeinkommen	Bis 1000 Euro	6,7 %
	1001 - 2000 Euro	16,0 %
	2001 - 3000 Euro	16,0 %
	3001 - 4000 Euro	12,0 %
	4001 - 5000 Euro	11,3 %
	Mehr als 5000 Euro	20,0 %
	Keine Angabe	18,0 %

Die Güte der gemessenen Konstrukte wurde auf Basis der in Kapitel 4.2 beschriebenen Gütekriterien und Grenzwerte bewertet. In Tabelle 30 sind die Gütebeurteilungen nach Bereinigung zusammengefasst. Hier zeigt sich, dass die Grenzwerte der Gütekriterien für Reliabilität und Validität alle überschritten sind. Daher werden die Skalen für die folgenden Auswertungen verwendet.

Tabelle 30: Gütekriterien (FP II: E 3)

Konstrukt	Indikator	Faktorladung	KITK	Chronbachs Alpha	KMO	Erklärte Varianz
		> 0,4	> 0,5	> 0,7	> 0,6	> 50 %
Wahrgenommene Fairness	Fair_1	0,908	0,849	0,937	0,862	80,197
	Fair_2	0,891	0,825			
	Fair_3	0,852	0,774			
	Fair_4	0,928	0,882			
	Fair_5	0,896	0,835			
Vermutetes Motiv	InfMo_1	0,798	0,553	0,748	0,690	66,887
	InfMo_2	0,827	0,589			
	InfMo_3	0,828	0,595			
Wahrgenommener Betrug	PerBet_1	0,946	0,915	0,963	0,862	87,188
	PerBet_2	0,940	0,906			
	PerBet_3	0,934	0,895			
	PerBet_4	0,916	0,870			
	PerBet_5	0,932	0,892			
Wiederverwendung	Reu_1	0,846	0,744	0,869	0,863	65,994
	Reu_2	0,832	0,719			
	Reu_3	0,802	0,681			
	Reu_4	0,841	0,731			
	Reu_5	0,735	0,603			

5.4.3.4 Ergebnisse

Das Testen der Hypothesen erfolgt zweigeteilt. Zuerst wird mit einer zweifaktoriellen ANOVA überprüft, welchen Einfluss der Offenlegungszeitpunkt sowie die Lösungsmöglichkeit haben und dabei auch explizit auf die Interaktion dieser Einflussfaktoren geschaut. In einem zweiten Schritt wird mit einer moderierten Mediationsanalyse untersucht, ob der Offenlegungszeitpunkt die Wiederverwendungsabsicht über die Fairnesswahrnehmung, das vermutete Motiv sowie den wahrgenommene Betrug beeinflusst und ob die Lösungsmöglichkeit hier eine moderierende Rolle hat.

Zur Überprüfung von Hypothesen H_1 und H_2 wurden drei zweifaktorielle ANOVAs in SPSS gerechnet. Die Ergebnisse zeigen keinen Interaktionseffekt des Offenlegungszeitpunkts und der Lösungsmöglichkeit ($F_{\text{PerBet}}(3, 146) = 0,454, p > 0,05$), $F_{\text{InfMo}}(3, 146) = 0,763, p > 0,05$), $F_{\text{Fair}}(3, 146) = 2,499, p > 0,05$). Folglich können Hypothesen H_{2a} , H_{2b} und H_{2c} nicht bestätigt

werden. Die Lösungsmöglichkeit des Chatbots moderiert weder die Beziehung zwischen dem Zeitpunkt der Offenlegung und dem wahrgenommenen Betrug, dem vermuteten Motiv noch der Fairnesswahrnehmung. Die Betrachtung der Abbildung 9 zeigt allerdings, dass sich die Mittelwerte zwischen den Lösungsmöglichkeiten unterscheiden. Hier wurden der wahrgenommene Betrug und das vermutete negative Motiv höher bewertet, wenn ein Fehler aufgetreten ist, also der Chatbot keine Lösung für die Anfrage liefern konnte. Ein Fehler wirkt auf die Wahrnehmung der Fairness negativ.

Sowohl der Offenlegungszeitpunkt und die Lösungsmöglichkeit haben einen Einfluss auf die Fairnesswahrnehmung ($F_{\text{Disclose}}(1, 146) = 20,691$ $p < 0,01$, $F_{\text{Fail}}(1, 146) = 25,628$ $p < 0,01$) und das vermutete Motiv ($F_{\text{Disclose}}(1, 146) = 4,610$ $p < 0,05$, $F_{\text{Fail}}(1, 146) = 17,955$ $p < 0,01$). Der wahrgenommene Betrug wird nur von der Lösungsmöglichkeit ($F_{\text{Disclose}}(1, 146) = 2,131$ $p > 0,05$, $F_{\text{Fail}}(1, 146) = 21,990$ $p < 0,01$) beeinflusst. Folglich können Hypothesen H_{1b} und H_{1c} bestätigt werden. Hypothese H_{1a} wurde im Experiment nicht bestätigt. Eine Offenlegung nach der Interaktion führt zu einem höheren vermuteten negativen Motiv sowie einer unfairen Wahrnehmung der Kommunikation, dass es sich um einen Chatbot handelt.

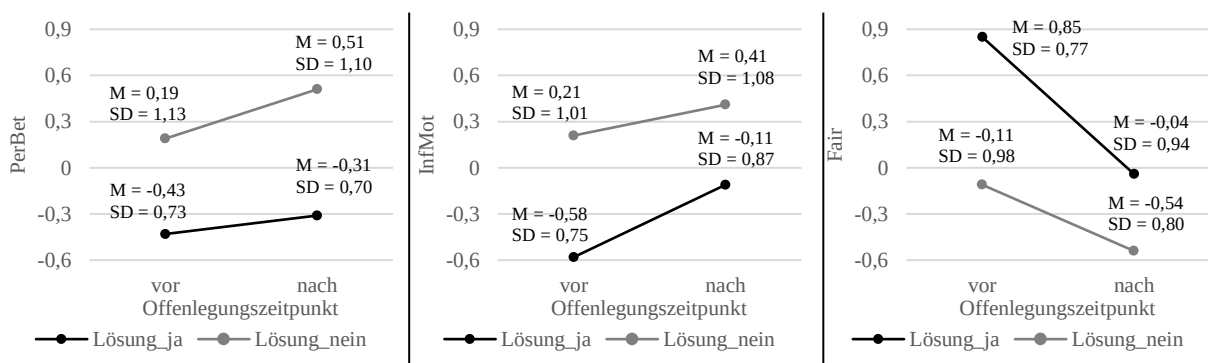


Abbildung 9: Ergebnisse der ANOVAs (FP II: E 3) – Effekte des Offenlegungszeitpunkts auf den wahrgenommenen Betrug, das vermutete Motiv und die Fairnesswahrnehmung

Zur Prüfung der Hypothesen H_2 und H_3 wird in einem zweiten Schritt die moderierte Mediation mithilfe des Modells 7 des PROCESS-Makros mit einem Bootstrap Sample von 5.000 durchgeführt (Hayes, 2022, S. 625). Hier gilt vorweg anzumerken, dass die Moderation wie aus den ANOVAs sichtbar nicht signifikant sein wird. Als unabhängige Variable wird der Offenlegungszeitpunkt verwendet. Die wahrgenommene Fairness, der wahrgenommene Betrug sowie das vermutete Motiv werden als Mediatorvariablen und die Wiederverwendungsabsicht als ab-

hängige Variable im PROCESS Modell verwendet. Die Lösungsmöglichkeit wird hier als Moderator genutzt. Das untersuchte moderierte Mediationsmodell ist in Abbildung 10 visualisiert. Als Referenzkategorie dient die Gruppe *Lösung_ja* und *Offenlegung_vor*.

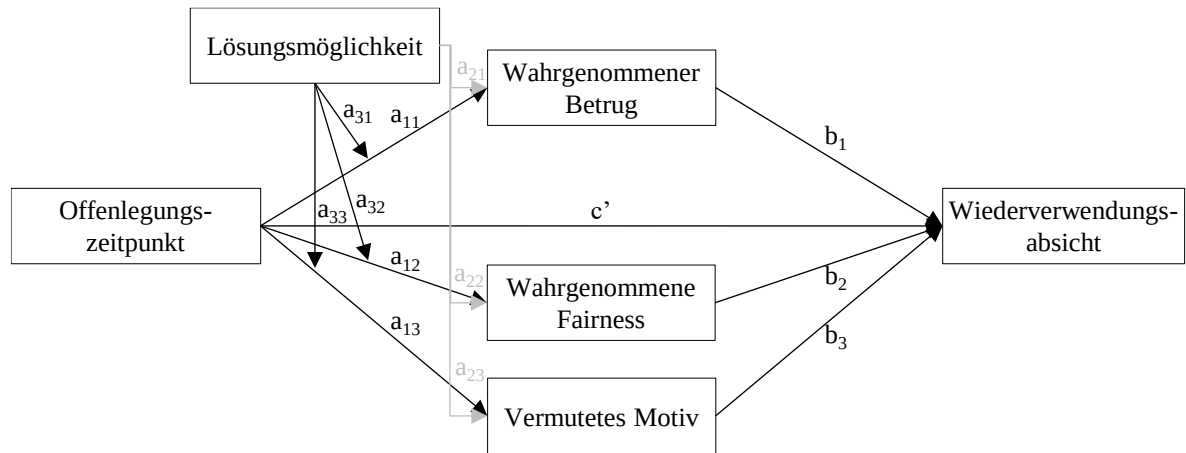


Abbildung 10: Moderiertes Mediationsmodell (FP II: E 3)

Die statistischen Kennzahlen der moderierten Mediation sind in Anlehnung an Hayes (2022, S. 463) in Tabelle 31 visualisiert. Der Offenlegungszeitpunkt beeinflusst die Wiederverwendungsabsicht indirekt über die wahrgenommene Fairness ($a_{12} = -0,89$, $p < 0,001$), die wiederum die Wiederverwendungsabsicht positiv beeinflusst ($b_2 = 0,42$, $p < 0,001$). Auch der indirekte Einfluss des Offenlegungszeitpunkts über das vermutete Motiv ($a_{13} = 0,47$, $p < 0,05$) auf die Wiederverwendungsabsicht ($b_3 = -0,28$, $p < 0,01$) konnte im Experiment gezeigt werden. Lediglich die Mediation über den wahrgenommenen Betrug ist nicht signifikant. Weder der Offenlegungszeitpunkt hat einen Einfluss auf den wahrgenommenen Betrug ($a_{11} = 0,12$, $p > 0,05$) noch der wahrgenommene Betrug auf die Wiederverwendungsabsicht ($b_1 = 0,06$, $p > 0,05$). Folglich können Hypothesen H_{1b} und H_{1c} sowie H_{3b} und H_{3c} bestätigt werden. Hypothesen H_{1a} und H_{3a} finden keine Bestätigung. Wie schon durch die ANOVAs abzusehen war, sind auch die Moderationspfade ($a_{31} = 0,21$, $p > 0,05$, $a_{32} = 0,46$, $p > 0,05$, $a_{33} = -0,27$, $p > 0,05$) nicht signifikant. Dies zeigt, dass die Lösungsmöglichkeit des Chatbots den Einfluss der Offenlegung weder auf den wahrgenommenen Betrug, die wahrgenommene Fairness, noch das vermutete Motiv moderiert. Daher können Hypothesen H_{2a} , H_{2b} , H_{2c} nicht bestätigt werden. Die bestätigten und nicht bestätigten Hypothesen sind in Tabelle 32 nochmal übersichtlich zusammengefasst.

Tabelle 32: Zusammenfassung der Hypothesen (FP II: E 3)

Hypothese H _i		Ergebnis
H _{1a}	Kunden fühlen sich eher betrogen, wenn die Identität nach der Interaktion offengelegt wird, als wenn dies vor der Interaktion geschieht.	✗
H _{1b}	Die Art der Kommunikation, dass es sich um einen Chatbot handelt, wird als weniger fair wahrgenommen, wenn die Identität nach der Interaktion offengelegt wird, als wenn die Identität vor der Interaktion offengelegt wird.	✓
H _{1c}	Das vermutete Motiv hinter der Offenlegung ist schlechter, wenn die Identität nach der Interaktion offengelegt wird, als wenn dies vor der Interaktion geschieht.	✓
H _{2a}	Die Lösungsmöglichkeit (Lösung vs. keine Lösung) moderiert die Beziehung zwischen Offenlegungszeitpunkt und wahrgenommenem Betrug, sodass der wahrgenommene Betrug größer ist, wenn der Chatbot keine Lösung liefert.	✗
H _{2b}	Die Lösungsmöglichkeit (Lösung vs. keine Lösung) moderiert die Beziehung zwischen Offenlegungszeitpunkt und Fairnesswahrnehmung, sodass die Fairnesswahrnehmung geringer ist, wenn der Chatbot keine Lösung liefert.	✗
H _{2c}	Die Lösungsmöglichkeit (Lösung vs. keine Lösung) moderiert die Beziehung zwischen Offenlegungszeitpunkt und vermutetem Motiv, sodass das vermutete negative Motiv hinter der Offenlegung höher ist, wenn der Chatbot keine Lösung liefert.	✗
H _{3a}	Eine hohe Wahrnehmung von Betrug führt zu einer verringerten Absicht der Nutzer, den Chatbot erneut zu verwenden.	✗
H _{3b}	Eine geringere Fairnesswahrnehmung verringert die Absicht der Nutzer, den Chatbot erneut zu verwenden.	✓
H _{3c}	Ein schlechtes vermutetes Motiv hinter der Offenlegung verringert die Absicht der Nutzer, den Chatbot erneut zu verwenden.	✓
✓ = bestätigt; ✗ = nicht bestätigt		

5.4.3.5 Diskussion der Ergebnisse

Die Ergebnisse von *FP II: E 3* zeigen, dass eine späte Offenlegung mit negativen Auswirkungen verbunden ist. So wird diese als weniger fair wahrgenommen und das vermutete Motiv hinter der Offenlegung ist schlechter als bei einer Offenlegung vor der Interaktion mit dem Chatbot. Vermuten Kunden ein schlechtes Motiv des Händlers, ist deren Wiederverwendungsabsicht geringer. Ähnlich ist es in Bezug auf die Fairnesswahrnehmung. Beurteilen Kunden die Fairness der Kommunikation, dass es sich um einen Chatbot handelt als unfair, sind sie weniger gewillt den Chatbot nochmals zu verwenden. Ob der Chatbot die Anfrage lösen kann oder nicht, hat keinen moderierenden Einfluss, so unterscheiden sich weder der wahrgenommene Betrug, die Fairnesswahrnehmung, noch das vermutete Motiv für die unterschiedlichen Offenlegungszeitpunkte. Die Reaktion der Kunden auf den unterschiedlichen Offenlegungszeitpunkt unterscheidet sich in Bezug auf die untersuchten Variablen nicht in Abhängigkeit des Ergebnisses der

Chatinteraktion. Die nicht signifikanten Ergebnisse in Bezug auf das Konstrukt des wahrgenommenen Betrugs könnten darauf zurückzuführen sein, dass das Verschweigen der Identität eines Chatbots im Kontext einer Nachfrage nach einer verzögerten Lieferung als zu unwichtig wahrgenommen wurde, um als Betrug eingestuft zu werden.

Durch eine späte Offenlegung kann der von Mozafari et al. (2022) beschriebenen positive Einfluss der Offenlegung verringert werden. Eine späte Offenlegung führt indirekt zu einer geringeren Wiederverwendungsabsicht, da Kunden ein schlechtes Motiv dahinter vermuten und die Kommunikation als weniger fair ansehen. Dies steht in Einklang mit den Ergebnissen, die Offenlegung allgemein einen negativen Effekt zuschreiben (Castillo et al., 2021; Ishowo-Oloko et al., 2019; Jin et al., 2024).

5.4.4 Abschließende Diskussion – Forschungsprojekt II

FP II untersucht den Einfluss des Offenlegungszeitpunkts auf die Wiederverwendungsabsicht. Dabei wurden in den drei Experimenten unterschiedliche Mediatoren und Moderatoren betrachtet. Als Faktoren, die einen direkten Einfluss auf die Wiederverwendungsabsicht haben, wurden die Soziale Präsenz des Chatbots, die Fairnesswahrnehmung sowie das vermutete Unternehmensmotiv hinter der Offenlegung identifiziert. Der Einfluss des Offenlegungszeitpunkts auf die Wiederverwendungsabsicht konnte nur in *FP II: E 1* gezeigt werden. In den beiden anderen Experimenten war der negative Einfluss nicht signifikant. Allerdings wurde die Wiederverwendungsabsicht bei allen Experimenten als geringer beurteilt, wenn die nicht-menschliche Identität des Chatbots erst nach der Chatinteraktion offengelegt wurde.

Der Zeitpunkt der Offenlegung konnte als Einflussfaktor auf die Soziale Präsenz nur in *FP II: E 1* bestätigt werden. Dabei hat eine späte Offenlegung die Wahrnehmung als sozial präsent erhöht. Eine höhere Soziale Präsenz führt zu einer höheren Wiederverwendungsabsicht. Dieser Mediationspfad ist in *FP II: E 1* dem direkten Effekt des Offenlegungszeitpunkts auf die Wiederverwendungsabsicht entgegengerichtet. Dieses Ergebnis konnte allerdings in *FP II: E 2* nicht bestätigt werden. Darüber hinaus wurde für die Kommunikationsplattform kein Interaktionseffekt in Bezug zum Offenlegungszeitpunkt gefunden. Hier war der Einfluss des Offenlegungszeitpunkts nicht von der Art der Kommunikationsplattform abhängig.

Im Rahmen von *FP II: E 3* konnte ein Einfluss des Offenlegungszeitpunkts auf die Fairnesswahrnehmung der Art und Weise der Kommunikation, dass es sich um einen Chatbot handelt,

gefunden werden. Dabei war die wahrgenommene Fairness geringer für eine Offenlegung nach der Interaktion. Des Weiteren wurde hinter einer späten Offenlegung negative Motive des Unternehmens vermutet. Durch geringere Fairnesswahrnehmung und die Beurteilung der Unternehmensmotive als negativ wird die Wiederverwendungsabsicht verringert. Allgemein scheinen die negativen Einflüsse der Offenlegung nach der Interaktion zu überwiegen. Bis auf die Erhöhung der Sozialen Präsenz war eine späte Offenlegung mit negativen Auswirkungen verbunden.

Bezüglich der ersten Forschungsfrage konnte keine eindeutige Antwort geliefert werden, da in den Experimenten unterschiedliche Ergebnisse herauskamen. So scheint es, dass der Zeitpunkt der Offenlegung einen positiven Einfluss auf die Soziale Präsenz hat und diese für die Offenlegung nach der Interaktion höher ist. In Bezug auf die Wiederverwendungsabsicht scheint diese geringer zu sein, wenn der Chatbot erst nach der Interaktion offengelegt wird. Allerdings war dieser Einfluss nicht signifikant. Diese Einflüsse konnten somit nicht abschließend gezeigt werden. In Bezug auf die zweite Forschungsfrage konnte kein Einfluss der Kommunikationsplattform gefunden werden. Hier scheinen Nutzer indifferent zu sein, was die entsprechende Wahrnehmung als sozial präsent angeht, wenn der Zeitpunkt der Offenlegung variiert wird. Bezüglich Forschungsfrage drei konnte ein Einfluss des Offenlegungszeitpunkts auf die Fairnesswahrnehmung, sowie das vermutet Motiv gemessen werden. Allerdings war dieser Einfluss unabhängig vom Erfolg oder Misserfolg der Anfrage.

Implikationen für die Praxis

Aufgrund der häufig negativen Auswirkungen einer späten Offenlegung sollten Unternehmen beim Einsatz von Chatbots direkt von Beginn an ehrlich und offen mit den Nutzern umgehen und die Identität des Chatbots vor der Interaktion offenlegen. Hier muss natürlich beachtet werden, dass die Offenlegung allgemein häufig mit negativen Kundenreaktionen einhergeht (Castillo et al., 2021; Ishowo-Oloko et al., 2019; Jin et al., 2024). Nachträgliche Offenlegung kann allerdings auch zu negativeren Reaktionen führen.

Sollte es in einem Kontext von besonderer Bedeutung sein, dass der Chatbot als sozialer Akteur wahrgenommen wird, kann es allerdings auch sinnvoll sein, die Identität zuerst zu verschweigen. Nutzer wissen dann während der Interaktion nicht, dass sie mit einem Chatbot interagieren, was zu einer höheren Wahrnehmung als sozial präsent führen kann aber nicht muss, wie *FP II: E 1* und *FP II: E 2* gezeigt haben.

An dieser Stelle sollte jedoch darauf geachtet werden, ob es rechtliche Vorschriften gibt, die eine Offenlegung verlangen. Wenn die Offenlegung der nicht-menschlichen Identität verpflichtend ist, kann der Offenlegungszeitpunkt eine Möglichkeit sein, negative Reaktionen der Nutzer zu vermeiden oder abzuschwächen. Gerade um Nutzer zu schützen und ehrlich mit ihnen umzugehen sollte die nicht-menschliche Identität immer offen kommuniziert und nicht verheimlicht werden.

Limitationen und Forschungsausblick

Die Experimente von *FP II* weisen einige Limitationen auf. So wurde in den drei Experimenten hypothetische Szenarien genutzt, in die sich die Teilnehmer hineinversetzen sollten, um den Zeitpunkt der Offenlegung mit den entsprechenden Mediatoren und Moderatoren zu untersuchen. Die Szenario-Technik ist eine in der Marketing-Forschung und auch explizit in der Forschung zu Chatbots verbreitete Methode zur Überprüfung von Hypothesen. Es gilt zu erwähnen, dass in den Szenarien auf Realitätsnähe geachtet wurde. Allerdings könnten Feldexperimente dazu beitragen, die Effekte zu bestätigen.

Des Weiteren wurden die Experimente in Deutschland durchgeführt und die Stichproben sind für die Grundgesamtheit der Deutschen zu jung und haben ein sehr hohes Bildungsniveau. Darüber hinaus ist auch die Verteilung der Geschlechter in *FP II: E 1* und *FP II: E 2* unausgeglich. An dieser Stelle könnte bei der Teilnehmerakquise auf das Alter, Geschlecht sowie den Bildungsgrad geachtet werden, um allgemeingültigere Aussagen treffen zu können. Es ist auch interessant, die Ergebnisse in anderen Ländern zu überprüfen, da Deutschland im Hinblick auf die Nutzung von Chatbots hinter Ländern wie der USA oder Indien zurückliegt (Gohil, 2025). In Ländern, in denen Chatbots etablierter sind, könnten unterschiedliche Ergebnisse auftreten.

Teilweise mussten viele Teilnehmer ausgeschlossen werden, da sie die Manipulationscheckfragen falsch beantworteten. Daher musste mit kleinen Gruppen gerechnet werden. Folglich sollte mit Verallgemeinerungen der Ergebnisse vorsichtig umgegangen werden. Es stellt sich die Frage, warum die Teilnehmer die Manipulationscheckfragen falsch beantworteten. Möglicherweise haben sie den Hinweis nicht wahrgenommen, oder es ist ihnen egal, ob sie in einem solchen Experiment mit einem Chatbot „interagieren“. Auch das genutzte Szenario mit dem jeweiligen Kontext könnte eine Limitation darstellen. Je nachdem könnte die Information zur Offenlegung für die Nutzer auch irrelevant sein, weshalb sie sich nicht daran erinnern. Weitere Forschung könnte dazu beitragen, die Relevanz dieser Information zu ermitteln. Hier könnte

nach der Wichtigkeit des Szenarios gefragt werden, beziehungsweise ob sich die Teilnehmer mit dem Kontext des Szenarios identifizieren können. Darüber hinaus könnte in Pre-Tests ein Szenario ausgewählt werden, welches dieses Kriterium erfüllt.

Wie zuvor beschrieben, konnte lediglich in einem von drei Experimenten der direkte Einfluss des Offenlegungszeitpunkts auf die Wiederverwendungsabsicht gezeigt werden. Hier sollten die Ergebnisse in weiterer Forschung erweitert und überprüft werden, ob ein direkter Einfluss des Offenlegungszeitpunkts vorliegt. Auch die verwendete Kommunikationsplattform stellt einen interessanten Ansatzpunkt für weitere Forschung dar. Auch die Kommunikationsplattform könnten in weitere Forschung als direkter Einflussfaktor betrachtet werden. Dort könnte ein stärkerer Fokus auf den Aspekt der Privatsphäre gelegt werden, da Technologien in einem privaten Kontext eingesetzt werden. Es könnten Privatsphäreverletzungen der Nutzer auftreten, wenn sie erst nach der Interaktion über die nicht-menschliche Identität informiert werden. Der Zeitpunkt der Offenlegung stellt ein interessantes Themengebiet dar, dem bisher nur wenig Beachtung geschenkt wurde. Hier könnte weitere Forschung dazu beitragen, den Einfluss des Offenlegungszeitpunkts auf die Wiederverwendungsabsicht sowie die Rolle relevanter Mediatoren und Moderatoren zu klären.

Seit Anfang 2025 wird auf WhatsApp automatisch ein intelligenter Chatbot für alle Nutzer bereitgestellt. Dieser unterstützt diese beispielsweise bei Web-Recherchen oder der Beantwortung von Alltagsanfragen (Klein, 2025). Daher ist davon auszugehen, dass die Präsenz von Chatbots auf WhatsApp zunehmend stärker ins Bewusstsein der Nutzer rückt. In der Herleitung der Hypothesen von *FP II: E 2* wurde darauf hingewiesen, dass WhatsApp überwiegend für die private Kommunikation genutzt wird und Chatbots dort eher unüblich sind. Die neue Funktion könnte jedoch dazu beitragen, diese Annahme zu relativieren. Allerdings ist der neue WhatsApp-Chatbot nicht im Kundenservice von Unternehmen integriert, über den ein direkter Kontakt zu einem Unternehmen möglich wäre, wie bei den in dieser Arbeit untersuchten Chatbots.

5.5 Forschungsprojekt III: Übergabe der Chatbot-Interaktion an einen Menschen

Wie in *FP II: E 3* dargelegt, sind Chatbots mitunter nicht in der Lage, bestimmte Anfragen adäquat zu beantworten, was nicht selten zu einem vorzeitigen Abbruch der Interaktion führt. Eine in der Praxis teilweise angewendete Möglichkeit, eine solche Interaktion zu retten, stellt

die Übergabe an einen Mitarbeiter dar (Meyer et al., 2022, S. 2501). Allerdings kann es sein, dass auch dieser die Anfrage nicht zur Zufriedenheit des Kunden lösen kann. Dadurch können Nutzerreaktion sehr negativ ausfallen und die Nutzer werden den Chatbot voraussichtlich nicht wiederverwenden. Die Nutzer könnten die Übergabe als überflüssig ansehen. Bisher gibt es keine bekannte Forschung, die sich mit dem Thema beschäftigt. Also, welche Auswirkung ein Fehler in Bezug auf die gesamte Anfrage trotz des Wechsels zu einem Mitarbeiter hat.

Es stellt sich die Frage, wer der Überbringer schlechter oder unzufriedenstellender Nachrichten sein soll. Dieser Forschungsfrage wird in diesem Forschungsprojekt nachgegangen. In diesem Fall könnte es sinnvoller sein, wenn der Chatbot die schlechte Nachricht überbringt und nicht der Mensch, da so auch ein möglicherweise unnötiger Prozessschritt, nämlich der Wechsel zu einem Mitarbeiter, vermieden wird. Darüber hinaus sind aktuell die Erwartungen an Chatbots häufig geringer als an Menschen. Ist für Unternehmen zu Beginn der Anfrage schon klar, dass die Nutzeranfrage auch von einem Menschen nicht gelöst werden kann, könnte es sinnvoll sein, auf die Übergabe zu verzichten. Daher wird in diesem Forschungsprojekt der Einfluss der Lösung und die moderierende Rolle des Wechsels zu einem Menschen untersucht.

5.5.1 Herleitung der Forschungshypothesen

Im Falle, dass der Chatbot die Anfrage eines Nutzers nicht lösen kann und dadurch Nutzererwartungen nicht erfüllt werden können, kann dies den Nutzer verärgern. Allerdings haben Nutzer häufig geringe Erwartungen an einen Chatbot und gehen aktuell eher davon aus, dass dieser nicht fehlerfrei funktioniert und nicht alle Anfragen beantworten kann (Sheehan et al., 2020, S. 14). Gerade im Kundenservice wird von Mitarbeitern häufig verlangt, dass diese die Anfrage lösen können. Erwartungen an einen Mitarbeiter sind höher als an einen Chatbot. Daher ist es weniger problematisch, wenn der Chatbot einen Fehler macht oder eine Anfrage nicht zur Zufriedenheit des Nutzers löst, als es bei einem menschlichen Mitarbeiter der Fall wäre. Manche Chatbots haben die Möglichkeit, im Verlauf einer Chatinteraktion diese an einen Mitarbeiter zu übergeben. Die Übergabe an einen Mitarbeiter kann als Wiederherstellungsstrategie gesehen werden, wenn der Chatbot nicht mehr weiterkommt und die Anfrage nicht eigenständig lösen kann (Song et al., 2022, S. 2). Dabei kann die Übergabe an einen Menschen die Verantwortungszuschreibung der Nutzer verändern. So kann eine erfolgreiche Wiederherstellung dazu führen, dass Nutzer den Fehler eher intern bei sich selbst suchen und nicht beim Kundenservice

(Meyer et al., 2022, S. 2499). Auch kann durch eine Übergabe die Zufriedenheit erhöht werden, wenn sich der Mitarbeiter entschuldigt (Choi et al., 2021, S. 365).

Was passiert allerdings, wenn der Mitarbeiter nach der Übergabe die Anfrage trotzdem nicht lösen kann? Wenn also der Chatbot nicht weiterkommt, anschließend an einen Mitarbeiter übergibt, der die Anfrage lösen soll und dieser Mitarbeiter auch keine Lösung für die Anfrage hat. Dies kann sich negativ auf die Nutzer und deren Reaktion auswirken. Denn der Nutzer hat gegebenenfalls eine kurze Wartezeit in Kauf genommen oder der Chatbot hat möglicherweise schon darauf hingewiesen, dass der menschliche Mitarbeiter die Anfrage lösen kann. Dadurch können Erwartungen höher sein. Kommt es dann allerdings auch zu einer negativen Erfahrung mit dem menschlichen Mitarbeiter, können dadurch die negativen Erfahrungen der gesamten Serviceinteraktion erhöht werden und in Folge die Wiederverwendungsabsicht eines Kundenservice beziehungsweise eines Chatbots verringert werden.

Nach der Attributionstheorie können negative oder unerwartete Ergebnisse einer Interaktion das Attributionsdenken fördern und Kunden eher dazu bewegen, kausale Erklärungsansätze zu finden (Kanazawa, 1992, S. 659; Weiner, 1985a, S. 77). Wenn Nutzer somit höhere Erwartungen an einen menschlichen Mitarbeiter haben, kann das Nicht-Erfüllen dieser die Ursachensuche verstärken und gegebenenfalls zu negativen Konsequenzen führen. Diese sind voraussichtlich höher als bei einem Chatbot, da Nutzer dort geringere Erwartungen haben. Folglich könnte es sinnvoller sein, wenn ein Chatbot sowie der Mitarbeiter die Anfrage nicht lösen können, die schlechte Nachricht durch den Chatbot zu überbringen. Daher sollte bei Anfragen, bei denen von Beginn klar ist, dass sie nicht zur Zufriedenheit des Nutzers bearbeitet werden können, auf einen Wechsel verzichtet werden. Denn die Wiederverwendungsabsicht bei einem Fehler kann geringer sein, wenn zuvor ein Wechsel zu einem menschlichen Mitarbeiter stattgefunden hat. Allgemein kann davon ausgegangen werden, dass eine erfolgreiche Bearbeitung einer Nutzeranfrage zu einer größeren Wiederverwendungsabsicht führt, sei es mit oder ohne Wechsel. Aus diesen Überlegungen ergeben sich die folgenden Hypothesen:

H₁: Die Lösung hat einen direkten Einfluss auf die Wiederverwendungsabsicht, sodass Nutzer, deren Problem gelöst wurde (*Lösung_ja*), eine höhere Wiederverwendungsabsicht zeigen als Nutzer, deren Problem nicht gelöst wurde (*Lösung_nein*).

H₂: Die Wirkung der Lösung auf die Wiederverwendungsabsicht wird durch die Wechselmöglichkeit moderiert, sodass die Wiederverwendungsabsicht der Nutzer, die zu einem Menschen wechseln (*Wechsel_ja*) geringer ist, wenn ihr Problem nicht gelöst wird (*Lösung_nein*) als bei Nutzern, die nicht zu einem Menschen wechseln (*Wechsel_nein*) und die gleiche Lösungssituation erfahren.

5.5.2 Methodik

In *FP III* wird ein hypothetisches Kundenservice-Szenario genutzt. Hierbei sollen sich die Teilnehmer in die Situation versetzen, dass sie ihr Smartphone verloren haben und ihre SIM-Karte sperren möchten. Hierfür nutzten sie das Smartphone eines Bekannten. Auf der Webseite des Anbieters gibt es die Möglichkeit eines Online-Kundenservice. Die Teilnehmer sollen sich vorstellen, diesen zu nutzen. Anschließend an die Szenariobeschreibung wird den Teilnehmern die Chatinteraktion mit diesem Kundenservice mithilfe von Screenshots gezeigt. Dabei wurde manipuliert, ob im Verlauf des Chats eine Lösung geliefert werden konnte (*Lösung_ja* und *Lösung_nein*) und ob der Chatbot an einen menschlichen Mitarbeiter übergeben hat (*Wechsel_ja* und *Wechsel_nein*). Die Stimuli der vier Gruppen sind in Tabelle 33 dargestellt. Hierbei ist der gesamte Chatverlauf gezeigt. Im Szenario wurden die Gesprächsabschnitte den Teilnehmern nach und nach gezeigt. Beide Gruppen haben zu Beginn einen Hinweis erhalten, dass die Interaktion mit einem Chatbot stattfinden wird. In der Gruppen *Wechsel_ja* wurde der Hinweis nach der Übergabe herausgenommen.

Tabelle 33: Szenariobeschreibung (FP III)

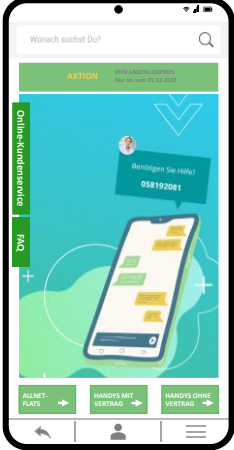
Einführung
Bitte versetzen Sie sich in die Situation, dass Ihnen ihr Handy gestohlen wurde. Da Sie zahlreiche Accounts mit Ihrer Handynummer verbunden haben, bevorzugen Sie es Ihre Nummer zu behalten. Sie möchten Ihre SIM-Karte umgehend sperren. Hierfür nutzen Sie das Handy eines Bekannten und besuchen die Website des Mobilfunkanbieters, um nach einer Kontaktmöglichkeit zu suchen. Auf der folgenden Seite ist die Webseite Ihres Mobilfunkanbieters abgebildet.

Sie sehen am linken Bildschirmrand die Möglichkeit eines Online-Kundenservice. Sie entscheiden sich, diesen Online-Kundenservice zu nutzen. Auf der nächsten Seite wird Ihnen der Gesprächsverlauf mit dem Online-Kundenservice gezeigt.

Tabelle 33: Szenariobeschreibung (FP III) (Fortsetzung)

Gruppe	
Chatinteraktion	
Wechsel_nein / Lösung_ja	Wechsel_ja / Lösung_ja
- Hallo, ich bin Chris, Ihr virtueller Assistent. Ich beantworte Ihnen Fragen zu Ihrem Vertrag. Wie kann ich Ihnen helfen? - Mein Handy wurde gestohlen. - Das tut mir leid! Haben Sie die SIM-Karte gesperrt? - Die Karte ist noch nicht gesperrt. - Um auf Nummer sicher zu gehen, sollten Sie ihre SIM-Karte sperren. Soll ich diese sperren? - Ja, bitte sperren Sie meine SIM-Karte. Handynummer: 0179 59760559 - Ok, ich habe Ihre SIM-Karte erfolgreich gesperrt. - Vielen Dank! - Sollten Sie noch ein weiteres Anliegen haben, können Sie mich gerne jederzeit kontaktieren. Ich wünsche Ihnen einen schönen Tag.	- Hallo, ich bin Chris, Ihr virtueller Assistent. Ich beantworte Ihnen Fragen zu Ihrem Vertrag. Wie kann ich Ihnen helfen? - Mein Handy wurde gestohlen. - Das tut mir leid! Haben Sie die SIM-Karte gesperrt? - Die Karte ist noch nicht gesperrt. - Um auf Nummer sicher zu gehen, sollten Sie ihre SIM-Karte sperren. Soll ich diese sperren? - Ja, bitte sperren Sie meine SIM-Karte! Hierfür muss ich Sie an einen Mitarbeiter weiterleiten. Bitte warten Sie einen Moment. - Hallo, ich bin Melanie. Sie möchten Ihre SIM-Karte sperren? - Ja, bitte sperren Sie meine SIM-Karte. Handynummer: 0179 59760559 - Ok, ich habe Ihre SIM-Karte erfolgreich gesperrt. - Vielen Dank! - Sollten Sie noch ein weiteres Anliegen haben, können Sie mich gerne jederzeit kontaktieren. Ich wünsche Ihnen einen schönen Tag.
	

Tabelle 33: Szenariobeschreibung (FP III) (Fortsetzung)

Wechsel_nein / Lösung_nein	Wechsel_ja / Lösung_nein
- Hallo, ich bin Chris, Ihr virtueller Assistent. Ich beantworte Ihnen Fragen zu Ihrem Vertrag. Wie kann ich Ihnen helfen? - Mein Handy wurde gestohlen. - Das tut mir leid! Haben Sie die SIM-Karte gesperrt? - Die Karte ist noch nicht gesperrt. - Um auf Nummer sicher zu gehen, sollten Sie ihre SIM-Karte sperren. Soll ich diese sperren? - Ja, bitte sperren Sie meine SIM-Karte. Handynummer: 0179 59760559 - Leider kann ich Ihre SIM-Karte nicht sperren. Hierfür müssen Sie sich telefonisch an die Hotline wenden: 058192081. Sollten Sie noch ein weiteres Anliegen haben, können Sie mich gerne jederzeit kontaktieren. Ich wünsche Ihnen einen schönen Tag.	- Hallo, ich bin Chris, Ihr virtueller Assistent. Ich beantworte Ihnen Fragen zu Ihrem Vertrag. Wie kann ich Ihnen helfen? - Mein Handy wurde gestohlen. - Das tut mir leid! Haben Sie die SIM-Karte gesperrt? - Die Karte ist noch nicht gesperrt. - Um auf Nummer sicher zu gehen, sollten Sie ihre SIM-Karte sperren. Soll ich diese sperren? - Ja, bitte sperren Sie meine SIM-Karte! - Hierfür muss ich Sie an einen Mitarbeiter weiterleiten. Bitte warten Sie einen Moment. - Hallo, ich bin Melanie. Sie möchten Ihre SIM-Karte sperren? - Ja, bitte sperren Sie meine SIM-Karte. Handynummer: 0179 59760559 - Leider kann ich Ihre SIM-Karte nicht sperren. Hierfür müssen Sie sich telefonisch an die Hotline wenden: 058192081. Sollten Sie noch ein weiteres Anliegen haben, können Sie mich gerne jederzeit kontaktieren. Ich wünsche Ihnen einen schönen Tag.
	

Im Anschluss an das Szenario wurden die Teilnehmer zur Wiederverwendungsabsicht des Online-Kundenservice (Moriuchi, 2021, S. 27-28) befragt. Tabelle 34 zeigt einen Überblick der verwendeten Skalen mit den übersetzten Indikatoren. Darüber hinaus wurde die Manipulation mithilfe von zwei Manipulationscheckfragen, jeweils mit einer Einfachauswahl überprüft. Abschließend wurden den Teilnehmern Fragen zu soziodemographischen Merkmalen und deren Nutzungshäufigkeit von Chatbots sowie ihrer Expertise zu Chatbots (Ohanian, 1990, S. 50) gestellt. Die Skala zur Expertise ist identisch mit der aus vorherigen Experimenten.

Tabelle 34: Operationalisierung der Variablen und Manipulationscheck (FP III)

Kürzel	Indikator	Skala
Manipulationscheck		
Von wem wurde die SIM-Karte gesperrt?		
MC_11	Die SIM-Karte wurde von einem Chatbot gesperrt.	Einfachauswahl
MC_12	Die SIM-Karte wurde von einem Menschen gesperrt.	
MC_13	Die SIM-Karte wurde nicht gesperrt.	
Mit wem haben Sie geschrieben?		
MC_21	Ich habe nur mit einem Chatbot geschrieben.	Einfachauswahl
MC_22	Ich habe nur mit einem Menschen geschrieben.	
MC_23	Ich habe mit einem Chatbot und einem Menschen geschrieben.	
Wiederverwendung (Moriuchi, 2021, S. 27-28)		
Inwiefern stimmen Sie den folgenden Aussagen zu?		
Reu_1	Wenn ich Zugang zu einem Online-Kundenservice hätte, würde ich ihn immer nutzen.	7-Punkt-Likert-Skala: 1 = Stimme überhaupt nicht zu, 7 = Stimme vollkommen zu
Reu_2	Ich denke, es wird sich für mich lohnen,einen Online-Kundenservice zu verwenden, wenn einer verfügbar ist.	
Reu_3	Ich würde meinen Freunden und meiner Familie die Nutzung eines Online-Kundenservice empfehlen.	
Reu_4	Wenn ich die Möglichkeit hätte mehr über Online-Kundenservice mit Unternehmen zu kommunizieren, würde ich dies tun.	
Reu_5	Ich nutze lieber einen Online-Kundenservice, um mein Anliegen zu klären, als eine andere Person zu fragen.	

5.5.3 Stichprobe und Gütebeurteilung

Nach dem Prüfen der Manipulationscheckfragen wurden 20 Teilnehmer ausgeschlossen und die Stichprobe auf 195 Teilnehmer reduziert. Diese verteilen sich wie in Tabelle 35 dargestellt auf die vier Experimentalgruppen. Die grün markierten Felder sind die, die die jeweilige Manipulationscheckfrage richtig beantwortet haben. Die rechte Spalte „Gesamt nach MC“ gibt die Gruppengröße nach Bereinigung an.

Tabelle 35: Ergebnisse des Manipulationschecks (FP IV: E 2)

Gruppe	MC_11	MC_12	MC_13		MC_21	MC_22	MC_23		Gesamt nach MC
Lösung_ja/Wechsel_nein	53	2	1		52	1	3		52
Lösung_ja/Wechsel_ja	3	48	0		4	1	46		46
Lösung_nein/Wechsel_nein	0	1	51		51	1	0		50
Lösung_nein/Wechsel_ja	2	2	52		6	1	49		47

68,2 % der Teilnehmer sind weiblich, 31,3 % männlich und 0,5 % machten keine Angabe zu ihrem Geschlecht. Die Teilnehmer sind im Mittel 28,7 Jahre alt ($SD = 8,12$). Zwei Teilnehmer beantworteten die Frage nach ihrem Alter nicht. Eine detaillierte Stichprobenbeschreibung wird in Tabelle 36 gezeigt. 93,3 % der Teilnehmer haben in den letzten zwölf Monaten mindestens einmal mit einem Chatbot interagiert. 29,2 % der Teilnehmer geben an, sogar häufiger als sechsmal in diesem Zeitrahmen einen Chatbot genutzt zu haben. Die Verteilung der Nutzungshäufigkeit von Chatbots in den letzten zwölf Monaten der Teilnehmer ist in Anhang 1 zusammengefasst. Darüber hinaus haben sich die Teilnehmer als erfahren im Umgang mit Chatbots eingestuft. Als Durchschnitt geben sie $M = 4,34$ ($SD = 1,42$). Die Werte der Expertise wurden mithilfe der Mean-Funktion berechnet.

Tabelle 36: Soziodemographische Merkmale (FP III)

Merkmal	Ausprägung	Prozentualer Anteil
Geschlecht	Weiblich	68,2 %
	Männlich	31,3 %
	Divers	0,0 %
	Keine Angabe	0,5 %
Alter	M = 28,7 (SD = 8,12), 2 kein Angabe	
Haushaltsgröße	1 Person	21,0 %
	2 Personen	53,3 %
	3-5 Personen	22,6 %
	Mehr als 5 Personen	1,5 %
	Keine Angabe	1,5 %
Höchster Bildungsabschluss	Noch in schulischer Ausbildung	0,5 %
	Volks-/Hauptschulabschluss	1,0 %
	Realschulabschluss	5,1 %
	Fachschulabschluss	4,6 %
	Abitur	19,0 %
	Fachhoch-/Hochschulabschluss	69,7 %
	Keine Angabe	0,0 %
Monatliches Haushaltsnettoeinkommen	Bis 1000 Euro	20,5 %
	1001 - 2000 Euro	25,6 %
	2001 - 3000 Euro	19,0 %
	3001 - 4000 Euro	6,7 %
	4001 - 5000 Euro	10,3 %
	Mehr als 5000 Euro	10,8 %
	Keine Angabe	7,2 %

Die Gütebeurteilung der gemessenen Konstrukte erfolgt auch in diesem Experiment auf Basis der in Kapitel 4.2 beschriebenen Gütekriterien und Grenzwerte. Deren Ergebnis ist in Tabelle 37 zusammengefasst. Hier zeigt sich, dass die Grenzwerte der Gütekriterien für Reliabilität und Validität alle überschritten sind. Daher werden die Skalen für die folgende Auswertung verwendet.

Tabelle 37: Gütekriterien (FP III)

Konstrukt	Indikator	Faktorladung	KITK	Chronbachs Alpha	KMO	Erklärte Varianz
		> 0,4	> 0,5	> 0,7	> 0,6	> 50 %
WoM	Wom_1	0,967	0,939	0,958	0,860	91,153
	Wom_2	0,966	0,938			
	Wom_3	0,952	0,914			
	Wom_4	0,934	0,884			
Wiederverwendung	Reu_1	0,854	0,766	0,893	0,862	70,573
	Reu_2	0,856	0,758			
	Reu_3	0,823	0,716			
	Reu_4	0,873	0,786			
	Reu_5	0,791	0,682			

5.5.4 Ergebnisse

Zuerst wird zur Überprüfung der Hypothesen eine zweifaktorielle ANOVA gerechnet. Hierbei tritt ein Interaktionseffekt auf ($F(1, 191) = 6,86, p < 0,01$), so dass der signifikante Unterschied zwischen den beiden Gruppen *Lösung_nein* und *Lösung_ja* ($F(1, 191) = 20,28, p < 0,001$) in Bezug auf die Wiederverwendungsabsicht nicht direkt interpretiert werden kann. Der zweite direkte Effekt des Wechsels (*Wechsel_nein* und *Wechsel_ja*) ist nicht signifikant ($F(1, 191) = 0,13, p > 0,05$). Abbildung 11 visualisiert die Ergebnisse der zweifaktoriellen ANOVA.

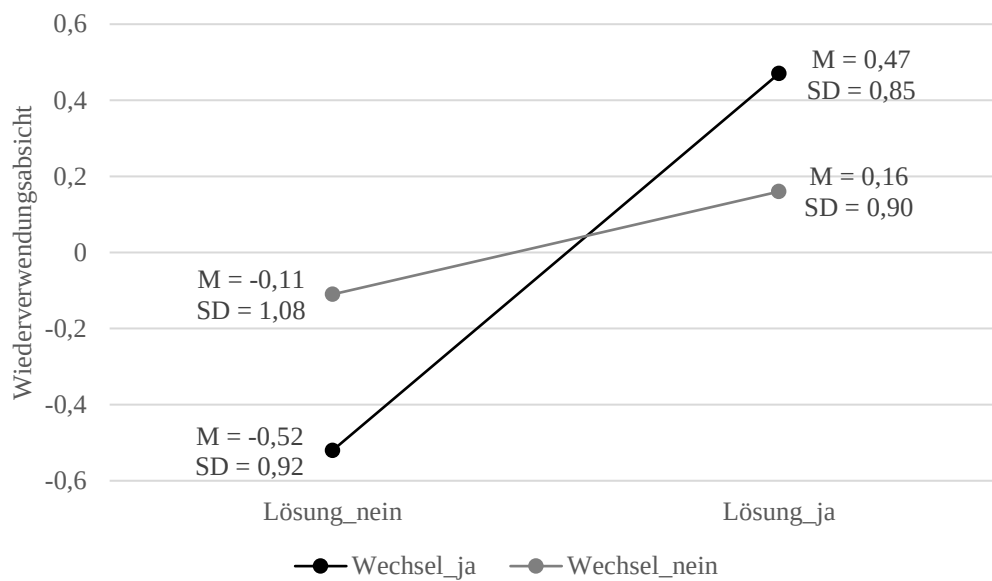


Abbildung 11: Ergebnisse der zweifaktorielle ANOVA (FP III) – Effekte der Lösung und des Wechsels auf die Wiederverwendungsabsicht

Der signifikante Haupteffekt sollte nicht einfach analysiert werden. Allerdings bietet die Simple-Effects-Analyse eine Möglichkeit Haupteffekte bei vorliegender signifikanter Interaktion zu prüfen. Die Simple-Effects-Analyse betrachtet dabei den Effekt der unabhängigen Variablen in Bezug auf die individuellen Ausprägungen der anderen unabhängigen Variablen (Field, 2018, 629, 635). Dabei zeigt sich, dass die Lösung des Problems für das Szenario *Wechsel_ja* einen signifikanten Einfluss hat ($F(1, 191) = 25,25, p < 0,001$). Hierbei werden Nutzer den Chat-Service eher nicht mehr nutzen, wenn das Problem trotz Wechsel nicht gelöst werden kann ($M = -0,52, SD = 0,92$). Wohingegen die Wiederverwendungsabsicht am höchsten ist, wenn der Chat an einen Menschen übergeben und das Problem gelöst wurde ($M = 0,47, SD = 0,85$). Für das Szenario *Wechsel_nein* ist der Einfluss der Lösung auf die Wiederverwendungsabsicht nicht signifikant ($F(1, 191) = 2,11, p > 0,05$). Zusammenfassend lässt sich sagen, dass die Lösung des Problems einen signifikanten Einfluss auf die Wiederverwendungsabsicht eines Online-Kundenservice mit einem Chatbot hat, allerdings nur wenn die Konversation mit dem Chatbot an einen Menschen übergeben wird. Wenn die Konversation nicht übergeben wird, hat die Lösung des Problems keinen signifikanten Einfluss auf die Wiederverwendungsabsicht.

Dieses Ergebnis kann auch in der Moderationsanalyse nach Hayes bestätigt werden. Dabei wurde mit den Daten eine Moderationsanalyse in SPSS mithilfe des PROCESS Model 1 mit

einem Bootstrap Sample von 5.000 gerechnet (Hayes, 2022, S. 625). Als unabhängige Variable wird die Lösung, als abhängige Variable die Wiederverwendungsabsicht und als Moderatorvariable die Wechselmöglichkeit verwendet. Abbildung 12 veranschaulicht das untersuchte Moderationsmodell. Als Referenzkategorie dient die Gruppe *Lösung_nein* und *Wechsel_nein*.

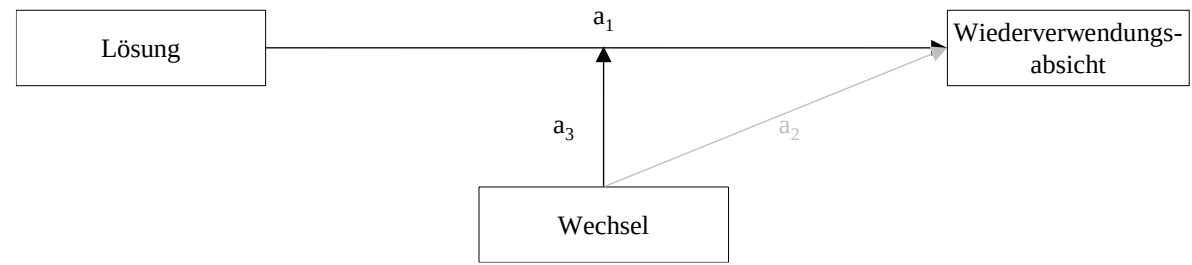


Abbildung 12: Moderationsmodell (FP III)

Im Folgenden werden die wichtigsten Erkenntnisse beschrieben. Die gesamten statistischen Kennzahlen der Moderation sind in Anlehnung an Hayes (2022, S. 463) in Tabelle 38 zusammengefasst.

Tabelle 38: Regressionskoeffizienten der Moderationsanalyse (FP III)

Wiederverwendungsabsicht				
Vorgänger	Symbol	Koeff.	SE	t-Wert
Lösung	a_1	0,27 ^{ns}	0,19	1,45
Wechsel	a_2	-0,40*	0,19	-2,11
Lösung x Wechsel	a_3	0,71**	0,27	2,63
Konstante	i_Y	0,11 ^{ns}	0,13	-0,83
$R^2 = 0,13$; $F(3, 191) = 9,18$; $p < 0,001$				
Konditionale Effekte auf die Wiederverwendungsabsicht				
		Effekt	SE	t-Wert
Wechsel_nein		0,27 ^{ns}	0,19	1,45
Wechsel_ja		0,98***	0,20	5,02
***: $p < 0,001$; **: $p < 0,01$; *: $p < 0,05$; ns: $p > 0,05$				

Der allgemeine Einfluss der Lösung auf die Wiederverwendungsabsicht kann nicht bestätigt werden ($a_1 = 0,27$, $p > 0,05$). Folglich wird Hypothese H_1 von den Ergebnissen dieser Analyse nicht unterstützt. Aufgrund des signifikanten Interaktionsterms ($a_3 = 0,71$, $p < 0,01$) kann eine Moderation hier beschrieben werden. Die Analyse der bedingten Effekte ergab, dass der positive Effekt einer gelösten Anfrage auf die Wiederverwendungsabsicht nur dann signifikant ist, wenn ein Wechsel zu einem menschlichen Mitarbeiter stattgefunden hat (Effekt = 0,98,

SE = 0,20, $p < 0,001$). Folglich ist die Wiederverwendungsabsicht geringer, wenn ein Wechsel stattfindet und keine Lösung des Problems gefunden wird. Ohne einen solchen Wechsel zeigte sich kein signifikanter Effekt (Effekt = 0,27, SE = 0,19, $p > 0,05$). Die Ergebnisse der Simple-Effects-Analyse sowie der Moderationsanalyse nach Hayes liefern teilweise Bestätigung für Hypothese H₂.

Tabelle 39: Zusammenfassung der Hypothesen (FP III)

Hypothese H _i		Ergebnis
H ₁	Die Lösung hat einen direkten Einfluss auf die Wiederverwendungsabsicht, sodass Nutzer, deren Problem gelöst wurde (Lösung_ja), eine höhere Wiederverwendungsabsicht zeigen als Nutzer, deren Problem nicht gelöst wurde (Lösung_nein).	✗
H ₂	Die Wirkung der Lösung auf die Wiederverwendungsabsicht wird durch die Wechselmöglichkeit moderiert, sodass die Wiederverwendungsabsicht der Nutzer, die zu einem Menschen wechseln (Wechsel_ja) geringer ist, wenn ihr Problem nicht gelöst wird (Lösung_nein) als bei Nutzern, die nicht zu einem Menschen wechseln (Wechsel_nein) und die gleiche Lösungssituation erfahren.	(✓)
✓ = bestätigt; ✗ = nicht bestätigt; (✓) = teilweise bestätigt		

5.5.5 Diskussion der Ergebnisse

Im Rahmen von FP III konnte kein allgemeiner Einfluss der Lösung auf die Wiederverwendungsabsicht gefunden werden. Die Ergebnisse der Moderationsanalyse deuten darauf hin, dass eine Lösung durch einen Chatbot allein nicht ausreicht, um die Wiederverwendungsabsicht zu steigern. Erst in Kombination mit einem Wechsel zu einem menschlichen Ansprechpartner wird der Effekt der Problemlösung sichtbar. Der Effekt der Lösung der Anfrage auf die Wiederverwendungsabsicht ist stärker für die Situation eines Wechsels zu einem menschlichen Mitarbeiter. Allerdings ist der Effekt für die Gruppe *Wechsel_nein* nicht signifikant. Der stärkere Effekt für die Gruppe *Wechsel_ja* hingegen ist signifikant.

Wenn folglich ein Wechsel zu einem Menschen stattfindet, sollte sichergestellt werden, dass die Anfrage auch gelöst werden kann. Denn ansonsten könnten die Auswirkungen auf den Nutzer negativer sein, als wenn der Chatbot die Anfrage nicht lösen kann und die Anfrage danach beendet wird. Ist gegebenenfalls im Vorfeld schon klar, dass die Nutzeranfrage auch von einem Menschen nicht gelöst werden kann, kann es sinnvoll sein, auf die Übergabe zu verzichten. Denn ohne einen Wechsel konnte kein signifikanter Einfluss der Lösung auf die Wiederverwendungsabsicht gefunden werden.

Bezüglich der Forschungsfrage kann gesagt werden, dass eher der Chatbot der Überbringer einer schlechten Nachricht sein sollte.

Implikationen für die Praxis

Die Wechselmöglichkeit sollte von Unternehmen bewusst gestaltet und genutzt werden. Unternehmen sollten den Wechsel zu einem Menschen als mögliche Wiederherstellungsstrategie nicht automatisch als positiv ansehen. Wenn auch nach dem Wechsel keine Lösung angeboten werden kann, kann ein Wechsel vielleicht enttäuschend wirken und die Wiederverwendungsabsicht sogar verringern. Die menschliche Unterstützung sollte gezielt eingesetzt werden. Besonders bei lösbaren Anliegen sollte die Möglichkeit bestehen, dass ein Mensch übernimmt. Das kann die Wiederverwendungsabsicht steigern. Unternehmen könnten intelligente Systeme entwickeln, die nur dann automatisch einen Wechsel veranlassen, wenn Aussicht auf Lösung der Nutzeranfrage besteht.

Limitationen und Forschungsausblick

FP III hat einige Limitationen. Das Experiment basiert auf einem fiktiven Kundenservice-Szenario, in dem sich die Teilnehmer in die Situation versetzen sollten, den Verlust ihrer SIM-Karte ihrem Mobilfunkanbieter zu melden. Des Weiteren wurden die Chatverläufe in Form von Screenshots gezeigt. Interaktive Dynamiken, wie sie in echten Kundengesprächen entstehen, blieben unberücksichtigt. Durch dieses hypothetische Szenario kann die externe Validität eingeschränkt sein. In realen Situationen sind Emotionen und Konsequenzen oft stärker ausgeprägt. An dieser Stelle könnte ein Feldexperiment genutzt werden.

Im Rahmen des Forschungsprojekts wurde ein Fokus auf die Wiederverwendungsabsicht gelegt. Weitere relevante Variablen wie Vertrauen, Zufriedenheit oder Fairnesswahrnehmung wurden nicht erhoben, könnten aber relevante Mediatoren beziehungsweise alternative Erklärungen darstellen.

Zudem wurden die Experimente in Deutschland durchgeführt, wobei die Stichprobe im Vergleich zur Gesamtbevölkerung der Deutschen zu jung ist, ein sehr hohes Bildungsniveau aufweist und etwa ein Drittel Frauen umfasst. An dieser Stelle könnte bei der Teilnehmerakquise auf das Alter, Geschlecht sowie den Bildungsgrad geachtet werden, um allgemeingültigere Aussagen treffen zu können.

Im Rahmen der Herleitung der Hypothesen wurde die Erwartungshaltung angesprochen. Diese wurde im Rahmen des Experiments allerdings nicht abgefragt. Darin kann eine mögliche Limitation gesehen werden. In Folgeexperimenten sollte die Erwartungshaltung aufgenommen und als möglicher Mediator betrachtet werden, ähnlich wie in *FP IV*, in dem das Confirmation/Disconfirmation-Paradigma genutzt wird. An dieser Stelle könnte auch die Kommunikation der Fähigkeiten des Chatbots oder des Menschen als Moderator betrachtet werden. Eine klare Kommunikation der Möglichkeiten des Chatbots oder des Menschen könnte helfen, Erwartungen anzupassen und möglicherweise negative Reaktionen zu vermeiden, wenn keine Lösung geliefert werden kann. Im Rahmen der Herleitung wurde die Attributionstheorie genannt. Hier könnte durch weitere Forschung Mediatoren untersucht werden, die im Falle des Wechsels die Beziehung zwischen Lösung sowie der Wiederverwendungsabsicht medieren. Also welche kausalen Erklärungsansätze nutzen die Nutzer, wenn trotz Wechsel keine Lösung gefunden wird.

Ein weiterer interessanter Forschungsausblick stellt die Tatsache dar, dass eine Übergabe auch an einen anderen Chatbot stattfinden kann (Koch et al., 2021, S. 1274). Hier könnten sich unterschiedliche Ergebnisse zwischen Menschen und Chatbots als Übergabepartner ergeben. Es stellt sich die Frage, ob die Lösung der Anfrage im Rahmen einer Übergabe an einen Chatbot einen Einfluss auf die Wiederverwendungsabsicht hat.

5.6 Forschungsprojekt IV: Bewertungen von Chatbots

In der Praxis werden auch bei Chatbots Nutzerbewertungen verwendet. Durch solche Nutzerbewertungen kann beispielsweise das Vertrauen erhöht (Gavilan et al., 2018, S. 58) oder Risiken beim Onlinekauf verringert werden (Bae & Lee, 2011, S. 262; Hong et al., 2017, S. 1). Gerade bei Produkten, die nicht greifbar sind oder bei Dienstleistungen können Erfahrungen Anderer und daraus resultierende Bewertungen ein guter Anhaltspunkt für die eigene Entscheidung sein. Im Online-Kontext können Nutzer Produkte nicht anfassen und sind auf Bilder oder Angaben der Händler angewiesen. Oder sie stehen vor einer großen Anzahl unterschiedlicher Anbieter oder ähnlicher Produkte. Daraus das richtige Produkt oder den richtigen Anbieter auszuwählen, kann schwierig sein. In solchen Situationen können Bewertungen helfen, Unsicherheiten zu verringern und Entscheidungen zu erleichtern (Bae & Lee, 2011, S. 262; Hong et al., 2017, S. 1).

Bei Technologien wie Chatbots können Bewertungen auch als Anker dienen und die Erwartungen oder die Nutzungsabsicht an den Chatbot beeinflussen. Im Rahmen der Literaturrecherche dieser Arbeit wurden zum Thema Bewertungen im Kontext von Chatbots keine Erkenntnisse gefunden. Daher werden in diesem Forschungsprojekt grundlegende Effekte von Bewertungen bei Chatbots untersucht. *FP IV* beschäftigt sich dabei mit den folgenden Forschungsfragen:

1. Welchen Einfluss haben Bewertungen des Chatbots auf die Wiederverwendungsabsicht?
2. Werden durch Bewertungen des Chatbots auch die Erwartungen an den Chatbot beeinflusst?
3. Welchen Einfluss haben Erwartungen an den Chatbot auf die Nutzungsabsicht sowie die Zufriedenheit mit dem Service?

Diese Fragen werden in *FP IV* in zwei Experimenten untersucht. Dabei wird in Experiment eins nur ein Rating verwendet und die Auswirkungen auf die Erwartungen, die Nutzungsabsicht, die Zufriedenheit und die Wiederverwendung untersucht. In Experiment zwei wird zusätzlich zum Rating noch ein Review genutzt, um den Stimulus der Bewertung zu erhöhen und sicherzustellen, dass die positive oder negative Bewertung auf jeden Fall von den Teilnehmern wahrgenommen wird. In diesem Forschungsprojekt werden die Hypothesen für beide Experimente zusammen hergeleitet. Hier werden in beiden Experimenten die gleichen Hypothesen untersucht. Daher erfolgt die abschließende Diskussion der Ergebnisse gemeinsam.

5.6.1 Herleitung der Forschungshypothesen

Bewertungen stellen eine Möglichkeit dar, Erwartungen an ein Produkt oder eine Dienstleistung zu beeinflussen. Bewertungen können dabei als Anker oder Startpunkte für weitere Überlegungen und Entscheidungen dienen (Lee & Kim, 2020, S. 593–594). Nutzer beurteilen Situationen auf Basis eines solchen Ankers. Die Beurteilung der Situation ist dabei in Richtung des Startwerts verzerrt (Tversky & Kahneman, 1974, S. 1128). Dabei können positive Bewertungen zu höheren Erwartungen führen als negative Bewertungen (Nath et al., 2018, S. 130). Auch für Chatbots können Erwartungen an den Chatbot durch Bewertungen in Form von Ratings oder Reviews beeinflusst werden und als Anker dienen. Daher ist anzunehmen, dass die Erwartungen an den Chatbot bei guten Bewertungen höher sind als bei schlechten. Aus dieser Überlegung ergibt sich die erste Hypothese:

H₁: Nutzer haben an einen Chatbot mit einer guten Bewertung (Rating beziehungsweise Rating/Review) höhere Erwartungen, dass der Chatbot die Anfrage lösen kann, als an einen Chatbot mit einer schlechten Bewertung.

Hohe Erwartungen an den Chatbot können zu einer größeren Nutzungsabsicht führen. Da davon ausgegangen werden kann, dass der Chatbot die Anfrage auch lösen kann und die Nutzung des Chatbots keine Zeitverschwendung darstellt. Daraus ergibt sich die zweite Hypothese:

H₂: Die Nutzungsabsicht ist bei hohen Erwartungen an einen Chatbot größer als bei geringen Erwartungen.

Bewertungen können als Quelle von Zufriedenheit oder Unzufriedenheit gesehen werden (Xu, 2019, S. 143). Nach dem Confirmation/Disconfirmation-Paradigma hängt Zufriedenheit von der Abweichung zwischen den Erwartungen und der tatsächlichen Erfahrung ab. Zufriedenheit ergibt sich aus dem Vergleich der tatsächlichen Erfahrung bei Nutzung einer Leistung (Ist-Leistung) mit einem bestimmten Vergleichsstandard (Soll-Leistung). Dieser Vergleichsstandard kann beispielsweise durch vorherige Erfahrungen oder Nutzererfahrungen Anderer und deren Bewertungen beeinflusst werden (Homburg, 2020, S. 46). Werden Erwartungen enttäuscht, sind die Kunden unzufrieden. Dabei ist dieser negative Effekt auf die Zufriedenheit um ein Vielfaches höher als der positive Effekt, der beim Erfüllen der Erwartung auftritt. Folglich sind Nutzer von Chatbots voraussichtlich zufriedener mit der erhaltenen Leistung, wenn diese besser ausfällt als zuvor erwartet. Da negative Bewertungen die Erwartungen an den Chatbot voraussichtlich verringern, dürften Nutzer, die vor der Interaktion eine negative Bewertung gesehen haben, mit der Erfüllung ihrer Anfrage zufriedener sein als Nutzern, die eine positive Bewertung im Vorfeld der Interaktion gesehen haben. Bei Erfüllung der Aufgabe kann es zu einer positiven Diskonfirmation kommen. Dabei kann die Zufriedenheit der Nutzer höher sein als bei reiner Konfirmation. Konfirmation kann auftreten, wenn bei höheren Erwartungen die Anfrage ebenfalls erfüllt wird. Hieraus ergibt sich die dritte Hypothese:

H₃: Die Zufriedenheit ist höher, wenn die Nutzer niedrigere Erwartungen an einen Chatbot haben und dieser trotzdem die Anfrage lösen kann als bei hohen Erwartungen.

Zufriedenheit kann positive Kundenreaktionen hervorrufen. So sind zufriedene Kunden in der Regel dem Unternehmen gegenüber loyaler. Loyalität kann sich beispielsweise in Form von Wiederkauf oder Wiederverwendung äußern (Homburg, 2020, S. 47; Szymanski & Henard, 2001, S. 19). Folglich können zufriedene Nutzer eher gewillt sein den Chatbot nochmal zu

verwenden als unzufriedene. Nutzer, die ursprünglich eine hohe Nutzungsabsicht haben, werden den Chatbot auch in Zukunft eher nutzen wollen. Gerade wenn die Chatbot Erfahrung positiv war und die Anfrage gelöst werden konnte. Daraus folgen die vierte und fünfte Hypothese:

H₄: Zufriedene Nutzer sind eher gewillt, den Chatbot wiederzuverwenden als unzufriedene Nutzer.

H₅: Haben Nutzer eine höherer Nutzungsabsicht ist deren Wiederverwendungsabsicht höher als bei Nutzern, die eine geringe Nutzungsabsicht aufweisen.

5.6.2 Experiment 1: Rating

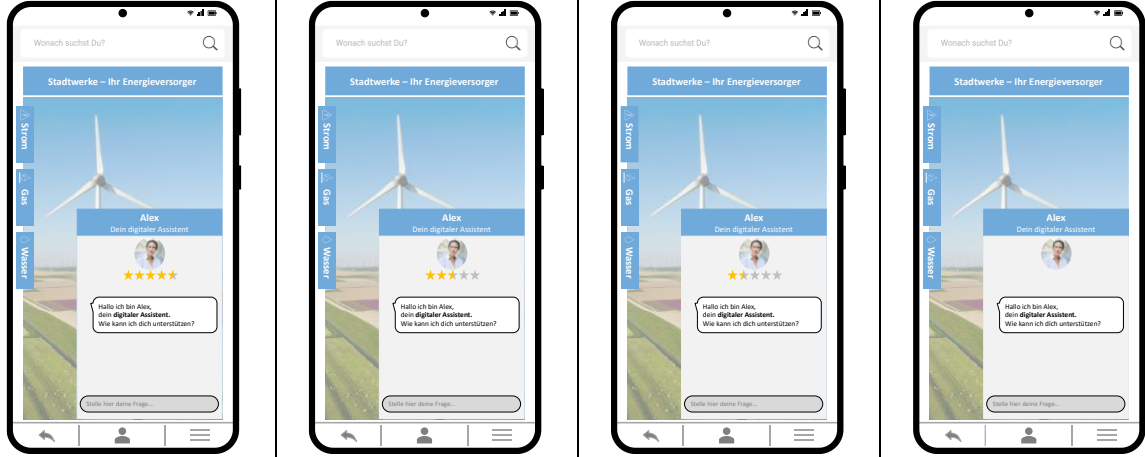
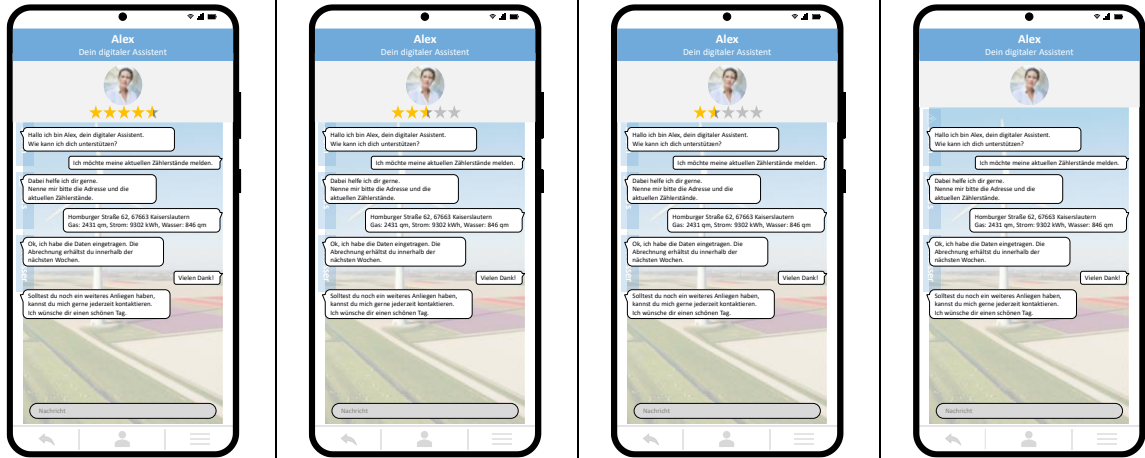
Das erste Experiment des Forschungsprojekts IV (*FP IV: E 1*) untersucht die Auswirkung eines Ratings eines anthropomorph gestalteten Chatbots auf die Wiederverwendungsabsicht. Die Erwartungen an den Chatbot, die Zufriedenheit sowie die Nutzungsabsicht werden in diesem Experiment als Mediatoren betrachtet. In diesem Experiment wird zwischen einem guten, einem mittleren und einem schlechten Rating unterschieden. Des Weiteren wird eine Referenzgruppe verwendet, der kein Rating angezeigt wird.

5.6.2.1 Methodik

Auch in *FP IV: E 1* wurden die Hypothesen mithilfe eines szenariobasierten online Experiments überprüft. Dabei sollen sich die Teilnehmer in die Situation versetzen, ihre Zählerstände dem Energieversorger melden zu wollen. Hierfür steht auf der Webseite des Energieversorgers ein Online-Kundenservice in Form eines Chatbots zur Verfügung. Auch in diesem Experiment wurden den Teilnehmern nach der Szenariobeschreibung nach und nach Screenshots der Chatinteraktion gezeigt. Es wurde ein einfaktorielles Design mit Zufallsgruppenbildung verwendet, bei dem das Rating des Chatbots manipuliert wurde. Der Zwischensubjektfaktor Rating umfasst vier Ausprägungen. Die erste Experimentalgruppe sah ein gutes Rating (*Rating_gut*), die Zweite ein mittleres Rating (*Rating_mittel*), die Dritte ein schlechtes Rating (*Rating_schlecht*) und die vierte Gruppe dient als Referenzkategorie und sah kein Rating (*Rating_nein*). Dabei wurde jeweils eine Sternebewertung gezeigt (4,5 Sterne, 2,5 Sterne, 1,5 Sterne oder kein Rating). Auch hier wurde durch das Design eine realitätsnahe Situation dargestellt. Der Chatbot wurde ähnlich dargestellt wie in den vorherigen Experimenten. Zuerst sahen die Teilnehmer die erste Seite des Chatbots. Auf dieser Seite war die Internetseite sowie der Chatbot mit dem

Rating und einer ersten Nachricht des Chatbots an den Nutzer abgebildet. Der Chatbot fragt den Nutzer, wie er ihn unterstützen kann. Anschließend wurden die Teilnehmer zu ihrer Nutzungsabsicht (Venkatesh et al., 2003, S. 460) sowie der Erwartung an den Chatbot befragt. Darüber hinaus wurde an dieser Stelle schon die Manipulation mithilfe einer Einfachauswahl aus vier möglichen Antworten überprüft. Im Anschluss wurde den Teilnehmern nach und nach die Chatinteraktion gezeigt. Am Ende der Chatinteraktion konnten die Zählerstände erfolgreich gemeldet werden. Diese Chatinteraktion war für alle vier Gruppen die Gleiche. In Tabelle 40 ist das Szenario übersichtlich dargestellt.

Tabelle 40: Szenariobeschreibung (FP IV: E 1)

Einführung			
Bitte versetzen Sie sich in die Situation, dass Sie Ihre aktuellen Zählerstände (Gas, Strom und Wasser) Ihrem Energieversorger (Stadtwerke) melden möchten. Auf der Webseite der Stadtwerke haben Sie die Möglichkeit, mit einem Online-Kundenservice in Form eines digitalen Assistenten zu schreiben. Auf der folgenden Seite ist die Webseite der Stadtwerke abgebildet.			
Gruppe			
Rating_gut	Rating_mittel	Rating_schlecht	Rating_nein
			
Sie entscheiden sich, diesen digitalen Assistenten zu nutzen. Auf der nächsten Seite wird Ihnen der Gesprächsverlauf mit dem digitalen Assistenten gezeigt.			
Chatinteraktion			
- Hallo ich bin Alex, dein digitaler Assistent. Wie kann ich dich unterstützen? - Ich möchte meine aktuellen Zählerstände melden. - Dabei helfe ich dir gerne. Nenne mir bitte die Adresse und die aktuellen Zählerstände. - Homburger Straße 62, 67663 Kaiserslautern, Gas: 2431 qm, Strom: 9302 kWh, Wasser: 846 qm - Ok, ich habe die Daten eingetragen. Die Abrechnung erhältst du innerhalb der nächsten Wochen. - Vielen Dank! - Solltest du noch ein weiteres Anliegen haben, kannst du mich gerne jederzeit kontaktieren. Ich wünsche dir einen schönen Tag.			
			

Nachdem die Teilnehmer das Szenario gesehen haben, wurden sie nach ihrer Zufriedenheit mit dem Service (Voss et al., 1998, S. 59), der Wiederverwendungsabsicht (Moriuchi, 2021, S. 27-28), ihrer Expertise in Bezug auf Chatbots (Ohanian, 1990, S. 50) und ihrer Einstellung zu Bewertungen (Khare et al., 2011, S. 115) befragt. Hierbei wurde auf Skalen aus der bisherigen Forschung zurückgegriffen, die auf Basis des in Kapitel 5.2 vorgestellten Prozess ins Deutsche übersetzt wurden. In Tabelle 41 sind die verwendeten Konstrukte mit ihren Skalen und den übersetzten Indikatoren übersichtlich dargestellt. Abschließend wurde neben den Soziodemografika die Nutzungshäufigkeit von Chatbots in den letzten zwölf Monaten abgefragt, um die Stichprobe beschreiben zu können.

Tabelle 41: Operationalisierung der Variablen und Manipulationscheck (FP IV: E 1)

Kürzel	Indikator	Skala
Manipulationscheck: Wie war der digitale Assistent bewertet (Sternebewertung)?		
MC_1	Der digitale Assistent war mit mehr als 4 Sternen bewertet.	Einfachauswahl
MC_2	Der digitale Assistent war mit 2,5 Sternen bewertet.	
MC_3	Der digitale Assistent war mit weniger als 2 Sternen bewertet.	
MC_4	Der digitale Assistent hatte keine Bewertung.	
Erwartung an den Chatbot		
Inwiefern stimmen Sie den folgenden Aussagen zu?		
ReExp	Ich erwarte, dass der digitale Assistent alles in seiner Macht stehende tut, um mir zu helfen.	7-Punkt-Likert-Skala: 1 = Stimme überhaupt nicht zu, 7 = Stimme vollkommen zu
Nutzungsabsicht (Venkatesh et al., 2003, S. 460)		
Inwiefern stimmen Sie den folgenden Aussagen zu?		
IntUse_1	Ich beabsichtige den digitalen Assistenten zu nutzen.	7-Punkt-Likert-Skala: 1 = Stimme überhaupt nicht zu, 7 = Stimme vollkommen zu
IntUse_2	Ich schätze, dass ich den digitalen Assistenten nutzen werde.	
IntUse_3	Ich plane den digitalen Assistenten zu nutzen.	
Zufriedenheit (Voss et al., 1998, S. 59)		
Inwiefern stimmen Sie den folgenden Aussagen zu?		
Sat_1	Ich war mit dem Service zufrieden.	7-Punkt-Likert-Skala: 1 = Stimme überhaupt nicht zu, 7 = Stimme vollkommen zu
Sat_2	Ich war über den Service erfreut.	
Sat_3 ^R	Ich war mit dem Service unzufrieden	

Tabelle 41: Operationalisierung der Variablen und Manipulationscheck (FP IV: E 1) (Fortsetzung)

Kürzel	Indikator	Skala
Wiederverwendung (Moriuchi, 2021, S. 27-28)		
Inwiefern stimmen Sie den folgenden Aussagen zu?		
Reu_1	Wenn ich Zugang zu einem digitalen Assistenten hätte, würde ich ihn immer nutzen.	7-Punkt-Likert-Skala: 1 = Stimme überhaupt nicht zu, 7 = Stimme vollkommen zu
Reu_2	Ich denke, es wird sich für mich lohnen, einen digitalen Assistenten zu verwenden, wenn einer verfügbar ist.	
Reu_3	Ich würde meinen Freunden und meiner Familie die Nutzung eines digitalen Assistenten empfehlen.	
Reu_4	Wenn ich die Möglichkeit hätte mehr über digitale Assistenten mit Unternehmen zu kommunizieren, würde ich dies tun.	
Reu_5	Ich nutze lieber einen digitalen Assistenten, um meine Anliegen zu klären, als eine andere Person zu fragen.	
Einstellung zu Bewertungen (Khare et al., 2011, S. 116)		
Inwiefern stimmen Sie den folgenden Aussagen zu?		
AttRa_1	Ich bin mit dem Lesen von Online-Bewertungen vertraut.	7-Punkt-Likert-Skala: 1 = Stimme überhaupt nicht zu, 7 = Stimme vollkommen zu
AttRa_2	Ich habe Online-Bewertungen genutzt, um eine Entscheidung über ein Produkt oder eine Dienstleistung zu treffen.	
AttRa_3	In der Vergangenheit wurden meine Entscheidungen von Bewertungen beeinflusst, die ich im Internet gesehen habe.	
AttRa_4	Ich tausche mich gerne mit anderen über meine Erfahrungen mit Produkten und Dienstleistungen aus.	
AttRa_5	Ich möchte mehr über die Erfahrungen anderer mit Produkten und Dienstleistungen erfahren.	
AttRa_6	Insgesamt helfen Online-Bewertungen den Verbrauchern, bessere Entscheidungen zu treffen.	
R = Invers formulierter Indikator		

5.6.2.2 Stichprobe und Gütebeurteilung

Nach Prüfung der Manipulationscheckfragen wurde die Stichprobe auf 117 Teilnehmer reduziert. 82 Teilnehmer mussten aus der Stichprobe ausgeschlossen werden, da diese die Manipulationscheckfrage falsch beantworteten (Tabelle 42). Die Stichprobe wurde somit auf die Teilnehmer der grün markierten Felder reduziert.

Tabelle 42: Ergebnisse des Manipulationschecks (FP IV: E 1)

Gruppe	MC_1	MC_2	MC_3	MC_4	Summe
Rating_gut	24	2	0	25	51
Rating_mittel	8	24	2	15	49
Rating_schlecht	6	13	25	5	49
Rating_nein	5	1	0	44	50

64,0 % der Teilnehmer sind weiblich, 32,5 % männlich, 2,6 % machen keine Angabe zu ihrem Geschlecht und 0,9 % der Teilnehmer geben Divers an. Das Durchschnittsalter der Teilnehmer liegt bei 30,07 Jahre ($SD = 11,535$). Tabelle 43 beschreibt die Stichprobe detailliert. 88,9 % der Teilnehmer haben in den letzten zwölf Monaten mindestens einmal mit einem Chatbot interagiert. 18,8 % der Teilnehmer geben an, sogar häufiger als sechsmal in diesem Zeitrahmen einen Chatbot genutzt zu haben. Die Verteilung der Nutzungshäufigkeit von Chatbots in den letzten zwölf Monaten der Teilnehmer ist in Anhang 1 zusammengefasst. Darüber hinaus haben sich die Teilnehmer als erfahren um Umgang mit Chatbots eingestuft. Als Durchschnitt geben sie $M = 4,61$ ($SD = 1,41$) auf einer Skala von 1 bis 7 an. Die Einstellung der Teilnehmer zu Bewertungen ist positiv ($M = 5,65$, $SD = 1,00$). Sowohl die Werte für die Erfahrung mit Chatbots als auch für die Einstellung zu Bewertungen wurde mithilfe der Mean-Funktion berechnet.

Tabelle 43: Soziodemographische Merkmale (FP IV: E 1)

Merkmal	Ausprägung	Prozentualer Anteil
Geschlecht	Weiblich	64,0 %
	Männlich	32,5 %
	Divers	0,9 %
	Keine Angabe	2,6 %
Alter	M = 30,07 (SD = 11,535), 2 kein Angabe	
Haushaltsgröße	1 Person	21,4 %
	2 Personen	35,0 %
	3-5 Personen	36,8 %
	Mehr als 5 Personen	3,4 %
	Keine Angabe	3,4 %
Höchster Bildungsabschluss	Noch in schulischer Ausbildung	0,0 %
	Volks-/Hauptschulabschluss	0,0 %
	Realschulabschluss	1,7 %
	Fachschulabschluss	6,0 %
	Abitur	27,4 %
	Fachhoch-/Hochschulabschluss	64,1 %
	Keine Angabe	0,9 %
Monatliches Haushaltsnettoeinkommen	Bis 1000 Euro	12,8 %
	1001 - 2000 Euro	29,9 %
	2001 - 3000 Euro	10,3 %
	3001 - 4000 Euro	12,0 %
	4001 - 5000 Euro	9,4 %
	Mehr als 5000 Euro	11,1 %
	Keine Angabe	14,5 %

Die Gütebeurteilung der gemessenen Konstrukte erfolgt auch für *FP IV: E1* auf Basis der in Kapitel 4.2 beschriebene Gütekriterien und Grenzwerte. Der Indikator AttRa_4 der Einstellung zur Bewertung wurde aus der Analyse ausgeschlossen, da dieser den Grenzwert der KITK unterschritten hat (AttRa_4 = 0,39). In Tabelle 37 ist die Gütebeurteilung nach Bereinigung der Konstrukte zusammengefasst. Hier zeigt sich, dass die Grenzwerte der Gütekriterien für Reliabilität und Validität alle erreicht sind. Daher werden die Skalen für die folgende Auswertung verwendet.

Tabelle 44: Gütekriterien (FP IV: E 1)

Konstrukt	Indikator	Faktorladung	KITK	Chronbachs Alpha	KMO	Erklärte Varianz
		> 0,4	> 0,5	> 0,7	> 0,6	> 50 %
Nutzungsabsicht	IntUse_1	0,950	0,885	0,934	0,748	88,301
	IntUse_2	0,915	0,816			
	IntUse_3	0,953	0,891			
Zufriedenheit	Sat_1	0,923	0,803	0,826	0,654	74,999
	Sat_2	0,859	0,672			
	Sat_3 ^R	0,812	0,600			
Wiederverwendung	Reu_1	0,883	0,804	0,907	0,886	73,671
	Reu_2	0,843	0,747			
	Reu_3	0,892	0,819			
	Reu_4	0,896	0,825			
	Reu_5	0,770	0,661			
Erfahrung	Exp_1	0,848	0,762	0,915	0,835	74,624
	Exp_2	0,862	0,777			
	Exp_3	0,836	0,741			
	Exp_4	0,851	0,762			
	Exp_5	0,920	0,864			
Einstellung zu Bewertungen	AttRa_1	0,755	0,610	0,848	0,849	62,855
	AttRa_2	0,855	0,743			
	AttRa_3	0,824	0,702			
	AttRa_5	0,739	0,600			
	AttRa_6	0,785	0,652			

5.6.2.3 Ergebnisse

Zur Überprüfung der Auswirkung eines Ratings auf die Nutzungsabsicht des Chatbots wurde in einem ersten Schritt eine einfaktorielle ANOVA mit Post-hoc-Test nach Bonferroni ausgeführt. An dieser Stelle soll auch der Unterschied zur Kontrollgruppe, die kein Rating hat, überprüft werden. Der Zwischensubjektfaktor Rating wird in der ANOVA als unabhängige Variable und die Nutzungsabsicht als abhängige Variable genutzt. Die Gruppe *Rating_gut* hat die höchste Nutzungsabsicht des Chatbots ($M = 0,63$, $SD = 0,93$) gefolgt von der Gruppe *Rating_mittel* ($M = -0,16$, $SD = 0,82$). Eine ähnliche Nutzungsabsicht weist die Kontrollgruppe auf ($M = 0,03$, $SD = 0,95$). Die Gruppe mit dem schlechtesten Rating (*Rating_schlecht*) weist die geringste Nutzungsabsicht auf ($M = -0,50$, $SD = 1,03$). Der Post-hoc-Test nach Bonferroni

zeigt, dass sich die Gruppen *Rating_gut* von *Rating_schlecht* und *Rating_mittel* signifikant unterscheiden ($p < 0,05$). Die Ergebnisse der ANOVA sind in Abbildung 13 grafisch dargestellt. Dort sind auch die Signifikanzniveaus angegeben.

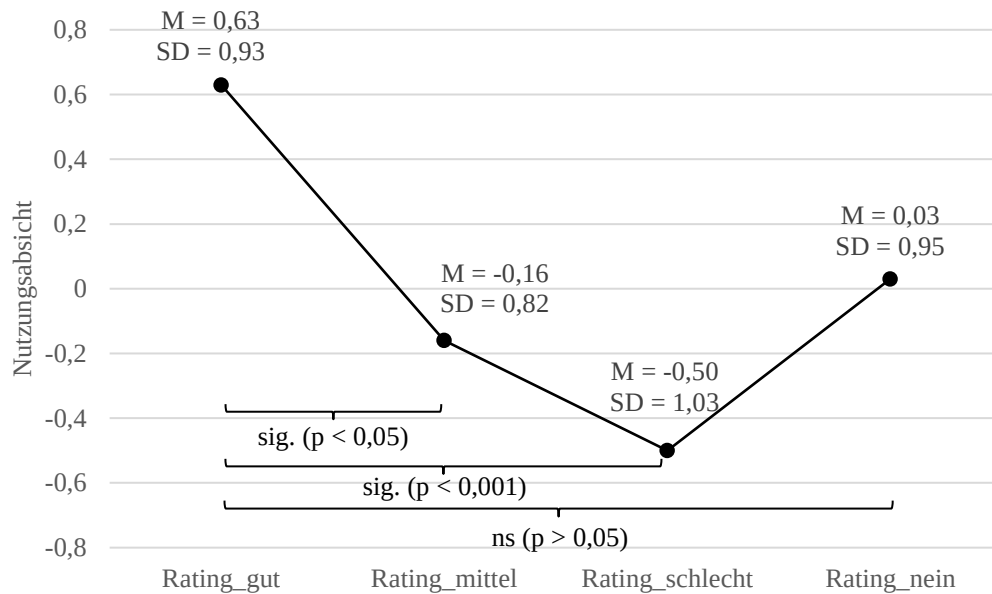


Abbildung 13: Ergebnisse der ANOVA (FP IV: E 1) – Effekte des Ratings auf die Nutzungsabsicht

Für die folgenden Analysen werden die Gruppen mit den Extremausprägungen (*Rating_schlecht* und *Rating_gut*) verwendet. Diese unterscheiden sich hinsichtlich der Nutzungsabsicht auch signifikant. Die Mediationsanalyse wird mithilfe des Mediationsmodells 81 des PROCESS-Makros mit einem Bootstrap Sample von 5.000 durchgeführt (Hayes, 2022, S. 622). Als unabhängige Variable wird das Rating verwendet. Die Erwartung, die Zufriedenheit sowie die Nutzungsabsicht werden als Mediatorvariable und die Wiederverwendungsabsicht als abhängige Variable genutzt. Das Mediationsmodell mit den entsprechenden Regressionskoeffizienten ist in Abbildung 14 dargestellt. Als Referenzkategorie dient hier die Experimentalgruppe *Rating_schlecht*.

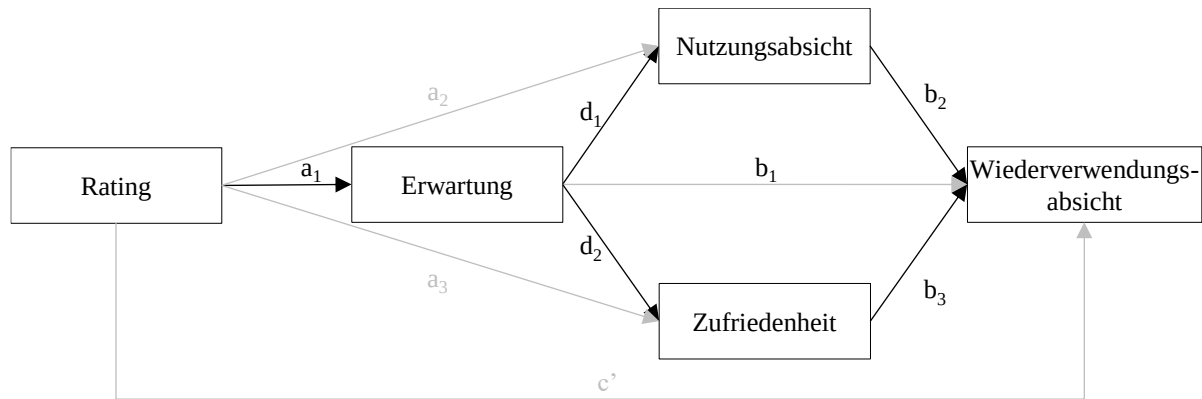


Abbildung 14: Mediationsmodell (FP IV: E 1)

Die Analyse zeigt einen Einfluss des Ratings auf die Erwartungen an den Chatbot. So sind die Erwartungen an den Chatbot bei einem guten Rating höher als bei einem schlechten Rating. ($a_1 = 0,738$, $p < 0,01$). Dieses Ergebnis liefert Bestätigung für Hypothese H₁. Die Nutzungsabsicht des Chatbots ist höher, wenn die Erwartungen hoch sind ($d_1 = 0,215$, $p < 0,01$), dies ist in Übereinstimmung mit Hypothese H₂. Niedrige Erwartungen sollten die Zufriedenheit erhöhen, da der Chatbot die Anfrage lösen konnte. Hierfür konnte allerdings keine Bestätigung gefunden werden. Hier war die Zufriedenheit größer, wenn die Erwartungen schon hoch waren. Dieser Unterschied war allerdings nicht signifikant ($d_2 = 0,055$, $p > 0,05$). Hypothese H₃ kann somit nicht bestätigt werden. Die Wiederverwendungsabsicht von Kunden ist sowohl für eine höhere initiale Nutzungsabsicht ($b_2 = 0,497$, $p < 0,001$) sowie eine höhere Zufriedenheit ($b_3 = 0,374$, $p < 0,01$) größer. Dies bestätigt die Hypothesen H₄ und H₅. Folglich mediieren die Erwartung sowie die Nutzungsabsicht die Beziehung zwischen Rating und der Wiederverwendungsabsicht seriell ($a \times d_1 \times b_2 = 0,079$, 95 % KI [0,016, 0,167]). Tabelle 45 zeigt einen Überblick der Ergebnisse der gesamten Mediationsanalyse. Die überprüften Hypothesen mit den Ergebnissen sind in Tabelle 46 zusammengefasst.

Tabelle 45: Regressionskoeffizienten der Mediationsanalyse (FP IV: E 1)

	Symbol	Koeff.	KI _u	KI _o
Effekte auf die Wiederverwendungsabsicht				
Rating	c'	-0,027 ^{ns}	-0,279	-0,226
Erwartung	b ₁	-0,035 ^{ns}	-0,165	0,096
Nutzungsabsicht	b ₂	0,497 ^{***}	0,260	0,735
Zufriedenheit	b ₃	0,374 ^{**}	0,133	0,614
Effekt auf die Erwartung				
Rating	a ₁	0,738 ^{**}	0,228	1,248
Effekte auf die Nutzungsabsicht				
Rating	a ₂	0,407 ^{**}	0,121	0,693
Erwartung	d ₁	0,215 ^{**}	0,063	0,366
Effekte auf die Zufriedenheit				
Rating	a ₃	0,294 [*]	0,011	0,576
Erwartung	d ₂	0,055 ^{ns}	-0,094	0,205
Rating x Erwartung	a ₁ x b ₁	-0,026	-0,125	0,072
Rating x Nutzungsabsicht	a ₂ x b ₂	0,202	0,043	0,428
Rating x Zufriedenheit	a ₃ x b ₃	0,110	0,011	0,263
Rating x Erwartung x Nutzungsabsicht	a ₁ x d ₁ x b ₂	0,079	0,016	0,167
Rating x Erwartung x Zufriedenheit	a ₁ x d ₂ x b ₃	0,015	-0,036	0,058
***: p < 0,001; **: p < 0,01; *: p < 0,05; ns: p > 0,05				

Tabelle 46: Zusammenfassung der Hypothesen (FP IV: E 1)

Hypothese H _i	Ergebnis
H ₁ Nutzer haben an einen Chatbot mit einer guten Bewertung (Rating beziehungsweise Rating/Review) höhere Erwartungen, dass der Chatbot die Anfrage lösen kann, als an einen Chatbot mit einer schlechten Bewertung.	✓
H ₂ Die Nutzungsabsicht ist bei hohen Erwartungen an einen Chatbot größer als bei geringen Erwartungen.	✓
H ₃ Die Zufriedenheit ist höher, wenn die Nutzer niedrigere Erwartungen an einen Chatbot haben und dieser trotzdem die Anfrage lösen kann als bei hohen Erwartungen.	✗
H ₄ Zufriedene Nutzer sind eher gewillt, den Chatbot wiederzuverwenden als unzufriedene Nutzer.	✓
H ₅ Haben Nutzer eine höherer Nutzungsabsicht ist deren Wiederverwendungsabsicht höher als bei Nutzern, die eine geringe Nutzungsabsicht aufweisen.	✓
✓ = bestätigt; ✗ = nicht bestätigt	

5.6.3 Experiment 2: Rating und Review

Das zweite Experiment aus Forschungsprojekt IV (*FP IV: E 2*) betrachtet die Auswirkung einer positiven oder negativen Bewertung eines anthropomorph gestalteten Chatbots auf die Wiederverwendungsabsicht. Dabei wird den Teilnehmern ergänzend zu dem Rating wie in *FP IV: E 1* noch ein positives beziehungsweise negatives Review gezeigt. Die Erwartungen an den Chatbot, die Zufriedenheit sowie die Nutzungsabsicht werden auch in diesem Experiment als Mediatoren betrachtet.

5.6.3.1 Methodik

Um den Stimulus zu erhöhen sowie die Wirkung von Rating und Review zu untersuchen, wird den Teilnehmer in diesem Experiment neben dem Rating ein Review gezeigt. Auch in diesem Experiment sehen die Teilnehmer ein hypothetisches Szenario in Form von Screenshots der Chatinteraktion. Das Kundenservice-Szenario war identisch mit *FP IV: E 1*. Es wird ein zweifaktorielles Design mit Zufallsgruppenbildung verwendet, mit Rating und Review als Zwischensubjektfaktoren. Dabei wurde jeweils die Ausprägung Rating und Review manipuliert (*Rating_gut*, *Rating_schlecht* und *Review_gut*, *Review_schlecht*). Nachdem die Teilnehmer den Chatbot mit dem Rating gesehen haben und dieser gefragt hat, wie er sie unterstützen kann, wurde den Teilnehmern ein Review gezeigt. Anschließend wurden sie zu ihrer Nutzungsabsicht (Venkatesh et al., 2003, S. 460) und ihren Erwartungen an den Chatbot befragt. Darüber hinaus wurde an dieser Stelle schon die Manipulation überprüft. Im Anschluss klickten sich die Teilnehmer durch die Chatinteraktion und konnten die Zählerstände erfolgreich melden. Das Szenario ist in Tabelle 47 dargestellt.

Tabelle 47: Szenariobeschreibung (FP IV: E 2)

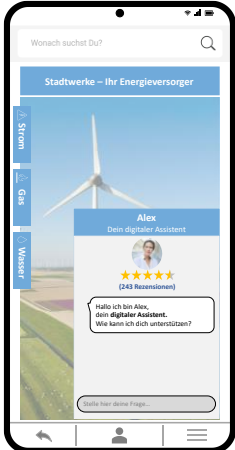
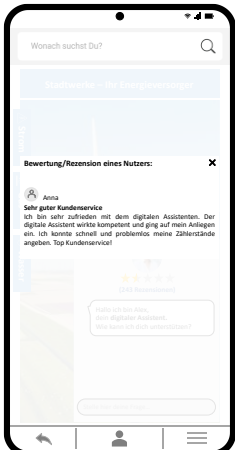
Einführung	
Bitte versetzen Sie sich in die Situation, dass Sie Ihre aktuellen Zählerstände (Gas, Strom und Wasser) Ihrem Energieversorger (Stadtwerke) melden möchten. Auf der Webseite der Stadtwerke haben Sie die Möglichkeit, mit einem Online-Kundenservice in Form eines digitalen Assistenten zu schreiben. Auf der folgenden Seite ist die Webseite der Stadtwerke abgebildet.	
Gruppe	
Rating_gut	Rating_schlecht
	
Schauen Sie sich die Webseite genau an und klicken Sie anschließend auf "Weiter". Danach wird Ihnen eine Rezension des digitalen Assistenten gezeigt.	
Review_gut	Review_schlecht
Sehr guter Kundenservice Ich bin sehr zufrieden mit dem digitalen Assistenten. Der digitale Assistent wirkte kompetent und ging auf mein Anliegen ein. Ich konnte schnell und problemlos meine Zählerstände angeben. Top Kundenservice!	Nicht zu empfehlen Ich bin sehr unzufrieden mit dem digitalen Assistenten. Der digitale Assistent wirkte inkompetent und konnte mein Anliegen nicht lösen. Die Zählerstände musste ich anschließend nochmals postalisch melden. Reine Zeitverschwendung dieser digitale Kundenservice!
	
Sie entscheiden sich, diesen digitalen Assistenten zu nutzen. Auf der nächsten Seite wird Ihnen der Gesprächsverlauf mit dem digitalen Assistenten gezeigt.	

Tabelle 47: Szenariobeschreibung (FP IV: E 2) (Fortsetzung)

Chatinteraktion	
• Hallo ich bin Alex, dein digitaler Assistent. Wie kann ich dich unterstützen? • Ich möchte meine aktuellen Zählerstände melden. • Dabei helfe ich dir gerne. Nenne mir bitte die Adresse und die aktuellen Zählerstände. • Homburger Straße 62, 67663 Kaiserslautern, Gas: 2431 m³, Strom: 9302 kWh, Wasser: 846 m³ • Ok, ich habe die Daten eingetragen. Die Abrechnung erhältst du innerhalb der nächsten Wochen. • Vielen Dank! • Solltest du noch ein weiteres Anliegen haben, kannst du mich gerne jederzeit kontaktieren. Ich wünsche dir einen schönen Tag.	
	

Im Anschluss an das Szenario wurde die Zufriedenheit der Teilnehmer mit dem Service abgefragt (Voss et al., 1998, S. 59). Des Weiteren wurde die Wiederverwendungsabsicht (Moriuchi, 2021, S. 27-28) sowie ihrer Expertise zu Chatbots (Ohanian, 1990, S. 50) und ihrer Einstellung zu Bewertungen (Khare et al., 2011, S. 115) abgefragt. Zur Messung der Nutzungsabsicht, der Erwartung an den Chatbot, der Zufriedenheit, der Wiederverwendung, der Erfahrung mit dem Chatbot sowie der Einstellung zu Bewertungen wurden die Konstrukte mit ihren Skalen aus *FP IV: E 1* genutzt. Die verwendeten Skalen mit den übersetzten Indikatoren sind in Tabelle 41 und Tabelle 48 übersichtlich dargestellt. Abschließend wurden neben den Soziodemografika die Nutzungshäufigkeit von Chatbots in den letzten zwölf Monaten abgefragt, um die Stichprobe beschreiben zu können.

Tabelle 48: Manipulationscheck (FP IV: E 2)

Kürzel	Indikator	Skala
Manipulationscheck		
Wie war der digitale Assistent Alex bewertet (Sternebewertung)?		
MC_11	Der digitale Assistent war mit mehr als 4 Sternen bewertet.	Einfachauswahl
MC_12	Der digitale Assistent war mit weniger als 2 Sternen bewertet.	
Wie war die Rezension des digitalen Assistenten Alex (Bewertungstext)?		
MC_21	Die schriftliche Rezension war positiv .	Einfachauswahl
MC_22	Die schriftliche Rezension war negativ .	

5.6.3.2 Stichprobenbeschreibung und Gütebeurteilung

Insgesamt haben 193 Teilnehmer an dem Experiment teilgenommen. Nach Bereinigung der Stichprobe hat sich die Stichprobe auf 131 Teilnehmer reduziert. Die restlichen 62 Teilnehmer haben die Manipulationscheckfragen falsch beantwortet (Tabelle 49). 56,5 % der Teilnehmer sind weiblich, 40,5 % männlich, 2,2 % Divers und 0,8 % machen keine Angabe zu ihrem Geschlecht. Die Teilnehmer sind im Mittel 29,88 Jahre alt ($SD = 9,53$). Details zur Stichprobe können Tabelle 50 entnommen werden. 87,8 % der Teilnehmer haben in den letzten zwölf Monaten mindestens einmal mit einem Chatbot interagiert. 22,2 % der Teilnehmer geben an, sogar häufiger als sechsmal in diesem Zeitraum einen Chatbot genutzt zu haben. Die Verteilung der Nutzungshäufigkeit von Chatbots in den letzten zwölf Monaten der Teilnehmer ist in Anhang 1 zusammengefasst. Die Teilnehmer sollten ihre Erfahrung im Umgang mit Chatbots selbst beurteilen. Dabei haben die Teilnehmer im Durchschnitt $M = 4,59$ ($SD = 1,41$) auf einer Skala von 1 bis 7 angegeben. Bewertungen werden von den Teilnehmern allgemein als positiv gesehen ($M = 5,59$, $SD = 0,92$). Sowohl für die Erfahrung mit Chatbots als auch die Einstellung zu Bewertungen wurden die Werte mithilfe der Mean-Funktion berechnet.

Tabelle 49: Ergebnisse des Manipulationschecks (FP IV: E 2)

Gruppe	MC_11	MC_12		MC_21	MC_22		Gesamt nach MC
Rating_gut/Review_gut	45	2		46	1		44
Rating_gut/Review_schlecht	27	23		3	47		24
Rating_schlecht/Review_gut	24	23		45	2		21
Rating_schlecht/Review_schlecht	7	42		5	44		42

Tabelle 50: Soziodemographische Merkmale (FP IV: E 2)

Merkmals	Ausprägung	Prozentualer Anteil
Geschlecht	Weiblich	56,5 %
	Männlich	40,5 %
	Divers	2,2 %
	Keine Angabe	0,8 %
Alter	M = 29,88 (SD = 9,53), 6 kein Angabe	
Haushaltsgröße	1 Person	20,6 %
	2 Personen	40,5 %
	3-5 Personen	33,6 %
	Mehr als 5 Personen	1,5 %
	Keine Angabe	3,8 %
Höchster Bildungsabschluss	Noch in schulischer Ausbildung	0,0 %
	Volks-/Hauptschulabschluss	1,5 %
	Realschulabschluss	2,3 %
	Fachschulabschluss	5,3 %
	Abitur	19,1 %
	Fachhoch-/Hochschulabschluss	68,7 %
	Keine Angabe	3,1 %
Monatliches Haushaltsnettoeinkommen	Bis 1000 Euro	17,6 %
	1001 - 2000 Euro	22,1 %
	2001 - 3000 Euro	17,6 %
	3001 - 4000 Euro	11,5 %
	4001 - 5000 Euro	10,5 %
	Mehr als 5000 Euro	9,9 %
	Keine Angabe	10,7 %

Die Gütebeurteilung der gemessenen Konstrukte erfolgt auch in *FP IV: E 2* auf Basis der in Kapitel 5.2 beschriebene Gütekriterien und Grenzwerte. Hier musste der Indikator AttRa_4 der Einstellung zur Bewertung aus der Analyse ausgeschlossen werden, da dieser den Grenzwert der KITK unterschritten hat (AttRa_4 = 0,43). Die Ergebnisse sind in Tabelle 51 nach Bereinigung der Konstrukte zusammengefasst. Hier zeigt sich, dass die Grenzwerte der Gütekriterien für Reliabilität und Validität alle überschritten sind. Daher werden die Skalen für die Auswertung und Überprüfung der Hypothesen verwendet.

Tabelle 51: Gütekriterien (FP IV: E 2)

Konstrukt	Indikator	Faktorladung	KITK	Chronbachs Alpha	KMO	Erklärte Varianz
		> 0,4	> 0,5	> 0,7	> 0,6	> 50 %
Nutzungsabsicht	IntUse_1	0,963	0,917	0,963	0,780	93,130
	IntUse_2	0,960	0,909			
	IntUse_3	0,965	0,920			
Zufriedenheit	Sat_1	0,928	0,803	0,824	0,649	75,472
	Sat_2	0,872	0,676			
	Sat_3 ^R	0,802	0,594			
Wiederverwendung	Reu_1	0,897	0,817	0,896	0,863	71,363
	Reu_2	0,871	0,782			
	Reu_3	0,816	0,706			
	Reu_4	0,892	0,818			
	Reu_5	0,737	0,616			
Erfahrung	Exp_1	0,866	0,797	0,948	0,897	82,471
	Exp_2	0,932	0,889			
	Exp_3	0,937	0,896			
	Exp_4	0,913	0,861			
	Exp_5	0,904	0,848			
Einstellung zu Bewertungen	AttRa_1	0,801	0,661	0,839	0,847	61,826
	AttRa_2	0,857	0,741			
	AttRa_3	0,833	0,707			
	AttRa_5	0,672	0,525			
	AttRa_6	0,754	0,610			

5.6.3.3 Ergebnisse

Zur Überprüfung der Auswirkung eines Ratings sowie eines Reviews auf die Erwartungen an den Chatbot wird in einem ersten Schritt zwei zweifaktorielle ANOVAs ausgeführt. Darüber hinaus wird an dieser Stelle auch der Interaktionseffekt der beiden Gruppen betrachtet. Die Zwischensubjektfaktoren Rating und Review werden als feste Faktoren und die Erwartung als abhängige Variable verwendet. In einem weiteren Schritt wurde als fester Faktor ebenfalls wie in *FP IV: E 1* die Nutzungsabsicht betrachtet.

Die Ergebnisse der zweifaktoriellen ANOVA zeigen einen Einfluss des Ratings auf die Erwartungen ($F(3, 127) = 7,26, p < 0,01$). Für das Review wurde keine Auswirkung auf die Erwartungen gefunden ($F(3, 127) = 0,47, p > 0,05$). Ebenso war der Interaktionseffekt ($F(3, 127) = 0,03, p > 0,05$) nicht signifikant.

Zur Überprüfung der Auswirkung eines Ratings sowie eines Reviews auf die Nutzungsabsicht wird in einem zweiten Schritt eine weitere zweifaktorielle ANOVA ausgeführt. Die Ergebnisse dieser zweifaktoriellen ANOVA zeigen einen Einfluss des Ratings auf die Nutzungsabsicht ($F(3, 127) = 8,93, p < 0,01$). Für das Review wurde keine Auswirkung auf die Nutzungsabsicht gefunden ($F(3, 127) = 3,41, p > 0,05$). Ebenso war der Interaktionseffekt ($F(3, 127) = 0,75, p > 0,05$) nicht signifikant. Folglich hat das Rating allein einen positiven Einfluss sowohl auf die Erwartungen als auch auf die Nutzungsabsicht. Für die zweifaktorielle ANOVA wurde die Mean-Skala für die Nutzungsabsicht verwendet. Dadurch sind die Werte einheitlich zu den Werten der Erwartung. Die deskriptiven Statistiken der beiden ANOVA sind in Abbildung 15 visualisiert.

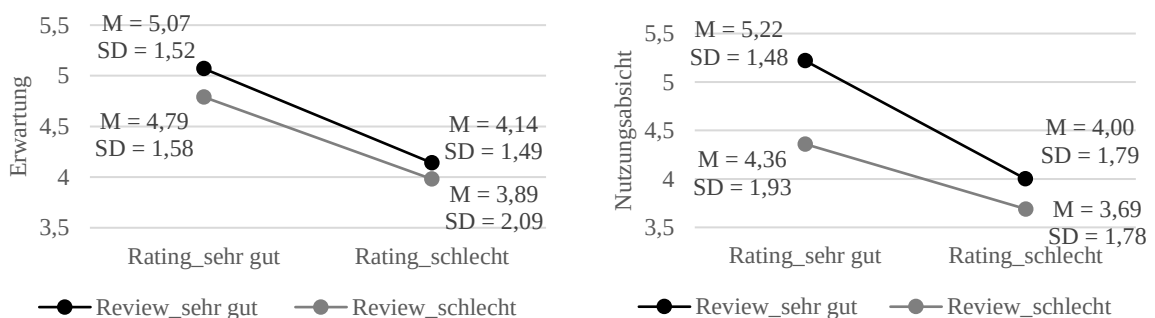


Abbildung 15: Ergebnisse der zweifaktoriellen ANOVAs (FP VI: E 2) – Effekte des Ratings und Review auf die Erwartung sowie die Nutzungsabsicht

Im Rahmen der Mediationsanalyse werden die Gruppe *Bewertung_gut* (*Rating_gut*, *Review_gut*) sowie die Gruppe *Bewertung_schlecht* (*Rating_schlecht*, *Review_schlecht*) verwendet. Diese unterscheiden sich hinsichtlich der Nutzungsabsicht sowie der Erwartung auch signifikant. Der Mittelwert der Nutzungsabsicht der Gruppe *Bewertung_gut* ist $M = 5,22$ ($SD = 1,48$), wohingegen die Gruppe *Bewertung_schlecht* die Nutzungsabsicht im Mittel mit $M = 3,69$ ($SD = 1,78$) beurteilt (95 % KI $[-2,233, -0,825]$, $t(80) = -4,32, p < 0,001$). Der Mittelwert der Erwartungen der Gruppe *Bewertung_gut* ist $M = 5,07$ ($SD = 1,52$). Die Gruppe *Bewertung_schlecht* beurteilt die Erwartungen im Mittel mit $M = 3,89$ ($SD = 2,09$) (95 % KI $[-1,872, -0,312]$, $t(84) = -2,784, p < 0,01$).

Die Mediationsanalyse zur Überprüfung der Forschungshypothesen wird mithilfe des Mediationsmodells 81 des PROCESS-Makros mit einem Bootstrap Sample von 5.000 durchgeführt (Hayes, 2022, S. 622). Als unabhängige Variable werden die beiden Gruppen *Bewertung_schlecht* und *Bewertung_gut* mit jeweils dem guten Rating und Review sowie dem schlechten Rating und Review verwendet. Die Erwartung, die Zufriedenheit sowie die Nutzungsabsicht werden als Mediatorvariablen und die Wiederverwendungsabsicht als abhängige Variable verwendet. Das Mediationsmodell mit den entsprechenden Regressionskoeffizienten ist in Abbildung 16 dargestellt. Als Referenzkategorie dient die Experimentalgruppe *Bewertung_schlecht*.

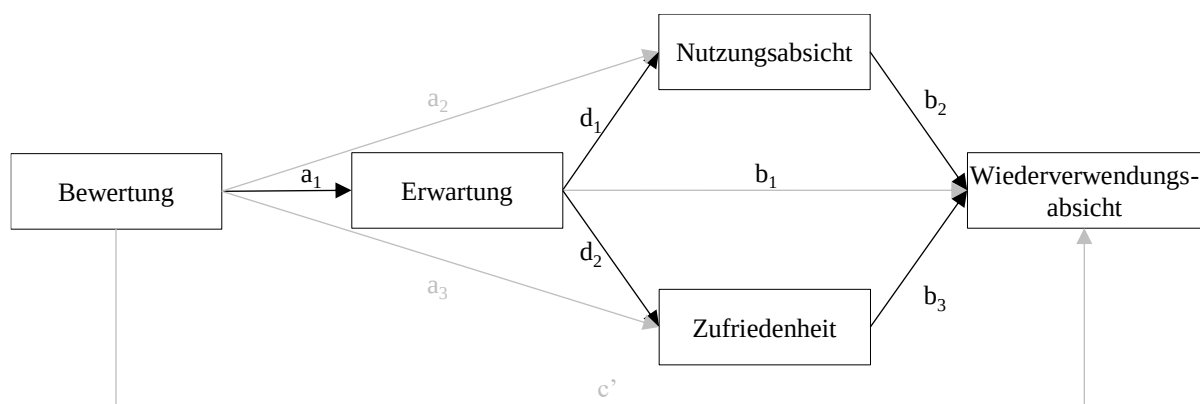


Abbildung 16: Mediationsmodell (FP IV: E 2)

Die Analyse zeigt einen Einfluss der Bewertung auf die Erwartungen an den Chatbot. Eine gute Bewertung führt zu höheren Erwartungen an den Chatbot als eine schlechte Bewertung ($a_1 = 0,364$, $p < 0,01$). Dies liefert Bestätigung für Hypothese H₁. Hohe Erwartungen führen zu einer größeren Nutzungsabsicht ($d_1 = 0,248$, $p < 0,001$). Folglich kann Hypothese H₂ bestätigt werden. Niedrige Erwartungen sollten in diesem Szenario einen positiven Einfluss auf die Zufriedenheit haben, was in dieser Analyse jedoch nicht bestätigt werden konnte. Der Einfluss der Erwartung auf die Zufriedenheit war nicht signifikant ($d_2 = 0,065$, $p > 0,05$). In diesem Experiment konnte somit keine Bestätigung für Hypothese H₃ gefunden werden. Eine höhere Nutzungsabsicht beeinflusst die Wiederverwendungsabsicht positiv ($b_2 = 0,219$, $p < 0,05$). Auch eine hohe Zufriedenheit führt zu einer höheren Absicht den Chatbot wiederzuverwenden ($b_3 = 0,453$, $p < 0,001$). Hypothesen H₄ und H₅ können somit bestätigt werden. Folglich medieren die Erwartung sowie die Nutzungsabsicht die Beziehung zwischen Bewertung und der

Wiederverwendungsabsicht seriell ($a \times d_1 \times b_2 = 0,020$, 95 % KI [0,011, 0,052]). Die Ergebnisse der gesamten Mediationsanalyse sind in Tabelle 52 zusammengefasst. Tabelle 53 gibt einen Überblick über die Hypothesen.

Tabelle 52: Regressionskoeffizienten der Mediationsanalyse (FP IV: E 2)

	Symbol	Koeff.	KI _u	KI _o
Effekte auf die Wiederverwendungsabsicht				
Bewertung	c'	-0,059 ^{ns}	-0,175	0,057
Erwartung	b ₁	-0,042 ^{ns}	-0,141	0,057
Nutzungsabsicht	b ₂	0,219*	0,011	0,428
Zufriedenheit	b ₃	0,453***	0,273	0,633
Effekt auf die Erwartung				
Bewertung	a ₁	0,364**	0,104	0,624
Effekte auf die Nutzungsabsicht				
Bewertung	a ₂	0,191**	0,074	0,308
Erwartung	d ₁	0,249***	0,155	0,343
Effekte auf die Zufriedenheit				
Bewertung	a ₃	-0,075 ^{ns}	-0,211	0,061
Erwartung	d ₂	0,065 ^{ns}	-0,044	0,174
	Symbol	Effekt	KI_u	KI_o
Bewertung x Erwartung	a ₁ x b ₁	-0,015	-0,059	0,021
Bewertung x Nutzungsabsicht	a ₂ x b ₂	0,042	0,003	0,102
Bewertung x Zufriedenheit	a ₃ x b ₃	0,034	-0,089	0,030
Bewertung x Erwartung x Nutzungsabsicht	a ₁ x d ₁ x b ₂	0,020	0,011	0,052
Bewertung x Erwartung x Zufriedenheit	a ₁ x d ₂ x b ₃	0,011	-0,007	0,032
***: p < 0,001; **: p < 0,01; *: p < 0,05; ns: p > 0,05				

Tabelle 53: Zusammenfassung der Hypothesen (FP IV: E 2)

Hypothese H _i		Ergebnis
H ₁	Nutzer haben an einen Chatbot mit einer guten Bewertung (Rating beziehungsweise Rating/Review) höhere Erwartungen, dass der Chatbot die Anfrage lösen kann, als an einen Chatbot mit einer schlechten Bewertung.	✓
H ₂	Die Nutzungsabsicht ist bei hohen Erwartungen an einen Chatbot größer als bei geringen Erwartungen.	✓
H ₃	Die Zufriedenheit ist höher, wenn die Nutzer niedrigere Erwartungen an einen Chatbot haben und dieser trotzdem die Anfrage lösen kann als bei hohen Erwartungen.	✗
H ₄	Zufriedene Nutzer sind eher gewillt, den Chatbot wiederzuverwenden als unzufriedene Nutzer.	✓
H ₅	Haben Nutzer eine höherer Nutzungsabsicht ist deren Wiederverwendungsabsicht höher als bei Nutzern, die eine geringe Nutzungsabsicht aufweisen.	✓
✓ = bestätigt; ✗ = nicht bestätigt		

5.6.4 Diskussion der Ergebnisse – Forschungsprojekt IV

FP IV untersucht den Einfluss von Bewertungen auf die Wiederverwendungsabsicht eines Chatbots. Dies geschieht im ersten Experiment unter Betrachtung eines Ratings und im zweiten Experiment unter Betrachtung einer gesamten Bewertung, die aus Rating und Review besteht. Dabei wurden in den beiden Experimenten die Erwartung, die Nutzungsabsicht sowie die Zufriedenheit als mögliche Mediatoren betrachtet.

Die Ergebnisse beider Experimente verdeutlichen, dass die Ausprägung eines Ratings einen Einfluss auf die Nutzungsabsicht des Chatbots sowie die Erwartungen an den Chatbot hat. Die Gruppenunterschiede zur Kontrollgruppe (*Rating_nein*) sind allerdings im ersten Experiment nicht signifikant. Diese Ergebnisse erweitern und bestätigen die Erkenntnisse von Arief et al. (2023), Nath et al. (2018), Ye et al. (2009) im Kontext von Chatbots. Für die Reviews kann im Rahmen der zweifaktoriellen ANOVAs kein Einfluss auf die Erwartungen und die Nutzungsabsicht gefunden werden. Hier können Ergebnisse bisheriger Forschung, die sich mit der Kauf- und Nutzungsabsicht beschäftigt haben, nicht bestätigt werden (Arief et al., 2023; Ye et al., 2009). Nutzer nutzen an dieser Stelle vielleicht eher die einfach und intuitiv zu verarbeitenden numerischen Informationen. Das Szenario könnte zu unwichtig sein, dass Nutzer Reviews lesen möchten und sie vertrauen viel mehr dem einfach zu verarbeitenden Rating. Dies ist in Einklang mit den Ausführungen von Kahneman (2012, S. 33), der im Rahmen der Informationsverarbeitung zwei kognitive Systeme unterscheidet: System eins, das automatisch und schnell arbeitet und weitgehend mühelos einfache Informationen verarbeitet und System zwei, das die Auf-

merksamkeit auf anstrengende mentale Aktivitäten lenkt. Reviews benötigen eine höhere kognitive Anstrengung als Ratings. Diese müssen gelesen und verstanden werden. Wohingegen Ratings einfach und automatisch verarbeitet werden können.

Die Ergebnisse der Mediationsanalyse beschreiben die Wirkzusammenhänge zwischen Bewertungen und der Wiederverwendungsabsicht genauer. Erwartungsgemäß führen positive Ratings oder Bewertungen zu höheren Erwartungen an den Chatbot. Diese Erkenntnis stützt die Annahme, dass Nutzer häufig unter Zuhilfenahme externer Informationen ein Urteil bilden und diese als Referenz genutzt werden. Hohe Erwartungen an einen Chatbot erhöhen die Nutzungsabsicht. Entgegen der Annahme, die auf Grundlage des Confirmation/Disconfirmation-Paradigma getroffen wurde, ist die Zufriedenheit mit dem Chatbot, der in diesem Szenario die Anfrage lösen kann, nicht höher, wenn die Erwartungen gering waren. Es deutet sich eher an, dass hohe Erwartungen auch zu einer leicht höheren Zufriedenheit führen. Dieser Einfluss ist allerdings nicht signifikant.

Die positiven Effekte der Nutzungsabsicht sowie der Zufriedenheit auf die Wiederverwendungsabsicht zeigen die Bedeutung dieser Konstrukte für die langfristige Nutzung von Chatbots. Gerade die serielle Mediation des Ratings beziehungsweise der Bewertung über die Erwartung und die Nutzungsabsicht auf die Wiederverwendungsabsicht zeigt, dass Bewertungen nicht nur einen direkten, kurzfristigen Einfluss auf die Nutzungsabsicht haben, sondern auch langfristig die Wiederverwendungsabsicht erhöhen können. Dies gilt zumindest für den Fall, wenn die Nutzeranfrage gelöst werden kann.

Für die erste Forschungsfrage konnte im Rahmen dieses Experiments eine mögliche Erklärung gefunden werden. Hier kann ein Rating beziehungsweise eine Bewertung über die Erwartung und die Nutzungsabsicht einen indirekten Einfluss auf die Wiederverwendungsabsicht haben. Ein direkter Einfluss wurde in diesem Forschungsprojekt nicht gefunden. Bezüglich der zweiten Forschungsfrage kann bestätigt werden, dass Ratings und Bewertungen einen Einfluss auf die Erwartungen an einen Chatbot haben können. Der Einfluss von Reviews hingegen konnte nicht bestätigt werden. Die dritte Forschungsfrage konnte im Rahmen dieses Forschungsprojekts auch beantwortet werden. Erwartungen erhöhen die Nutzungsabsicht. Die Zufriedenheit wird durch Erwartungen nicht beeinflusst.

Implikationen für die Praxis

Insgesamt zeigen die Ergebnisse, dass Ratings nicht nur als Informationsquelle wirken, sondern auch Erwartungen und indirekt die Nutzungsabsicht sowie die Wiederverwendungsabsicht beeinflussen. Für die Praxis bedeutet dies, dass Unternehmen, die Chatbots einsetzen, der Sichtbarkeit von Bewertung hohe Relevanz beimessen sollten. Gute Bewertungen können als Hebel zur Aktivierung von Nutzern dienen, während schlechte Bewertungen möglicherweise zu nachhaltigen Nutzungsverweigerungen führen, selbst dann, wenn der funktionale Nutzen des Systems gegeben ist. Bewertungen können gegebenenfalls helfen, skeptische Kunden zu erreichen und sie zur Nutzung von Chatbots zu bewegen. Unternehmen können auch aktiv Nutzer zur Abgabe von Bewertungen auffordern, um die Anzahl der Bewertungen zu erhöhen und dadurch mehr Glaubwürdigkeit zu erreichen. Je nachdem kann es auch ausreichen im Rahmen von Chatbots, die lediglich für einfache, unkritische Aufgaben verwendet werden, lediglich Ratings einzusetzen und auf komplexere Reviews zu verzichten.

Limitationen und Forschungsausblick

Auch *FP IV* weist gewisse Limitationen auf. Die Experimente basieren auf einem fiktiven Kundenservice-Szenario, in dem sich die Teilnehmer in die Situation versetzen sollten, ihre Zählerstände dem Energieversorger melden zu wollen. Des Weiteren wurden die Chatverläufe in Form von Screenshots gezeigt. Interaktive Dynamiken, wie sie in echten Kundengesprächen entstehen, blieben unberücksichtigt. In realen Situationen sind Emotionen und Konsequenzen oft stärker ausgeprägt. An dieser Stelle könnte ein Feldexperiment genutzt werden. Dadurch kann die externe Validität erhöht werden.

Eine weitere Limitation stellt das Design des hypothetischen Szenarios dar. Hier wurde nur ein schlechtes und ein gutes Review gezeigt. Es könnten mehrere schlechte oder gute Reviews genutzt werden, um den Stimulus zu verstärken und die Glaubwürdigkeit zu erhöhen. Darüber hinaus könnte es auch interessant sein, welche Gestaltungsmerkmale von Reviews im Kontext von Chatbots relevant sind. Des Weiteren wurden für das gute und das schlechte Rating nicht die Extremausprägungen (5 Sterne und 1 Stern) verwendet. Um den Stimulus zu verstärken, könnten diese allerdings genutzt werden.

Als Einflussfaktor auf die Zufriedenheit wurde in diesem Forschungsprojekt nur die Erwartung an den Chatbot genutzt. Dabei kann Zufriedenheit auch von anderen Faktoren abhängen, wie beispielsweise von der Gestaltung des Chatbots (Chang-Chien et al., 2022; Chen et al., 2024),

oder der Art der Serviceanfrage (Yu et al., 2024). Daher kann es sinnvoll sein, andere Mediatoren zu untersuchen, die einen Einfluss auf die Zufriedenheit haben können. Ein weiterer Ansatzpunkt, um die Ergebnisse in Bezug auf die Auswirkungen auf die Zufriedenheit nochmals genauer zu untersuchen, könnte die Betrachtung der Lösung als zusätzliche Variable sein. Dadurch wäre das Design noch näher an den Annahmen des Confirmation/Disconfirmation-Paradigmas. Es könnte besser überprüft werden, ob es zu Konfirmation oder Diskonfirmation kommt.

Eine weitere Limitation stellt die Tatsache dar, dass andere Faktoren, die die Erwartungen beeinflussen könnten, wie beispielsweise vorherigen Erfahrungen oder Informationen, die der Chatbot zu Beginn der Konversation gibt, nicht betrachtet wurden. Diese Limitation könnte in Folgeexperimenten adressiert werden. So könnten vorherige Erfahrungen genauer abgefragt werden und nicht nur die Chatbotnutzung in den letzten zwölf Monaten. Auch die zuvor beschriebene Aussage, dass Reviews vermehrt System zwei der kognitiven Verarbeitung zuzuordnen sind und dabei auch der Kontext eine Rolle spielen könnte, kann ein Ausgangspunkt für weiteren Forschungsarbeiten darstellen.

Weitere Limitationen sind auf die Stichprobe zurückzuführen. Hier mussten in Experiment 1 viele Teilnehmer aus der Stichprobe ausgeschlossen werden. Dies hat zu kleinen Gruppengrößen geführt. Gerade beim Rating wurden die Stimuli scheinbar nicht richtig wahrgenommen. Hier könnten die Ratings mehr hervorgehoben werden. Außerdem wurden die Experimente in Deutschland durchgeführt und die Stichproben sind für die Grundgesamtheit der Deutschen zu jung und haben ein sehr hohes Bildungsniveau.

6 Schlussbetrachtung

Dieses Kapitel fasst zunächst die zentralen Ergebnisse der Arbeit zusammen. Anschließend werden die gewonnenen Erkenntnisse genutzt, um Implikationen für Wissenschaft und unternehmerische Praxis abzuleiten. Abschließend folgt ein Ausblick auf den potenziellen zukünftigen Forschungsbedarf.

6.1 Zusammenfassung

Chatbots werden immer häufiger von Unternehmen im Kundenservice eingesetzt, um Nutzeranfragen in Echtzeit zu beantworten und Mitarbeiter zu entlasten. Perspektivisch werden Chatbots für viele Bereiche im Kundenservice zum Standard werden. Neben dem erheblichen Anstieg von Chatbot Anwendungen hat auch das Forschungsinteresse in den letzten Jahren deutlich zugenommen (Mariani et al., 2023, S. 4; Robinson et al., 2020, S. 373). Diese Arbeit behandelt vier interessante Themenfelder von anthropomorph gestalteten Chatbots: Anthropomorphismus, Offenlegungszeitpunkt, Übergabe an Menschen und Bewertungen.

Gerade im Bereich Anthropomorphismus und Soziale Präsenz sind hier in den letzten Jahren einige neue Forschungsbeiträge herausgekommen. Im Rahmen dieser Arbeit wird zu Beginn in *FP I* ein weiterer Beitrag zu diesem Themenfeld geliefert. Anschließend wird ein Fokus auf den Offenlegungszeitpunkt (*FP II*), die Übergabe an einen Menschen (*FP III*) sowie den Einfluss von Bewertungen (*FP IV*) im Kontext zu Chatbots gelegt. Hierzu finden sich bisher nur wenige wissenschaftliche Erkenntnisse.

Als erstes Ziel dieser Arbeit sollen konzeptionelle Grundlagen beschrieben und der empirische Stand der Forschung in den vier Bereichen aufgezeigt werden. Hierzu werden zunächst Grundlagen zu Chatbots beschrieben, indem unterschiedliche Definitionen vergleichend betrachtet werden und abschließend eine Arbeitsdefinition formuliert wird. Anschließend werden Kategorisierungsmöglichkeiten sowie mögliche Einsatzgebiete von Chatbots im Kundenservice betrachtet. Die konzeptionellen Grundlagen werden mit einem Kapitel zu Bewertungen abgeschlossen. In einem weiteren Schritt werden empirische Erkenntnisse zu den vier Themenfeldern beschrieben. Hierbei wird zuerst das Thema Anthropomorphismus betrachtet, anschließend das Thema Offenlegung, darauffolgend das Thema der Übergabe an einen Menschen und

abschließend das Thema Bewertungen. Hierbei wurde sichtbar, dass das Thema Anthropomorphismus von Chatbots zwischenzeitlich einige Forschungsbeiträge hervorgebracht hat. *FP I* bestätigt und erweitert bisherige Ergebnisse zu Anthropomorphismus. Zur Offenlegung von Chatbots finden sich auch einige Forschungsbeiträge. Allerdings stellt der Zeitpunkt der Offenlegung ein wenig untersuchtes Themenfeld dar. Dieser wurde in *FP II* adressiert. Verhaltenswissenschaftliche Erkenntnisse zum Thema der Übergabe an einen Menschen liegen auch wenige vor. Hier liefert *FP III* einen weiteren Beitrag. Das Themenfeld Bewertungen im Kontext von Chatbots stellt eine weitere Forschungslücke dar. *FP IV* erweitert die wissenschaftlichen Erkenntnisse zu Bewertungen im Kontext von Chatbots. Das zweite Ziel dieser Arbeit wird durch die Experimente der Forschungsprojekte adressiert. Diese liefern interessante wissenschaftliche Erkenntnisse zu den vier Themenfeldern. Dabei erweitern und bestätigen sie bisherige Ergebnisse im Kontext von Chatbots.

FP I beschäftigt sich in einem Experiment mit dem Einfluss von Anthropomorphismus auf die Wiederverwendungsabsicht. Hier wird die Soziale Präsenz und das Vertrauen als Mediatoren betrachtet. In *FP I* wird kein direkter Einfluss von Anthropomorphismus auf die Wiederverwendungsabsicht gefunden, allerdings ein indirekter. Dabei erhöht Anthropomorphismus die Soziale Präsenz, was wiederum zu einer höheren Wiederverwendungsabsicht und einem höheren Vertrauen führt. Ein hohes Vertrauen erhöht die Wiederverwendungsabsicht eines Chatbots. Hier liegt eine serielle Mediation vor.

FP II betrachtet die Rolle des Offenlegungszeitpunkts. Denn gerade durch die schnellen technologischen Fortschritte von Chatbots ist es immer schwieriger diese von menschlichen Mitarbeitern zu unterscheiden und der Punkt der Offenlegung wird immer wichtiger. Dabei wird *FP II: E 1* und *FP II: E 2* ein Fokus auf die Soziale Präsenz gelegt. Im ersten Experiment wurde die Wahrnehmung als sozial präsent durch eine späte Offenlegung erhöht. Dies konnte in Experiment zwei allerdings nicht bestätigt werden. Soziale Präsenz hat in beiden Experimenten die Wiederverwendungsabsicht erhöht. In *FP II: E 2* wurde die moderierende Rolle der Kommunikationsplattform untersucht. Hier konnte kein Einfluss der Kommunikationsplattform nachgewiesen werden. Der Einfluss des Offenlegungszeitpunkts auf die Soziale Präsenz ist nicht von der Kommunikationsplattform abhängig. *FP II: E 3* zeigt, dass mit einer Offenlegung nach der Interaktion negativere Auswirkungen verbunden sind, im Vergleich zu einer Offenlegung vor der Interaktion. Nutzer empfinden eine Offenlegung nach der Interaktion als weniger fair

und vermuten eher eigennützige Motive des Unternehmens dahinter. Diese negativen Auswirkungen der späten Offenlegung führen zu einer geringeren Wiederverwendungsabsicht. Eine moderierende Rolle, ob der Chatbot die Anfrage lösen kann oder nicht, wurde im Rahmen dieses Experiments nicht gefunden.

FP III untersucht den Einfluss einer Lösung der Anfrage auf die Wiederverwendungsabsicht von Chatbots. Hier wurde kein allgemeiner Einfluss gefunden. Allerdings unter Betrachtung der Wechselemöglichkeit zu einem Menschen konnte ein signifikanter Einfluss festgestellt werden. Wenn der Chatbot an einen Menschen übergibt und die Anfrage auch von diesem nicht gelöst werden kann, hat dies einen negativen Einfluss auf die Wiederverwendungsabsicht. Daher sollte sichergestellt werden, dass bei einem Wechsel die Nutzeranfrage gelöst werden kann.

Die Ergebnisse von *FP IV* zeigen, dass Bewertungen nicht nur kurzfristigen Einfluss auf die Erwartungen an den Chatbot und die Nutzungsabsicht eines Chatbots haben, sondern auch langfristig über die Nutzungsabsicht zu einer höheren Wiederverwendungsabsicht führt. In *FP IV* wurde eine serielle Mediation von Bewertungen über die Erwartungen und die Nutzungsabsicht auf die Wiederverwendungsabsicht gezeigt. Hier konnte der Chatbot die Nutzeranfrage lösen. So ist bei einer positiven Bewertung oder einem positiven Review die Erwartung an den Chatbot höher. In der Folge steigt die Nutzungsabsicht und eine höhere Nutzungsabsicht führt zu einer größeren Wiederverwendungsabsicht.

Die Forschungsprojekte dieser Arbeit verdeutlichen welche Faktoren die Wiederverwendungsabsicht von Chatbots beeinflussen können und dadurch zu einem erfolgreichen Einsatz dieser beitragen können.

6.2 Implikationen für Forschung und Praxis

Diese Arbeit liefert einige Implikationen für die wissenschaftliche Forschung sowie die unternehmerische Praxis. Diese werden im Folgenden genauer beschrieben.

Implikationen für die wissenschaftliche Forschung

Zunächst wird der Begriff Chatbots definiert. Dabei werden unterschiedliche Definitionen vergleichend betrachtet, die bestehende Literatur aufgearbeitet und anschließend eine Arbeitsdefinition geliefert. Darüber hinaus wird der wissenschaftliche Stand der empirischen Forschung in

den für diese Arbeit relevanten Themengebieten aufgearbeitet und die wissenschaftlichen Erkenntnisse übersichtlich dargestellt. An dieser Stelle der Arbeit zeigt sich weiterer Forschungsbedarf, der mithilfe der Forschungsprojekte untersucht wird.

FP I unterstützt die wissenschaftliche Literatur, indem die Auswirkungen von Anthropomorphismus auf die Soziale Präsenz sowie die Wiederverwendungsabsicht untersucht werden. Dabei können bisherige Forschungsergebnisse bestätigt werden (Gnewuch et al., 2018, S. 11; Jiang et al., 2023, S. 8; Liu et al., 2024, S. 12; Morana et al., 2020, S. 11). Darüber hinaus stehen die Ergebnisse im Einklang mit der Media Equation Theorie sowie der Social Response Theorie. Auch das Vertrauen spielt bei Chatbots eine wichtige Rolle. Hier zeigen die Ergebnisse einen positiven Einfluss der Sozialen Präsenz auf das Vertrauen. Das Vertrauen wiederum beeinflusst die Wiederverwendungsabsicht positiv. Auch an dieser Stelle kann durch die Bestätigung der bisherigen Ergebnisse (De Cicco et al., 2020, S. 1224; Janson, 2023, S. 10; Lei et al., 2021, S. 3987; Liu et al., 2024, S. 11) im Bereich von Chatbots ein Beitrag zur wissenschaftlichen Forschung geleistet werden. Die serielle Mediation des Anthropomorphismus auf die Wiederverwendungsabsicht über die Soziale Präsenz sowie das Vertrauen erweitert die bisherige Forschung im Bereich von Chatbots.

FP II befasst sich mit dem Zeitpunkt der Offenlegung von Chatbots. Hier wird die bestehende Literatur ergänzt, die sich hauptsächlich mit der Offenlegung allgemein und weniger dem Zeitpunkt beschäftigt. Der Einfluss des Offenlegungszeitpunkts auf die Soziale Präsenz sowie die Wiederverwendungsabsicht kann im Rahmen dieses Forschungsprojekts nicht abschließend bestätigt oder abgelehnt werden. Die Kommunikationsplattform, auf der der Chatbot eingesetzt wird, hatte keine moderierende Funktion. Die Ergebnisse von *FP II* zeigen, dass eine späte Offenlegung die Fairnesswahrnehmung negativ beeinflusst und dass Nutzer negative Unternehmensmotive dahinter vermuten. In der Folge nimmt die Wiederverwendungsabsicht ab. Bis auf die Erhöhung der Sozialen Präsenz war eine späte Offenlegung mit negativen Auswirkungen verbunden.

FP III liefert erste wissenschaftliche Erkenntnisse zum Wechsel von einem Chatbot zu einem Menschen sowie der Lösung der Anfrage. Hier konnte kein direkter Einfluss der Lösung auf die Wiederverwendungsabsicht nachgewiesen werden. Allerdings zeigen die Ergebnisse eine moderierende Funktion der Wechselmöglichkeit. Hier werden bisherige wissenschaftliche Ergebnisse ergänzt und ein erster Ansatz zur Betrachtung dieses Forschungsfeldes geliefert.

FP IV adressiert erstmals in der Literatur Bewertungen im Kontext von Chatbots. Dabei wird das Confirmation/Disconfirmation-Paradigma als möglicher Erklärungsansatz verwendet. Weiterhin wurden die Annahmen, die auf Grundlage des Confirmation/Disconfirmation-Paradigmas getroffen wurden, nicht bestätigt. Allerdings haben Bewertungen auch im Kontext von Chatbots einen Einfluss auf die Nutzungs- und Wiederverwendungsabsicht. Hiermit können bisherige Ergebnisse (Arief et al., 2023; Cui et al., 2012; Ye et al., 2009; Zhao et al., 2015) im Kontext von Chatbots erweitert werden. Entgegen der Aussage von Nazlan et al. (2018, S. 457), die eine Bewertung als Kombination aus Rating und Review empfehlen, deuten die Ergebnisse von *FP IV* darauf hin, dass bei Chatbots auch Ratings alleine einen Einfluss auf die Erwartungen und das Nutzungsverhalten haben. Im Kontext von Chatbots und einfachen Nutzeranfragen könnten Ratings ausreichen, um die Nutzer zu beeinflussen. Möglicherweise bevorzugen Nutzer bei der Interaktion mit Chatbots eine einfache und schnelle Form der Bewertung, die lediglich auf den quantitativen Ratings basiert.

Implikationen für die unternehmerische Praxis

Soziale Präsenz scheint für Nutzer ein wichtiger Einflussfaktor auf die Wiederverwendungsabsicht zu sein. Anthropomorphismus erhöht die Wahrnehmung als sozial präsent. Daher sollten Chatbots anthropomorph gestaltet werden. Gerade bei Interaktionen, bei denen Nutzer großen Wert auf menschlichen Kontakt legen oder für Nutzer, denen soziale Interaktionen sehr wichtig sind, kann eine anthropomorphe Gestaltung helfen. Perspektivisch könnten Chatbots personalisiert eingesetzt und abhängig vom jeweiligen Nutzer gestaltet werden. Hier könnten durch die Verwendung von Nutzerprofilen personalisierte Chatbots eingesetzt werden, ähnlich wie beispielsweise bei personalisierter Werbung, die auf dem vorherigen Nutzerverhalten basiert. Chatbots könnten in der Folge bezüglich des Grads an anthropomorphen Gestaltungsmerkmalen an die Vorlieben des Nutzers angepasst werden.

Unternehmen sollten Chatbots zu Beginn der Kommunikation offenlegen und somit von Anfang an ehrlich mit dem Nutzer umgehen. Eine Offenlegung nach der Interaktion kann vermehrt zu negativen Nutzerreaktionen führen als eine Offenlegung vor der Interaktion. In einem Kontext, in dem es besonders wichtig ist, dass der Chatbot als sozialer Akteur wahrgenommen wird, könnte es jedoch auch sinnvoll sein, seine Identität zunächst zu verbergen. In diesem Fall wissen die Nutzer während der Interaktion nicht, dass sie mit einem Chatbot kommunizieren, was die Wahrnehmung einer Sozialen Präsenz erhöhen kann, aber nicht zwangsläufig erhöhen muss,

wie die Ergebnisse von FP II: E 1 und FP II: E 2 gezeigt haben. An dieser Stelle müssen Unternehmen mögliche Regulierungen beachten und regel-, wie auch gesetzkonform agieren.

Unternehmen sollten die Wechsellmöglichkeit zu einem Menschen bewusst einsetzen und diesen Wechsel als mögliche Wiederherstellungsstrategie nicht automatisch als positiv ansehen. Denn wenn auch nach dem Wechsel keine Lösung der Anfrage erzielt werden kann, kann die Wiederverwendungsabsicht geringer sein. Wenn ein Wechsel zu einem Menschen stattfindet, sollte sichergestellt werden, dass die Anfrage auch gelöst werden kann, denn ansonsten könnten die Auswirkungen auf den Nutzer negativ sein. Ist gegebenenfalls im Vorfeld schon klar, dass die Nutzeranfrage auch von einem Menschen nicht gelöst werden kann, kann es sinnvoll sein, auf die Übergabe zu verzichten.

Unternehmen, die Chatbots einsetzen, sollten Bewertungen verwenden, um ihre Kunden zur Nutzung des Chatbots zu bewegen. Dabei sollten Bewertungen gut sichtbar dargestellt werden. Allerdings führen schlechte Bewertungen auch bei Chatbots zu negativen Nutzerreaktionen. Daher ist es von besonderer Bedeutung die Funktionalität der Chatbots sicherzustellen. Dies sollte mit technologischen Fortschritten auch in der breiten Unternehmenspraxis in Zukunft möglich sein. Mitarbeiter können dadurch entlastet werden und für andere Tätigkeiten im Unternehmen eingesetzt werden, die menschliche Intelligenz erfordern. Unternehmen sollten ihre Nutzer zukünftig auch um Bewertungen ihrer Chatboterfahrung bitten, um die Anzahl an Bewertungen zu erhöhen und dadurch eine bessere Grundlage zu bekommen. Allerdings muss an dieser Stelle angemerkt werden, dass in dieser Arbeit eine Vielzahl der Nutzer die Bewertungen nicht richtig wahrgenommen haben. Dadurch wird die Wichtigkeit solcher Bewertungen etwas in Frage gestellt. Im Rahmen von Chatbots und einfachen Nutzeranfragen könnte es auch ausreichen, lediglich Ratings von Nutzern zu verwenden und auf die kognitiv schwer zu verarbeitenden Reviews zu verzichten.

6.3 Forschungsausblick

Alle Experimente dieser Arbeit basieren auf szenariobasierten Studien. Diese werden in der wissenschaftlichen Forschung zwar häufig angewandt, allerdings schränken sie die Verallgemeinerbarkeit der Ergebnisse ein. Daher wurde auch in den Diskussionsteilen dieser Arbeit häufig der Konjunktiv verwendet. Des Weiteren sind die Stichproben der Experimente nicht repräsentativ für die Grundgesamtheit, denn das Alter, der Bildungsstand sowie das Geschlecht

repräsentieren diese nicht ausreichend. Diese Tatsache ist teilweise auf die Verwendung von Studierendensamples zurückzuführen, denn die Akquise wurde im Rahmen von Forschungsprojekten an der RPTU durchgeführt. Um die externe Validität der Ergebnisse dieser Arbeit zu verstärken, könnten die Experimente in angepasster Form als Feldexperimente durchgeführt und bei der Teilnehmerakquise ein Fokus auf eine repräsentative Gruppe gelegt werden.

Im Rahmen dieser Arbeit wurden relativ einfache Szenarien verwendet, die für Nutzer voraussichtlich keine hohe Priorität haben oder für die alternative Lösungen zur Verfügung stehen. So wäre beispielsweise beim Melden von Zählerständen ein Onlineformular eine naheliegende Option. Oder auch im Fall von *FP III* nutzen die Nutzer trotzdem ein Handy, um den Verlust ihres eigenen Handys zu melden. Zwar wurde im Szenario beschrieben, dass dies das Handy eines Freundes ist, allerdings könnte dieses Szenario auch nicht ganz der Realität entsprechen oder zu Verwirrung bei Teilnehmern geführt haben. Hier könnte durch die Nutzung eines realen Chatbots in weiteren Forschungsvorhaben die Ergebnisse unterstützt werden. Darüber hinaus wurde in allen Forschungsprojekten nur eine Interaktion an einem Handy simuliert. Diese Simulation ist auch statisch und wenig interaktiv. Zwar wurde ab *FP II: E 2* den Teilnehmern die Screenshots der Interaktion sukzessive gezeigt und dadurch versucht ein realitätsnahes Szenario darzustellen, allerdings verfügen Chatbots heutzutage durch den Einsatz künstlicher Intelligenz und der zunehmenden Menschenähnlichkeit über eine gewisse Reflexivität. Dadurch können sich Chatbots an die Nutzer und die entsprechende Situation dynamisch anpassen (Pütz et al., 2021, S. 379). Gerade auch durch die schnelle Entwicklung und den Einsatz von Chatbots für die Allgemeinheit (z. B. ChatGPT) haben sich in den letzten Jahren die Fähigkeiten und das Nutzungsverhalten von Chatbots verändert. Dieser Tatsache sollte in zukünftiger Forschung Beachtung geschenkt werden.

Im Rahmen dieser Arbeit wird ein Fokus auf die Wiederverwendungsabsicht gelegt. Weitere Faktoren könnten als Indikatoren für den Erfolg eines Chatbots untersucht werden. Hier könnte beispielsweise die Akzeptanz oder die Kundenzufriedenheit eher in den Fokus gesetzt werden.

Teilweise mussten viele Teilnehmer in den Experimenten ausgeschlossen werden, da sie die Manipulationscheckfragen falsch beantworteten. In der Folge wurde die statistische Auswertung dann mit kleinen Gruppen ausgeführt. Folglich sollte mit Verallgemeinerungen der Ergebnisse vorsichtig umgegangen werden. Hier stellt sich die Frage, warum die Teilnehmer die Ma-

nipulationscheckfragen so häufig falsch beantworteten. In den Forschungsprojekten zur Offenlegung haben die Teilnehmer den Hinweis möglicherweise nicht wahrgenommen oder überlesen. Oder es spielte für sie keine Rolle, ob sie in einem solchen Experiment mit einem Chatbot „interagieren“. In diesem Fall könnte die Offenlegungsinformation für die Nutzer irrelevant gewesen sein, was erklären würde, warum sie sich nicht daran erinnerten. Weitere Untersuchungen könnten helfen, die Bedeutung dieser Information für die Nutzer zu klären.

Der Einfluss des Offenlegungszeitpunkts auf die Soziale Präsenz konnte in *FP II* nicht abschließend geklärt werden. Daher sollte der Offenlegungszeitpunkt durch weitere Forschung überprüft werden. Allgemein hat die Offenlegung einen Einfluss auf die Soziale Präsenz (Hendriks et al., 2020; Mozafari et al., 2021b; Van der Goot et al., 2024). Folglich sollte sich die Soziale Präsenz auch für unterschiedliche Zeitpunkte der Offenlegung unterscheiden. Denn eine späte Offenlegung sollte in Bezug auf die Soziale Präsenz ähnlich bewertet werden, wie keine Offenlegung. In beiden Fällen ist während der Interaktion unklar, ob es sich um einen Chatbot oder einen menschlichen Interaktionspartner handelt.

In dieser Arbeit wurde auf die Wichtigkeit von Ratings im Bereich von Chatbots hingewiesen. In zukünftigen Experimenten sollte überprüft werden, ob Nutzer bei der Interaktion mit Chatbots, die einfache Nutzeranfragen bearbeiten, eine vereinfachte und schnellere Bewertungsform bevorzugen, die ausschließlich auf quantitativen Ratings basiert, während kognitiv anspruchsvollere Reviews nur für komplexere oder besonders wichtige Anliegen in Anspruch genommen werden. Darüber hinaus könnte auch der Einfluss der anthropomorphen Gestaltung von Chatbots bei der Nutzung von Ratings oder Reviews untersucht werden.

Diese Arbeit nutzt das Confirmation/Disconfirmation-Paradigma, um Zufriedenheit im Kontext von Chatbots zu beschreiben. Die Annahmen des Paradigmas konnten allerdings nicht bestätigt werden. Hier wurde allerdings auch nur die positive Diskonfirmation im Vergleich zur Konfirmation geprüft. Daher könnte in weiterer Forschung das gesamte Konstrukt geprüft werden. Kommt es beispielsweise zu negativer Diskonfirmation, wenn durch hohe Bewertungen die Erwartungen hoch sind, der Chatbot die Anfrage allerdings nicht lösen kann?

Durch die schnelle Entwicklung im Bereich der künstlichen Intelligenz basierten Chatbots werden Fehler, wie die in *FP II: E 3* und *FP III* dargestellter Form, immer seltener. Denn moderne, intelligente Chatbots liefern eigentlich fast immer eine Antwort und können die Anfrage scheinbar lösen. Hier können allerdings auch Fehlinformationen geliefert werden, auf die sich die

Nutzer verlassen. Weitere Forschung könnte sich daher mit dem Umgang von Aussagen von intelligenten Chatbots beschäftigen und wie diese von Nutzern beurteilt werden oder welche Faktoren Nutzer zum kritischen Denken in Bezug auf die Antworten des Chatbots anregen. In diesem Zusammenhang wäre auch ein Vergleich zwischen regelbasierten und auf künstlicher Intelligenz basierenden Chatbots von Interesse, insbesondere im Hinblick darauf, inwieweit sich bisherige Forschungsergebnisse auf Chatbots mit künstlicher Intelligenz übertragen lassen.

Das Verhalten von Menschen im Umgang mit Chatbots hat sich schon durch die Einführung von ChatGPT geändert. So werden Informationen teilweise nicht mehr gegoogelt, sondern ChatGPT wird gefragt. Dabei vertrauen Nutzer teilweise blind den Informationen, die der Chatbot liefert. Auch im Kundenservice sind solche Entwicklungen zu erwarten und positive Erfahrungen mit Chatbots werden überwiegen. Daher wird Forschung im Bereich von Chatbots ein dynamisches und interessantes Themengebiet bleiben.

Anhang

Anhang 1: Nutzungshäufigkeit von Chatbots in den letzten zwölf Monaten.....	172
---	-----

Anhang 1: Nutzungshäufigkeit von Chatbots in den letzten zwölf Monaten

	FP I	FP II: E 1	FP II: E 2	FP II: E 3	FP III	FP IV: E 1	FP IV: E 2
0 mal	25,2 %	10,9 %	13,0 %	16,0 %	6,7 %	11,1 %	12,2 %
1-3 mal	36,0 %	50,5 %	35,0 %	37,3 %	39,0 %	40,2 %	38,2 %
4-6 mal	18,9 %	20,8 %	30,0 %	19,3 %	25,1 %	29,9 %	27,5 %
7-9 mal	6,3 %	8,9 %	6,0 %	14,0 %	11,3 %	7,7 %	9,2 %
> 9 mal	13,5 %	8,9 %	16,0 %	13,3 %	17,9 %	11,1 %	13,0 %

Literaturverzeichnis

- Adam, M., Wessel, M. & Benlian, A. (2021). AI-Based Chatbots in Customer Service and Their Effects on User Compliance. *Electronic Markets*, 31(2), 427-445. <https://doi.org/10.1007/s12525-020-00414-7>
- Adamopoulou, E. & Moussiades, L. (2020a). Chatbots: History, Technology, and Applications. *Machine Learning with Applications*, 2, 1-18. <https://doi.org/10.1016/j.mlwa.2020.100006>
- Adamopoulou, E. & Moussiades, L. (2020b). An Overview of Chatbot Technology. In I. Maglogiannis, L. Iliadis & E. Pimenidis (Hrsg.), *Artificial Intelligence Applications and Innovations* (S. 373-383). Springer International Publishing. https://doi.org/10.1007/978-3-030-49186-4_31
- Al Farisi, R., Ferdiana, R. & Adj, T. B. (2022). The Effect of Anthropomorphic Design Cues on Increasing Chatbot Empathy. In *Proceedings of the 1st International Conference on Information System & Information Technology (ICISIT)* (S. 370-375). <https://doi.org/10.1109/ICISIT54091.2022.9873008>
- Alba, J. W. & Hutchinson, J. W. (1987). Dimensions of Consumer Expertise. *Journal of Consumer Research*, 13(4), 411-454. <https://doi.org/10.1086/209080>
- Anderson, E. W. (1998). Customer Satisfaction and Word of Mouth. *Journal of Service Research*, 1(1), 5-17. <https://doi.org/10.1177/109467059800100102>
- Araujo, T. (2018). Living up to the Chatbot Hype: The Influence of Anthropomorphic Design Cues and Communicative Agency Framing on Conversational Agent and Company Perceptions. *Computers in Human Behavior*, 85, 183-189. <https://doi.org/10.1016/j.chb.2018.03.051>
- Arief, M., Mustikowati, R. I. & Chrismardani, Y. (2023). Why Customers Buy an Online Product? The Effects of Advertising Attractiveness, Influencer Marketing and Online Customer Reviews. *Journal of Management & Research*, 21(1), 81-99. <https://doi.org/10.1108/LBSJMR-09-2022-0052>
- Arsenijevic, U. & Jovic, M. (2019). Artificial Intelligence Marketing: Chatbots. In *Proceedings of the International Conference on Artificial Intelligence: Applications and Innovations (IC-AIAI)* (S. 19-22). <https://doi.org/10.1109/IC-AIAI48757.2019.00010>
- Backhaus, K., Erichson, B., Gensler, S., Weiber, R. & Weiber, T. (2021). *Multivariate Analysemethoden*. Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-32425-4>
- Bae, S. & Lee, T. (2011). Product Type and Consumers' Perception of Online Consumer Reviews. *Electronic Markets*, 21(4), 255-266. <https://doi.org/10.1007/s12525-011-0072-0>
- Baek, H., Ahn, J. & Choi, Y. (2012). Helpfulness of Online Consumer Reviews: Readers' Objectives and Review Cues. *International Journal of Electronic Commerce*, 17(2), 99-126. <https://doi.org/10.2753/JEC1086-4415170204>
- Bala, K., Kumar, M., Hulawale, S. & Pandita, S. (2017). Chat-Bot for College Management System Using A.I. *International Journal of Engineering and Technology*, 4(11), 2030-2033.

- Bansal, H. S., Taylor, S. F. & James, Y. S. (2005). „Migrating“ to New Service Providers: Toward a Unifying Framework of Consumers' Switching Behaviors. *Journal of the Academy of Marketing Science*, 33(1), 96-115. <https://doi.org/10.1177/0092070304267928>
- Baron, R. M. & Kenny, D. A. (1986). The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, and Statistical Considerations. *Journal of Personality and Social Psychology*, 51(6), 1173-1182. DOI: 10.1037//0022-3514.51.6.1173
- Bearden, W. O., Netemeyer, R. G. & Teel, J. E. (1989). Measurement of Consumer Susceptibility to Interpersonal Influence. *Journal of Consumer Research*, 15(4), 473-481. <https://doi.org/10.1086/209186>
- Belanche, D., Belk, R. W., Casaló, L. V. & Flavián, C. (2024). The Dark Side of Artificial Intelligence in Services: The Service Industries Journal, 44(3-4), 149-172. <https://doi.org/10.1080/02642069.2024.2305451>
- Belanche, D., Casaló, L. V., Flavián, C. & Schepers, J. (2020). Robots or Frontline Employees? Exploring Customers' Attributions of Responsibility and Stability after Service Failure or Success. *Journal of Service Management*, 31(2), 267-289. <https://doi.org/10.1108/JOSM-05-2019-0156>
- Bendel, O. (2024). Chatbot: Was ist „Chatbot“? <https://wirtschaftslexikon.gabler.de/definition/chatbot-54248>, 16.04.2025
- Benner, D., Elshan, E., Schöbel, S. & Kanson, A. (2021). What Do You Mean? A Review on Recovery Strategies to Overcome Conversational Breakdowns of Conversational Agents. In *Proceedings of the International Conference on Information Systems (ICIS)* (S. 1-18). https://aisel.aisnet.org/icis2021/hci_robot/hci_robot/13
- Blut, M., Wang, C., Wunderlich, N. V. & Brock, C. (2021). Understanding Anthropomorphism in Service Provision: A Meta-Analysis of Physical Robots, Chatbots, and Other AI. *Journal of the Academy of Marketing Science*, 49(4), 632-658. <https://doi.org/10.1007/s11747-020-00762-y>
- Böhler, H., Germelmann, C. C., Baier, D. & Woratschek, H. (2021). *Marktforschung* (4. Auflage). Kohlhammer Verlag. <https://doi.org/10.17433/978-3-17-032249-3>
- Braggaar, A., Verhagen, J., Martijn, G. & Liebrecht, C. (2024). Conversational Repair Strategies to Cope with Errors and Breakdowns in Customer Service Chatbot Conversations. In A. Følstad, T. Araujo, S. Papadopoulos, E. L.-C. Law, E. Luger, M. Goodwin, S. Hobert & P. B. Brandtzaeg (Hrsg.), *Chatbot Research and Design* (S. 23-41). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-54975-5_2
- Browning, V., So, K. K. F. & Sparks, B. (2013). The Influence of Online Reviews on Consumers' Attributions of Service Quality and Control for Service Standards in Hotels. *Journal of Travel & Tourism Marketing*, 30(1-2), 23-40. <https://doi.org/10.1080/10548408.2013.750971>
- Cai, D., Li, H. & Law, R. (2022). Anthropomorphism and Ota Chatbot Adoption: A Mixed Methods Study. *Journal of Travel & Tourism Marketing*, 39(2), 228-255. <https://doi.org/10.1080/10548408.2022.2061672>
- Cascio, C. N., O'Donnell, M. B., Bayer, J., Tinney, F. J. & Falk, E. B. (2015). Neural Correlates of Susceptibility to Group Opinions in Online Word-of-Mouth Recommendations. *Journal of Marketing Research*, 52(4), 559-575. <https://doi.org/10.1509/jmr.13.0611>

- Castelo, N., Boegershausen, J., Hildebrand, C. & Henkel, A. P. (2023). Understanding and Improving Consumer Reactions to Service Bots. *Journal of Consumer Research*, 50(4), 848-863. <https://doi.org/10.1093/jcr/ucad023>
- Castillo, D., Canhoto, A. I. & Said, E. (2021). The Dark Side of Ai-Powered Service Interactions: Exploring the Process of Co-Destruction from the Customer Perspective. *The Service Industries Journal*, 41(13-14), 900-925. <https://doi.org/10.1080/02642069.2020.1787993>
- Chang-Chien, Y.-T., Cheng, J. C., Hsu, J. S.-C. & Yen, Y. W. (2022). Anthropomorphism of AI-based Intelligent Customer Service, and Its Affective and Behavioral Consequences. In *Proceedings of the 55th Hawaii International Conference on System Sciences (HICSS) (S. 1-10)*. DOI:10.24251/HICSS.2022.241
- Chen, J. V., Lin, C., Yen, D. C. & Linn, K.-P. (2011). The Interaction Effects of Familiarity, Breadth and Media Usage on Web Browsing Experience. *Computers in Human Behavior*, 27(6), 2141-2152. <https://doi.org/10.1016/j.chb.2011.06.008>
- Chen, J., Guo, F., Ren, Z., Li, M. & Ham, J. (2024). Effects of Anthropomorphic Design Cues of Chatbots on Users' Perception and Visual Behaviors. *International Journal of Human-Computer Interaction*, 40(14), 3636-3654. <https://doi.org/10.1080/10447318.2023.2193514>
- Cheng, X., Bao, Y., Zarifis, A., Gong, W. & Mou, J. (2022a). Exploring Consumers' Response To Text-Based Chatbots in E-Commerce: The Moderating Role of Task Complexity and Chatbot Disclosure. *Internet Research*, 32(2), 496-517. <https://doi.org/10.1108/INTR-08-2020-0460>
- Cheng, X., Zhang, X., Cohen, J. & Mou, J. (2022b). Human vs. AI: Understanding the Impact of Anthropomorphism on Consumer Response to Chatbots from the Perspective of Trust and Relationship Norms. *Information Processing & Management*, 59(3), 102940. <https://doi.org/10.1016/j.ipm.2022.102940>
- Cheng, Y. & Jiang, H. (2020). How Do AI-driven Chatbots Impact User Experience? Examining Gratifications, Perceived Privacy Risk, Satisfaction, Loyalty, and Continued Use. *Journal of Broadcasting & Electronic Media*, 64(4), 592-614. <https://doi.org/10.1080/08838151.2020.1834296>
- Chira, I., Adams, M. & Thornton, B. (2008). Behavioral Bias Within the Decision Making Process. *Journal of Business & Economics Research*, 6(8), 11-20. <https://doi.org/10.19030/jber.v6i8.2456>
- Choi, S., Mattila, A. S. & Bolton, L. E. (2021). To Err Is Human(-oid): How Do Consumers React to Robot Service Failure and Recovery? *Journal of Service Research*, 24(3), 354-371. <https://doi.org/10.1177/1094670520978798>
- Chong, T., Yu, T., Keeling, D. I. & Ruyter, K. de (2021). Ai-Chatbots on the Services Frontline Addressing the Challenges and Opportunities of Agency. *Journal of Retailing and Consumer Services*, 63, 102735. <https://doi.org/10.1016/j.jretconser.2021.102735>
- Ciechanowski, L., Przegalinska, A., Magnuski, M. & Gloor, P. (2019). In the Shades of the Uncanny Valley: An Experimental Study of Human-Chatbot Interaction. *Future Generation Computer Systems*, 92, 539-548. <https://doi.org/10.1016/j.future.2018.01.055>

- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2. Auflage). Lawrence Erlbaum Associates. <https://doi.org/10.4324/9780203771587>
- Crolic, C., Thomaz, F., Hadi, R. & Stephen, A. T. (2022). Blame the Bot: Anthropomorphism and Anger in Customer–Chatbot Interactions. *Journal of Marketing*, 86(1), 132-148. <https://doi.org/10.1177/00222429211045687>
- Cui, G., Lui, K.-K. & Guo, X (2012). The Effect of Online Consumer Reviews on New Product Sales. *International Journal of Electronic Commerce*, 17(1), 39-57. <https://doi.org/10.2753/JEC1086-4415170102>
- Dale, R. (2016). The Return of the Chatbots. *Natural Language Engineering*, 22(5), 811-817. <https://doi.org/10.1017/S1351324916000243>
- Davis, N., Olsen, N., Perry, V. G., Stewart, M. M. & White, T. B. (2023). I’m Only Human? The Role of Racial Stereotypes, Humanness, and Satisfaction in Transactions with Anthropomorphic Sales Bots. *Journal of the Association for Consumer Research*, 8(1), 47-58. <https://doi.org/10.1086/722703>
- De Cicco, R., Da Costa e Silva, S. C. L. & Palumbo, R. (2021). Should a Chatbot Disclose Itself? Implications for an Online Conversational Retailer. In A. Følstad, T. Araujo, S. Papadopoulos, E. L.-C. Law, E. Luger, M. Goodwin & P. B. Brandtzaeg (Hrsg.), *Chatbot Research and Design* (S. 3-15). Springer International Publishing. https://doi.org/10.1007/978-3-030-68288-0_1
- De Cicco, R., Silva, S. C. & Alparone, F. R. (2020). Millennials’ Attitude Toward Chatbots: An Experimental Study in a Social Relationship Perspective. *International Journal of Retail & Distribution Management*, 48(11), 1213-1233. <https://doi.org/10.1108/IJRD-12-2019-0406>
- Diederich, S., Brendel, A. B. & Kolbe, L. M. (2020). Designing Anthropomorphic Enterprise Conversational Agents. *Business & Information Systems Engineering*, 62(3), 193-209. <https://doi.org/10.1007/s12599-020-00639-y>
- Donath, T. (2024). Verbraucher offen für die Nutzung intelligenter Chatbots im Kundenservice: Einsatz von KI-Chatbots im Kundenservice. <https://www.nordlight-research.com/de/publikationen/presse/presse-detail/verbraucher-offen-fuer-die-nutzung-intelligenter-chatbots-im-kundenservice.html>, 16.04.2025
- Duan, W., Gu, B. & Whinston, A. B. (2008). Do Online Reviews Matter? – An Empirical Investigation of Panel Data. *Decision Support Systems*, 45(4), 1007-1016. <https://doi.org/10.1016/j.dss.2008.04.001>
- Ebert, T. A. E. & Raithel, S. (2009). Operationalisierung latenter Variablen. *Wirtschaftswissenschaftliches Studium*, 38(3), 125-130. <https://doi.org/10.15358/0340-1650-2009-3-125>
- Elwalda, A., Lü, K. & Ali, M. (2016). Perceived Derived Attributes of Online Customer Reviews. *Computers in Human Behavior*, 56, 306-319. <https://doi.org/10.1016/j.chb.2015.11.051>
- Ennew, C. T., Banerjee, A. K. & Li, D. (2000). Managing Word of Mouth Communication: Empirical Evidence from India. *International Journal of Bank Marketing*, 18(2), 75-83. <https://doi.org/10.1108/02652320010322985>

- Epley, N., Akalis, S., Waytz, A. & Cacioppo, J. T. (2008a). Creating Social Connection Through Inferential Reproduction: Loneliness and Perceived Agency in Gadgets, Gods, and Greyhounds. *Psychological Science*, 19(2), 114-120. <https://doi.org/10.1111/j.1467-9280.2008.02056.x>
- Epley, N., Waytz, A., Akalis, S. & Cacioppo, J. T. (2008b). When We Need A Human: Motivational Determinants of Anthropomorphism. *Social Cognition*, 26(2), 143-155. <https://doi.org/10.1521/soco.2008.26.2.143>
- Epley, N., Waytz, A. & Cacioppo, J. T. (2007). On Seeing Human: A Three-Factor Theory of Anthropomorphism. *Psychological Review*, 114(4), 864-886. <https://doi.org/10.1037/0033-295X.114.4.864>
- Europäisches Parlament. (2023). KI-Gesetz: Erste Regulierung der künstlichen Intelligenz. <https://www.europarl.europa.eu/topics/de/article/20230601STO93804/ki-gesetz-erste-regulierung-der-kunstlichen-intelligenz#nchste-schritte-5>, 16.04.2025
- Fan, H., Han, B. & Wang, W. (2024). Aligning (In)Congruent Chatbot–Employee Empathic Responses with Service-Recovery Contexts for Customer Retention. *Journal of Travel Research*, 63(8), 1870-1893. <https://doi.org/10.1177/00472875231201505>
- Feine, J., Morana, S. & Gnewuch, U. (2019). Measuring Service Encounter Satisfaction with Customer Service Chatbots Using Sentiment Analysis. In *Proceedings of the 14th International Conference on Wirtschaftsinformatik* (S. 1-17). Siegen, Deutschland.
- Ferner, C. & Lienbacher, E. (2023). Chatbots an Physischen Touchpoints: Anwendungsbeispiele und Anforderungen aus Perspektive der Informationstechnologie. In G. Schuster & B. Wecke (Hrsg.), *Marketingtechnologien* (S. 159-172). Springer Fachmedien Wiesbaden.
- Field, A. (2018). *Discovering Statistics Using IBM SPSS Statistics* (5. Auflage). Sage, Newbury Park.
- Filieri, R. (2016). What Makes an Online Consumer Review Trustworthy? *Annals of Tourism Research*, 58, 46-64. <https://doi.org/10.1016/j.annals.2015.12.019>
- Floh, A., Koller, M. & Zauner, A. (2013). Taking a Deeper Look at Online Reviews: The Asymmetric Effect of Valence Intensity on Shopping Behaviour. *Journal of Marketing Management*, 29(5-6), 646-670. <https://doi.org/10.1080/0267257X.2013.776620>
- Folkes, V. S. (1984). Consumer Reactions to Product Failure: An Attributional Approach. *Journal of Consumer Research*, 10(4), 398-409. <https://doi.org/10.1086/208978>
- Folkes, V. S. (1988). Recent Attribution Research in Consumer Behavior: A Review and New Directions. *Journal of Consumer Research*, 14(4), 548-565. <https://doi.org/10.1086/209135>
- Fu, J., Mouakket, S. & Sun, Y. (2023). The Role of Chatbots' Human-Like Characteristics in Online Shopping. *Electronic Commerce Research and Applications*, 61, 101304. <https://doi.org/10.1016/j.elerap.2023.101304>
- Funke, N., Stadler, K., Vakkuri, H., Wagner, A., Lunkenheimer, M. & Kracklauer, A. H. (2023). „Your Conversational Partner Is a Chatbot“ – An Experimental Study on the Influence of Chatbot Disclosure and Service Outcome on Trust and Customer Retention in the Fashion Industry. *Journal of Applied Interdisciplinary Research*. 1-18. <https://doi.org/10.25929/JAIR.V11I1.113>

- Gambino, A., Fox, J. & Ratan, R. (2020). Building a Stronger CASA: Extending the Computers Are Social Actors Paradigm. *Human-Machine Communication*, 1, 71-86. <https://doi.org/10.30658/hmc.1.5>
- Garbarino, E. & Maxwell, S. (2010). Consumer Response to Norm-Breaking Pricing Events in E-Commerce. *Journal of Business Research*, 63(9-10), 1066-1072. <https://doi.org/10.1016/j.jbusres.2008.12.010>
- Gavilan, D., Avello, M. & Martinez-Navarro, G. (2018). The Influence of Online Ratings and Reviews on Hotel Booking Consideration. *Tourism Management*, 66, 53-61. <https://doi.org/10.1016/j.tourman.2017.10.018>
- Gentsch, P. (2018). Künstliche Intelligenz für Sales, Marketing und Service. Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-19147-4>
- Gentsch, P. (2019). Conversational AI: How (Chat)Bots Will Reshape the Digital Experience. In P. Gentsch (Hrsg.), *AI in Marketing, Sales and Service* (S. 81-125). Springer International Publishing. https://doi.org/10.1007/978-3-319-89957-2_4
- Gerbing, D. W. & Anderson, J. C. (1988). An Updated Paradigm for Scale Development Incorporating Unidimensionality and Its Assessment. *Journal of Marketing Research*, 25(2), 186-192. <https://doi.org/10.1177/002224378802500207>
- Gershgor, D. (2018). A California Law Now Means Chatbots Have to Disclose They're Not Human. <https://qz.com/1409350/a-new-law-means-californias-bots-have-to-disclose-theyre-not-human>, 16.04.2025
- Gethmann, C. F., Buxmann, P., Distelrath, J., Humm, B. G., Lingner, S., Nitsch, V., Schmidt, J. C. & Spiecker, I. (2022). Künstliche Intelligenz in der Forschung (Bd. 48). Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-662-63449-3>
- Giacalone, D. & Jaeger, S. R. (2016). Better the Devil You Know? How Product Familiarity Affects Usage Versatility of Foods and Beverages. *Journal of Economic Psychology*, 55, 120-138. <https://doi.org/10.1016/j.joep.2016.02.005>
- Gnewuch, U., Morana, S., Adam, M. & Maedche, A. (2018). Faster is Not Always Better: Understanding the Effect of Dynamic Response Delays in Human-Chatbot Interaction. In *Proceedings of the European Conference on Information Systems (ECIS)* (S.1-18). Portsmouth, UK.
- Gnewuch, U., Morana, S., Hinz, O., Kellner, R. & Maedche, A. (2024). More Than a Bot? The Impact of Disclosing Human Involvement on Customer Interactions with Hybrid Service Agents. *Information Systems Research*, 35(3), 936-955. <https://doi.org/10.1287/isre.2022.0152>
- Gnewuch, U., Morana, S. & Maedche, A. (2017). Towards Designing Cooperative and Social Conversational Agents for Customer Service. In *Proceedings of the International Conference on Information Systems (ICIS)* (S. 1-14). <https://aisel.aisnet.org/icis2017/HCI/Presentations/1>
- Go, E. & Sundar, S. S. (2019). Humanizing Chatbots: The Effects of Visual, Identity and Conversational Cues on Hu-Manness Perceptions. *Computers in Human Behavior*, 97, 304-316. <https://doi.org/10.1016/j.chb.2019.01.020>
- Gohil, K. (2025). 35+ Chatbot Statistics That Need Your Immediate Attention. <https://meetanshi.com/blog/chatbot-statistics/>, 16.04.2025.

- Gong, L. (2008). How Social Is Social Responses to Computers? The Function of the Degree of Anthropomorphism in Computer Representations. *Computers in Human Behavior*, 24(4), 1494-1509. <https://doi.org/10.1016/j.chb.2007.05.007>
- Graff, B. (2013). Rassistischer Chat-Roboter: Mit falschen Werten bombardiert. *Süddeutsche Zeitung*. <https://www.sueddeutsche.de/wirtschaft/microsoft-programm-tay-rassistischer-chat-roboter-mit-falschen-werten-bombardiert-1.2928421>, 16.04.2025
- Grégoire, Y. & Fisher, R. J. (2008). Customer Betrayal and Retaliation: When Your Best Customers Become Your Worst Enemies. *Journal of the Academy of Marketing Science*, 36(2), 247-261. <https://doi.org/10.1007/s11747-007-0054-0>
- Grimes, G. M., Schuetzler, R. M. & Giboney, J. S. (2021). Mental Models and Expectation Violations in Conversational AI Interactions. *Decision Support Systems*, 144, 113515. <https://doi.org/10.1016/j.dss.2021.113515>
- Guha Majumder, M., Dutta Gupta, S. & Paul, J. (2022). Perceived Usefulness of Online Customer Reviews: A Review Mining Approach Using Machine Learning & Exploratory Data Analysis. *Journal of Business Research*, 150, 147-164. <https://doi.org/10.1016/j.jbusres.2022.06.012>
- Han, E., Yin, D. & Zhang, H. (2022). Chatbot Empathy in Customer Service: When It Works and When It Backfires. In *Proceedings of the PRE_ICIS Workshop SIGHCI* (S. 1-8). <https://aisel.aisnet.org/sighci2022/1>
- Han, M. C. (2021). The Impact of Anthropomorphism on Consumers' Purchase Decision in Chatbot Commerce. *Journal of Internet Commerce*, 20(1), 46-65. <https://doi.org/10.1080/15332861.2020.1863022>
- Hassanein, K. & Head, M. (2007). Manipulating Perceived Social Presence Through the Web Interface and Its Impact on Attitude Towards Online Shopping. *International Journal of Human-Computer Studies*, 65(8), 689-708. <https://doi.org/10.1016/j.ijhcs.2006.11.018>
- Hatwar, N., Patil, A. & Gondane, D. (2016). AI Based Chatbot. *International Journal of Emerging Trends in Engineering and Basic Sciences*, 3(2), 85-87.
- Haugeland, I. K. F., Følstad, A., Taylor, C., & Bjørkli, C. A. (2022). Understanding the User Experience of Customer Service Chatbots: An Experimental Study of Chatbot Interaction Design. *International Journal of Human-Computer Studies*, 161, S. 1-16, <https://doi.org/10.1016/j.ijhcs.2022.102788>
- Haupt, M., Rozumowski, A., Freidank, J. & Haas, A. (2023). Seeking Empathy or Suggesting a Solution? Effects of Chatbot Messages on Service Failure Recovery. *Electronic Markets*, 33(1), 1-22. <https://doi.org/10.1007/s12525-023-00673-0>
- Hayes, A. F. (2015). An Index and Test of Linear Moderated Mediation. *Multivariate Behavioral Research*, 50(1), 1-22. <https://doi.org/10.1080/00273171.2014.962683>
- Hayes, A. F. (2022). *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach* (3. Auflage). The Guilford Press.
- Hayes, A. F., Montoya, A. K. & Rockwood, N. J. (2017). The Analysis of Mechanisms and Their Contingencies: PROCESS versus Structural Equation Modeling. *Australasian Marketing Journal*, 25(1), 76-81. <https://doi.org/10.1016/j.ausmj.2017.02.001>

- Heerink, M., Kröse, B., Evers, V. & Wielinga, B. (2008). The Influence of Social Presence on Acceptance of a Companion Robot by Older People. *Journal of Physical Agents*, 2(2), 33-40. <https://doi.org/10.14198/JoPha.2008.2.2.05>
- Heerink, M., Kröse, B., Evers, V. & Wielinga, B. (2010). Relating Conversational Expressiveness to Social Presence and Acceptance of an Assistive Social Robot. *Virtual Reality*, 14(1), 77-84. <https://doi.org/10.1007/s10055-009-0142-1>
- Heider, F. (1958). *Psychology of Interpersonal Relations*. John Wiley & Sons Inc. <https://doi.org/10.1037/10628-000>
- Hendriks, F., Ou, C. X., Khodabandeh Amiri, A. & Bockting, S. (2020). The Power of Computer-Mediated Communication Theories in Explaining the Effect of Chatbot Introduction on User Experience. In *Proceedings of the 53rd Hawaii International Conference on System Sciences* (S. 1-9). <https://doi.org/10.24251/HICSS.2020.034>
- Ho-Dac, N. N., Carson, S. J. & Moore, W. L. (2013). The Effects of Positive and Negative Online Customer Reviews: Do Brand Strength and Category Maturity Matter? *Journal of Marketing*, 77(6), 37-53. <https://doi.org/10.1509/jm.11.0011>
- Homburg, C. (2020). *Marketingmanagement*. Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-29636-0>
- Homburg, C. & Giering, A. (1996). Konzeptualisierung und Operationalisierung komplexer Konstrukte: Ein Leitfaden für die Marketingforschung. *Marketing ZFP - Journal of Research and Management*, 18(1), 5-24. DOI:10.15358/0344-1369-1996-1-5
- Homburg, C. & Koschate, N. (2005). Behavioral Pricing-Forschung im Überblick: Teil 1: Grundlagen, Preisinformationsaufnahme und Preisinformationsbeurteilung. *Zeitschrift für Betriebswirtschaftslehre*, 75(4), 383-423.
- Hong, H., Di Xu, Wang, G. A. & Fan, W. (2017). Understanding the determinants of online review helpfulness: A meta-analytic investigation. *Decision Support Systems*, 102, 1-11. <https://doi.org/10.1016/j.dss.2017.06.007>
- Hoyer, W. D., Kroschke, M., Schmitt, B., Kraume, K. & Shankar, V. (2020). Transforming the Customer Experience Through New Technologies. *Journal of Interactive Marketing*, 51, 57-71. <https://doi.org/10.1016/j.intmar.2020.04.001>
- Hsiao, K.-L. & Chen, C.-C. (2022). What Drives Continuance Intention to Use a Food-Ordering Chatbot? An Examination of Trust and Satisfaction. *Library Hi Tech*, 40(4), 929-946. <https://doi.org/10.1108/LHT-08-2021-0274>
- Hu, Y. & Sun, Y. (2023). Understanding the Joint Effects of Internal and External Anthropomorphic Cues of Intelligent Customer Service Bot on User Satisfaction. *Data and Information Management*, 7(3), 100047. <https://doi.org/10.1016/j.dim.2023.100047>
- Huang, J., Zhou, M. & Yang, D. (2007). Extracting Chatbot Knowledge from Online Discussion Forums. In *Proceedings of the 20th International Joint Conference on Artificial Intelligence* (S. 423-428) Hyderabad, India.
- Huang, J.-W. & Lin, C.-P. (2011). To Stick or Not to Stick: The Social Response Theory in the Development of Continuance Intention from Organizational Cross-Level Perspective. *Computers in Human Behavior*, 27(5), 1963-1973. <https://doi.org/10.1016/j.chb.2011.05.003>

- Huang, Y.-S. & Dootson, P. (2022). Chatbots and Service Failure: When Does It Lead to Customer Aggression. *Journal of Retailing and Consumer Services*, 68, 103044. <https://doi.org/10.1016/j.jretconser.2022.103044>
- Huseynov, F. (2023). Chatbots in Digital Marketing: Enhanced Customer Experience and Reduced Customer Service Costs. In E. Li, A. S. Munna, M. S. I. Shaikh & B. U. Kazi (Hrsg.), *Contemporary Approaches of Digital Marketing and the Role of Machine Intelligence* (S. 46-72). <https://doi.org/10.4018/978-1-6684-7735-9.ch003>
- Io, H. N. & Lee, C. B. (2017). Chatbots and Conversational Agents: A Bibliometric Analysis. In *Proceedings of the International Conference on Industrial Engineering and Engineering Management (IEEM)* (S. 215-219). <http://ieeexplore.ieee.org/servlet/opac?punumber=8279571>
- Ischen, C., Araujo, T., van Noort, G., Voorveld, H. & Smit, E. (2020a). “I Am Here to Assist You Today”: The Role of Entity, Interactivity and Experiential Perceptions in Chatbot Persuasion. *Journal of Broadcasting & Electronic Media*, 64(4), 615-639. <https://doi.org/10.1080/08838151.2020.1834297>
- Ischen, C., Araujo, T., Voorveld, H., Noort, G. V. & Smit, E. (2020b). Privacy Concerns in Chatbot Interactions. In A. Følstad, T. Araujo, S. Papadopoulos, E. L.-C. Law, O.-C. Granmo, E. Luger & P. B. Brandtzaeg (Hrsg.), *Lecture Notes in Computer Science. Chatbot Research and Design* (Bd. 11970, S. 34-48). Springer International Publishing. https://doi.org/10.1007/978-3-030-39540-7_3
- Ishowo-Oloko, F., Bonnefon, J.-F., Soroye, Z., Crandall, J., Rahwan, I. & Rahwan, T. (2019). Behavioural Evidence for a Transparency–Efficiency Tradeoff in Human–Machine Cooperation. *Nature Machine Intelligence*, 1(11), 517-521. <https://doi.org/10.1038/s42256-019-0113-5>
- Izadi, S. & Forouzanfar, M. (2024). Error Correction and Adaptation in Conversational AI: A Review of Techniques and Applications in Chatbots. *AI*, 5(2), 803-841. <https://doi.org/10.3390/ai5020041>
- Jain, M., Kumar, P., Kota, R. & Patel, S. N. (2018). Evaluating and Informing the Design of Chatbots. In *Proceedings of the 2018 Designing Interactive Systems Conference* (S. 895-906). <https://doi.org/10.1145/3196709.3196735>
- Janson, A. (2023). How to Leverage Anthropomorphism for Chatbot Service Interfaces: The Interplay of Communication Style and Personification. *Computers in Human Behavior*, 149, 107954. <https://doi.org/10.1016/j.chb.2023.107954>
- Jiang, Y., Yang, X. & Zheng, T. (2023). Make Chatbots More Adaptive: Dual Pathways Linking Human-Like Cues and Tailored Response to Trust in Interactions with Chatbots. *Computers in Human Behavior*, 138, 107485. <https://doi.org/10.1016/j.chb.2022.107485>
- Jin, J., Walker, J. & Reczek, R. W. (2024). Avoiding Embarrassment Online: Response to and Inferences About Chatbots When Purchases Activate Self-Presentation Concerns. *Journal of Consumer Psychology*. <https://doi.org/10.1002/jcpy.1414>
- Jin, S. V. & Youn, S. (2023). Social Presence and Imagery Processing as Predictors of Chatbot Continuance Intention in Human-AI-Interaction. *International Journal of Human–Computer Interaction*, 39(9), 1874-1886. <https://doi.org/10.1080/10447318.2022.2129277>

- Johnson, E. J. & Russo, E. J. (1984). Product Familiarity and Learning New Information. *Journal of Consumer Research*, 11(1), 542-550. DOI:10.1086/208990
- Jurafsky, D. & Martin, J. H. (2024). *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition with Language Models* (3. Auflage). Pearson.
- Kahneman, D. (2012). *Schnelles Denken, langsames Denken* (2. Auflage). Siedler.
- Kaiser, H. F. & Rice, J. (1974). Little Jiffy, Mark IV. *Educational and Psychological Measurement*, 34(1), 111-117. <https://doi.org/10.1177/001316447403400115>
- Kanazawa, S. (1992). Outcome or Expectancy? Antecedent of Spontaneous Causal Attribution. *Personality and Social Psychology Bulletin*, 18(6), 659-668. <https://doi.org/10.1177/0146167292186001>
- Kayeser Fatima, J., Khan, M. I., Bahmannia, S., Chatrath, S. K., Dale, N. F. & Johns, R. (2024). Rapport With a Chatbot? The Underlying Role of Anthropomorphism in Socio-Cognitive Perceptions of Rapport and E-word of Mouth. *Journal of Retailing and Consumer Services*, 77, 103666. <https://doi.org/10.1016/j.jretconser.2023.103666>
- Kelley, H. H. (1973). The Process of Casual Attribution. *American Psychologist*, 28(2), 107-128. <https://doi.org/10.1037/h0034225>
- Kelley, H. H. & Michela, J. L. (1980). Attribution Theory and Research. *Annual Review of Psychology*, 31(1), 457-501. <https://doi.org/10.1146/annurev.ps.31.020180.002325>
- Keyser, A. de, Köcher, S., Alkire, L., Verbeeck, C. & Kandampully, J. (2019). Frontline Service Technology Infusion: Conceptual Archetypes and Future Research Directions. *Journal of Service Management*, 30(1), 156-183. <https://doi.org/10.1108/JOSM-03-2018-0082>
- Khanna, A., Pandey, B., Vashishta, K., Kalia, K., Pradeepkumar, B. & Das, T. (2015). A Study of Today's A.I. through Chatbots and Rediscovery of Machine Intelligence. *International Journal of u- and e-Service, Science and Technology*, 8(7), 277-284. <https://doi.org/10.14257/ijunesst.2015.8.7.28>
- Khare, A., Labrecque, L. I. & Asare, A. K. (2011). The Assimilative and Contrastive Effects of Word-of-Mouth Volume: An Experimental Examination of Online Consumer Ratings. *Journal of Retailing*, 87(1), 111-126. <https://doi.org/10.1016/j.jretai.2011.01.005>
- Kim, K. J., Park, E. & Shyam Sundar, S. (2013). Caregiving Role in Human–Robot Interaction: A Study of the Mediating Effects of Perceived Benefit and Social Presence. *Computers in Human Behavior*, 29(4), 1799-1806. <https://doi.org/10.1016/j.chb.2013.02.009>
- Kirchgeorg, M. (2018). Kundendienst. <https://wirtschaftslexikon.gabler.de/definition/kundendienst-40742/version-264120>, 16.04.2025
- Klein, K. & Martinez, L. F. (2023). The Impact of Anthropomorphism on Customer Satisfaction in Chatbot Commerce: An Experimental Study in the Food Sector. *Electronic Commerce Research*, 23(4), 2789-2825. <https://doi.org/10.1007/s10660-022-09562-8>
- Klein, O. (2025). Start der Meta-KI in Deutschland: Was hinter dem neuen Ring bei WhatsApp steckt. <https://www.zdf.de/nachrichten/panorama/whatsapp-meta-ki-kreis-chatbot-100.html>, 05.05.2025

- Klein, S. H., Papies, D. & Utz, S. (2024). How Interaction Mechanism and Error Responses Influence Users' Responses to Customer Service Chatbots. *International Journal of Human-Computer Interaction*, 41(7) 4300-4318. <https://doi.org/10.1080/10447318.2024.2351707>
- Koch, C., Linnik, B., Prlzel, F., Sultanow, E., Welter, S. & Cox, S. (2021). A Reference Architecture for On Premises Chatbots in Banks and Public Institutions. *Gesellschaft für Informatik*, 1265-1281. DOI: 10.18420/informatik2021-106.
- Konya-Baumbach, E., Biller, M. & Janda, S. von (2023). Someone Out There? A Study on the Social Presence of Anthropomorphized Chatbots. *Computers in Human Behavior*, 139, 107513. <https://doi.org/10.1016/j.chb.2022.107513>
- Koschate, N. (2008). Experimentelle Marktforschung. In A. Herrmann, C. Homburg & M. Klarmann (Hrsg.), *Handbuch Marktforschung: Methoden, Anwendungen, Praxisbeispiele* (3. Auflage) (S. 107-121). Gabler.
- Koschate-Fischer, N. & Schandelmeier, S. (2014). A Guideline for Designing Experimental Studies in Marketing Research and a Critical Discussion of Selected Problem Areas. *Journal of Business Economics*, 84(6), 793-826. <https://doi.org/10.1007/s11573-014-0708-6>
- Kostyk, A., Leonhardt, J. M. & Niculescu, M. (2017). Simpler Online Ratings Formats Increase Consumer Trust. *Journal of Research in Interactive Marketing*, 11(2), 131-141. <https://doi.org/10.1108/JRIM-06-2016-0062>
- Kotler, P., Keller, K. L. & Opresnik, M. O. (2017). *Marketing-Management: Konzepte - Instrumente - Unternehmensfallstudien* (15. Auflage). Pearson.
- Koumaras V., Foteas, A., Kapari, M., Sakkas, C. & Koumaras H. (2018). 5G Performance Testing of Mobile Chatbot Applications. In *Proceedings of the 23rd International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)* (S. 1-6). <http://ieeexplore.ieee.org/servlet/opac?punumber=8500732>
- Kreis, H., Wildner, R. & Kuß, A. (2021). *Marktforschung*. Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-32459-9>
- Kruse Brandão, T. & Wolfram, G. (2018). *Digital Connection*. Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-18759-0>
- Kukar-Kinney, M., Xia, L. & Monroe, K. B. (2007). Consumers' Perceptions of the Fairness of Price-Matching Refund Policies. *Journal of Retailing*, 83(3), 325-337. <https://doi.org/10.1016/j.jretai.2007.03.005>
- Kumar, V., Leone, R. P., Aaker, D. A. & Day, G. S. (2019). *Marketing Research* (13. Auflage). Wiley.
- Lackermair, G., Kailer, D. & Kanmaz, K. (2013). Importance of Online Product Reviews from a Consumer's Perspective. *Advances in Economics and Business*, 1(1), 1-5. <https://doi.org/10.13189/aeb.2013.010101>
- Lackes, R. & Siepermann, M. (2018). Künstliche Intelligenz (KI). <https://wirtschaftslexikon.gabler.de/definition/kuenstliche-intelligenz-ki-40285/version-263673>, 16.04.2025

- LaCour, M. & Serra, M. J. (2022). Average User Ratings Prompt Disparate Decision Strategies in Online Retail Shopping. *Journal of Consumer Behaviour*, 21(2), 231-244. <https://doi.org/10.1002/cb.1995>
- Langhe, B. de, Fernbach, P. M. & Lichtenstein, D. R. (2016). Navigating by the Stars: Investigating the Actual and Perceived Validity of Online User Ratings. *Journal of Consumer Research*, 42(6), 817-833. <https://doi.org/10.1093/jcr/ucv047>
- Lee, J. & Kim, Y. K. (2020). Online Reviews of Restaurants: Expectation-Confirmation Theory. *Journal of Quality Assurance in Hospitality & Tourism*, 21(5), 582-599. <http://dx.doi.org/10.1080/1528008X.2020.1712308>
- Lee, J., Park, D.-H. & Han, I. (2008). The Effect of Negative Online Consumer Reviews on Product Attitude: An Information Processing View. *Electronic Commerce Research and Applications*, 7(3), 341-352. <https://doi.org/10.1016/j.elerap.2007.05.004>
- Lee, K. M. (2008). Media Equation Theory. In W. Donsbach (Hrsg.), *The International Encyclopedia of Communication*. Wiley. <https://doi.org/10.1002/9781405186407.wbiecm035>
- Lei, S. I., Shen, H. & Ye, S. (2021). A Comparison Between Chatbot and Human Service: Customer Perception and Reuse Intention. *International Journal of Contemporary Hospitality Management*, 33(11), 3977-3995. <https://doi.org/10.1108/IJCHM-12-2020-1399>
- Leitherer, J. (2024). Chatbots kommen im Kundenservice gut an. <https://www.springerprofessional.de/kundenservice/chatbots/chatbots-kommen-im-kundenservice-gut-an/26685448>, 16.04.2025
- Lemon, K. N. & Verhoef, P. C. (2016). Understanding Customer Experience Throughout the Customer Journey. *Journal of Marketing*, 80(6), 69-96. <https://doi.org/10.1509/jm.15.0420>
- Lester, J., Branting, K. & Mott, B. (2004). Conversational Agents. In P. Singh (Hrsg.), *The Practical Handbook of Internet Computing* (S. 220-240). Chapman & Hall. <https://doi.org/10.1201/9780203507223>
- Li, B., Yao, R. & Nan, Y. (2023a). How Do Friendship Artificial Intelligence Chatbots (FAIC) Benefit the Continuance Using Intention and Customer Engagement? *Journal of Consumer Behaviour*, 22(6), 1376-1398. <https://doi.org/10.1002/cb.2218>
- Li, J., Wu, L., Qi, J., Zhang, Y., Wu, Z. & Hu, S. (2023b). Determinants Affecting Consumer Trust in Communication With AI Chatbots. *Journal of Organizational and End User Computing*, 35(1), 1-24. <https://doi.org/10.4018/JOEUC.328089>
- Li, Y., Gan, Z. & Zheng, B. (2023c). How do Artificial Intelligence Chatbots Affect Customer Purchase? Uncovering the Dual Pathways of Anthropomorphism on Service Evaluation. *Information Systems Frontiers*, 1-18. <https://doi.org/10.1007/s10796-023-10438-x>
- Lim, W. M., Kumar, S., Verma, S., & Chaturvedi, R. (2022). Alexa, What Do We Know About Conversational Commerce? Insights From a Systematic Literature Review. *Psychology & Marketing*, 39(6), 1129-1155. <https://doi.org/10.1002/mar.21654>
- Lis, B. & Korchmar, S. (2013). *Digitales Empfehlungsmarketing*. Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-01008-9>

- Liu, D., Li, C., Zhang, J. & Huang, W. (2023a). Robot Service Failure and Recovery: Literature Review and Future Directions. *International Journal of Advanced Robotic Systems*, 20(4). <https://doi.org/10.1177/17298806231191606>
- Liu, W., Jiang, M., Li, W. & Mou, J. (2024). How Does the Anthropomorphism of AI Chatbots Facilitate Users' Reuse Intention in Online Health Consultation Services? The Moderating Role of Disease Severity. *Technological Forecasting and Social Change*, 203, 123407. <https://doi.org/10.1016/j.techfore.2024.123407>
- Liu, Y., Hu, B., Yan, W. & Lin, Z. (2023b). Can Chatbots Satisfy Me? A Mixed-Method Comparative Study of Satisfaction with Task-Oriented Chatbots in Mainland China and Hong Kong. *Computers in Human Behavior*, 143, 107716. <https://doi.org/10.1016/j.chb.2023.107716>
- Lüken, J. & Schimmelpfennig, H. (2020). Bestimmung und Darstellung von Faktorwerten. *Planung & Analyse*, 3/2020, S. 19.
- Luo, X., Tong, S., Fang, Z. & Qu, Z. (2019). Frontiers: Machines vs. Humans: The Impact of Artificial Intelligence Chatbot Disclosure on Customer Purchases. *Marketing Science*, 38(6), 937-947. <https://doi.org/10.1287/mksc.2019.1192>
- Mariani, M. M., Hashemi, N. & Wirtz, J. (2023). Artificial Intelligence Empowered Conversational Agents: A Systematic Literature Review and Research Agenda. *Journal of Business Research*, 161, 113838. <https://doi.org/10.1016/j.jbusres.2023.113838>
- Mauldin, M. L. (1994). ChatterBots, TinyMuds, and the Turing Test Entering the Loebner Prize Competition. In *Proceedings of the 12th National Conference on Artificial Intelligence (AAAI-94)* (S. 16-21). Seattle Washington.
- McCammon, I. (2002). Evidence of Heuristic Traps in Recreational Avalanche Accidents. In *Proceedings of the International Snow Science Workshop* (S. 1-8). Penticton, British Columbia.
- McDonnell, M. & Baxter, D. (2019). Chatbots and Gender Stereotyping. *Interacting with Computers*, 31(2), 116-121. <https://doi.org/10.1093/iwc/iwz007>
- McGuire, J., Cremer, D. de, Hesselbarth, Y., Schutter, L. de, Mai, K. M. & Van Hiel, A. (2023). The Reputational and Ethical Consequences of Deceptive Chatbot Use. *Scientific Reports*, 13(1), 1-16. <https://doi.org/10.1038/s41598-023-41692-3>
- McTear, M. F. (2017). The Rise of the Conversational Interface: A New Kid on the Block. In J. F. Quesada, F.-J. Martín Mateos & T. López Soto (Hrsg.), *Future and Emerging Trends in Language Technology. Machine Learning and Big Data* (S. 38-49). Springer International Publishing.
- Melián-González, S., Bulchand-Gidumal, J. & González López-Valcárcel, B. (2013). Online Customer Reviews of Hotels. *Cornell Hospitality Quarterly*, 54(3), 274-283. <https://doi.org/10.1177/1938965513481498>
- Mende, M., Scott, M. L., Van Doorn, J., Grewal, D. & Shanks, I. (2019). Service Robots Rising: How Humanoid Robots Influence Service Experiences and Elicit Compensatory Consumer Responses. *Journal of Marketing Research*, 56(4), 535-556. <https://doi.org/10.1177/0022243718822827>

- Meng, L., Li, T., Shi, X. & Huang, X. (2023). Double-Sided Messages Improve the Acceptance of Chatbots. *Annals of Tourism Research*, 102, 103644. <https://doi.org/10.1016/j.annals.2023.103644>
- Meyer, N., Schwede, M., Hammerschmidt, M. & Weiger, W. H. (2022). Users Taking the Blame? How Service Failure, Recovery, and Robot Design Affect User Attributions and Retention. *Electronic Markets*, 32(4), 2491-2505. <https://doi.org/10.1007/s12525-022-00613-4>
- Meyer von Wolff, R., Hobert, S. & Schumann, M. (2019). How May I Help You? - State of the Art and Open Research Questions for Chatbots at the Digital Workplace. In *Proceedings of the 52nd Hawaii International Conference on System Sciences* (S. 95-104). Hawaii.
- Meyer von Wolff, R. & Schumann, M. (2018). Einsatz von Chatbots am digitalen Büroarbeitsplatz der Zukunft: Grundlagen, Stand der Forschung und Forschungsagenda (Arbeitsberichte des Instituts für Wirtschaftsinformatik, Professur für Anwendungssysteme und E-Business, Universität Göttingen No. 1). Göttingen.
- Moon, Y. (2000). Intimate Exchanges: Using Computers to Elicit Self-Disclosure from Consumers. *Journal of Consumer Research*, 26, 323-339. <https://doi.org/10.1086/209566>
- Moore, S. G. (2015). Attitude Predictability and Helpfulness in Online Reviews: The Role of Explained Actions and Reactions. *Journal of Consumer Research*, 42(1), 30-44. <https://doi.org/10.1093/jcr/ucv003>
- Morana, S., Gnewuch, U. & Jung, D. (2020). The Effect of Anthropomorphism on Investment Decision-Making with Robo-Advisor Chatbots. In *Proceedings of the 28 European Conference on Information Systems (ECIS)* (S. 1-20). https://aisel.aisnet.org/ecis2020_rp/63
- Mori, M. (1970). The Uncanny Valley. *Energy*, 7(4), 33-35.
- Mori, M., MacDorman, K. & Kageki, N. (2012). The Uncanny Valley. *Robotics & Automation Magazine*, 19(2), 98-100. <https://doi.org/10.1109/MRA.2012.2192811>
- Moriuchi, E. (2021). An Empirical Study on Anthropomorphism and Engagement With Disembodied Ais and Consumers' Reuse Behavior. *Psychology & Marketing*, 38(1), 21-42. <https://doi.org/10.1002/mar.21407>
- Mozafari, N., Weiger, W. H. & Hammerschmidt, M. (2021a). Resolving the Chatbot Disclosure Dilemma: Leveraging Selective Self-Presentation to Mitigate the Negative Effect of Chatbot Disclosure. In *Proceedings of the 54th Hawaii International Conference on System Sciences* (S. 2916-2923).
- Mozafari, N., Weiger, W. H. & Hammerschmidt, M. (2021b). That's so Embarrassing! When not to Design for Social Presence in Human-Chatbot Interactions. In *Proceedings of the 42nd International Conference on Information Systems* (S. 1-17). <https://ssrn.com/abstract=4305288>
- Mozafari, N., Weiger, W. H. & Hammerschmidt, M. (2022). Trust Me, I'm a Bot – Repercussions of Chatbot Disclosure in Different Service Frontline Settings. *Journal of Service Management*, 33(2), 221-245. <https://doi.org/10.1108/JOSM-10-2020-0380>
- Mozafari, N., Weiger, W. H. & Hammerschmidt, M. (2020). The Chatbot Disclosure Dilemma: Desirable and Undesirable Effects of Disclosing the Non-Human Identity of Chatbots.

- In Proceedings of the International Conference on Information Systems (S. 1415). https://aisel.aisnet.org/icis2020/hci_artintel/hci_artintel/6
- Mucci, T. (2024). Chatbot-Beispiele: Ein Leitfaden für Einsteiger. <https://www.ibm.com/de-de/think/topics/chatbot-use-cases>, 02.05.2025
- Mudambi, S. M. & Schuff, D. (2010). What Makes a Helpful Online Review? A Study of Customer Reviews on Amazon.com. *MIS Quarterly*, 34(1), 185-200. <https://ssrn.com/abstract=2175066>
- Murgia, A., Janssens, D., Demeyer, S. & Vasilescu, B. (2016). Among the Machines: Human-Bot Interaction on Social Q&A Websites. In Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (S. 1272-1279). <https://doi.org/10.1145/2851581.2892311>
- Murtarelli, G., Gregory, A. & Romenti, S. (2021). A Conversation-Based Perspective for Shaping Ethical Human–Machine Interactions: The Particular Challenge of Chatbots. *Journal of Business Research*, 129, 927-935. <https://doi.org/10.1016/j.jbusres.2020.09.018>
- Nadkarni, P. M., Ohno-Machado, L. & Chapman, W. W. (2011). Natural Language Processing: An Introduction. *Journal of the American Medical Informatics Association*, 18(5), 544-551. <https://doi.org/10.1136/amiajnl-2011-000464>
- Nass, C. & Moon, Y. (2000). Machines and Mindlessness: Social Responses to Computers. *Journal of Social Issues*, 56(1), 81-103. <https://doi.org/10.1111/0022-4537.00153>
- Nass, C., Moon, Y., Fogg, B., Reeves, B. & Dryer, C. (1995). Can Computer Personalities be Human Personalities? In Proceedings of the Conference Companion on Human Factors in Computing Systems (S. 228-229). DOI:10.1145/223355.223538
- Nass, C., Moon, Y. & Green, N. (1997). Are Machines Gender Neutral? Gender-Stereotypic Responses to Computers with Voices. *Journal of Applied Social Psychology*, 27(10), 864-876. <https://doi.org/10.1111/j.1559-1816.1997.tb00275.x>
- Nass, C., Steuer, J., Henriksen, L. & Dryer, C. (1994a). Machines, Social Attributions, and Ethopoeia: Performance Assessments of Computers Subsequent to ‘Self-‘ or ‘Other-‘ Evaluations. *International Journal of Human-Computer Studies*, 40, 543-559. <https://doi.org/10.1006/ijhc.1994.1025>
- Nass, C., Steuer, J. & Tauber, E. R. (1994b). Computers are Social Actors. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (S.72-78). <https://doi.org/10.1145/191666.191703>
- Nath, P., Devlin, J. & Reid, V. (2018). The Effects of Online Reviews on Service Expectations: Do Cultural Value Orientations Matter? *Journal of Business Research*, 90, 123-133. <https://doi.org/10.1016/j.jbusres.2018.05.001>
- Nazlan, N. H., Tanford, S. & Montgomery, R. (2018). The Effect of Availability Heuristics in Online Consumer Reviews. *Journal of Consumer Behaviour*, 17(5), 449-460. <https://doi.org/10.1002/cb.1731>
- Nguyen, M., Casper Ferm, L.-E., Quach, S., Pontes, N. & Thaichon, P. (2023). Chatbots in Frontline Services and Customer Experience: An Anthropomorphism Perspective. *Psychology & Marketing*, 40(11), 2201-2225. <https://doi.org/10.1002/mar.21882>

- Nunnally, J. C. & Bernstein, I. H. (1994). *Psychometric Theory* (3. Auflage). McGraw-Hill. <https://doi.org/10.1177/01466216950190030>
- Ohanian, R. (1990). Construction and Validation of a Scale to Measure Celebrity Endorsers' Perceived Expertise, Trustworthiness, and Attractiveness. *Journal of Advertising*, 19(3), 39-52. <https://doi.org/10.1080/00913367.1990.10673191>
- Ozuem, W., Ranfagni, S., Willis, M., Salvietti, G. & Howell, K. (2024). Exploring the Relationship Between Chatbots, Service Failure Recovery and Customer Loyalty: A Frustration–Aggression Perspective. *Psychology & Marketing*, 41(10), 2253-2273. <https://doi.org/10.1002/mar.22051>
- Park, D.-H., Lee, J. & Han, I. (2007). The Effect of On-Line Consumer Reviews on Consumer Purchasing Intention: The Moderating Role of Involvement. *International Journal of Electronic Commerce*, 11(4), 125-148. <https://doi.org/10.2753/JEC1086-4415110405>
- Park, G., Lee, S. & Chung, J. (2023). Do Anthropomorphic Chatbots Increase Counseling Satisfaction and Reuse Intention? The Moderated Mediation of Social Rapport and Social Anxiety. *Cyberpsychology, Behavior and Social Networking*, 26(5), 357-365. <https://doi.org/10.1089/cyber.2022.0157>
- Peter, P. J. (1979). Reliability: A Review of Psychometric Basics and Recent Marketing Practices. *Journal of Marketing Research*, 16(1), 6-17. <https://doi.org/10.1177/002224377901600102>
- Pisu, F. (2024). WhatsApp für Unternehmen: 7 sinnvolle Einsatzbereiche und Tipps. <https://www.userlike.com/de/blog/whatsapp-fuer-unternehmen>, 16.04.2025
- Poser, M., Hackbarth, T. & Bittner, E. A. (2022). Don't Throw It Over the Fence! Toward Effective Handover from Conversational Agents to Service Employees: User Experience and Behavior. In M. Kurosu (Hrsg.), *Human-Computer Interaction: User Experience and Behavior* (S. 531-545). Springer International Publishing. <https://doi.org/10.1007/978-3-031-05412-9>
- Poser, M., Hörhold, H. K. & Bittner, E. A. (2023). May the Guide be with you: CA-Facilitated Information Elicitation to Prevent Service Failure. In *Proceedings of the 31st European Conference on Information Systems (ECIS)* (S. 1-17). Kristiansand, Norway.
- Poser, M., Singh, S. & Bittner, E. A. (2021). Hybrid Service Recovery: Design for Seamless Inquiry Handovers between Conversational Agents and Human Service Agents. *Proceedings of the 54th Hawaii International Conference on System Sciences* (S. 1-10). <http://hdl.handle.net/10125/70756>
- Pütz, C., Düppre, S., Roth, S. & Weiss, W. (2021). Akzeptanz und Nutzung von Chat-/Voicebots. In M. Bruhn & K. Hadwich (Hrsg.), *Künstliche Intelligenz im Dienstleistungsmanagement* (S. 362-384). Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-34326-2>
- Puzakova, M., Kwak, H. & Rocereto, J. F. (2013). When Humanizing Brands Goes Wrong: The Detrimental Effect of Brand Anthropomorphization Amid Product Wrongdoings. *Journal of Marketing*, 77(3), 81-100. <https://doi.org/10.1509/jm.11.0510>

- Qiu, L. & Benbasat, I. (2009). Evaluating Anthropomorphic Product Recommendation Agents: A Social Relationship Perspective to Designing Information Systems. *Journal of Management Information Systems*, 25(4), 145-181. <https://doi.org/10.2753/MIS0742-1222250405>
- Racherla, P. & Friske, W. (2012). Perceived 'Usefulness' of Online Consumer Reviews: An Exploratory Investigation Across Three Services Categories. *Electronic Commerce Research and Applications*, 11(6), 548-559. <https://doi.org/10.1016/j.elerap.2012.06.003>
- Rajaobelina, L., Prom Tep, S., Arcand, M., & Ricard, L. (2021). Creepiness: Its Antecedents and Impact on Loyalty When Interacting with a Chatbot. *Psychology & Marketing*, 38(12), 2339-2356. <https://doi.org/10.1002/mar.21548>
- Reeves, B. & Nass, C. (1996). *The Media Equation: How People Treat Computers Television, and New Media Like Real People and Places*. Cambridge University Press.
- Reichelt, J. (2013). *Informationssuche und Online Word-of-Mouth*. Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-01373-8>
- Rese, A., Ganster, L. & Baier, D. (2020). Chatbots in Retailers' Customer Communication: How to Measure Their Acceptance? *Journal of Retailing and Consumer Services*, 56, 1-14. <https://doi.org/10.1016/j.jretconser.2020.102176>
- Robinson, S., Orsingher, C., Alkire, L., Keyser, A. de, Giebelhausen, M., Papamichail, K. N., Shams, P. & Temerak, M. S. (2020). Frontline Encounters of the AI Kind: An Evolved Service Encounter Framework. *Journal of Business Research*, 116, 366-376. <https://doi.org/10.1016/j.jbusres.2019.08.038>
- Roy, R. & Naidoo, V. (2021). Enhancing Chatbot Effectiveness: The Role of Anthropomorphic Conversational Styles and Time Orientation. *Journal of Business Research*, 126, 23-34. <https://doi.org/10.1016/j.jbusres.2020.12.051>
- Russell, S. J. & Norvig, P. (2022). *Artificial Intelligence: A Modern Approach* (4. Auflage). Pearson.
- Sands, S., Campbell, C., Plangger, K. & Pitt, L. (2022). Buffer Bots: The Role of Virtual Service Agents in Mitigating Negative Effects When Service Fails. *Psychology & Marketing*, 39(11), 2039-2054. <https://doi.org/10.1002/mar.21723>
- Sands, S., Ferraro, C., Campbell, C. & Tsao, H.-Y. (2021). Managing the Human–Chatbot Divide: How Service Scripts Influence Service Experience. *Journal of Service Management*, 32(2), 246-264. <https://doi.org/10.1108/JOSM-06-2019-0203>
- Savage, M. (2022). Was bedeutet Kundenservice heute? <https://news.sap.com/germany/2022/01/was-ist-kundenservice-heute/>, 16.04.2025
- Schanke, S., Burtch, G. & Ray, G. (2021). Estimating the Impact of "Humanizing" Customer Service Chatbots. *Information Systems Research*, 32(3), 736-751. <https://doi.org/10.1287/isre.2021.1015>
- Schindler, R. M. & Bickart, B. (2012). Perceived Helpfulness of Online Consumer Reviews: The Role of Message Content and Style. *Journal of Consumer Behaviour*, 11(3), 234-243. <https://doi.org/10.1002/cb.1372>
- Schmitt, B. (2019). From Atoms to Bits and Back: A Research Curation on Digital Technology and Agenda for Future Research. *Journal of Consumer Research*, 46(4), 825-832. <https://doi.org/10.1093/jcr/ucz038>

- Seeger, A.-M., Pfeiffer, J. & Heinzl, A. (2021). Texting with Humanlike Conversational Agents: Designing for Anthropomorphism. *Journal of the Association for Information Systems*, 22(4), 931-967. <https://doi.org/10.17705/1jais.00685>
- See-To, E. W. & Ho, K. K. (2014). Value Co-Creation and Purchase Intention in Social Network Sites: The Role of Electronic Word-of-Mouth and Trust—A Theoretical Analysis. *Computers in Human Behavior*, 31, 182-189. <https://doi.org/10.1016/j.chb.2013.10.013>
- Serban, I. V., Lowe, R., Henderson, P., Charlin, L. & Pineau, J. (2015). A Survey of Available Corpora for Building Data-Driven Dialogue Systems. <http://arxiv.org/pdf/1512.05742v3>
- Shankar, V. (2018). How Artificial Intelligence (AI) is Reshaping Retailing. *Journal of Retailing*, 94(4), vi-xi. [https://doi.org/10.1016/S0022-4359\(18\)30076-9](https://doi.org/10.1016/S0022-4359(18)30076-9)
- Shawar, B. A. & Atwell, E. (2007). Chatbots: Are They Really Useful? *Forum der Gesellschaft für Linguistische Datenverarbeitung*, 22(1), 29-49. <https://doi.org/10.21248/jlcl.22.2007.88>
- Sheehan, B., Jin, H. S. & Gottlieb, U. (2020). Customer service chatbots: Anthropomorphism and adoption. *Journal of Business Research*, 115, 14-24. <https://doi.org/10.1016/j.jbusres.2020.04.030>
- Shi, W., Wang, X., Oh, Y. J., Zhang, J, Sahay, S. & Yu, Z. (2020). Effects of Persuasive Dialogues: Testing Bot Identities and Inquiry Strategies. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (S. 1-13). <https://doi.org/10.1145/3313831.3376843>
- Short, J., Williams, E. & Christie, B. (1976). *The Social Psychology of Telecommunications*. Wiley.
- Silva, F. A., Shojaei, A. S. & Barbosa, B. (2023). Chatbot-Based Services: A Study on Customers' Reuse Intention. *Journal of Theoretical and Applied Electronic Commerce Research*, 18(1), 457-474. <https://doi.org/10.3390/jtaer18010024>
- Singh, J. P., Irani, S., Rana, N. P., Dwivedi, Y. K., Saumya, S. & Kumar Roy, P. (2017). Predicting the "Helpfulness" of Online Consumer Reviews. *Journal of Business Research*, 70, 346-355. <https://doi.org/10.1016/j.jbusres.2016.08.008>
- Skjuve, M., Haugstveit, I. M., Følstad, A. & Brandtzaeg, P. B. (2019). Help! Is My Chatbot Falling into the Uncanny Valley? An Empirical Study of User Experience in Human-Chatbot Interaction. *Human Technology*, 15(1), 30-54. <https://doi.org/10.17011/ht/urn.201902201607>
- Söderlund, M. & Oikarinen, E.-L. (2021). Service Encounters with Virtual Agents: An Examination of Perceived Humanness as a Source of Customer Satisfaction. *European Journal of Marketing*, 55(13), 94-121. <https://doi.org/10.1108/EJM-09-2019-0748>
- Song, M., Du, J., Xing, X. & Mou, J. (2022). Should the Chatbot "Save Itself" or "Be Helped by Others"? The Influence of Service Recovery Types on Consumer Perceptions of Recovery Satisfaction. *Electronic Commerce Research and Applications*, 55, 101199. <https://doi.org/10.1016/j.elerap.2022.101199>
- Song, S. W. & Shin, M. (2024). Uncanny Valley Effects on Chatbot Trust, Purchase Intention, and Adoption Intention in the Context of E-Commerce: The Moderating Role of Avatar

- Familiarity. *International Journal of Human-Computer Interaction*, 40(2), 441-456. <https://doi.org/10.1080/10447318.2022.2121038>
- Statista (2023). Wenn Sie einen Online-Kundenservice nutzen, wie zuversichtlich sind Sie, wenn überhaupt, dass Sie den Unterschied zwischen einem menschlichen Kundenberater und einem KI-Chatbot erkennen? <https://de.statista.com/statistik/daten/studie/1461828/umfrage/kundenberatung-durch-kuenstliche-intelligenz-chatbots-laender-vergleich/>, 16.04.2025
- Statista (2025). Weltweites Marktvolumen von Chatbots im Jahr 2022 und Prognose bis 2032. <https://de.statista.com/statistik/daten/studie/1373729/umfrage/weltweites-marktvolumen-chatbots/>, 16.04.2025
- Statistisches Bundesamt. (2025). Bevölkerung nach dem Gebietsstand und Durchschnittsalter 1990 bis 2023. <https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Bevoelkerungsstand/Tabellen/bevoelkerungsstand-gebietsstand-werte.html>, 16.04.2025
- Straits Research. (2024). Chatbot-Marktgröße, Nachfrage, Wachstum, Bericht, Trends, Prognose bis 2032. <https://straitsresearch.com/de/report/chatbot-market#:~:text=Markt%C3%BCbersicht,von%2023%2C1%20%25%20wachsen.>, 16.04.2025
- Stucki, T., D'Onofrio, S. & Portmann, E. (2018). Chatbot – Der digitale Helfer im Unternehmen: Praxisbeispiele der Schweizerischen Post. *HMD Praxis der Wirtschaftsinformatik*, 55(4), 725-747. <https://doi.org/10.1365/s40702-018-0424-8>
- Stucki, T., D'Onofrio, S. & Portmann, E. (2020). Chatbots gestalten mit Praxisbeispielen der Schweizerischen Post. Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-28586-9>
- Sun, Y., Chen, J. & Sundar, S. S. (2024). Chatbot Ads With a Human Touch: A Test of Anthropomorphism, Interactivity, and Narrativity. *Journal of Business Research*, 172, 114403. <https://doi.org/10.1016/j.jbusres.2023.114403>
- Szymanski, D. M. & Henard, D. H. (2001). Customer Satisfaction: A Meta-Analysis of the Empirical Evidence. *Journal of the Academy of Marketing Science*, 29(1), 16-35. <https://doi.org/10.1177/0092070301291002>
- Thomaz, F., Salge, C., Karahanna, E. & Hulland, J. (2020). Learning From the Dark Web: Leveraging Conversational Agents in the Era of Hyper-Privacy to Enhance Marketing. *Journal of the Academy of Marketing Science*, 48(1), 43-63. <https://doi.org/10.1007/s11747-019-00704-3>
- Toader, D.-C., Boca, G., Toader, R., Măcelaru, M., Toader, C., Ighian, D. & Rădulescu, A. T. (2020). The Effect of Social Presence and Chatbot Errors on Trust. *Sustainability*, 12(1), 1-24. <https://doi.org/10.3390/su12010256>
- Tuzovic, S. & Paluch, S. (2018). Conversational Commerce - A New Era for Service Business Development. In M. Bruhn & K. Hadwich (Hrsg.), *Service Business Development: Methoden - Erlösmodelle - Marketinginstrumente*. (Band 1, S. 81-100). Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-22426-4>
- Tversky, A. & Kahneman, D. (1974). Judgment Under Uncertainty: Heuristics and Biases. *Science*, 185(4157), 1124-1131. <https://doi.org/10.1126/science.185.4157.1124>

- Ukpabi, D. C., Aslam, B. & Karjaluoto, H. (2019). Chatbot Adoption in Tourism Services: A Conceptual Exploration. In S. Ivanov & C. Webster (Hrsg.), *Robots, Artificial Intelligence, and Service Automation in Travel, Tourism and Hospitality* (Bd. 7, S. 105-121). Emerald Publishing Limited. <https://doi.org/10.1108/978-1-78756-687-320191006>
- Utz, S., Kerkhof, P. & van den Bos, J. (2012). Consumers Rule: How Consumer Reviews Influence Perceived Trustworthiness of Online Stores. *Electronic Commerce Research and Applications*, 11(1), 49-58. <https://doi.org/10.1016/j.elerap.2011.07.010>
- Van der Goot, M. J., Koubayová, N. & Van Reijmersdal, E. A. (2024). Understanding Users' Responses to Disclosed vs. Undisclosed Customer Service Chatbots: A Mixed Methods Study. *AI & Society*, 39(6), 2947-2960. <https://doi.org/10.1007/s00146-023-01818-7>
- Van Doorn, M. & Duivestijn, S. (2016). The bot Effect: 'Friending Your Brand': Machine Intelligence Applied Innovation Exchange, Report 2 of 4, 1-60.
- Venkatesh, V., Morris, M. G., Davis, G. B. & Davis, F. D. (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*, 27(3), 425-478. <https://doi.org/10.2307/30036540>
- Ventre, I. & Kolbe, D. (2020). The Impact of Perceived Usefulness of Online Reviews, Trust and Perceived Risk on Online Purchase Intention in Emerging Markets: A Mexican Perspective. *Journal of International Consumer Marketing*, 32(4), 287-299. <https://doi.org/10.1080/08961530.2020.1712293>
- Verhagen, T., van Nes, J., Feldberg, F. & van Dolen, W. (2014). Virtual Customer Service Agents: Using Social Presence and Personalization to Shape Online Service Encounters. *Journal of Computer-Mediated Communication*, 19(3), 529-545. <https://doi.org/10.1111/jcc4.12066>
- Völkle, C. & Planing, P. (2019). Digital Automation of Customer Contact Processes – An Empirical Research on Customer Acceptance of Different Chatbot Use-Cases. In A. Lochmahr, P. Müller, P. Planing & T. Popović (Hrsg.), *Digitalen Wandel gestalten* (S. 217-229). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-24651-8_6.4
- Voss, G. B., Parasuraman, A. & Grewal, D. (1998). The Roles of Price, Performance, and Expectations in Determining Satisfaction in Service Exchanges. *Journal of Marketing*, 62(4), 46-61. <https://doi.org/10.2307/1252286>
- Waytz, A., Heafner, J. & Epley, N. (2014). The Mind in the Machine: Anthropomorphism Increases Trust in an Autonomous Vehicle. *Journal of Experimental Social Psychology*, 52, 113-117. <https://doi.org/10.1016/j.jesp.2014.01.005>
- Wei, P.-S. & Lu, H.-P. (2013). An Examination of the Celebrity Endorsements and Online Customer Reviews Influence Female Consumers' Shopping Behavior. *Computers in Human Behavior*, 29(1), 193-201. <https://doi.org/10.1016/j.chb.2012.08.005>
- Weiber, R. & Sarstedt, M. (2021). *Strukturgleichungsmodellierung*. Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-32660-9>
- Weiner, B. (1985a). 'Spontaneous' Causal Thinking. *Psychological Bulletin*, 97(1), 74-84. <https://doi.org/10.1037/0033-2909.97.1.74>
- Weiner, B. (1985b). An Attributional Theory of Achievement Motivation and Emotion. *Psychological Review*, 92(4), 548-573. <https://doi.org/10.1037/0033-295X.92.4.548>

- Weiner, B. (1994). *Motivationspsychologie* (3. Auflage). Beltz Psychologie-VerlagsUnion.
- Weizenbaum, J. (1966). Eliza – a Computer Program for the Study of Natural Language Communication Between Man and Machine. *Communications of the ACM*, 9(1), 36-45. <https://doi.org/10.1145/365153.365168>
- Wernsmann, C. (2022). Automatisierung dialoggeführter Prozesse mit Chatbots. In C. Feldmann (Hrsg.), *Praxishandbuch Robotic Process Automation (RPA)* (S. 359-382). Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-38379-4>
- Wintersberger, P., Klotz, T. & Riener, A. (2020). Tell Me More: Transparency and Time-Fillers to Optimize Chatbots' Waiting Time Experience. In *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society* (S. 1-6). <https://doi.org/10.1145/3419249.3420170>
- Wirtz, J., Patterson, P. G., Kunz, W. H., Gruber, T., Lu, V. N., Paluch, S. & Martins, A. (2018). Brave New World: Service Robots in the Frontline. *Journal of Service Management*, 29(5), 907-931. <https://doi.org/10.1108/JOSM-04-2018-0119>
- Wu, C., Che, H., Chan, T. Y. & Lu, X. (2015). The Economic Value of Online Reviews. *Marketing Science*, 34(5), 739-754. <https://doi.org/10.1287/mksc.2015.0926>
- Wuenderlich, N. V. & Paluch, S. (2017). A Nice and Friendly Chat with a Bot: User Perceptions of AI-Based Service Agents. In *Proceedings of the International Conference on Information Systems (ICIS)*. <http://aisel.aisnet.org/icis2017/ServiceScience/Presentations/11/>
- Xie, Y., Liang, C., Zhou, P. & Jiang, L. (2024). Exploring the Influence Mechanism of Chatbot-Expressed Humor on Service Satisfaction in Online Customer Service. *Journal of Retailing and Consumer Services*, 76, 103599. <https://doi.org/10.1016/j.jretconser.2023.103599>
- Ye, Q., Law, R. & Gu, B. (2009). The Impact of Online User Reviews on Hotel Room Sales. *International Journal of Hospitality Management*, 28(1), 180-182. <https://doi.org/10.1016/j.ijhm.2008.06.011>
- Yen, C. & Chiang, M.-C. (2021). Trust Me, if You Can: A Study on the Factors That Influence Consumers' Purchase Intention Triggered by Chatbots Based on Brain Image Evidence and Self-Reported Assessments. *Behaviour & Information Technology*, 40(11), 1177-1194. <https://doi.org/10.1080/0144929X.2020.1743362>
- Yu, R., Feng, J., Wang, K., Yang, L. & Feng, J. (2024). Who Are You Talking To? How Chatbot Identity Disclosure Affects Service Satisfaction. *Journal of Travel & Tourism Marketing*, 41(8), 1090-1106. <https://doi.org/10.1080/10548408.2024.2369755>
- Yuan, J., Peng, X., Liu, Y & Wang, Q. (2025). 'Don't Look At Me! ': The Role of Avatars' Presentation Style and Gaze Direction in Social Chatbot Design. *Computers in Human Behavior*, 164, 108501. <https://doi.org/10.1016/j.chb.2024.108501>
- Zarouali, B., van den Broeck, E., Walrave, M. & Poels, K. (2018). Predicting Consumer Responses to a Chatbot on Facebook. *Cyberpsychology, Behavior and Social Networking*, 21(8), 491-497. <https://doi.org/10.1089/cyber.2017.0518>
- Zhang, J., Lu, X., Zheng, W. & Wang, X. (2024). It's Better Than Nothing: The Influence of Service Failures on User Reusage Intention in AI Chatbot. *Electronic Commerce Research and Applications*, 67, 101421. <https://doi.org/10.1016/j.elerap.2024.101421>

- Zhao, X., Wang, L., Guo, X. & Law, R. (2015). The Influence of Online Reviews to Online Hotel Booking Intentions. *International Journal of Contemporary Hospitality Management*, 27(6), 1343-1364. <https://doi.org/10.1108/IJCHM-12-2013-0542>
- Zheng, L. (2021). The Classification of Online Consumer Reviews: A Systematic Literature Review and Integrative Framework. *Journal of Business Research*, 135, 226-251. <https://doi.org/10.1016/j.jbusres.2021.06.038>
- Zhu, F. & Zhang, X. (2010). Impact of Online Consumer Reviews on Sales: The Moderating Role of Product and Consumer Characteristics. *Journal of Marketing*, 74(2), 133-148. <https://doi.org/10.1509/jm.74.2.133>

CURRICULUM VITAE

Persönliche Informationen

Name	Sebastian Düppre
E-Mail	sebastian.dueppre@rptu.de
Telefon	+49 631 205 5248

Schulbildung und akademischer Werdegang

10/2019-heute	Wissenschaftlicher Mitarbeiter und Doktorand Rheinland-Pfälzische Technische Universität Kaiserslautern-Landau Lehrstuhl für Marketing, Prof. Dr. Stefan Roth
10/2016-02/2019	Masterstudium Wirtschaftsingenieurwesen (Fachrichtung Maschinenbau), Abschluss: M. Sc. Technische Universität Kaiserslautern
10/2012-01/2017	Bachelorstudium Wirtschaftsingenieurwesen (Fachrichtung Maschinenbau), Abschluss: B. Sc. Technische Universität Kaiserslautern
08/2004-06/2012	Abitur Gymnasium am Krebsberg

Beruflicher Werdegang

10/2019-heute	Wissenschaftlicher Mitarbeiter und Doktorand Rheinland-Pfälzische Technische Universität Kaiserslautern-Landau Lehrstuhl für Marketing, Prof. Dr. Stefan Roth
03/2019-09.2019	Vertriebsmitarbeiter VENSYS Energy AG