

Lehrstuhl für Wirtschaftsinformatik, insb. IT-Sicherheit und Datenschutz
Fakultät III – Universität Siegen
Prof. Dr. Gunnar Stevens

BACHELORARBEIT

Alexa außer Kontrolle - Aneignung von Sprachassistenten bei
steigender Aufgabenkomplexität als nutzerzentrierte
Gestaltungsherausforderung

Siegen den 25.05.2021

Erstkorrektor: Prof. Dr. Gunnar Stevens
Zweitkorrektor: Univ.-Prof. Dr. Volker Wulf

Name, Vorname:	Lobit, Jan
Anschrift:	██████████ ██████████
Studiengang:	Wirtschaftsinformatik
Tel.Nr.:	██████████
E-Mail:	████████████████████
Matr.-Nr.:	██████
Fachsemester:	█

„In dieser Arbeit wird aus Gründen der besseren Lesbarkeit das generische Maskulinum verwendet. Weibliche und anderweitige Geschlechteridentitäten werden dabei ausdrücklich mitgemeint, soweit es für die Aussage erforderlich ist.“

Tabellenverzeichnis

1	Übersicht der wichtigsten Ergebnisse, Erkenntnisse und Thesen aus der Literatur	13
2	Die Tagesaufgaben und deren Schwierigkeitsgrad	16
3	Übersicht der Probanden	18
4	Gegenüberstellung der Ergebnisse der vorliegenden Studie mit den Ergebnissen und Erkenntnissen aus der Literatur	29

Inhaltsverzeichnis

1	Einleitung	1
2	Forschungsstand	3
2.1	Aneignung von Sprachassistenten	3
2.2	Chatbots und Dialogkomplexität	9
2.3	Aufgabenkomplexität und Kooperation mit Sprachassistenten	11
3	Methode	14
3.1	Einführung und Probanden	14
3.2	Auftaktinterview	14
3.3	Tagebuch	14
3.4	Postinterview	17
3.5	Qualitative Inhaltsanalyse	17
3.6	Deskriptive Statistik	17
4	Ergebnisse	19
4.1	Erwartungen und Einrichtung	19
4.2	Erste Berührungspunkte mit der komplexen Interaktion	20
4.3	Aneignung der Befehlsform	20
4.4	Standort und Nutzungsbereich	21
4.5	Zufriedenheit mit Qualität der Antworten und Ergebnisse	21
4.6	Nutzungsveränderung und Nutzungserfahrung durch Steigerung der Aufgabenkomplexität	22
4.7	Zufriedenheit mit der allgemeinen Nutzung	23
4.8	Einfluss der Usability in der Zusammenarbeit	24
5	Diskussion	27
6	Fazit	30
	Literaturverzeichnis	33

1 Einleitung

Heutzutage sind digitale Medien aus dem Alltag kaum wegzudenken. Von der Smart-Watch über Smart-TV, bis hin zu intelligenten Haushaltsgeräten, haben sich einige Geräte etabliert. Mit Abstand am häufigsten verbreitet ist das Smartphone. Mittlerweile besitzt fast jeder zweite Mensch ein Mobilgerät[38]. Während Smartphones schon längere Zeit in den Alltag integriert wurden, sind intelligente Sprachassistenten erst in den letzten Jahren im privaten Umfeld populär geworden. Da viele Smartphones standardmäßig mit einem integrierten intelligenten Assistenten auf den Markt gebracht werden, beeinflussten diese auch die Verbreitung Letzterer. Sprachassistenten haben sich über die Zeit in den unterschiedlichsten Bereichen und Endgeräten integriert[5]. In vielen modernen Geräten, vor allem in den anfangs genannten, aber auch in modernen Fahrzeugen, sind Sprachassistenten mittlerweile vertreten. Am häufigsten sind sie allerdings neben den Smartphones in den speziell entworfenen intelligenten Lautsprechern vorhanden[5]. Diese werden unter anderem von Tech-Riesen Amazon oder Google vermarktet. Der Sprachassistent von Amazon, „Alexa“, wird beispielsweise als Echo Dot verkauft. Dieser ist ein kleiner Lautsprecher, ohne weitere Bedienschnittstelle als die der Sprache. Anders ist der Echo Show. Im Vergleich zum Echo Dot, ist Echo Show mit einem Display ausgestattet und ist dementsprechend größer sowie vielfältiger in der Nutzung. Beide, beziehungsweise die meisten Sprachassistenten, bieten eine Schnittstelle im „Smart Home“. In diesem Kontext werden die kompatiblen, intelligenten Geräte folglich vom Sprachassistenten angesteuert. Um eine gute Erreichbarkeit und Nutzung zu erhalten, hat der Sprachassistent meist einen festen Platz im Haushalt[5]. Das Besondere, die Interaktion, findet hauptsächlich über Sprachbefehle statt und werden auch mit menschlicher Stimme beantwortet. Diese Art der Kommunikation mit einer Maschine kann zu Vertrauen führen, sogleich kann eine zu hohe Erwartungshaltung durch die folglich vermenschlichte Maschine entwickelt werden [18, 17, 30]. Es ist demnach wissenswert, wie und ob sich diese Erwartungshaltung auch bestätigt oder inwieweit sich Nutzer den Bedingungen und Möglichkeiten bewusst sind, da es zur Ernüchterung oder Enttäuschung führen kann[30]. Die Nutzung von Sprachassistenten zu optimieren sowie neuen Nutzern einen einfachen frustfreien Start zu ermöglichen, gehören zu wichtigen Bereichen der aktuellen Forschung. Dazu wird unter anderem die Nutzung über verschiedene Zeiträume untersucht oder durch Online-Umfragen erfasst. Dabei stellt sich häufig die schlechte Gebrauchstauglichkeit bei der Nutzung von komplexen Anforderungen an den Sprachassistenten als ein größeres Problem heraus[17]. Zwar sind die Applikationen, bei Amazons Alexa „Skills“ genannt, welche von Drittanbietern zur Verfügung gestellt werden können, durch eine oft schlechte

Umsetzung ein Grund für Frust beim Nutzer[30]. Allerdings sind auch andere Faktoren verantwortlich für schlechte Erfahrungen, denn die Nutzung beläuft sich häufig auf einfach zu bedienende Skills, wobei sie bereits gut angepasst wurden und Fehler nur noch sehr geringfügig vorkommen[36, 30]. Am häufigsten werden dabei Radio und Musik Skills genutzt, aber auch im Bereich des Smart-Home werden oft Befehle zum Licht ein- und ausschalten verwendet[5]. Neben solchen Skills werden ebenfalls sehr viele Informationen abgefragt[5]. Kompliziertere Aufgaben an den Sprachassistenten werden statistisch gesehen weniger genutzt, aus den genannten Gründen, sowie der Schwierigkeit, eine komplexe Interaktion zu gestalten[17, 30].

Im Rahmen dieser Bachelorarbeit wurde in einer dreiwöchigen Tagebuchstudie, mit 15 Probanden, auf solche Aspekte eingegangen. Es wird demnach der Frage nachgegangen, inwieweit Sprachassistenten einen Mehrwert in der alltäglichen Nutzung darstellen. Die Nutzer sollten im Laufe der Studie Aufgaben mit dem Sprachassistenten, mit steigender Komplexität, lösen. Dabei wird darauf geachtet, wie die Nutzer mit den aufkommenden Problemen umgegangen sind und auch wie sich die allgemeine Nutzung über die Laufzeit verändert hat und inwieweit die Erwartungen nicht erfüllt wurden. Im Fokus steht besonders, im Hinblick auf die steigende Aufgabenkomplexität, die Verschiebung des Interaktionskontextes, denn bei den einfacheren Aufgaben ist die Interaktion darauf ausgelegt, dass der Sprachassistent die Befehle lediglich ausführt, ohne dabei den Nutzer speziell zu integrieren. Bei komplexeren Aufgaben fordert es eine Zusammenarbeit von Sprachassistent und Nutzer, um zum Erfolg zu kommen. Dadurch stellt sich die Frage wie sich die Nutzung während des verschobenen Interaktionskontextes verändert und welche Unterschiede währenddessen bei den Nutzern, innerhalb der Interaktion und Zusammenarbeit mit Alexa, vorkommen. Diesbezüglich geht es um komplexe Dialoge, die nötig sein könnten und nicht um einzeilige Befehle, um Alexa einfache Anweisungen ausführen zu lassen. Darum wird auch die Dialogkomplexität im Zusammenhang der Zusammenarbeit betrachtet und inwiefern andere Arten von intelligenten Dialogsystemen mit komplexen Aufgaben umgehen können. Mit der steigenden Aufgabenkomplexität, und auch in Kombination mit einem integrierten Bildschirm, ist zudem darauf zu achten, wie ein Mehrwert in der Zusammenarbeit mit dem Sprachassistenten entsteht oder ob die Interaktion stattdessen scheitert und wo dessen Ursache liegt.

2 Forschungsstand

In den letzten Jahren ist das öffentliche Interesse an Sprachassistenten rasant gestiegen. Demzufolge sind zunehmend mehr Endgeräte im privaten Besitz. Dadurch steigt auch der Stellenwert von Studien und Designoptimierung. Um den Nutzen der Sprachassistenten sowie die Gebrauchstauglichkeit zu ermitteln und zu erhöhen, gewinnen besonders Studien zur Aneignung und Nutzung gleichermaßen an Bedeutung. Sehr aufschlussreich können Aneignungsstudien sein. Solche wurden bereits einige mit diversen Sprachassistenten durchgeführt. Diese dienen zur Untersuchung der Nutzung mit Sprachassistenten und der aufkommenden Probleme. Außerdem wird auf Gründe des Scheiterns geachtet und inwiefern Lösungsversuche von Nutzern dem Erfolg positiv beigetragen haben. Dadurch wird es möglich, auch Einblicke in die tatsächliche Nutzung des Sprachassistenten zu bekommen, welche von jener abweicht, wie es ursprünglich bei der Entwicklung vorgesehen war[30]. Dabei spielt die Aufgabenkomplexität in der Nutzung, Interaktion und Zusammenarbeit mit Sprachassistenten eine entscheidende Rolle. Des Weiteren sind Studien zur Dialogkomplexität, da diese einen wichtigen Aspekt in der Zusammenarbeit darstellt[15], zu berücksichtigen.

2.1 Aneignung von Sprachassistenten

Es haben sich bereits einige Studien mit der Aneignung von Sprachassistenten befasst. Diese fanden unter verschiedenen Bedingungen und Herausforderungen statt, um verschiedene Aspekte untersuchen zu können. Bei der Aneignung soll insbesondere die Benutzerfreundlichkeit (Usability) und Nutzungserfahrung (User-Experience) in Fokus genommen werden sowie die Nutzung im Allgemeinen. Für einen effizienten und effektiven Gebrauch von Sprachassistenten sowie gelingende Interaktion sind diese entscheidend. Die Usability hängt häufig davon ab, inwiefern das Front-End, also die Benutzeroberfläche bzw. Schnittstelle zwischen Computer und Maschine, umgesetzt ist, aber auch von der Software und des nutzerzentrierten Designs[26]. Bezüglich der Usability wird bei Aneignungsstudien darauf zu achten sein, ob und wie leicht die Interaktion mit dem Sprachassistenten gelingt, wie häufig Fehler auftauchen und wie der Umgang mit solchen stattfindet. Aber auch die Zufriedenheit zur Nutzung und des Designs fließen mit ein[26]. Bei der Entwicklung und Weiterentwicklung ist es auch wichtig, die User-Experience zu berücksichtigen. Diese zeigt auf, wie sich die Nutzung auf den Nutzer auswirkt, wie die Eindrücke sowie die Erlebnisse mit der Interaktion sind. Die Nutzungserfahrung zeigt letztlich, worauf bei der Entwicklung zu achten ist, wenn es nach den Bedürfnissen des Nutzers gehen soll[25]. Dementsprechend haben User-Experience und Usability einen hohen Stellenwert in

der Forschung. Ein guter und umfangreicher Einblick in diese Aspekte kann mit verschiedenen Methoden erreicht werden. Folglich werden diverse Vorgehensweisen, welche die Anwendung von Nutzern untersuchen und analysieren, umgesetzt. Zu den Methoden gehören unter anderem Umfragen. Diese ermöglichen es, eine standardisierte Datengrundlage zu erschaffen. Teilnehmer können durch kurze Antworten schnell auf Fragen reagieren. So werden in kürzester Zeit viele Informationen erhoben. Durch Online-Umfragen können zudem ohne großen Aufwand eine Vielzahl an Menschen aus der Zielgruppe gefunden werden. Diese Methode haben Pyae und Joelsson mithilfe einer internetbasierten Umfrage zur Usability und User-Experience innerhalb eines sozialen Netzwerkes zum Thema Smart Home durchgeführt[34]. Der Fokus der Studie lag auch auf der Untersuchung der Herausforderungen, die mit der Usability während der Interaktion aufkamen[34]. Die Ergebnisse lassen einige Schlüsse über Einflüsse einer reibungslosen Nutzung daraus ziehen. Dazu wurden die Usability betreffenden Punkte zusammengefasst und mit einer qualitativen Analyse ausgewertet. Die Ergebnisse der Umfrage zeigen ein häufig wiederkehrendes Problem. Schwierigkeiten mit dem Lösen von Aufgaben gab es häufiger, sodass Befehle mehrfach ausgesprochen werden mussten. Auch andere Interaktionsprobleme, wie Verständnisprobleme, verursacht durch dem sprachlichen Input, wurden aufgefasst. Jedenfalls zeigte sich trotz unterschiedlicher Probleme, ein recht positives Gesamtbild. Mit Hilfe der Umfrage lässt sich als Fazit der Studie sagen, dass die Teilnehmer trotz der genannten und anderen Schwierigkeiten zufrieden mit dem Sprachassistenten sind. Demnach sei die Usability für die genutzten Zwecke ausreichend, das Gerät sehr leicht zu bedienen und es verspricht einen großen Unterhaltungswert[34]. Andere Studien zeigen auf, dass Probleme mit der Interaktion durchaus zur Enttäuschung führen können oder Verständnisprobleme zu genervten Nutzern führen kann[7], was wiederum dazu führt, dass der Sprachassistent weniger bis gar nicht mehr genutzt wird[36, 17]. Cho befasste sich dazu mit verschiedenen Modellen, um die Handhabung mit Sprachassistenten zu optimieren, indem er Nutzer auf die Art und Weise der Interaktion hin beobachtet hat[7]. Unter anderem sollen einfache Sätze und Minderung von Umgangssprache zum Ziel führen und eine zufriedenstellende Interaktion zu ermöglichen[7]. Allerdings hilft während der Nutzung die sich einstellende Routine[26, 30, 17]. Um eine funktionierende, routinierte Nutzung zu erschaffen, benötigt es viel Eigenleistung. Ein Nutzer muss sich also das meiste im Umgang mit dem Sprachassistenten selbst aneignen. Vor allem die richtige Befehlsform muss erlernt werden[26]. Durch die Routinen werden über längere Zeit immer mehr Skills oder andere Arten von Interaktion nicht mehr genutzt[17, 30]. Zusätzlich werden auf dem Weg zur funktionierenden Routine Fehler begangen oder es treten Probleme auf, welche somit in der Zukunft vermieden werden können[26]. Das führt auf der einen Seite zu einer weniger enttäuschenden User-Experience, auf der anderen Seite

werden dadurch die Interaktionen nicht unbedingt weiterentwickelt. Schlussfolgernd ist die Usability, wenn es um Interaktionen geht, die tiefergehend sind, weniger vorhanden[26]. Umso wichtiger ist es, die User-Experience unter schwierigeren Bedingungen zu betrachten, bei denen der Nutzer auf eigene Lösungen kommen muss, um daraus Designvorschläge machen zu können, die bei der zukünftigen Entwicklung helfen können. Aber auch gewöhnliche oder häufige Probleme und Fehler, die im Laufe der Interaktion aufkommen, sollten gesammelt werden und letztendlich dazu beitragen, dass sie in Zukunft minimiert werden können. Zur direkten Verbesserung der User-Experience und dem Sprachinterface befassten sich daher auch einige Studien, wie solche von Reeves et al.[35] und Moore et al.[23]. Moore et al. haben Vorschläge zur Verbesserung des Sprachinterface genannt. Dabei geht es überwiegend um die Wahrnehmung des Sprachassistenten, der sich momentan in der Interaktion mit einem Menschen noch zu sehr von diesem unterscheidet[23]. Um die User-Experience diesbezüglich zu optimieren, wären manchmal einfache, menschlichere Antworten, bzw. Einwüfe wie ein „Oh“ von der Maschine nötig[23], was sich mit anderen Studien deckt[30, 33]. Hinzu kommt, dass sich einige Nutzer einen persönlicheren Gesprächspartner vom Sprachassistenten wünschen[33]. Pollmann et al. haben sich mit der User-Experience von Sprachassistenten im Vergleich mit einem Roboter befasst[31]. Dabei wurden Einflüsse und Eindrücke bei der Nutzung von Unterhaltungs-Skills gesammelt[31]. Die Ergebnisse zeigten auf, dass der Unterhaltungswert bei dem modifizierten Roboter höher als beim Sprachassistenten Echo Show war. Die Nutzer hatten mit dem Sprachassistenten weniger Spaß und eine schlechtere Erfahrung beim Spiel. Andere Studien bestätigen, dass dem Sprachassistenten menschlichere Charakterzüge zur besseren User-Experience fehlen könnten [33, 23, 31].

Die Betrachtung des einzelnen Probanden und einer individuellen Befragung, verknüpft mit standardisierten Interviews, trägt der Analyse und Erfassung gestaltungsfördernder Daten ideal bei. Eine solche Methode haben Pins et al. genutzt[30]. Allerdings zeigte sich auch dadurch, dass die Teilnehmer der Studie größtenteils Langzeitnutzer waren, dass die Probanden bereits mit den Stärken und Schwächen der Sprachassistenten überwiegend vertraut waren. Mit der Interaktion waren die Teilnehmer, ähnlich wie in der erwähnten Studie von Pyae und Joelsson, insgesamt recht zufrieden[34, 30]. Erwähnenswert sind die Entwicklungen, welche zur Zufriedenheit geführt haben. Präzisere und knappere Befehle haben demnach zu einem positiven Einfluss auf das Verständnis des Sprachassistenten geführt[30]. Aber auch die sich mit der Zeit eingependelte Routine, wie oben genannt, hat dazu beigetragen, dass sich weniger Fehler eingeschlichen haben[30]. Auch die Möglichkeit dem Sprachassistenten durch die sprachliche Interaktion einen menschlichen Charakter zuzuschreiben, ist durch die knappen Dialoge oder Befehle nicht, beziehungsweise

nur geringfügig eingetroffen[30]. Während der Nutzung ist auch aufgefallen, dass sich der Sprachassistent durch manche Probleme nicht allzu flexibel zeigt. Besonders auffallend sind die eher seltenen, komplexen Interaktionen. Schlechtes Verständnis oder nicht zufriedenstellende Antworten, seitens des Sprachassistenten, führten zur Minderung der komplexen Interaktionen. Dahingehend beeinflusst die Art und Weise, wie die Konversationen gehalten werden, ob dem Gerät des Öfteren menschlichere Charakteristiken fehlen, sowie bei unklaren oder komplexen Befehlen Rückfragen zu stellen[30]. In manchen Situationen sei es sinnvoll eine weitere Schnittstelle, wie einen Bildschirm oder ein Smartphone, zur Unterstützung zu haben, wenn die sprachliche Komponente nicht zu dem gewünschten Ergebnis führt[30].

Die Küche stellt in dem Fall ein besonderes Umfeld dar. Da es in der Küche häufig dazu kommt, dass man keine freien Hände hat, könnte dieser Aspekt hilfreich sein, besonders da Nutzer Sprachassistenten gerne aus diesem Grund verwenden[19]. In der Küche wurde dabei mit dem Skill von Chefkoch versucht komplexe Interaktionen zu führen, unter anderem Rezepte zu suchen und mit diesen zu kochen. Allerdings hat sich in dem Kontext ein Mangel an Usability des Skills gezeigt[30]. Dennoch kann dort bei Aktivitäten, wie beim Kochen, Backen oder auch beim Erstellen von Einkaufslisten ein Sprachassistent behilflich sein[29]. Speziell in der Küche könnte eine weitere Interaktionsoberfläche, wie ein bereits erwähnter Bildschirm dienlich sein, um eine Visualisierung zu ermöglichen. Diese könnte dem Erlebnis positiv beitragen und die Usability zu erhöhen[29, 30]. Inwieweit ein Bildschirm dazu beisteuert, eine bessere Interaktion zu schaffen, ist auch Bestandteil der im Rahmen dieser Arbeit durchgeführten Studie. Unter anderem wurde dies während Tätigkeiten innerhalb der Küche, aber auch im wesentlichen Kontext der Nutzung im Allgemeinen ermittelt. Des Weiteren sind Fehler und Interaktionsschwierigkeiten im Bereich der komplexen Nutzung zu erfassen. Im Hinblick auf diese und auch komplizierte Aufgaben sowie tiefergehenden Konversationen, sind Fehler häufiger entstanden, wenn der Sprachassistent nicht der Logik des Nutzers gefolgt ist[30]. Unter anderem mussten sich die Teilnehmer an solche und ähnliche Schwächen dementsprechend so anpassen, dass diese nicht zu Problemen führten[30]. Anpassung ist nicht in jeglicher Form förderlich für eine verbesserte User-Experience. Fehler werden zwar vermieden, aber die Nutzung, wie sie eventuell gewünscht wird, bleibt damit aus[30]. Folglich scheint sich etwa die Nutzungshäufigkeit zu verringern[30]. Mit der unter anderem auch zeitlich bedingten Abnahme der allgemeinen Nutzung haben sich Bentley et al.[4] oder auch Sciuto et al.[36] befasst. Bentley et al. haben dabei nicht auf die Probleme oder Ursachen geachtet, sondern lediglich Log-Daten ausgewertet, um die tägliche Nutzung und das Nutzungsverhalten zu erfassen[4]. Gezeigt hat sich dadurch eine Abnahme diverser Interaktionen und Skills. Dagegen haben Sciuto

et al. ähnlich wie Pins et al.[30] eine interviewgestützte Studie mit den Log-Daten der Sprachassistenten verbunden[36]. So konnte individuell festgehalten werden, welche Anwendungen und Nutzungsbereiche mit dem Sprachassistenten genutzt wurden und welche Probleme zu Fehlern und Änderungen im Nutzungsverhalten geführt haben. Auch die Häufigkeit und ab- und Zunahme der Nutzung über die Zeit konnte erfasst werden. Auffällig dabei war wie erwähnt die abnehmende Nutzung, erkennbar an der ebenfalls abnehmenden Anzahl an Skills über die Zeit sowie eine einstellende Routine, wodurch weniger ausprobiert wird[36]. In der Studie von Sciuto et al. wurde auch der Einfluss von Kindern der Probanden mit der Nutzung von Alexa beobachtet[36]. Dabei konnten die Kinder sinnvolle Informationen, wie Wörter buchstabieren oder mathematische Rechnungen, erledigen lassen[36]. Zusätzlich wurden die Kinder mit Alexa an die Interaktion und die Art der Kommunikation gewöhnt und konnten in späteren Nutzungen mit technologischen Geräten schnelle Erfolge bezüglich der Interaktion vorzeigen[36]. Wobei Kinder dem Sprachassistenten gerne längere Antworten geben[2, 36]. Das kann dazu führen, dass der Sprachassistent die Antwort in kleinere Fragmente einteilt und somit nicht der Logik von Kindern folgen kann[22]. Sofern die Interaktion allerdings verstanden und erfolgreich genutzt werden konnte, wurde auch die Möglichkeit missbraucht den Sprachassistenten gegen die Eltern zu nutzen, indem zum Beispiel Lieder gegen den Willen der Eltern gewechselt wurden[36]. Grundsätzlich ist ein Sprachassistent im Rahmen der familiären Nutzung dennoch nicht fehl am Platz. Denn sowohl Kinder als auch deren Eltern nutzen den Sprachassistenten auf ihre Art und Weise, während das Gerät selbst als ein neutraler Vermittler zwischen ihnen agiert[3]. So können zusammen und während der Interaktion der Kinder erzieherische Möglichkeiten in die Nutzung mit einfließen[3]. Durch die Interaktion mit Sprache wird die Erziehung in verschiedenen Bereichen erweitert[3]. Jedoch birgt sich neben den Vorteilen der sprachlichen Kommunikation auch ein Nachteil. Durch diese Integration eines Sprachassistenten in der zwischenmenschlichen Beziehung zeigt sich die vermenschlichende Komponente, die auch dazu führen kann, dass die Nutzer zu viel von dem Sprachassistenten erwarten[30], was unter anderem durch eine Personifizierung hervorgerufen werden könnte[17]. Wobei eine Personifizierung des Sprachassistenten nicht häufig vorkommt, was teilweise an den zu einfachen Antworten oder eher oberflächlichen Interaktionen liegt[17]. Eine Personifizierung kann eine Zufriedenheit hervorrufen, welche trotz mangelnder Leistung und signifikanter technischer Probleme nicht bestehen dürfte[33].

Lopatovska et al. führten eine Tagebuchstudie zum Verständnis der Nutzerperspektiven und zur Beobachtung der Personifizierung[17] durch. Die Teilnehmer haben vier Tage über ein online strukturiertes Tagebuch ihre Nutzung und Interaktion mit Alexa beschrieben. Eine Erfolgsquote zur Nutzung und eine Einschätzung zum

Gebrauch personifizierenden Worten sollte erfasst werden. Purington et al. befassten sich ebenfalls mit einer Personifizierung von Sprachassistenten[33]. Dafür analysierten sie Kundenrezensionen inklusive Erfahrungsberichten zu Amazons Alexa. Demnach hängt der Eindruck zur Umgänglichkeit von der Art der Nutzung ab und wie das Gerät von den Nutzern genannt wird. Dabei führen reine assistierende und Komfort-Funktionen, wie Musik abspielen, oder Einkäufe erledigen eher weniger zu einer Personifizierung[33].

Seymore hat sich mit aktuellen Sprachassistenten als Problem befasst[37]. Demzufolge sind Sprachassistenten dem Nutzer in mehreren Bereichen viel zu unflexibel. Für eine Besserung sollen Dialoge nicht nur einem gleichen Schema folgen. Zudem müssten Sprachassistenten einen respektvolleren Umgang mit den Nutzern pflegen. In Bezug auf Datensicherheit sollten Sprachassistenten außerdem loyaler sein[37]. Besonders, da ein Großteil der Nutzer Datensicherheit als wichtiges Anliegen sehen[9]. Die Idee sei, dass der Sprachassistent unter anderem Sicherheitslücken oder unerwünschte datensammelnde Dienste erkennt und dem Nutzer die Möglichkeit gibt, diese zu schließen[9].

Neben der Usability von Alexa und diverser Skills sowie der User-Experience während des Aneignungsprozesses befasst sich diese Arbeit um die Aufgabenkomplexität und dessen Auswirkungen auf die Interaktion und Kommunikation mit dem Sprachassistenten. Es zeigt sich, dass die Nutzung vom Sprachassistenten oft bei einfacheren Befehlen und Anwendungsbereichen bleibt, wie Wetterabfragen oder Ähnliches[10, 20], da unter anderem die Kapazität der Sprachassistenten bezüglich der Dialogkomplexität eher begrenzt ist[10, 30]. Dabei ist die Komplexität der Dialoge von der Aufgabenkomplexität abhängig. Je komplexer eine Aufgabe in Zusammenarbeit mit dem Sprachassistenten ist, desto tiefergehend sind die Dialoge. Kompliziertere oder komplexe Aufgaben sollten demnach auch durch komplexe Dialoge gelöst werden können[16, 28, 21].

Im Folgenden wurden daher Studien und deren resultierenden Erkenntnisse zum Themenbereich der Dialogkomplexität und die damit oftmals verbundenen Chatbots, welche textbasierte Dialogsysteme bezeichnen, berücksichtigt und erläutert. Sprachassistenten und Chatbots sind sich in einigen Punkten sehr ähnlich und verfolgen je nach Aufgabenbereich auch dieselben Ziele, dem Nutzer Aufgaben abzunehmen oder Informationen zu beschaffen[36, 10, 11].

2.2 Chatbots und Dialogkomplexität

In den meisten Fällen ist die Nutzung und User-Experience von Sprachassistenten, in Bezug auf Dialog- und Aufgabenkomplexität, durch fest geregelte Schemata und begrenzte Kapazitäten in der Durchführung von komplexen Dialogen eher ernüchternd[37, 11]. Ähnlich wie Sprachassistenten sind Chatbots Anwendungen die mithilfe von künstlicher Intelligenz mit Menschen kommunizieren können, zwar ohne die sprachliche Komponente, dafür mit der Möglichkeit deutlich tiefergehende Unterhaltungen zu führen[10]. Einige Chatbots können durchaus längere Gespräche führen. Bis hin zu mehreren 100 sinnvollen wechselseitigen Austausch sind in Konversationen möglich[10]. Dadurch stellen Chatbots ebenfalls eine Grundlage in der Forschung zur Dialogkomplexität im Bezug zur Aufgabenlösung und Interaktion dar. Allerdings gibt es viele Arten von Chatbots. Zum einen, die als meist reine Internet-Anwendung, sogenannte Virtual-Companion (Virtuelle Gefährten), schlüssige und komplexe Unterhaltungen führen können. Diese können teilweise durch maschinelles Lernen menschenähnlich auftreten[27]. Zum anderen gibt es aufgabenorientierte Chatbots. Letztere können im Gegensatz zum Virtual Companion keine Unterhaltungen selbstständig weiterführen, sondern arbeiten lediglich kurze Befehle oder einfachere Abläufe ab[10]. Daher werden diese unter anderem in Bereichen genutzt, in denen Serviceanfragen gestellt werden und mit dem Chatbot bearbeitet werden können. Aber auch als Unterstützung der Mitarbeiter in Unternehmen werden aufgabenorientierte Chatbots gerne eingesetzt, da sie dort effizient und fehlerfrei arbeiten[39]. So bieten Studien zu Chatbots einen Einblick in die Mensch-Computer-Interaktion aus verschiedenen Blickwinkeln.

Es haben sich bereits einige Studien zu Chatbots mitunter der Dialogkomplexität befasst[10, 11, 15]. Dabei wurden die Beobachtungen auf die Kapazitäten und die Möglichkeiten der Interaktion mit Chatbots fokussiert. Dialoge mit diesen, die allerdings auch Aufgaben lösen können, sind deutlich weniger komplex. Zwar können die aufgabenorientierten Chatbots mit drei bis sieben wechselseitigen Austausch mehr als die ein bis drei Austausch von Sprachassistenten liefern, sind dem Lösen von komplexeren Aufgaben, dementsprechend aber auch nur mit kurzen Antworten programmiert[10]. Dadurch sind die Konversationen mit Chatbots tiefergehend, jedoch ebenfalls nur mit festen Schemata. Dennoch sind die scheinbar einfachen Aufgaben, wie z.B. über einen Chatbot online ein Flugticket zu kaufen, durchaus komplexer als sie scheinen[11]. Komplexe Aufgaben können dementsprechend gelöst werden, wenn während dem Dialog keine abweichenden Fragen oder Antworten notwendig sind. Wichtig dabei ist die Unterscheidung der von Liao et al. eingebrachten Kategorien der Dialogkomplexität[15]. Dazu wurden verschiedene Aspekte während der Interaktion

von Nutzern mit einem Chatbot kategorisiert, welcher Fragen von Angestellten beantworten sollte. Zu den verschiedenen Kategorien zählt zum einen die lexikalische Komplexität, in der ausführlichere oder formellere Fragen eine Rolle spielen. Ein anderer Aspekt der Dialogkomplexität ist die informative Komplexität. Dabei werden Fragen mit mehr Input und mehr zu verarbeitende Informationen berücksichtigt. Die dritte Kategorie ist die strukturelle Komplexität. Demnach beinhaltet dieser Aspekt die Möglichkeit oder den Wunsch im Dialog auf mehrere wechselnde Austauschpartner zu setzen und die Frage nicht nur innerhalb eines Satzes zu beantworten. Aber auch Fragen und Teilfragen zu stellen und somit auf Antworten zu reagieren zählen dazu. Da manche Nutzer diese Art der Konversation als natürlicher empfanden, neigten diese zu tiefergehenden Unterhaltungen mit den Dialogsystemen, wie auch beim letzten Aspekt der Dialogkomplexität beschrieben worden ist[15]. Die letzte Kategorie, die Komplexität der Interaktion, zeigt sich in der Zusammenarbeit mit Nutzer und Maschine. Die Konversationen sollten auf tiefergehenden Dialogen und aufeinander abgestimmten Sätzen folgen[15]. Letztendlich stellten Liao et al. eine Lösung mit verschiedenen individuellen Profilen vor. Jeder Nutzer sollte je nach Bedürfnis und Situation ein zugeschnittenes Komplexitätsprofil zugeordnet bekommen, sodass die User-Experience verbessert werden kann[15]. Nutzer könnten mit komplexen Dialogen demnach auch die gewünschten Aufgaben mit dem Chatbot oder Sprachassistenten erledigen, welche zum jetzigen Standpunkt an der zu unausgereiften Komplexität noch scheitern würden[10].

Um die Nutzung von Chatbots zu untersuchen, haben Jain et al. eine Aneignungsstudie mit verschiedenen Chatbots durchgeführt[12]. Dazu wurden Nutzer über drei Tage mit unterschiedlichen Aufgabenbereichen beauftragt. Aber auch in der Studie waren die besten Erfahrungen mit einfacheren Aufgaben, da auch dort die Chatbots am besten funktionieren[12]. Zudem kann auch bei Chatbots die Einschätzung zur Intelligenz zu hoch eingeschätzt werden[1]. Aneja et al. befassten sich mit einer Fehleranalyse, in der sich die eingeschätzte Intelligenz des Sprachassistenten oder Chatbots durch vermehrt vorkommender Probleme verringerte[1]. Zusätzlich fällt die Sympathie mit jedem Fehler signifikant. Dabei spielen die Probleme bei Abzweigungen im Dialog eine besondere Rolle[1]. Diese Abzweigungen können unter Umständen nicht zurückgenommen werden und so muss das ganze Gespräch bei einem Fehler aus der Sicht des Nutzers neugestartet werden[6]. Im Vordergrund ist es deshalb wichtig die Fehler zu minimieren. Nach Aneja et al. scheint die Sympathie wiederum zu steigen, sofern die Fehler lediglich aus der Kohärenz herführen[1]. Dabei liegt es eher daran, dass die Nutzer es amüsant finden, wenn die Zusammenhänge von der Maschine nicht verstanden werden[1]. Einige grundlegende Probleme sind in Bezug auf eine funktionierende Zusammenarbeit dabei nicht gelöst. Aufgaben die auch für den Nutzer komplex scheinen,

sind somit schwieriger oder gar nicht zu lösen.

2.3 Aufgabenkomplexität und Kooperation mit Sprachassistenten

Die Aufgabenkomplexität stellt einen wichtigen Stellenwert in dieser Arbeit dar. Unter Aufgabenkomplexität werden die Art und Weise wie mit dem Sprachassistenten interagiert wird, um Aufgaben lösen zu können. So wird bei komplexen Aufgaben eine Zusammenarbeit mit dem Sprachassistenten erfordert, um ans Ziel zu kommen. Daher liegt der Fokus besonders auf Zusammenarbeit und gemeinsamer Lösungsfindung mit dem Computer als Gegenüber. Mit der steigenden Komplexität der Aufgaben steigt auch die Häufigkeit der Fehlerinteraktionen[17, 30]. Allerdings können durch komplexe Aufgaben Ergebnisse erzielt werden, welche mit einfachen Interaktionen nicht möglich wären[10]. Dementsprechend ist es sinnvoll den Mehrwert einer komplexen Zusammenarbeit mit dem Sprachassistenten zu ermitteln und Fehlerquellen zu erfassen. Möglich ist es, dass komplexe Aufgaben allgemein weniger angegangen werden, da es den Nutzern angenehmer ist, die Nutzung auf einfache Befehle zu beschränken[13]. Der Grund für den Verzicht von komplizierten Interaktionen kann am mangelnden Vertrauen zur Computer-Schnittstelle liegen[13]. Porcheron et al. stellten dementsprechend Ansätze vor, die eine Zusammenarbeit bezüglich einer sozialen Interaktion beleuchteten und Schwerpunkte der sprachlichen Komponente in der Kooperation mit Sprachassistenten erfassten[32].

Da eine Zusammenarbeit mit Sprachassistenten überwiegend an fehlgeschlagener Interaktion liegt, sind mehrere Wege zur Verbesserung dieser unumgänglich[2]. Beneteau et al. haben sich mit verschiedenen Lösungswegen zur Minderung von Verständnisproblemen und fehlschlagender Interaktion befasst[2]. Dabei wurden mehrere Methoden vorgestellt, die eine Interaktion so verbessern oder Probleme lösen könnten, sodass unter anderem auch eine Zusammenarbeit grundsätzlich verbessert werden kann[2].

Zu einer verbesserten Zusammenarbeit von Mensch und Computer befassten sich zudem Dinarelli et al.[8] und Winkler et al.[40]. Dinarelli et al. haben dazu zwei Gruppen verglichen. Eine Gruppe sollte Dialoge zwischen Mensch und Mensch führen und die andere Gruppe sollte mit einem Chatbot interagieren. Dabei wurden die Dialoge während der Bearbeitung von Aufgaben der Gruppen verglichen. Die erste Gruppe konnte mehr Varianz in den Konversationen durch Spontaneität erzielen, was dem Chatbot für eine bessere Interaktion gefehlt hatte[8]. Der Chatbot war auf

den einfachen Austausch von Informationen und kurzen, eindeutigen Befehlen beschränkt[8]. Dadurch war eine zufriedenstellende Zusammenarbeit mit dem Chatbot nicht möglich. Diese Varianz ist auch essentiell beim Lösen von komplexen Aufgaben mit Sprachassistenten. Winkler et al. haben ähnlich wie in der Studie von Dinarelli et al.[8] zwei Gruppen verglichen[40]. Für den Vergleich sollten Aufgaben in Zusammenarbeit mit Alexa erledigt werden oder mithilfe eines menschlichen Tutors. Dabei wurde festgehalten, dass die Zusammenarbeit mit Alexa besser funktioniert hat, als mit dem menschlichen Tutor[40]. Dies widerspricht sich mit vorangegangenen Studien in denen versuchte Kooperationen zu Fehlinteraktionen führten[13, 30, 36]. Das liegt aber an dem speziellen Fall und der für die Aufgaben entwickelter Skill, welcher eine hohe Usability in dem Bereich hatte[40]. Dadurch war die Zusammenarbeit mit dem Sprachassistent optimal vorbereitet, hat sich damit aber nicht in andere komplexe Anwendungsbereiche eingearbeitet.

Eine Anpassung an den Sprachassistenten führt häufig zu einer Vermeidung komplexer Aufgaben, da während des Anpassungsprozesses fehlgeschlagene Interaktionen wegfallen[30]. Das führt zur Minderung der vielfältigen Anwendung und weiteren Problemen bei erneuter Konfrontation komplexer Aufgaben. Daher hindert die Anpassung an die Art der Sprecherwechsel mit dem Sprachassistenten an der Zusammenarbeit[28]. Des Weiteren erschwert der deutliche Unterschied in der Kommunikation zwischen Mensch und einem Sprachassistenten, die Anpassung an die Interaktion bei komplexen Aufgaben[28]. Dialoge mit besserem Verständnis, seitens des Sprachassistenten, durch die Optimierung von menschenähnlicher Kommunikation und richtigen Interpretieren des Kontextes, können ebenfalls zu komplexeren Dialogen und damit zu einer besseren Zusammenarbeit führen[16, 21].

Zusammenfassend zeigt die Literatur einige Aspekte und Erkenntnisse in der Nutzung, Usability und User-Experience von Sprachassistenten auf. Außerdem werden Probleme und Hindernisse bei einer Zusammenarbeit von Mensch und Sprachassistent beleuchtet. In dieser Arbeit liegt der Fokus auf komplexen Interaktionen. Dazu sollen die Gründe für dessen Scheitern erfasst werden und die Ergebnisse der Studie in den Forschungsstand eingeordnet werden können, indem die wichtigsten Erkenntnisse aus der Nutzung und Zusammenarbeit (siehe Tabelle 1) repliziert, oder ergänzt werden.

Ergebnisse und Erkenntnisse	Literatur
Probleme beim Lösen von komplexen Aufgaben	[34]
Verständnisprobleme durch Interaktion mit Sprache	[34]
Usability ist ausreichend für meist genutzte Anwendungen	[34]
Enttäuschung führt zu einer Abnahme der Nutzung	[36, 17]
Eine routinierte Interaktion hilft bei der Nutzung	[30, 17, 26]
Viel Eigenleistung und Anpassung zu einer Routine benötigt	[26]
Die Usability ist nicht für tiefergehende Gespräche konzipiert	[26]
Sprachassistenten unterscheiden sich zu sehr von einem Menschen in der Konversation	[33, 30, 23]
Zufriedenheit der Nutzer mit Sprachassistenten ist insgesamt recht hoch	[17, 30, 26]
Weniger Verständnisprobleme durch kurze und knappe Befehle	[17, 30, 26]
Mangel an Flexibilität bei Sprachassistenten führt zu seltenen komplexen Interaktionen	[30, 23, 28, 36, 35, 37]
Feedback durch Rückfragen helfen bei Problemen mit komplexen Aufgaben	[30]
Sprachassistent kann bei längeren Sätzen nicht der Logik folgen	[22, 30]
Vermenschlichung des Sprachassistenten durch sprachliche Interaktion	[28]
Personifizierung des Sprachassistenten findet durch einfache Anwendungen nicht statt	[17, 33]
Die Dialogkomplexität ist bei Sprachassistenten begrenzt und erschwert eine Zusammenarbeit	[30, 15, 10, 28, 21]
Eine bessere strukturelle Dialogkomplexität und Interaktionskomplexität führen zu einer erfolgreichen Zusammenarbeit bei komplexen Aufgaben	[15, 1]
Zusammenarbeit mit Sprachassistenten scheitert überwiegend an fehlgeschlagener Interaktion	[2]
Spontanität und Varianz fehlen dem Sprachassistenten zur erfolgreichen Zusammenarbeit	[15, 8]

Tabelle 1: Übersicht der wichtigsten Ergebnisse, Erkenntnisse und Thesen aus der Literatur

3 Methode

3.1 Einführung und Probanden

Um die Usability und User-Experience mit Alexa unter den Bedingungen steigend komplexer Aufgaben zu untersuchen und das einhergehende Interaktionsverhalten bei verschiedenen Herausforderungen zu ermitteln, erfolgte eine qualitative Studie mit 15 Probanden aus verschiedenen Haushalten. Für die Erfassung vergleichbarer Daten diente ein Online-Tagebuch sowie leitfadengestützte Interviews. Für die Studie wurde jedem der Teilnehmer, welche überwiegend über das Schneeballsystem angeworben sind, eine Amazon Alexa zur Verfügung gestellt. Bei dem Gerät handelte es sich um einen Echo Show. Wie bereits beschrieben, ist die Besonderheit der integrierte Touchscreen, bzw. Display. Die Teilnehmer waren zwischen 20 und 70 Jahre alt. Die Probanden hatten größtenteils keine weitere Erfahrung mit Sprachassistenten, wobei es ein paar wenige Ausnahmen gab. So nutzten zwei Haushalte bereits Sprachassistenten, allerdings nicht den Echo Show. In den Haushalten der 15 Probanden lebten teilweise Kinder, Partner oder andere Mitbewohner. Weitere Personen, abseits der Probanden, welche ebenfalls mit Alexa interagieren konnten, wurden dabei nicht mitgezählt, aber in den Interviews berücksichtigt. Die Tagebuchstudie erstreckte sich über drei Wochen, wobei die Probanden gestaffelt eingeführt wurden. Für eine Ermittlung von Unterschieden zwischen der einfachen Nutzung und der komplexen Zusammenarbeit mit Alexa, sollten die Teilnehmer jeden Tag eine Aufgabe erledigen und die Nutzung und Erfahrung beschreiben.

3.2 Auftaktinterview

Zu Beginn wurde jeder Proband mit einem leitfadengestützten Interview zum Erfahrungsstand und bisherigen Nutzung bezüglich Sprachassistenten und digitalen Medien befragt. In den Interviews wurden auch die Erwartungen an den Sprachassistenten sowie an die Studie erfasst.

3.3 Tagebuch

Nach dem Interview am ersten Tag wurde das Gerät zusammen eingerichtet und die erste Aufgabe erteilt sowie ein erster Skill installiert, um die Vorgehensweise dazu kennen zu lernen. Dazu wurde der Chefkoch-Skill gewählt, da dieser unter anderem für manche Aufgaben benötigt wurde. Um den Skill zu installieren, wurde zudem

die Echo-App auf einem Smartphone heruntergeladen, da diese App eine Übersicht der verbundenen Skills, sowie die Möglichkeit neue Skills hinzuzufügen oder alte zu entfernen, bietet. Für die Lösung der ersten Aufgabe sollte mithilfe von Alexa ein Ei gekocht werden. Das Kochen stellt eine komplexe Tätigkeit dar, da der Proband zur Lösung der Aufgabe mit dem Sprachassistenten Zusammenarbeiten sollte. Daher war dies eine herausfordernde Aufgabe für den Beginn. Gewählt wurde diese, da währenddessen noch dem Probanden unter Umständen direkt geholfen werden konnte. Während der ersten Nutzung sollte möglichst nur beobachtet werden und ein Eingriff nur bei mehrfachen Fehlversuchen bei nicht zustande kommenden Interaktion stattfinden. Dadurch konnten erste Probleme mit der Interaktion festgestellt werden und trotzdem unter besagten Umständen eingelenkt werden. Die darauffolgenden Tage wurden täglich neue Aufgaben über eine virtuelle Tagebuch-Anwendung hochgeladen. Dafür wurde Ethos (EthOSApp) genutzt. Dort konnten die Teilnehmer die Tagesaufgabe abrufen und im Laufe des Tages eine Nachricht hochladen. Diese sollte enthalten, wie der Sprachassistent genutzt wurde und wie das Lösen der Aufgabe funktioniert hat sowie auch ein Feedback zur Interaktion mit Alexa, ob diese zufriedenstellend war, oder was eher enttäuschend war. Die Nachrichten konnten frei verfasst werden, sollten sich allerdings an dem vorgegebenen Schema orientieren.

Die Aufgaben wurden in drei Schwierigkeitsstufen eingeteilt und dementsprechend über die drei Wochen verteilt. Dabei wurden die Schwierigkeitsstufen nach Komplexität der geforderten Dialoge und der Schwierigkeit der Aufgabe an sich eingestuft. Die erste Stufe (Leicht) bezog sich auf einfache Interaktionen, die der ersten Annäherung dienlich sein sollten. Dadurch konnten die Teilnehmer sich an das Gerät gewöhnen und Berührungsängste abbauen. Zusätzlich sollten Probanden zur Interaktion mit dem Sprachassistenten durch einfache Aufgaben motiviert werden. Während der leichten Aufgaben konnte zudem erstmal die Befehlsform beobachtet werden, da die Konversationen nur aus einzelnen Sätzen bestanden und nur kurze Dialoge nötig waren, um Erfolge zu erzielen. Die erste Steigerung der Komplexität (Mittelkomplex) sollte die Interaktion mit Alexa vertiefen, indem z.B. mehrere, also mehr als ein wechselseitiger Austausch oder gleichzeitige Nutzung verschiedener Skills gefordert wurde. Neben der Erhöhung der Dialogtiefe sollten auch komplizierte Aufgaben gelöst werden. Dazu wurden Interaktionen mit wenig aufeinanderfolgenden Befehlen gefordert, allerdings mit Bedingungen, welche mit einer bereits angepassten Befehlsform gut lösbar waren. Abschließend wurde in der dritten Woche überwiegend komplexe Aufgaben gestellt, in denen das angeeignete Wissen und eine Zusammenarbeit mit Alexa nötig wurde. Dabei wurde die Interaktion mit gezielten und komplexen Suchen gefordert. Die Aufgaben provozierten teilweise bewusst Fehlerquellen, indem unter anderem mehrere Tätigkeiten gleichzeitig durchgeführt werden sollten. Bei den komplexen Aufgaben konzentriert

sich die Nutzung hauptsächlich auf eine Zusammenarbeit und gemeinsames Lösen der Probleme.

Tag	Schwierigkeitsgrad	Aufgabe	Kriterium für erfolgreiche Ausführung	Ziel bzw. was soll getestet werden
1	Komplex/Kernenlernaufgabe unter Anweisung	Koche ein Ei	Kochen mit Hilfe von Alexa	Nutzer soll erste Eindrücke beim Kochen mit Alexa sammeln
2	leicht	Mache ein Foto vom Standort mit Alexa	Man soll erkennen, wo der Standort räumlich ist	
3	leicht	Frage Alexa nach dem Wetter	Richtige Befehlsform	Sind die Formen der Befehle einfach?
4	leicht	Probiere 3 Tipps aus, die dir Alexa vorschlägt	Richtige Befehlsform	Sind die Formen der Befehle einfach?
5	leicht	Wünschen einen "Guten Morgen"	Deutliche Aussprache	Reaktion auf die Antwort von Alexa beobachten
6	leicht	Lass dir den Witz des Tages erzählen	Richtige Befehlsform	Neue Befehle kennen lernen
7				
8	Mittel Komplex	Frage Alexa nach einem Restaurant in deiner Nähe und dessen Öffnungszeiten	Mit Alexa mehrere aufeinander folgenden Befehle umsetzen	Probleme bei aufeinanderfolgenden Befehlen erkennen
9	Mittel Komplex	Lass dir Lieder von deinen 2 Lieblings Musikern abspielen	Aufeinanderfolgende Befehle werden richtig erkannt	Gibt es Probleme nach dem ersten Befehl?
10	Mittel Komplex	Erkundige dich über Sylvester		
11	Mittel Komplex	Aktiviere den Skill Trivial Pursuit und teste ihn	Spiel einmal erfolgreich gespielt	Gibt es Probleme beim Installieren und Nutzen von Skill?
12	Mittel Komplex	Suche nach einem Winter-Ratatouille Rezept bei Chefkoch und setze die Sachen auf eine Einkaufsliste	Alexa versteht den Befehl und bleibt nicht im bisherigen Skill	Gibt es Probleme beim Wechseln wegen Befehlen?
13				
14				
15	Komplex	Suche nach einem Pfannengericht und besorge die Zutaten (bis Tag 19)	Mit Chefkoch richtige Menge an Zutaten bestimmen	Umgang mit Chefkoch und Alexa
16	Komplex	Richte einen Radiosender ein	Richtige Befehlsform	Neue Skills und Befehle kennen lernen
17	Komplex	(Höre Radio und frage Alexa nach einem beliebigen Rezept)	Richtige Lautstärke	Probleme mit mehreren Skills zur selben Zeit finden
18	Komplex	Erstelle eine Playlist mit deinen 5 Lieblings Liedern und lass diese abspielen		
19	Komplex	Koche das Pfannengericht mit Alexa	Nutzen von Chefkoch während der Zubereitung	Handhabung von Alexa beim Kochen
20				
21	leicht	Mache ein Foto von dem Standort mit Alexa	Man soll erkennen, wo der Standort räumlich ist	Soll mit erstem Bild verglichen werden

Tabelle 2: Die Tagesaufgaben und deren Schwierigkeitsgrad

3.4 Postinterview

Nach den drei Wochen wurde ein Post-Interview durchgeführt, indem die Einträge im Tagebuch noch einmal besprochen wurden und das Nutzungsverhalten reflektiert wurde. Dabei sind die genutzten Skills sowie eine mögliche Vernachlässigung erfasst worden. Zum Abschluss wurde noch die Zufriedenheit mit der Interaktion des Sprachassistenten sowie mit dem Echo Show im Allgemeinen erfragt und wie die Aussichten der weiteren Nutzung mit Sprachassistenten sind.

3.5 Qualitative Inhaltsanalyse

Nach der Durchführung aller Interviews wurden die Daten zusammengetragen. Die Tagebucheinträge konnten mithilfe einer Übertragung in Excel-Tabellen übersichtlich dargestellt werden. Die Daten aus den Interviews und Tagebucheinträgen sind dann mittels Qualitativer Inhaltsanalyse nach Kuckartz[14] ausgewertet worden. Dazu wurden die Interviews mit der Software MAXQDA zur computergestützten Daten- und Textanalyse transkribiert und Deduktiv/Induktiv Codiert. Die transkribierten Interviews konnten dazu in Hauptkategorien eingeteilt und zunächst deduktive Codes erzeugt werden. Darauffolgend wurde induktiv Codiert indem die Segmente der Hauptkategorien in Unterkategorien feincodiert oder ersetzt wurden. Die Codes wurden Kategorisiert und anschließend Ausgewertet. Mit Fokus auf die Aufgabenkomplexität und den Umgang unter erschwerten Bedingungen mit Alexa, wurden die Codes entsprechend auf die Interaktion bezüglich der Aufgaben und Nutzung gewählt.

3.6 Deskriptive Statistik

Mit den ausgewerteten Interviews und Tagebucheinträgen, konnte unter anderem abgelesen werden, wie sich die Häufigkeit der Nutzung über die drei Wochen verändert hat, welche Probleme besonders häufig vorkamen oder welchen Vorteil bzw. Nachteil der Bildschirm bei der Nutzung und Interaktion dargestellt hat. Letzteres wurde vor allem in Bezug auf die steigende Aufgabenkomplexität betrachtet. Aus den Tagebucheinträgen wurden verschiedene Aspekte bezüglich der Interaktion und Nutzung von Alexa kategorisiert und deskriptiv statistisch in geeignete Kennzahlen überführt. Diese helfen einen Überblick der allgemeinen Nutzung und der User-Experience zu erlangen.

Im Folgenden sollen die Ergebnisse dieser Studie dargestellt und in Hinblick auf die Forschungsfrage eingeordnet werden.

Probanden-ID	Geschlecht	Alter	Erfahrung	Branche/Beruf	Haushaltsmitglieder
P1	m	21	Viel	Praktikant	3 Erw.
P2	w	50	Keine	Bürofachangestellte	2 Erw.
P3	w	31	Wenig	Erzieherin	2 Erw.
P4	m	27	Viel	Projektleiter in Industrie	2 Erw.
P5	w	56	Wenig	Sozialarbeiterin	2 Erw.
P6	w	57	Keine	Erzieherin	2 Erw.
P7	w	33	Wenig	Erzieherin	2 Erw., 1 Ki.
P8	m	20	Wenig	Sozialwesen	3 Erw.
P9	w	42	Keine	Erzieherin	1 Erw., 1 Ki.
P10	w	60	Keine	Rentnerin	2 Erw.
P11	w	29	Viel	Erzieherin	2 Erw., 1 Ki.
P12	w	33	Keine	Erzieherin	2 Erw., 1 Ki.
P13	w	21	Wenig	Studentin	3 Erw.
P14	w	25	Keine	Erzieherin	1 Erw., 1 Ki.
P15	w	70	Keine	Rentnerin	2 Erw.

Tabelle 3: Übersicht der Probanden

4 Ergebnisse

4.1 Erwartungen und Einrichtung

Die Teilnehmer waren zu Beginn zum größten Teil keine Nutzer von Alexa. Drei der Teilnehmer haben bereits Alexa genutzt, beziehungsweise besitzen einen Echo Dot oder Sonos Lautsprecher mit integrierter Alexa. In drei weiteren Haushalten bestand die Möglichkeit, Alexa durch Sonos Lautsprecher oder durch den Amazon TV-Stick zu nutzen, allerdings wurde es nur sehr selten bis gar nicht in Anspruch genommen. Sprachassistenten über Smartphones wurden von sechs der Probanden genutzt oder ausprobiert.

Anfangs waren die Erwartungen teilweise noch sehr hoch. Da der Echo Show für jeden der Teilnehmer unangetastetes Gebiet war, war auch für die mit Sprachassistenten erfahreneren Probanden die Erwartung höher. Die erwarteten oder geplanten Nutzungsbereiche waren recht vielfältig. Zum einen konnten Probanden sich vorstellen, dass Alexa sich gut im Geschehen in der Küche integrieren könnte, um z.B. beim Kochen oder Backen einen Timer zu stellen, wenn die Hände nicht frei seien. Zum anderen wäre Alexa gut für Hintergrundmusik oder als Smartphone-Ersatz, für schnelle Internetsuchen.

Wenn Sprachassistenten schon mal genutzt oder die Interaktion beobachtet wurde, waren die bisherigen Erfahrungen mit diesen eher negativ geprägt, mit Abweichungen der Probanden, die im Besitz von Alexa waren. Unter anderem fiel bei den ersten Erfahrungen auf, dass Alexa häufig nicht verstanden hat, was man möchte oder die Fragen oft wiederholt werden mussten.

Bedenken gegenüber der Datensicherheit und Überwachung waren unwesentlich, wobei das Thema durchaus bei manchen Teilnehmern mit Skepsis gesehen wurde. Im Großen und Ganzen wurde dennoch der Technik und dem Sprachassistenten vertraut, dass die persönlichen Daten sicher seien. Nach dem Interview verlief die Einrichtung von Alexa weitgehend problemlos, mit kleinen Ausnahmen, welche aber nicht mit dem Sprachassistenten, sondern aus den erforderlichen Amazon Account-Daten herführten.

Bei der Einrichtung fiel einigen Nutzen die Tastatur auf, die bei einigen ungewöhnlich und unhandlich empfunden wurde. Als das Gerät Updates herunterladen musste, was je nach Internet den Probanden als sehr lang empfunden wurde, konnte nebenbei auf dem Smartphone die Echo-App und der erste Anwendung installiert werden. Nachdem

der Chefkoch-Skill und Alexa einsatzbereit waren, wurde den Teilnehmern die erste Aufgabe gestellt.

4.2 Erste Berührungspunkte mit der komplexen Interaktion

Die erste Aufgabe, die unter Beobachtung stattfand, bereitete den meisten Teilnehmern Schwierigkeiten. Dabei machte es keinen Unterschied, ob bereits Erfahrungen mit Alexa vorhanden waren. Die Befehlsform wurde recht schnell angepasst, allerdings gab es Probleme, die Frage zur Lösung der Aufgabe so zu stellen, dass Alexa eine zufriedenstellende Antwort gegeben hat. Kleinere Abweichungen in der Fragestellung führten zu fehlerhaften Rückmeldungen. Für die Aufgabe gab es mehrere Lösungswege, allerdings waren diese -auch aufgrund der nach Meinung mancher Probanden eigentlich trivialen Aufgabenstellung- erst nach einigen Versuchen geglückt. Der vorher installierte Chefkoch Skill wurde dazu bei manchen Probanden genutzt, wobei dies zu den ersten, größeren Enttäuschungen und zu keiner Lösung geführt hat. Wenn davon abgesehen wurde, die Aufgabe mit dem Skill zu lösen, musste erstmal überlegt werden, was eigentlich benötigt werde, bzw. wie genau Alexa für die Lösung der Aufgabe beitragen könne. Dennoch konnte durch die erste Aufgabe ein Gefühl für die kommende Nutzung aufkommen. In den meisten Fällen wurde das Ei letztendlich mit einem standardmäßig integrierten Timer unterstützt.

4.3 Aneignung der Befehlsform

Tatsächlich verlief die darauffolgende Interaktion recht problemlos. Die Befehlsform für die einfacheren Interaktionen konnte schnell angeeignet werden und die ersten leichten Aufgaben konnten damit gut bearbeitet werden. Bis auf zwei Ausnahmen, bei denen die Befehlsform am Anfang bei manchen Interaktionen zu Frust geführt hat und trotz mehrfachen Umformulierungen keine gewünschten Ergebnisse bekamen. Neben den Tagesaufgaben wurde Alexa in den meisten Haushalten bereits viel ausprobiert und konnte sich innerhalb einer Woche in den alltäglichen Gebrauch integrieren.

„Am Anfang wurde noch viel ausprobiert“ (P10) um sich der Interaktion und Befehlsform anzunähern. Mit der Zeit wurden dann nur noch Funktionen benutzt, bei denen die Nutzer wussten, dass es mit Alexa funktioniert (P10). Dazu gehörten Musik, Wetter oder andere Informationsabfragen, welche nur kurze und knappe Befehle benötigten.

4.4 Standort und Nutzungsbereich

Der Standort des Echo Show war bei 13 der Haushalte im Essbereich, Wohnzimmer oder in der Küche. Teilweise wurde der Ort entsprechend des verfügbaren Platzes ausgesucht. Durch den Bildschirm war das Gerät größer als andere Sprachassistenten und konnte somit nicht immer an den gewünschten Ort platziert werden. Allerdings konnte der Sprachassistent jederzeit bewegt werden, falls sich eine Möglichkeit bot.

Dabei wurde des Öfteren bemängelt, dass das Gerät recht unflexibel im Wechsel des Standortes sei. So hätte ein Akku dazu beigetragen, dass das Gerät ohne lange Wartezeit, die für das Neustarten notwendig wäre, überall schnell einsetzbar gewesen wäre (P8), wodurch es auch aufgrund des Bildschirms als eine Art Tablet-Ersatz funktioniert hätte.

Der Bildschirm wurde von manchen Probanden dazu genutzt, um Filme oder Videos zu gucken. Sofern das Gerät auf einem Tisch stand, führte diese Nutzung zu keinerlei Problemen. Allerdings kam es zu Schwierigkeiten, wenn der Sprachassistent z.B. für eine Fitnessanwendung genutzt wurde. Dazu war unter anderem der verfügbare Platz im Raum nicht anwendungsgerecht oder der Abstand des Gerätes zum Nutzungsort zu hoch, um den Bildschirm einsehen zu können. Die erwähnte Dauer des Standortwechsels hatte am Ende dazu geführt, dass diverse Skills eher weniger genutzt wurden.

4.5 Zufriedenheit mit Qualität der Antworten und Ergebnisse

Obwohl die Befehlsform für die erfolgreiche Interaktion mit den leichten Aufgaben korrekt angewendet wurde, war die Zufriedenheit der Nutzer nicht durchgehend vorhanden. Ungefähr die Hälfte der Teilnehmer war öfter unzufrieden mit der Qualität der Ergebnisse, unabhängig von der erfolgreichen oder gescheiterten Interaktion. Unter anderem hat Alexa Fragen nicht richtig verstanden oder Befehle mit ähnlich klingenden Wörtern verwechselt und somit gänzlich andere Ergebnisse erzielt. Auch bei Fragen nach Musik-Bands, Schauspielern oder Städten wurden gleichnamige Ergebnisse aus anderen Kontexten gezeigt. Ähnlich klingende Namen haben zudem häufig zu dem gleichen Problem geführt. Zwar konnte Alexa in diesen Problemfällen bei erneuter, etwas umformulierte Frage das gewünschte Ziel finden, zu einem späteren Zeitpunkt wurden Nutzer allerdings mit demselben Problem seitens Alexa erneut konfrontiert.

4.6 Nutzungsveränderung und Nutzungserfahrung durch Steigerung der Aufgabenkomplexität

Mit der ersten Steigerung des Schwierigkeitsgrades der Aufgaben kamen im größeren Ausmaß als in der ersten vorangegangenen Woche Probleme auf. Gleichzeitig war die Neugierde der Teilnehmer gesunken. Weniger neue Skills wurden ausprobiert oder Interaktionsmöglichkeiten ausgetestet. Zudem war auch die allgemeine Nutzung im Vergleich schon geringer ausgefallen. Die häufig genutzten Anwendungen und Funktionen waren zum Zeitpunkt der zweiten Woche bereits in einer routinierten Nutzung und gaben dem Nutzer zufriedenstellende Ergebnisse. Durch die Steigerung der Komplexität der Aufgaben hatte sich die Nutzung allerdings durch mehrere Probleme zwangsläufig verändert. Die Probleme, die mit dem gesteigerten Schwierigkeitsgrad auftraten, waren dabei teilweise eher der verfügbaren Datengrundlage von Alexa zu Schulden gekommen. Dabei spielte es eine Rolle, dass Alexa bei Antworten nicht zwingend den eigenen Standort zur logischen Schlussfolgerung einbezogen hatte, sondern für die Probanden eine scheinbar zufällige Wahl bei der Präsentation von Lokalen, Bankautomaten oder Ähnlichem bekamen. Aber auch die Art und Weise der Interaktion führte zu Fehlern. Da sich die geforderte komplexe Interaktion mit Alexa mit der aus der ersten Woche unterschied, mussten die Befehlsform und Dialoge teilweise entsprechend angepasst werden. Bei denjenigen bei denen Probleme aufgetaucht waren, wurde die Anpassung nur langsam oder erst sehr spät angegangen. Dadurch sind zu Beginn häufiger komplexe Interaktionen und Aufgaben gescheitert.

Die schwerwiegendsten Interaktionsprobleme in der mittleren Phase der Studie kamen mit der Nutzung des Chefkoch-Skills auf. Obwohl eine korrekte Befehlsform Verwendung fand, wurde oftmals kein richtiges Ergebnis bei der Rezeptsuche angezeigt. Da allerdings bei anderen Probanden alles funktionierte, lag es an Verständnisproblemen und Schwierigkeiten die Fragen anzupassen. Bei der Zusatzaufgabe, die Zutaten eines Rezeptes auf eine separate Einkaufsliste zu setzen, kam teilweise der Skill an seine Grenzen. Der Skill war zu schnell oder war während der Begriffsaufzählung der Zutaten abgestürzt. Ein paar der Teilnehmer hatten mit viel Geduld versucht dennoch die Aufgabe zu lösen, aber nach längeren Scheitern aufgegeben. Da die Befehlsform grundsätzlich richtig angewendet wurde, wurde bei erneuten Schwierigkeiten der Befehl nicht besonders verändert. Dadurch blieben große Unterschiede in den Antworten aus. Aber auch bei Probanden, die versucht haben, die Befehlsform oder die Fragestellung stark zu verändern, kamen nicht immer auf eine Lösung, wenn das Problem letztendlich an dem Skill oder des Wissensstandes der Online Datenbank von Alexa lag.

Der letzte Abschnitt der Tagebuchstudie hat die Probanden durch die komplizierten

Aufgaben durchaus gefordert. Die Aufgaben der letzten Woche wurden von einem Großteil der Teilnehmer als zu schwer eingeschätzt. Eine Zusammenarbeit mit Alexa wurde durch Interaktionsprobleme und geringwertiger Usability diverser Skills stark eingeschränkt. So haben die Interaktion und die nicht förderlichen Antworten seitens Alexa den Probanden viel Geduld gefordert. Aber auch die wiederholten Fehlversuche und einiger nicht funktionierender Interaktionen wurde „irgendwann nur noch nervig“ (P3). Vor allem, weil Alexa nicht verstand, was gefordert wurde und nicht klar war, ob die Befehlsform das Problem sei oder ob Alexa die Informationen, bzw. die gewünschten Antworten nicht habe. Während der ersten zwei Wochen gab es, im Vergleich zur dritten, insgesamt weit weniger negative Bemerkungen zur Interaktion. Diesbezüglich gab es einige Tagebucheinträge mit den Hinweisen, dass die geforderten, komplexen Dialoge oder komplizierten Aufgaben und Interaktionen nicht möglich waren und vermehrt zur bereits erwähnten genervten Stimmung führten. Diese wiederum erschwerte, bzw. verhinderte ein weiteres Lösen der Aufgaben.

Zusätzlich wurde bemängelt, dass Alexa teilweise zu freundlich reagiert habe, nachdem der Sprachassistent beleidigt wurde, wenn der Nutzer ohnehin durch Fehlinteraktionen schon gereizt war (P9).

„Ein bisschen böser“ (P8) hätte Alexa demnach sein können, um sich der Stimmung des Nutzers anzupassen und die Interaktion somit etwas menschlicher zu gestalten. In anderen Fällen hat die bleibende Freundlichkeit oder „lustige Antworten“ (P12) bei Fehlinteraktionen dazu verholfen, die negative Situation aufzulösen.

4.7 Zufriedenheit mit der allgemeinen Nutzung

In der allgemeinen Nutzung hat sich dagegen ein positiveres Bild gezeigt. Die Nutzung der bekannten Anwendungen wurde mit der Zeit durch eine routinierte Interaktion optimiert, wodurch weniger Fehler und Verständnisschwierigkeiten aufgetreten waren. Zudem pendelte sich eine gewisse Routine in der Nutzung der Skills und Funktionen ein. So wurden nur noch Skills genutzt, die eine fehlerfreie Nutzung aufzeigten und zudem leicht handzuhaben waren. Gleichzeitig wurde die Nutzung unabhängig der Tagesaufgaben geringer. Zu den noch häufig -wie auch schon anfangs- genutzten Skills gehörten Musik oder Radio Skills. Hinzu kamen bei manchen Teilnehmern auch Entspannungs- oder bestimmte Geräusche-Skills die während des Abends oder der Nacht regelmäßig genutzt wurden.

4.8 Einfluss der Usability in der Zusammenarbeit

Nach den drei Wochen wurden die Post-Interviews mit den Probanden gestaffelt durchgeführt. Die Interviews haben zum größten Teil bestätigt, was durch die Tagebucheinträge erläutert wurde. Während der Nutzung hatte sich ein großer Teil der Probanden mehr Funktionen gewünscht. Insofern dass existierende Funktionen, oder Skills eher den Vorstellungen entsprächen. Aber auch eine auf den Benutzer entsprechend angepasste Alexa, welche Antworten und Konversationen auf einen Erfahrungspool herausleiten könne, wurde öfters angesprochen. Dies zeigte sich bei der Interaktion der routinierten Nutzung, insofern häufig genutzte Skills den jeweiligen Kontext besser erfassen und somit dem Nutzer keine vermeidbaren schlechten Antworten liefern würden. Häufig angewendete Skills haben zum einen dem Nutzer gefallen, zum anderen haben jedoch durch wiederkehrende, scheinbar willkürliche Antworten von Seiten Alexas zu Unverständnis geführt. Auch ein Ausweichen auf andere Geräte, wie Smartphones führte aus der zu häufig angeforderten Anpassung der Befehlsform heraus. So wurde des Öfteren in besonderen „Situationen eher zum Smartphone gegriffen, um etwas mit Google zu suchen“ (P12), da dadurch in viel kürzerer Zeit ein Ergebnis vorlag.

Kinder, die in den Haushalten gelebt haben, konnten ebenfalls Interesse am Sprachassistenten gewinnen. So wurden Filme oder kindergerechte Skills genutzt. Letztere waren allerdings eher bedürftig und zeigten einige Mängel in der Usability auf. Vor allem bei undeutlicher Sprache oder nicht korrekt angewendeter Befehlsform, welche von den Kindern nicht in allen Fällen beherrscht wurde, führte zu Fehlern und Problemen. Für die Kinder schien demnach der Echo Show als eine interessante Möglichkeit der Unterhaltung mit dem Vorbehalt, die Interaktion überwiegend den Eltern zu überlassen, sofern es sich nicht um sehr einfache Fragen gehandelt hat. Während Alexa Kindern Tipps gegen Langeweile gegeben hat, wurde bemängelt, dass „einem Erwachsenen auch langweilig sein kann“ (P2), und somit zur Abwechslung z.B. ein Buch vorgeschlagen werden sollte.

Neben den Probanden und deren Kindern, haben auch weitere im Haushalt lebende Personen mit Alexa interagiert. Diese Interaktionen waren dabei allerdings überwiegend einfache Konversationen, sodass kaum Probleme auftauchten. Dabei gab es auch Fälle, bei denen mehrere Personen gleichzeitig interagiert haben, oder bewusst Alexa genutzt wurde um andere zu ärgern.

Bei den ältesten Probanden führte die Interaktion mit Alexa häufig zu Schwierigkeiten. Allerdings in Situationen der Kontaktlosigkeit, in denen man durch Alexa „gefühl

[...] nicht alleine“ (P15) sei, könnte dieser Aspekt die Nachteile der Interaktionsschwierigkeiten überwiegen. Die Interaktion mit Alexa durch Sprachbefehle sei dennoch ungewohnt (P15) und sei nicht die gewünschte Art der Kommunikation über einen Computer (P10). Insgesamt war die Häufigkeit der Nutzung bei älteren stärker als bei jüngeren Probanden gesunken. Allerdings waren Interaktionsschwierigkeiten bei allen Altersgruppen gleichmäßig verteilt. Es bestand eine Abhängigkeit aus den Erfahrungen die bereits vor der Studie vorhanden waren, wodurch das Alter zumindest in der Häufigkeit der Fehlinteraktionen keinen wesentlichen Einfluss ausgeübt hat. Dennoch hatten ältere Probanden im Vergleich zu Jüngeren häufiger Probleme mit der Befehlsform und die damit verbundenen Schwierigkeiten. Auffällig war jedoch auch, dass jüngere Teilnehmer teilweise anders an Probleme herangetreten sind. Geduld und Verständnis mit dem Sprachassistenten haben dabei eine immer wiederkehrende Rolle gespielt, wonach ältere Probanden häufig geduldiger mit Alexa waren.

Es gab viele Probleme während der Bearbeitung der komplizierten und komplexen Aufgaben. Wie bereits in den Tagebucheinträgen genannt, wurden in den Interviews ergänzt, dass die Usability von Skills, wie die der für die Aufgaben benötigte Chefkoch-Skill, in vielen Situationen zur Verzweiflung führte. Schon bei der Rezeptsuche gab es viele Probleme. So wurde „nach einem Pfannengericht gefragt [...], und ganz andere Gerichte“ (P11) angezeigt. Eine „Suchfunktion mit dem Touchscreen“ (P3) wäre dahingehend möglicherweise hilfreich gewesen, die zum Zeitpunkt der Studie zumindest im Chefkoch-Skill nicht verfügbar war. In den drei Wochen kamen allerdings auch Updates für die Anwendung, welche ein paar der Probleme behoben. Dennoch blieb die einfache, frustfreie Nutzung aus. Eine Ausnahme bildeten die Probanden mit Vorkenntnissen. Bei diesen konnte der Chefkoch-Skill beim Kochen durch strukturelle Komplexität innerhalb der Tätigkeit überzeugen. Vor allem, dass „ohne Hände“ (P4) gekocht und in den „Schritten vor und zurück“ (P4) gegangen werden konnte, hatte für die User-Experience einen positiven Einfluss und hat die Zusammenarbeit zielführend beeinflusst. Zu Beginn konnten Schritte beim Kochen nicht nutzerfreundlich verfolgt werden, da die Updates des Skills erst im Laufe der Studie hinzukamen. Die einzelnen Schritte der Rezepte wurden viel zu schnell und ohne Pause vorgetragen und es keine Möglichkeit gab, diese zu wiederholen, ohne dass das Rezept neu aufzurufen. In diesem Aspekt haben die Verbesserungen des Skills der Zusammenarbeit mit Alexa dementsprechend vereinfacht. Allerdings hat auch in den Fällen der Bildschirm nur bedingt beigetragen. Bei einigen Probanden wurde erst im Nachhinein bewusst, bzw. wurde im Laufe der Zeit vergessen, dass das Gerät nicht nur einen Bildschirm hat, sondern dieser auch eine Touchscreen Funktion besitzt, welcher in manchen Situationen geholfen hätte. Unter anderem,

wenn die Spracheingabe gescheitert war, z.B. bei einer Rezeptsuche, hätte die Frage mit der „Tastatur eingegeben“ werden können „und da hätte es direkt gefunden“ (P11) werden können.

Gewünscht wurde von einigen Probanden zudem die weitere Entwicklung einiger Skills, um auch komplexe Aufgaben zu lösen und auch genutzte Skills besser nutzen zu können. Für die Zukunft konnte sich der Großteil der Teilnehmer eine Anschaffung des Echo Shows vorstellen, zumindest „wenn es ausgereifter wäre“ (P3). Allerdings würde ein Kauf hauptsächlich wegen der guten Soundqualität des Gerätes überlegt werden. Für die Besitzer eines Sonos Lautsprechers fällt dementsprechend ein Kauf des Echo Show weg, da der Bildschirm für sie keinen Mehrwert ergibt.

Insgesamt waren die meisten Probanden mit der Nutzung von Alexa, auch wegen dem Bildschirm zufrieden. Falls ein Echo Show gekauft werden sollte, würde allerdings auf bestimmte Skills verzichtet werden. Besonders diese, welche durch ein Smartphone schneller an das Ziel kommen können und dadurch auf eine schwere Bedienung verzichtet werden kann. Auch die Anpassung an die Erwartung und die Einstellung gegenüber dem Gerät, als „Maschine und nicht als Mensch“ (P5), führt in der Nutzung zu weniger Enttäuschungen. Das Gerät konnte vordergründig mit dem Lautsprecher punkten, wobei der Bildschirm dabei auch als eine abwechslungsreiche Interaktion erwiesen hat. Besonders bei den Kindern konnte die Visualisierung von Geschichten oder Wissensfragen das Interesse steigern. Bei der Zusammenarbeit für kompliziertere Aufgaben konnte der Bildschirm nach Aussage der Nutzer nur bedingt helfen. Wobei einige gut fanden, dass Alexas Antworten zusätzlich auf dem Display zu lesen waren. Die Dialoge und Anwendungsbereiche haben sich über die drei Wochen insofern verändert, dass die Nutzer bis auf die Aufgaben kompliziertere Anwendungen möglichst gemieden haben und Konversationen relativ kurzgehalten haben.

5 Diskussion

Durch die Studie wurde deutlich, dass der genutzte Sprachassistent, Echo Show, durch die Tonqualität überzeugen konnte. Der am häufigsten genannte Grund für eine geplante Anschaffung sei demnach der Lautsprecher als Ersatz einer Musikanlage. Der Bildschirm war dagegen eher zweitrangig und nicht immer förderlich in der Zusammenarbeit mit Alexa. Zwar fanden einige Probanden die Visualisierung von dem Gesprochenen als schöne Untermalung oder nutzten Alexa auch um Videos zu gucken, während komplexen Aufgaben diente der Bildschirm allerdings weniger dazu Aufgaben Zielführend zu bearbeiten. Besonders bei der Zusammenarbeit in der Küche wurde vom Bildschirm erwartet, das Kochen oder Backen durch die Visualisierung der Rezepte und Arbeitsschritte zu bereichern[29, 32]. Dadurch, dass der Bildschirm von den Teilnehmern nur in wenigen Fällen beim Kochen genutzt wurde, schließt jedoch nicht die These aus, dass ein Bildschirm in der Küche durchaus dem Sprachassistenten einen Mehrwert gewähren könnte[29]. Da die Skills dabei Probleme bereitet haben und nicht den Bedürfnissen der Nutzer entsprochen haben, war der Mehrwert in der Zusammenarbeit nicht sehr hoch. In diesem Zusammenhang sollte der Fokus in zukünftigen Forschungen auf der Usability von ausgewählten Skills, welche komplexe Aufgaben ermöglichen könnten, liegen. Denn letztendlich wurde durch die schlechte Usability der Skills der Bildschirm von Alexa in den meisten Fällen nur selten beim Kochen oder anderen komplexen Aufgaben erfolgreich integriert. Die Usability einiger Skills stellen demnach ein größeres Hindernis dar, um den Nutzern eine erfolgreiche Zusammenarbeit mit dem Sprachassistenten zu ermöglichen.

Durch die Studie konnten einige Erkenntnisse und Ergebnisse aus der Literatur repliziert und bestätigt werden (siehe Tabelle 4). Darunter hat die Tagebuchstudie bestätigt, dass die Nutzer mit Alexa in einfacheren Konversationen Interesse entwickeln, aber bei steigender Komplexität die Nutzung abweicht und nachlässt[30, 13]. Dabei hatte sich ähnlich wie in anderen Studien eine Routine in der Nutzung eingependelt[36, 30, 17, 26]. Durch eine Anpassung der Interaktion, Befehlen und Meidung von Funktionen die vermehrt zu Fehlinteraktionen geführt haben, kam eine Routine recht schnell in die alltägliche Nutzung[30]. Wie in vorangegangenen Studien[30, 34] waren die Probanden abgesehen von den komplexen Aufgaben überwiegend mit dem Sprachassistenten zufrieden. Demnach beschränkten sich die meisten Interaktionen auf einfache Befehlen, welche nur wenig wechselseitige Austausch erzeugten[8, 13]. Dabei wurden hauptsächlich Musik und Informationen abgerufen, wofür die Usability der Skills und auch des Gerätes vollkommen ausreichend war[34].

Während der Bearbeitung der komplizierten und komplexen Aufgaben musste die Befehlsform teilweise mehrfach angepasst werden, um dem Ziel näher zu kommen. Die fehlende Varianz in der Verarbeitung des sprachlichen Inputs hat somit, wie bereits erläutert, schnell zu Frust und Ungeduld geführt[15, 26]. Besonders wenn Nutzer über einen längeren Zeitraum an einem Befehl überlegen mussten, um ein befriedigendes Ergebnis zu erhalten, wurde die Bearbeitung späterer Aufgaben und die damit verbundene Interaktion negativ beeinflusst.

Die komplexen Aufgaben haben in der Anwendung gezeigt, dass die Dialogkomplexität seitens des Sprachassistenten zu gering ist, um diese fehler- und frustfrei anzugehen. Besonders die niedrige Komplexität der Interaktion[15] hatte sich bei einigen Aufgaben bemerkbar gemacht. Die steigende Aufgabenkomplexität hatte zudem die Probleme in der Flexibilität und Usability nochmals hervorgehoben[24]. Wie Sciuto et al. angedeutet haben, dass durch die Routine in der Nutzung davon abhält, neue Interaktionen zu erforschen, da Nutzer nur ungern freiwillig Interaktionen provozieren, welche unweigerlich zu Problemen führen[36], hat sich durch die Tagebuchstudie gezeigt, dass es kaum Abweichungen an der Routine gab, abseits der täglichen Aufgaben. Solange sich Sprachassistenten in der Interaktion und Konversation zu sehr von einem menschlichen Gegenüber unterscheiden, wird sich das möglicherweise auch nicht bald ändern[24, 31]. Da komplexe Aufgaben eine Zusammenarbeit mit dem Sprachassistenten erfordern, um erfolgreich bearbeitet zu werden, ist es unabdingbar, dass Alexa und Sprachassistenten im Allgemeinen sich dem Nutzer anpassen[30, 2]. Wie sich in den Tagebucheinträgen gezeigt hat, gab es mindestens einen Probanden ohne Probleme mit der Tagesaufgabe, womit es möglich gewesen sein muss, jede Aufgabe zu lösen. Dadurch bestätigt sich auch die Überlegung mit individuellen Komplexitätsprofilen[15], um den Nutzern entsprechend der Bedürfnisse sowie deren Art und Weise der Interaktion entgegenzukommen.

Die Studie hat ebenfalls bestätigt, dass die Nutzung in komplexen Aufgaben häufig nur mit sehr präzisen Befehlen gelingt, und dass durch kleine Fehler bereits die Konversation von vorne beginnen muss[1]. Allerdings hat sich dadurch nicht die Sympathie für den Sprachassistenten erhöht[1], sondern eher gebremst. Ein Problem innerhalb der Tagebuchstudie konnte aus Sicht der Nutzer erfasst werden. Da die Nutzung manche Probanden bei komplexen Aufgaben sehr frustriert hat, wurden teilweise Anschlussaufgaben nicht versucht oder viel früher aufgegeben.

Ergebnisse der vorliegenden Arbeit	Ergebnisse und Erkenntnisse	Literatur
Nicht bei allen Probanden führte das ausführen komplexer Aufgaben zu Problemen	Probleme beim Lösen von komplexen Aufgaben	[34]
Der sprachliche Input führte bei komplexen Aufgaben zu Verständnisproblemen	Verständnisprobleme durch Interaktion mit Sprache	[34]
Für einfache Anwendungen war Usability meistens zufriedenstellend	Usability ist ausreichend für meist genutzte Anwendungen	[34]
Sofern Skills zu Enttäuschungen geführt haben wurde diese weniger genutzt	Enttäuschung führt zu einer Abnahme der Nutzung	[36, 17]
In einer routinierten Nutzung kommen weniger Fehlinteraktionen vor	Eine routinierte Interaktion hilft bei der Nutzung	[30, 17, 26]
Nutzer mussten für eine Anpassung einige Fehler und Enttäuschungen erleben	Viel Eigenleistung und Anpassung zu einer Routine benötigt	[26]
Besonders bei einigen Skills führt die niedrige Usability zu kurzen Befehlen	Die Usability ist nicht für tiefergehende Gespräche konzipiert	[26]
Einige Probanden hatten Probleme durch die fehlenden menschlichen Reaktionen, z.B. bei Ungeduld oder Genervtheit	Sprachassistenten unterscheiden sich zu sehr von einem Menschen in der Konversation	[33, 30, 23]
In der allgemeinen Nutzung war die Zufriedenheit überwiegend hoch	Zufriedenheit der Nutzer mit Sprachassistenten ist insgesamt recht hoch	[17, 30, 26]
Kurze Befehle verringerten Verständnisprobleme, waren allerdings nicht immer Zielführend	Weniger Verständnisprobleme durch kurze und knappe Befehle	[17, 30, 26]
Bei diversen Skills konnte durch fehlende Flexibilität keine Zusammenarbeit geführt werden	Mangel an Flexibilität bei Sprachassistenten führt zu seltenen komplexen Interaktionen	[30, 23, 28, 36, 35, 37]
Feedback hätte besonders bei Verständnisproblemen weitergeholfen	Feedback durch Rückfragen helfen bei Problemen mit komplexen Aufgaben	[30]
Länge Sätze wurden von Alexa meistens nicht verstanden	Sprachassistent kann bei längeren Sätzen nicht der Logik folgen	[22, 30]
Bei wenigen Probanden kam das Gefühl eines menschlichen Gesprächspartner auf	Vermenschlichung des Sprachassistenten durch sprachliche Interaktion	[28]
Auch bei schweren Aufgaben fand keine Personifizierung statt	Personifizierung des Sprachassistenten findet durch einfache Anwendungen nicht statt	[17, 33]
Besonders bei einigen Skills führte die begrenzte Dialogkomplexität zum Scheitern einer Zusammenarbeit	Die Dialogkomplexität ist bei Sprachassistenten begrenzt und erschwert eine Zusammenarbeit	[30, 15, 10, 28, 21]
Nach Updates eines Skills der strukturellen Komplexität wurde eine bessere Zusammenarbeit festgestellt	Eine bessere strukturelle Dialogkomplexität und Interaktions-Komplexität führen zu einer erfolgreichen Zusammenarbeit bei komplexen Aufgaben	[15, 1]
Abhängig der Geduld von Probanden scheitert schon nach wenigen Interaktionsproblemen eine Zusammenarbeit	Zusammenarbeit mit Sprachassistenten scheitert überwiegend an fehlgeschlagener Interaktion	[2]
Nur präzise Befehle und bereits leichte Abweichungen erschwerten eine Zusammenarbeit	Spontanität und Varianz fehlen dem Sprachassistenten zur erfolgreichen Zusammenarbeit	[15, 8]

Tabelle 4: Gegenüberstellung der Ergebnisse der vorliegenden Studie mit den Ergebnissen und Erkenntnissen aus der Literatur

6 Fazit

Diese Bachelorarbeit hatte als Ziel die Unterschiede zwischen einfacher Nutzung und einer Kooperation mit Sprachassistenten zu untersuchen und inwiefern diese Einflüsse auf die Lösung komplexer Aufgaben ausüben. Außerdem sollten die dabei entstehenden Interaktionsschwierigkeiten und Vorteile der Zusammenarbeit erfasst werden und die allgemeine Nutzung mit Fokus auf die steigende Komplexität der Aufgaben beleuchtet werden. Dadurch konnte ein Zusammenhang der Dialogkomplexität mit der Komplexität der Aufgaben herausgearbeitet und näher beobachtet werden.

Um die Nutzung und Zusammenarbeit von und mit Sprachassistenten zu untersuchen, wurde eine Tagebuchstudie mit steigender Aufgabenkomplexität durchgeführt. Die dafür angeworbenen 15 Probanden konnten durch die individuellen Nutzungen einen guten Einblick in die Herangehensweisen komplexer Aufgaben bieten. Durch die Tagebuchstudie und der qualitativen Inhaltsanalyse konnte zudem ein Gesamtbild individueller Probleme während der komplexen Nutzung festgehalten werden. Wie bereits erläutert konnten vorangehende Studien größtenteils bestätigt werden und konnten durch den Fokus auf die Aufgabenkomplexität ergänzt werden.

Demnach konnten im Rahmen der Studie einige Erkenntnisse zur Interaktion mit Sprachassistenten und den Umgang mit komplexen Aufgaben und Gründe der Vermeidung tiefergehender Dialoge bestätigt werden[36, 30, 13]. Die Ergebnisse der Studie zeigen auch, dass die Zufriedenheit der Nutzer mit dem Sprachassistenten je nach Aufgabenbereich recht hoch war. Zum einen durch die routinierte Interaktion[30], als auch durch einen hohen Unterhaltungswert und wenig Probleme bei alltäglich genutzten Anwendungen. Durch die täglichen Aufgaben konnte gezeigt werden, dass eine Zusammenarbeit zwar möglich ist, sich diese allerdings in einigen Punkten von einer zufriedenstellenden Interaktion unterscheidet. Während einfache Anwendungen, wie Musik hören und Informationsabfragen und manche Visualisierungen durch den Bildschirm zur allgemeinen Zufriedenheit geführt haben, konnten komplexe Aufgaben nur durch klare und präzise Befehle gelöst werden und führten dennoch durch die fehlende Spontanität und Varianz seitens Alexa zu Fehlinteraktionen[30]. Die Nutzung fällt vor allem in der Küche unter gestressten Umständen, schweren Bedingungen und komplizierten Aufgaben schnell zu Frust und mindert das Interesse an weiteren Versuchen die Interaktion zu erweitern[30, 26]. Zwar konnte der Bildschirm in einigen Fällen einen Ansatz zur Zusammenarbeit vorzeigen, durch die fehlende Usability der genutzten Skills war der tatsächliche Mehrwert damit allerdings schwer zu erfassen. Einen großen Anteil am Scheitern einer fehlerfreien

Zusammenarbeit haben somit Skills und deren Usability zu tragen. Viele Skills die für einen Tätigkeitszweck komplexer Aufgaben entwickelt wurden, zeigen demnach einige Schwierigkeiten auf, eine Interaktion mit komplexen Dialogen zu führen, welche, wie erwähnt, signifikant für ein Gelingen der Zusammenarbeit sind. Während der durchgeführten Studie wurden teilweise Updates der Skills veröffentlicht, welche die Usability etwas erhöht haben. Dies zeigt auch, dass die Entwicklung in die richtige Richtung geht. Es besteht dennoch weiterer Forschungsbedarf zur Usability der Drittanbieter-Skills. Komplexe Aufgaben mit einer Zusammenarbeit zu lösen kann durch zielführende Konversationen und Nutzung kompatibler Skills zu einem Mehrwert führen. Außerdem könnte eine komplexe Interaktion optimiert werden, indem bei Fehlern, einzelne Schritte im Dialog vor oder zurückgesprungen werden könnten, ohne das Gespräch neu beginnen zu müssen[6]. Besonders bei Verständnisproblemen hat es einen solchen ausgeprägten Aspekt der Dialogkomplexität gefehlt.

Während der Nutzung der leichten Aufgaben hatte sich zusammenfassend eine zufriedenstellende Routine eingependelt. Abseits der komplexen Aufgaben konnte Alexa den meisten Nutzern eine unterhaltsame Zeit bieten. Manche Probanden haben durch die täglichen Aufgaben das Interesse an der Zusammenarbeit mit Alexa verloren. Das kam daher, dass komplexe Aufgaben deutlich mehr Anpassung und letztendlich Nerven gefordert haben. Nutzer, die es geschafft haben, mit Alexa zusammenzuarbeiten, hatten dadurch einen breiteren Anwendungsbereich. Allerdings wurden komplexe Aufgaben dennoch außerhalb der Tagesaufgaben gemieden, sofern die Lösung und Bearbeitung dieser zu lange gedauert hatte oder nur mit viel Geduld durch die genannten Hindernisse geschafft wurden.

Da einige Probleme von dem Sprachassistenten herführten, kann in manchen Bereichen nur auf eine Steigerung der Usability gewartet werden, oder an erwähnte Lösungswege gehalten werden. Eine strikte Anpassung an die Befehlsform innerhalb jeder Anwendung und Meidung komplexer Satzstrukturen oder aneinander folgende Sätze helfen in den meisten komplexen Aufgaben. In der Studie wurde auch die Nutzung mit dem Bildschirm genauer betrachtet. In einigen Anwendungsbereichen konnte die Visualisierung einen höheren Unterhaltungswert bieten, sowie bei einfachen Anwendungen, unter anderem durch eine Wetteranzeige, auch zweckgebundene Informationen ästhetischer gestalten als ein Gerät ohne Bildschirm. Außerdem wurde der Bildschirm auch in komplexen Aufgaben gefordert oder zwangsläufig integriert. Der Bildschirm kann so in der Küche durchaus behilflich oder nützlich sein. Durch die Visualisierung der Rezepte oder Zubereitungsschritte können komplexe Vorgänge und Zusammenarbeit mit Alexa unterstützt und verbessert werden. Diesbezüglich sind die Probleme innerhalb der Interaktion und der Skills allerdings noch ein Hindernis.

Daher könnte es helfen, wenn Nutzer vor der Interaktion einsehen könnten, welche Möglichkeiten der Nutzung und welche Hindernisse bestehen[30], bevor es zum Scheitern kommt. Des Weiteren führten Wechsel zwischen Skills häufig zu Problemen und schließlich zur Hinderung an der Lösung komplexer Aufgaben. Daher sollten solche Funktionen und die Vorgänge einfacher und nutzerfreundlicher gestaltet werden.

Im Rahmen der Zusammenarbeit von Nutzer und Sprachassistent besteht Verbesserungsbedarf der Usability und der Dialogkomplexität. Nutzer sollten die minderen Kapazitäten der Sprachassistenten im Bereich der menschlichen Interaktion nicht als Hindernis zu einer zufriedenstellenden Nutzung sehen. Um eine zielführende Zusammenarbeit, wie beispielsweise beim kochen, zu ermöglichen, muss sich zurzeit noch zu sehr an den Sprachassistenten angepasst werden und präzise auf die Befehlsform geachtet werden. In Zukunft könnten allerdings schon durch kleinere Verbesserungen der Skills Interaktionen in Kooperation mit Sprachassistenten deutlich benutzerfreundlicher stattfinden. Weiterer Forschungsbedarf ergibt sich dementsprechend aus der stetig weiterentwickelnden technischen Grundlage, wodurch sich die Interaktion und Zusammenarbeit durch Designentwicklungen diverser Skills verändern kann. Die vorgestellten Ergebnisse werfen die Frage auf, inwieweit die Nutzung durch verbesserte Usability tatsächlich zur Minderung der Fehler in der Zusammenarbeit führt.

Literaturverzeichnis

- [1] Aneja, D. et al. *Conversational Error Analysis in Human-Agent Interaction*. Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents - IVA '20. 2020.
- [2] Beneteau, E. et al. *Communication Breakdowns Between Families and Alexa*. Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI '19. 2019.
- [3] Beneteau, E. et al. *Parenting with Alexa: Exploring the Introduction of Smart Speakers on Family Dynamics*. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems - CHI '20. 2020.
- [4] Bentley, F. et al. *Understanding the Long-Term Use of Smart Speaker Assistants*. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies. 2018.
- [5] Bitkom. *Zukunft der Consumer Technology – 2019*. Studie. 2019. URL: https://www.bitkom.org/sites/default/files/2019-09/190903_ct_studie_2019_online.pdf.
- [6] Castle-Green, T. et al. *Decision Trees as Sociotechnical Objects in Chatbot Design*. Proceedings of the 2nd Conference on Conversational User Interfaces - CUI '20. 2020.
- [7] Cho, J. *Mental Models and Home Virtual Assistants*. Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems - CHI EA '18. 2018.
- [8] Dinarelli, M. et al. *Annotating Spoken Dialogs: from Speech Segments to Dialog Acts and Frame Semantics*. Proceedings of the 2nd Workshop on Semantic Representation of Spoken Language - SRSL '09. 2009.
- [9] Furini, M. et al. *On the Usage of Smart Speakers During the Covid-19 Coronavirus Lockdown*. Proceedings of the 6th EAI International Conference on Smart Objects and Technologies for Social Good - GoodTechs '20. 2020.
- [10] Grudin, J. and Jacques, R. *Chatbots, Humbots, and the Quest for Artificial General Intelligence*. Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI '19. 2019.
- [11] Jacques, R. et al. *Conversational Agents - Acting on the Wave of Research and Development*. Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems - CHI EA '19. 2019.
- [12] Jain, M. et al. *Evaluating and Informing the Design of Chatbots*. Proceedings of the 2018 Designing Interactive Systems Conference - DIS '18. 2018.

- [13] Kassel, J. and Rohs, M. *Talk to ME Intelligibly- Investigating An Answer Space to Match the Users Language in Visual Analysis*. Proceedings of the 2019 on Designing Interactive Systems Conference - DIS '19. 2019.
- [14] Kuckartz, U. *Qualitative Inhaltsanalyse*. Beltz Juventa Weinheim. 2012.
- [15] Liao, Q. V. et al. *Tailoring Conversational UX through the Lens of Dialogue Complexity*. In CHI Workshop on Conversational UX Design. 2017.
- [16] Liao, Q. V. et al. *All Work and No Play - Conversations with a Question-and-Answer Chatbot in the Wild*. Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18. 2018.
- [17] Lopatovska, I. and Williams, H. *Personification of the Amazon Alexa*. Proceedings of the 2018 Conference on Human Information Interaction and Retrieval - CHIIR '18. 2018.
- [18] Luger, E. and Sellen, A. *"Like Having a Really Bad PA": The Gulf between User-Expectation and Experience of Conversational Agents*. Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16. 2016.
- [19] Luria, M. et al. *Comparing Social Robot, Screen and Voice Interfaces for Smart-Home Control*. Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17. 2017.
- [20] Luria, M. et al. *Robotic Futures- Learning about Personally-Owned Agents through Performance*. Proceedings of the 2020 ACM Designing Interactive Systems Conference - DIS '20. 2020.
- [21] Mekni M. et al. *A Smart Virtual Assistant for Students*. Proceedings of the 3rd International Conference on Applications of Intelligent Systems - APPIS '20. 2020.
- [22] Monarca, I. et al. *Why Doesn't the Conversational Agent Understand Me - A language analysis of children speech*. Adjunct Proceedings of the 2020 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2020 ACM International Symposium on Wearable Computers - UbiComp-ISWC '20. 2020.
- [23] Moore, R. J. et al. *Conversational UX Design*. Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '17. 2017.
- [24] Moore, R. K. *Is Spoken Language All-or-Nothing? Implications for Future Speech-Based Human-Machine Interaction*. Lecture Notes in Electrical Engineering 427. 2017.

- [25] NielsenNormanGroup. *The Definition of User Experience (UX)*. <https://www.nngroup.com/articles/definition-user-experience/>. abgerufen am 21.04.2021.
- [26] NielsenNormanGroup. *Intelligent Assistants Have Poor Usability: A User Study of Alexa, Google Assistant, and Siri*. <https://www.nngroup.com/articles/intelligent-assistant-usability/>. abgerufen am 07.04.2021. 2018.
- [27] Paikari, E. and van der Hoek, A. *A Framework for Understanding Chatbots and their Future*. Proceedings of the 11th International Workshop on Cooperative and Human Aspects of Software Engineering - CHASE '18. 2018.
- [28] Pelikan, H. R. M. and Broth, M. *Why That Nao - How Humans Adapt to a Conventional Humanoid Robot in Taking Turns-at-Talk*. Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16. 2016.
- [29] Pins, D. et al. *Aneignung und Gebrauchstauglichkeit von Sprachassistenten im Anwendungsbereich Küche*. 14th International Conference on Wirtschaftsinformatik. 2019.
- [30] Pins, D. et al. *Miss Understandable - A study on how users appropriate voice assistants and deal with misunderstandings*. Proceedings of the Conference on Mensch und Computer - MuC '20. 2020.
- [31] Pollmann, K. et al. *Robot vs Voice Assistant: Is Playing with Pepper More Fun than Playing with Alexa?* Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction - HR '20. 2020.
- [32] Porcheron, M. et al. *Talking with Conversational Agents in Collaborative Action*. Companion of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing - CSCW '17. 2017.
- [33] Purington, A. et al. *“Alexa is my new BFF”: Social Roles, User Satisfaction, and Personification of the Amazon Echo*. Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '17. 2017.
- [34] Pyae, A. and Joelsson, T. N. *Investigating the Usability and UserExperiences of Voice User Interface: A Case of Google Home Smart Speaker*. Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct - MobileHCI '18. 2018.
- [35] Reeves, S. et al. *Voice-based Conversational UX Studies and Design*. Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems - CHI EA '18. 2018.

- [36] Sciuto, A. et al. *“Hey Alexa, What’s Up?”: Studies of In-Home Conversational Agent Usage*. Proceedings of the 2018 on Designing Interactive Systems Conference 2018 - DIS '18. 2018.
- [37] Seymour, W. *How loyal is your Alexa?: Imagining a Respectful Smart Assistant*. Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems - CHI EA '18. 2018.
- [38] Statista. *Anzahl der Smartphone-Nutzer weltweit von 2016 bis 2019 und Prognose bis 2023*. <https://de.statista.com/statistik/daten/studie/309656/umfrage/prognose-zur-anzahl-der-smartphone-nutzer-weltweit/>. abgerufen am 09.04.2021. 2021.
- [39] Von Wolff, R. M. et al. *Einsatz von Chatbots am digitalen Büroarbeitsplatz – Eine praxisorientierte Betrachtung von Einsatzbereichen, Wirkungen und Handlungsempfehlungen*. HMD 57. 2020.
- [40] Winkler, R. et al. *Alexa, Can You Help us Solve This Problem - How Conversations With Smart Personal Assistant Tutors Increase Task Group Outcomes*. Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems - CHI EA '19. 2019.

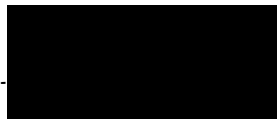
ERKLÄRUNG

Hiermit versichere ich, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe, insbesondere keine anderen als die Angegebenen Informationen aus dem Internet. Diejenigen Paragraphen der für mich geltenden Prüfungsordnungen, die etwaige Betrugsversuche betreffen, habe ich zur Kenntnis genommen.

Der Speicherung meiner Bachelor- bzw. Masterarbeit zum Zweck der Plagiatsprüfung stimme ich zu. Ich versichere, dass die elektronische Version mit der gedruckten Version inhaltlich übereinstimmt.

25.05.2021

(Datum)

A solid black rectangular box used to redact the signature of the author.

(Unterschrift)