Conversational Agents with Personality

Dissertation

an der Fakultät für Mathematik, Informatik und Statistik
der Ludwig-Maximilians-Universität München

vorgelegt von

Sarah Theres Völkel

M.Sc. Technical Communication

München, den 10. Mai 2022

Erstgutachter:	Prof. Dr. Albrecht Schmidt, Prof. Dr. Heinrich Hussmann *(verstorben)*
Zweitgutachter:	Prof. Dr. Elisabeth André
Drittgutachter:	Prof. Dr. Benjamin Cowan

Tag der mündlichen Prüfung: 14. Juli 2022

Abstract

Conversational agents (CAs) such as voice assistants and chatbots have permeated people's everyday lives. When interacting with these CAs, people automatically attribute a personality to them regardless of whether the CA designer intended it or not. This personality attribution fundamentally influences people's interaction behaviour and attitude towards the CA. By deliberately shaping the CA personality, designers have the opportunity to steer these automatic personality attributions in a desired direction. However, little information is available on how to design such a desired personality impression for a CA. Furthermore, in inter-human interaction, there is no such thing as a perfect personality. Nonetheless, today's commercial CAs have adopted a one-size-fits-all approach to their personality design, ignoring the potential benefits of adaptation.

These two insights, namely (1) that users assign a personality to CAs and (2) that there is no such thing as a perfect personality, motivate the vision of this thesis: To improve the interaction between users and CAs by deliberately imbuing CAs with personality and tailoring them to user preferences. This dissertation pursues two primary goals to realise this vision: (1) to develop methods to imbue CAs with personality systematically and (2) to examine user preferences for CA personalities.

To achieve the first goal, I introduce two approaches to imbue CAs with personality based on two underlying personality descriptions. The first approach adopts the human Big Five personality model as the theoretical basis for describing CA personality. This adoption allows me to transfer behaviour cues associated with human personality traits compiled from the psycholinguistic literature and my work to synthesise three levels of Agreeableness and Extraversion implemented in fully functional text-based CAs. An empirical evaluation of users' perceptions of these CAs after interacting with them demonstrates that human behaviour cues may be used to synthesise Agreeableness. However, they are insufficient to elicit the impression of low Extraversion or paint a complete picture of CA personality.

Due to this insufficiency, I develop a second approach in which I explore whether the human Big Five model can be used to describe CA personality. To this end, I apply the psycholexical approach, which yields ten personality dimensions that do not correspond with the human Big Five model. Consequently, I propose these ten dimensions as an alternative comprehensive way to describe CA personality and introduce a new method, Enactment-based Dialogue Design, to synthesise personality based on these ten dimensions.

To achieve the second goal, I present two approaches to examine user preferences for CA personality. Using a deductive approach, I investigate whether users prefer low, average, or high levels of four different personality dimensions in a CA in the context of different use cases. These investigations show that users have very individual preferences for the dimensions Extraversion and Social-Entertaining, whereas the majority prefer CAs that have a medium or high level of Agreeableness and a low level of Confrontational. I find the deductive approach to be useful for capturing users' evaluation of a personality-imbued

CA, but it is not effective in collecting user requirements and visions of a perfect CA. The second inductive approach, however, furnishes a novel pragmatic method to better engage users in developing CA personalities. In this context, I also examine the influence of users' personalities on their preferences for CA personality, but the effects are minimal.

In summary, this thesis makes the following contributions to imbuing CAs with personality: (1) theoretical clarity on the necessity of dedicated personality descriptions for CAs, (2) a set of verbal cues associated with human personality implemented in fully functional text-based CA artefacts, (3) an exploration of two methods for synthesising personality in CAs, and (4) a new method for eliciting users' vision of the perfect CA. I consolidate these methods into a user-centred design process for developing CAs with personality. Furthermore, I provide empirical evidence of diverging user preferences and discuss overarching patterns which CA designers may use to tailor their CA personalities to individual users. Finally, this thesis proposes a research agenda for future work, which addresses the challenges that emerged from the presented work.

Zusammenfassung

Conversational Agents (CAs) wie Sprachassistenten und Chatbots sind aus dem Alltag der Menschen nicht mehr wegzudenken. In der Interaktion mit CAs schreiben Benutzer:innen ihnen automatisch eine Persönlichkeit zu, unabhängig davon, ob die CA-Designer:innen dies beabsichtigten oder nicht. Diese Persönlichkeitszuschreibung beeinflusst grundlegend das Interaktionsverhalten und die Einstellung der Benutzer:innen gegenüber den CAs. Eine bewusste Gestaltung der CA-Persönlichkeit erlaubt Designer:innen, diese automatischen Persönlichkeitszuschreibungen in eine gewünschte Richtung zu lenken. Jedoch gibt es nur wenige Informationen darüber, wie eine solche gewünschte Persönlichkeit für einen CA gestaltet werden kann. Darüber hinaus gibt es in der zwischenmenschlichen Interaktion nicht die eine perfekte CA-Persönlichkeit, die allen Benutzer:innen gleichermaßen gefällt. Nichtsdestotrotz sind heutige kommerzielle CAs lediglich mit einer Persönlichkeit für alle Benutzer:innen ausgestattet und lassen somit die potenziellen Vorteile einer Anpassung an individuelle Präferenzen außer Acht.

Diese beiden Erkenntnisse, (1) dass Benutzer:innen CAs eine Persönlichkeit zuweisen und (2) dass es die eine perfekte Persönlichkeit nicht gibt, motivieren die Vision dieser Arbeit: Die Interaktion zwischen Benutzer:innen und CAs zu verbessern, indem CAs gezielt mit einer Persönlichkeit ausgestattet und an die Präferenzen der Benutzer:innen angepasst werden. Um diese Vision zu realisieren, verfolgt die vorliegende Dissertation zwei primäre Ziele: (1) die Entwicklung von Methoden, um CAs systematisch eine Persönlichkeit zu verleihen und (2) die Untersuchung von Präferenzen der Benutzer:innen für CA-Persönlichkeiten.

Um das erste Ziel zu erreichen, stelle ich zwei Ansätze zur Ausstattung von CAs mit Persönlichkeit vor, die auf der jeweiligen zugrunde liegenden Persönlichkeitsbeschreibung basieren. In dem ersten Ansatz verwende ich das menschliche Big Five Persönlichkeitsmodell als theoretische Grundlage für die Beschreibung von CA-Persönlichkeit. Diese Annahme ermöglicht es, Verhaltenshinweise, die mit menschlichen Persönlichkeitsmerkmalen assoziiert sind, in der psycholinguistischen Literatur sowie meiner eigenen Arbeit zu identifizieren. Diese Verhaltenshinweise übertrage ich dann auf CAs, um jeweils drei Ausprägungen von Verträglichkeit und Extraversion zu synthetisieren, die in vollständig funktionsfähigen text-basierten CAs implementiert sind. Eine empirische Untersuchung der Wahrnehmung dieser text-basierten CAs deutet darauf hin, dass menschliche Verhaltenshinweise genutzt werden können, um Verträglichkeit zu synthetisieren. Sie sind jedoch unzureichend, um den Eindruck von niedriger Extraversion zu vermitteln sowie die Persönlichkeit von CAs vollständig abzubilden.

Aufgrund der mangelnden Eignung der menschlichen Persönlichkeitsbeschreibung entwickle ich einen zweiten Ansatz, in dem ich untersuche, ob das menschliche Big Five Modell für die Beschreibung von CA-Persönlichkeit genutzt werden kann. Zu diesem Zweck wende ich den psycholexikalischen Ansatz an, aus dem zehn Persönlichkeitsdimensionen hervorgehen, die nicht mit dem menschlichen Big Five Modell übereinstimmen. Folglich schlage ich diese

zehn Dimensionen als eine alternative und vollständige Möglichkeit zur Beschreibung von CA-Persönlichkeit vor. Außerdem führe ich eine neue Methode, genannt Inszenierung-basiertes Dialogdesign, ein, die es ermöglicht, Persönlichkeit auf Grundlage dieser zehn Dimensionen zu synthetisieren.

Um das zweite Ziel zu erreichen, stelle ich zwei Ansätze zur Untersuchung der Präferenzen von Benutzer:innen für CA-Persönlichkeit vor. In einem deduktiven Ansatz untersuche ich zunächst, ob Benutzer:innen eine niedrige, durchschnittliche oder hohe Ausprägung von vier verschiedenen Persönlichkeitsdimensionen in einem CA im Kontext unterschiedlicher Anwendungsfälle bevorzugen. Diese Untersuchungen zeigen, dass die Benutzer:innen sehr individuelle Präferenzen für die Dimensionen Extraversion und Sozial-Unterhaltend haben, während die Mehrheit CAs bevorzugt, die eine mittlere oder hohe Ausprägung in Verträglichkeit sowie eine niedrige Ausprägung in Konfrontativ aufweisen. Obgleich der deduktive Ansatz nützlich für die Evaluierung von CA-Prototypen ist, ermöglicht dieser es nicht, Bedürfnisse und Vorstellungen der Benutzer:innen einzufangen. Im zweiten, induktiven Ansatz präsentiere ich daher eine neue pragmatische Methode, um die Benutzer:innen besser in die Entwicklung von CA-Persönlichkeiten einzubinden. In diesem Zusammenhang untersuche ich darüber hinaus den Einfluss der Persönlichkeit der Benutzer:innen auf ihre Präferenzen für die CA-Persönlichkeit, finde jedoch nur einen begrenzten Effekt.

Zusammenfassend leistet die vorliegende Arbeit die folgenden wissenschaftlichen Beiträge zur Ausstattung von CAs mit Persönlichkeit: (1) Theoretische Klarheit über die Notwendigkeit dedizierter Persönlichkeitsbeschreibungen für CAs, (2) eine Sammlung verbaler Verhaltenshinweise, die mit menschlicher Persönlichkeit assoziiert sind und in voll funktionsfähigen CA-Artefakten implementiert sind, (3) eine Exploration von zwei Methoden zur Synthese von Persönlichkeit in CAs und (4) eine neue Methode, um die Vision eines perfekten CAs von Benutzer:innen zu eruieren. Ich führe diese Methoden in einem benutzungszentrierten Designprozess für die Entwicklung von CA-Persönlichkeiten zusammen. Darüber hinaus liefere ich empirische Belege für divergierende Präferenzen der Benutzer:innen für CA-Persönlichkeit und erörtere übergreife Muster, die CA-Designer:innen anwenden können, um ihre CA-Persönlichkeiten auf individuelle Benutzer:innen zuzuschneiden. Abschließend wird eine Forschungsagenda für zukünftige Arbeiten präsentiert, welche die Herausforderungen diskutiert, die sich aus den vorgestellten Arbeiten ergeben.

Acknowledgements

First and foremost, I would like to express my sincerest gratitude to my supervisor, **Heinrich Hussmann**, for his guidance and support throughout my PhD. Thank you for providing me with the flexibility to pursue my own research interests, going to conferences before publishing my first papers, creating a warm and friendly atmosphere at our lab, and most importantly, always putting the human before the work. Your sudden death has left an irreplaceable void in our lab, and I am beyond sad that we could not complete this journey together. I would also like to thank **Albrecht Schmidt** for being willing to "inherit" me and supporting me as if I had been his PhD student from the beginning. I appreciate your enthusiasm for research, which creates an inspiring environment. Thank you for your career advice and for creating numerous opportunities to network. I extend my heartfelt thanks to my external examiner, **Elisabeth André**, for reviewing this thesis, for the valuable feedback, and for sharing fascinating insights into the development of conversational agents. I am grateful to **Benjamin Cowan** for reviewing this thesis, his infectious enthusiasm for conversational agents, and his initiative in organising workshops and conferences on conversational agents and inviting me to this community. Thank you for your valuable insights and the pleasant atmosphere when collaborating.

Furthermore, I would like to express my gratitude to the unique and inspiring people at our lab, who made pursuing this PhD an invaluable experience. First and foremost, I would like to thank three incredible and kind power women who were not only my most valuable colleagues but also became some of my closest friends. Thank you for celebrating my success more than I did and comforting me when something did not work out as planned. Thank you, **Ceenu George**, for eventually giving in and becoming my friend. I deeply admire how you successfully conduct excellent research while raising two kids. You are indeed a #truefriend; thank you for the plethora of advice you have shared with me, for encouraging me, and for making me send this thesis while sitting on the beach in the Bahamas. I hope that these last years are only the beginning of our work together. Thank you, **Malin Eiband**, for inviting me to collaborate with you, which marked the beginning of me enjoying this PhD. Thank you for your unlimited helpfulness, for always listening to my concerns, for throwing a fantastic 30's birthday party together, and for sticking by me through thick and thin. I could not have been luckier than being seated next to **Nađa Terzimehić** on my first day at LMU. Thank you for your kindness, food and life recommendations, and unlimited coffee supply. Your friendship has made a significant impact on my PhD experience.

Thank you, **Daniel Buschek**, for being a mentor and a friend. There is no one from whom I have learned more about writing papers than you. Thank you also for a lot of tea times and great trips back in the days this was still possible.

I was very lucky that many more colleagues became rather friends than colleagues. Thank you, **Beat Rossmy**, for inspiring me to think outside of the box by showing me different angles of HCI research, for your artsy hat contributions, for your interest in my and others' research,

and for many shared meals together. Thank you, **Carl Oechsner**, for your infinite helpfulness, kind and caring nature, and our joint efforts in creating both meaningful projects and fun lab parties. Thank you, **Christian Mai**, for being my brother in temperament. You have been deeply missed since you left the lab. Thank you, **Christina Schneegass**, for many morning chats that made the day better, your kind and open ear, your engagement in organising great conference parties, for improving my skiing skills, and for introducing me to the first dog I actually like. Thank you, **Fiona Draxler**, for being an amazingly uncomplicated yet diligent collaborator and for the funny comments when I least expect them. Thank you, **Hanna Schneider**, for being a great researcher, colleague, and friend, whose work and working style truly inspired me when joining the lab. Thank you, **Jingyi Li**, for pleasing me with your witty comments, for shared skiing adventures, and for introducing me to Chinese New Year. Thank you, **Kai Holländer**, for your strength and dark humour in the face of the biggest challenges. You are an inspiration. Thank you, **Linda Hirsch**, for all the hard work you do for our team, for contemplating together on how to improve the working environment, for interesting morning chats for which the door has to be closed, and for enjoyable days and nights out. Thank you, **Luke Haliburton**, for a great collaboration on the IUI class, awesome Canada road trip tips, and the joy and fun you bring to all our nights out. Thank you, **Mariam Hassib**, for an inspiring defence and your thoughtful gifts. Thank you, **Matthias Schmidmaier**, for putting equal emphasis on slide design, for one of the best emails I have ever got, and for being my partner in crime in the Hauptseminar. I hope we will finally have more of these coffees and beers together. Thank you, **Michi Braun**, for an unforgettable trip to the Highlands, for sharing plenty of fun nights at conferences and in Munich, and for switching to Standard German when I'm around. Thank you, **Mohamed "Mo" Khamis**, for your kindness and helpfulness to everyone and your inspiring, endless motivation. This was a blessing to me when I started in our lab and continues to be in our monthly online parTEAs. Thank you, **Renate Häuslschmid**, for inviting me as a friend when I needed one, for supporting me in my first research projects, and for a very memorable in-depth qualitative coding session on Madeira. Thank you, **Robin Welsch**, for sharing your statistics and psychology wisdom, for providing feedback on my thesis, and for private Robin who is just fun to be around (except for when he lost Timmy Ta-co). Thank you, **Sarah "Saragon" Aragon-Bartsch**, for your thoughtful gifts which lighten up my days and for sharing the joy and burden of being a perfectionist.

Thank you, **Alexander Wiethoff**, for initiating the UX Ukraine class. Thank you, **Annika Kaltenhauser**, it is always a joy to meet you at IDCs. Thank you, **Bastian Pfleging**, for advancing my skiing skills and useful tips and tricks about PhD life. Thank you, **Dennis Dietz**, for an amazing hiking and barbecue trip to your house. Thank you, **Florian Bemmann**, for taking over the PhoneStudy project from me. Thank you, **Flo Lachner**, for tackling our first CHI together and supporting me in getting an internship. I look forward to working with you again. Thank you, **Francesco Choissi**, for unlimited recommendations on great Italian food places. Thank you, **Jesse Grootjen**, for bearing with me during the snowstorm. Thank you, **Julian Rasch**, for making me aware again of how nice it is to look out of the window. Thank you, **Markus Wieland**, for your great efforts organising the Christmas party. Thank

you, **Martin Ebert**, for your help in proofreading papers. Thank you, **Maxi Windl**, for sharing my love for Harry Potter. Thank you, **Nick Ornstein**, for the nice times during CHI. Thank you, **Steeven Villa**, for your dance performances. Thank you, **Sylvia Rothe**, for spreading your joy for research in the offices and your kind-hearted nature. Thank you, **Thomas Weber**, for your help with the Christmas party and your creativity.

Furthermore, I would like to extend my thanks to my current and former colleagues for being an amazing team that looks out for one another and has each other's backs, shares successes, fun, and frustrations. Thank you, Heiko, Sören, Amy, Sophia, Cara, Tony, David, Jue, Yongma, Changkun, Daniel U., Axel, Michi C., Martin, Gesa, Sarah P., Lukas, Ville, Maria, Florian L., Florian M., Lewis, Tonja, Bettina, Sebastian, Jakob, and Matthias H.

Moreover, I express my deepest gratitude to the other professors in our lab. Thank you, **Sven Mayer**, for your constant help with strategic and community matters, and for encouraging us all to have after-work drinks together. Your company at the lab makes weekend and evening work much more enjoyable. Thank you, **Andreas Butz**, for believing in me and my skills early on and furthering my career. Thank you also for organising a great via ferrata trip and progressing my outdoor sports skills. Thank you, **Florian Alt**, for sharing your wisdom and approaches to getting papers accepted to CHI.

Special thanks go to the amazing administrative team at our lab, without whom we would truly be lost. Thank you, **Rainer Fink**, for always providing us with the needed technology and setting everything up smoothly, for your patience, and for the witty remarks you comment all this with. Thank you, **Franziska Schwamb**, for going beyond your work to support us and for stating the facts when no one else will. Thank you, **Christa Feulner**, for all your help and support, which is truly appreciated. Thank you, **Anja Mebus**, for managing our labs and speaking your mind.

A big thank you also goes out to my **amazing students and interns**, in particular Lale Kaya, Jelena Pranjic, Sophia Cook, Moutoshi Pal, Julia Graefe, Verena Winterhalter, Anna Werner, Patricia Kimm, Samantha Meindl, Penelope Kempf, Romy Gruber, Till Kleinhans, Maria Vazquez, Henrike Haase, Florian Bemmann, Daniela Becker, and Florian Lehmann. Your contributions to my research are deeply appreciated and it was a pleasure collaborating with you and learning from you.

Thank you also to my amazing co-authors from other universities and all the people from the **HCI community**I met at conferences and winter schools. I sincerely hope to regularly meet you again. A particular shout-out goes to the amazing and kindhearted **CUI community** and **Anna-Katharina Frison** who has become not only a valued co-author, but also a travel companion and a friend.

I am also deeply thankful to the company Audi for funding the first three years of my PhD and providing me with insights into industrial research. Thank you, **Stephan Hummel**, for giving us great flexibility to follow our own research interests. Thank you, **Anna Christ**, for fighting for us and for recognition within the company. Thank you, **Ramona Schödel**, for tackling the adventure of doing a PhD in cooperation with a company together, for onboarding me at the beginning, and for the continuous joy of collaborating interdisciplinary. Thank you, **Quay Au**, for your dry sense of humour in the light of challenges. Thank you, **Clemens Stachl**, for initiating this project and the PhoneStudy project and teaching me a lot about interdisciplinary research. Thank you to **Markus Bühner** and **Bernd Bischl** for their interdisciplinary feedback and support in this project.

Furthermore, I would also like to thank the **Google Account team** for the opportunity to do an internship and gain insights into UX research at Google. In particular, thank you to my manager **Tobi** for going above and beyond for ensuring that I have a productive and pleasant time, to **Anja** for teaching me a lot about management style and research, and to **Terhi** for taking a lot of time to show me which aspects to consider when conducting industrial research.

I would not have started a PhD if it were not for the following people that sparked my interest in HCI and research. Thank you, **Jan Borchers**, for chasing me after a mentoring session just to show me the FabLab because you thought I might be interested in it. Thank you for introducing me to the topic of HCI. Thank you, **Chat Wacharamanotham**, for supervising my Bachelor thesis, thereby sparking my enthusiasm and interest in doing research. Thank you for deeply integrating me into the research design and encouraging me to follow a PhD in the first place. Thank you, **Martina Ziefle**, for teaching me a lot about study designs, users, and myself. Furthermore, thank you for showing me what an amazing lab atmosphere looks like, a lot of which I also tried to implement at LMU.

I would also like to thank the people outside the research community that accompanied me on this way. Thank you to all my friends, both near and far, from school, from university, from semesters abroad and summer school and finally to Munich. You are a constant source of joy, happiness, and encouragement. Thank you for many adventures, laughs, and open ears. You know who you are but I want to highlight a few here. Thank you, **Katharina**, for being the first person I call when I need to talk. Your listening ear and words of wisdom have been a lifeline during challenging moments. Thank you, **Eva**, for being my personal career advisor and for doing amazing vacation trips to refill the battery. Thank you, **Laura**, for many memorable moments in Lauri-Sari-weekends, such as painting together whilst discussing life. These weekends always lifted me up. Thank you, **Wenzel**, for the great hikes and for not letting me forget that sometimes work can wait. Thank you to my roommates, **Lara** and **Jonas**, for bearing with me before the CHI deadline, taking over the cleaning shift, for fun brunches, and for sharing stories at a perfectly lit kitchen table. Thank you to all my **Aachen friends** who showed me the joy of collaboration, teamwork, my strengths and weaknesses, and what can be done with intrinsic motivation.

And last but definitely not least, I would like to thank my family. Thank you to my brother **Jan Gerrit** for being an amazing researcher and even more amazing person, for providing feedback on this thesis, for sharing the pains of PhD life, and for being a moral compass. Thank you to my other brother **Moritz** for your witty remarks on research, for your kind and unlimited willingness to help, and for sharing insights on Computational Linguistics for my thesis. Thank you to my sister **Anna Marlen** for your warm and kind-hearted nature, for always supporting me, for sometimes being the older sister, and for a great vacation in Canada when I definitely needed it. Finally, I would like to thank my mother **Dorothee** for being a role model, for early on instilling the confidence in me that girls can be equally good in Maths and tech, for your unconditional love and endless support and thank my father **Matthias** for showing me rigour and diligence by example, for always supporting me, for believing in me and my and skills, and for sparking my interest in science with giving me puzzles to solve.

TABLE OF CONTENTS

Publications

This dissertation is cumulative, consisting of previously published research projects. These comprise the main body of this thesis and contribute to its overarching narrative. When referring to them, the format "[Core i]" is used, with $i \in [1...7]$. Each of the core publications is accompanied by a website on which the resources used for the project, such as study material and implementation details, are made available so as to make the research accessible and transparent in the spirit of Open Science.

In addition, I present ten publications to complement the primary contributions. I refer to these with "[Pub i]", with $i \in [1...10]$. Specifically, the publications [Pub4, Pub5, Pub6, Pub7, Pub10] help to set the context of this work and provide background information. Conversely, the publications [Pub1, Pub2, Pub3, Pub8, Pub9] contribute to the overarching discussion beyond the scope of this thesis, especially with regard to challenges concerning the acceptance of personality-imbued conversational agents.

[Core4] received an *Honourable Mention Award* and [Pub2] an *Outstanding Paper Award* at the respective conferences. A predecessor of the publication [Core2] was honoured with a *Best Paper Award* at the Conversational User Interfaces workshop as part of the Intelligent User Interfaces conference. The journal article [Pub6] was awarded *Best Paper* from the German Society for Online Research in 2021.

Due to the collaborative nature of the research, all of my publications were completed in a joint effort with fellow researchers and students. Table A.1 clarifies the contributions of all co-authors. Throughout my thesis, I use the scientific term "we" when referring to projects and publications, whilst writing "I" when referring to this thesis.

Core Publications

[Core1] Völkel, Sarah Theres, Buschek, Daniel, Pranjic, Jelena and Hussmann, Heinrich. 'Understanding Emoji Interpretation through User Personality and Message Context'. In: *Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services*. MobileHCI '19. Project Website: `www.medien.ifi.lmu.de/personality-emojis`. New York, NY, USA: Association for Computing Machinery, 2019. DOI: `10.1145/3338286.3340114`.

[Core2] Völkel, Sarah Theres and Kaya, Lale. 'Examining User Preference for Agreeableness in Chatbots'. In: *CUI 2021 - 3rd Conference on Conversational User Interfaces*. CUI '21. Project Website: `www.medien.ifi.lmu.de/agreeableness-chatbots`. New York, NY, USA: Association for Computing Machinery, 2021. DOI: `10.1145/3469595.3469633`.

[Core3] Völkel, Sarah Theres, Schoedel, Ramona, Kaya, Lale and Mayer, Sven. 'User Perceptions of Extraversion in Chatbots after Repeated Use'. In: *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. CHI '22. Project Website: www.medien.ifi.lmu.de/extraversion-chatbots. New York, NY, USA: Association for Computing Machinery, 2022. DOI: 10.1145/3491102.3502058.

[Core4] Völkel, Sarah Theres, Schödel, Ramona, Buschek, Daniel, Stachl, Clemens, Winterhalter, Verena, Bühner, Markus and Hussmann, Heinrich. 'Developing a Personality Model for Speech-Based Conversational Agents Using the Psycholexical Approach'. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. CHI '20. Project Website: www.medien.ifi.lmu.de/personality-model. New York, NY, USA: Association for Computing Machinery, 2020. DOI: 10.1145/3313831.3376210.

[Core5] Völkel, Sarah Theres, Meindl, Samantha and Hussmann, Heinrich. 'Manipulating and Evaluating Levels of Personality Perceptions of Voice Assistants through Enactment-Based Dialogue Design'. In: *CUI 2021 - 3rd Conference on Conversational User Interfaces*. CUI '21. Project Website: www.medien.ifi.lmu.de/voice-assistant-personality. New York, NY, USA: Association for Computing Machinery, 2021. DOI: 10.1145/3469595.3469605.

[Core6] Völkel, Sarah Theres, Kempf, Penelope and Hussmann, Heinrich. 'Personalised Chats with Voice Assistants: The User Perspective'. In: *Proceedings of the 2nd Conference on Conversational User Interfaces*. CUI '20. Project Website: www.medien.ifi.lmu.de/users-va-dialogues. New York, NY, USA: Association for Computing Machinery, 2020. DOI: 10.1145/3405755.3406156.

[Core7] Völkel, Sarah Theres, Buschek, Daniel, Eiband, Malin, Cowan, Benjamin R and Hussmann, Heinrich. 'Eliciting and Analysing Users' Envisioned Dialogues with Perfect Voice Assistants'. In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Project Website: www.medien.ifi.lmu.de/envisioned-va-dialogues. New York, NY, USA: Association for Computing Machinery, 2021. DOI: 10.1145/3411764.3445536.

Complementing Publications

[Pub1] Eiband, Malin, Völkel, Sarah Theres, Buschek, Daniel, Cook, Sophia and Hussmann, Heinrich. 'A Method and Analysis to Elicit User-Reported Problems in Intelligent Everyday Applications'. In: *ACM Trans. Interact. Intell. Syst.* 10.4 (2020). DOI: 10.1145/3370927.

[Pub2] Eiband, Malin, Völkel, Sarah Theres, Buschek, Daniel, Cook, Sophia and Hussmann, Heinrich. 'When People and Algorithms Meet: User-Reported Problems in Intelligent Everyday Applications'. In: *Proceedings of the 24th International Conference on Intelligent User Interfaces*. IUI '19. New York, NY, USA: Association for Computing Machinery, 2019, pp. 96–106. DOI: 10.1145/3301275.3302262.

[Pub3] Rossmy, Beat, Völkel, Sarah Theres, Naphausen, Elias, Kimm, Patricia, Wiethoff, Alexander and Muxel, Andreas. 'Punishable AI: Examining Users' Attitude Towards Robot Punishment'. In: *Proceedings of the 2020 ACM Designing Interactive Systems Conference*. DIS '20. New York, NY, USA: Association for Computing Machinery, 2020, pp. 179–191. DOI: 10.1145/3357236.3395542.

[Pub4] Schoedel, Ramona, Au, Quay, Völkel, Sarah Theres, Lehmann, Florian, Becker, Daniela, Bühner, Markus, Bischl, Bernd, Hussmann, Heinrich and Stachl, Clemens. 'Digital Footprints of Sensation Seeking'. In: *Zeitschrift für Psychologie* 226 (2019), pp. 232–245. DOI: 10.1027/2151-2604/a000342.

[Pub5] Schoedel, Ramona, Pargent, Florian, Au, Quay, Völkel, Sarah Theres, Schuwerk, Tobias, Bühner, Markus and Stachl, Clemens. 'To Challenge the Morning Lark and the Night Owl: Using Smartphone Sensing Data to Investigate Day–Night Behaviour Patterns'. In: *European Journal of Personality* 34.5 (2020), pp. 733–752. DOI: 10.1002/per.2258.

[Pub6] Stachl, Clemens, Au, Quay, Schoedel, Ramona, Gosling, Samuel D, Harari, Gabriella M, Buschek, Daniel, Völkel, Sarah Theres, Schuwerk, Tobias, Oldemeier, Michelle, Ullmann, Theresa, Hussmann, Heinrich, Bischl, Bernd and Bühner, Markus. 'Predicting Personality from Patterns of Behavior Collected with Smartphones'. In: *Proceedings of the National Academy of Sciences* 117.30 (2020), pp. 17680–17687. DOI: 10.1073/pnas.1920484117.

[Pub7] Völkel, Sarah Theres. 'Voice User Interfaces'. In: *Mensch-Maschine-Interaktion*. Ed. by Andreas Butz, Antonio Krüger and Sarah Theres Völkel. De Gruyter Oldenbourg, 2022. DOI: 10.1515/9783110753325.

[Pub8] Völkel, Sarah Theres, Haeuslschmid, Renate, Werner, Anna, Hussmann, Heinrich and Butz, Andreas. 'How to Trick AI: Users' Strategies for Protecting Themselves from Automatic Personality Assessment'. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. CHI '20. New York, NY, USA: Association for Computing Machinery, 2020, pp. 1–15. DOI: 10.1145/3313831.3376877.

[Pub9] Völkel, Sarah Theres, Schneegass, Christina, Eiband, Malin and Buschek, Daniel. 'What is "Intelligent" in Intelligent User Interfaces? A Meta-Analysis of 25 Years of IUI'. In: *Proceedings of the 25th International Conference on Intelligent User Interfaces*. IUI '20. New York, NY, USA: Association for Computing Machinery, 2020, pp. 477–487. DOI: 10.1145/3377325.3377500.

[Pub10] Völkel, Sarah Theres, Schödel, Ramona, Buschek, Daniel, Stachl, Clemens, Au, Quay, Bischl, Bernd, Bühner, Markus and Hussmann, Heinrich. 'Opportunities and Challenges of Utilizing Personality Traits for Personalization in HCI'. In: *Personalized Human-Computer Interaction*. Ed. by Mirjam Augstein, Eelco Herder and Wolfgang Wörndl. Berlin, Germany: De Gruyter Oldenbourg, 2019. DOI: 10.1515/9783110552485.

1
INTRODUCTION

"There is no such thing as a voice user interface with no personality."

– Michael H. Cohen, James P. Giangola, Jennifer Balogh.
***Voice User Interface Design*. 2004.**

Conversational agents (CAs) have circulated as an idea since ELIZA, a psychotherapist chatbot developed in the 1960s [212]. Yet, it is only recent enhancements in natural language processing that have sparked a new wave of interest in CAs and commercially available products such as chatbots and voice assistants [147]. Unlike traditional graphical user interfaces, users interact with CAs through natural language. This type of interaction has many advantages, such as being hands-free and easy to learn [171]. However, communication via natural language is inherently social, even in human-computer interactions [97], causing a shift from interacting with computers as tools to computers as partners [17]. This paradigm shift inevitably shapes user expectations, which today's CAs often fall short of, leading to user frustration [51, 59, 135, 178]. It has been argued that the interaction with CAs may be significantly improved by gaining a deeper understanding of users' expectations of the social interaction with CAs [157]. One of the fundamental aspects of social interaction is the attribution of personality [91], which is the focus of this thesis.

When we meet someone for the first time, we involuntarily form an impression of their personality, which significantly influences our future behaviour and expectations towards this person [142]. It has long been established that we also automatically attribute CAs a personality, regardless of whether the CA designer intended this or not (cf. Figure 1.1) [52, 161, 182]. As in inter-human interaction, a growing body of literature has shown that this personality attribution influences user trust [27, 224], likeability [20, 27, 37, 158], engagement [200, 224], self-disclosure [85, 224], and purchase behaviour [200]. By shaping the CA personality, CA designers have the opportunity to steer these automatic personality attributions in a desired direction rather than leaving them to chance [157]. However, despite this opportunity, there is little information on *how* to design the personality of CAs deliberately [112, 179].

There is no such thing as an objectively perfect personality. Whilst another person's personality is a strong predictor of how much we would like to interact with them again, not all people have the same preferences [33]. Similarly, user preferences for certain CA personality types are unlikely to be homogeneous, with previous research suggesting that users often favour personalities that match their own [27, 157, 158]. Nonetheless, today's commercial CAs have adopted a one-size-fits-all approach to their personality design [206]. As a result, CAs,

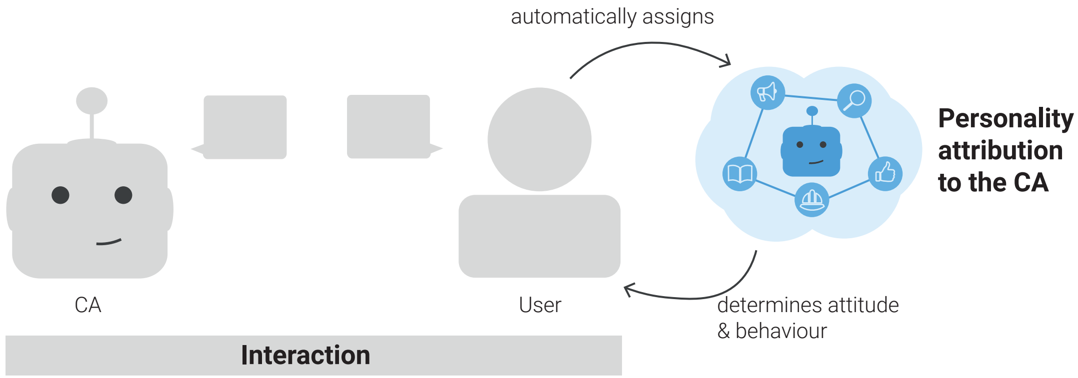

Figure 1.1: Similar to human-human interaction, people automatically attribute CAs a personality, which determines the user's interaction behaviour and attitude towards the CA.

such as Apple's Siri[1], Amazon's Alexa[2], or the Google Assistant[3], exhibit the same behaviour regardless of the individual user needs, ignoring the potential benefits of adaptation.

These two insights, (1) that users assign a personality to CAs and (2) that there is no such thing as a perfect personality, motivate the vision of this thesis:

> **Vision:** To improve the interaction between users and CAs by deliberately imbuing CAs with personality and tailoring them to user preferences.

In order to realise this vision, we must achieve two goals: First, we must identify suitable methods and tools for systematically imbuing CAs with different personalities. Second, we must examine user preferences for CA personalities to develop an adaptation mechanism [103]. Importantly, the two goals are mutually dependent: On the one hand, imbuing CAs with different personalities is primarily meaningful if users have different preferences. On the other hand, to examine which personalities users prefer, CAs must first be imbued with different personalities. In line with these goals, this thesis examines two guiding research questions:

RQ 1: *How to systematically imbue conversational agents with personality?*

RQ 2: *What preferences do users have for conversational agent personality?*

To answer the first research question, this thesis contributes two approaches to describe the CA's personality, and based on this, presents two methods to synthesise the CA personality. By *synthesising* a CA personality, I refer to the artificial generation of behaviour cues to emulate a personality [205]. In the first approach, we adopt a human personality model for describing CA personality and synthesise the personality by identifying and transferring

[1] `https://www.apple.com/siri/`, last accessed 10th May 2022
[2] `https://developer.amazon.com/alexa`, last accessed 10th May 2022
[3] `https://assistant.google.com`, last accessed 10th May 2022

human behaviour cues to CAs [Core1, Core2, Core3]. In the second approach, we develop ten dedicated CA personality dimensions [Core4] to describe CA personality and synthesise dialogues for them through a new method called Enactment-based Dialogue Design [Core5].

To answer the second research question, this thesis presents empirical evidence on user preferences for CA personality. Using a deductive approach, we collect users' likeability ratings of different CA personalities we created [Core2, Core3, Core5]. In an inductive approach, we introduce a new method to elicit user visions of a perfect CA [Core6, Core7]. Moreover, we investigate the influence of user personality on their preferences [Core2, Core3, Core5, Core6, Core7]. Based on these developed methods, a conceptual process to systematically endow CAs with personality is proposed.

In the following subsections, I present each of the two goals, elaborating on the theoretical motivation, the central research questions, and the contributions. I first introduce my approaches to imbuing CAs with personality and then present how these personalities impact user preferences.

1.1 Imbuing Conversational Agents with Personality

Although imbuing CAs with personality has been highlighted as a promising and powerful possibility to change how users will think, feel, and behave towards a CA [16, 28, 103, 123, 157], there is a lack of clear guidance on how to systematically design CA personalities [112, 179]. Whilst companies acknowledge the need for personality in CAs, they provide little information to designers as to how this personality can be infused in the CA. For example, the Google Assistant developer guide suggests listing four to six key adjectives that describe the CA's personality and then writing sample dialogues which best impersonate this personality.[4] This process neither gives the CA designer any insights into **which specific personality adjectives** are important to consider when devising a CA's personality nor does it explain **how the CA should behave** to convey the intended personality adjectives.

Today, personality design is most discernible in the popular voice assistants, for which teams of experts have spent years developing the personality of the respective agent [206]. Due to the increasing prevalence of CAs, I expect that new design teams and researchers will enter the development of CAs. In the light of the lack of CA design training in academia [156] and of accessible information about personality design [148], these new teams are likely to face several questions. First, what are adjectives that adequately and completely describe a CA personality? Listing arbitrary adjectives to define a personality could result in a lack of consistency of the CA personality, especially when multiple CA designers write dialogues that express the personality. As in inter-human communication, users expect consistent and hence predictable behaviour when interacting with CAs [100, 101, 123, 182]. For example,

[4] `https://developers.google.com/assistant/conversation-design/create-a-persona`, last accessed 10th May 2022

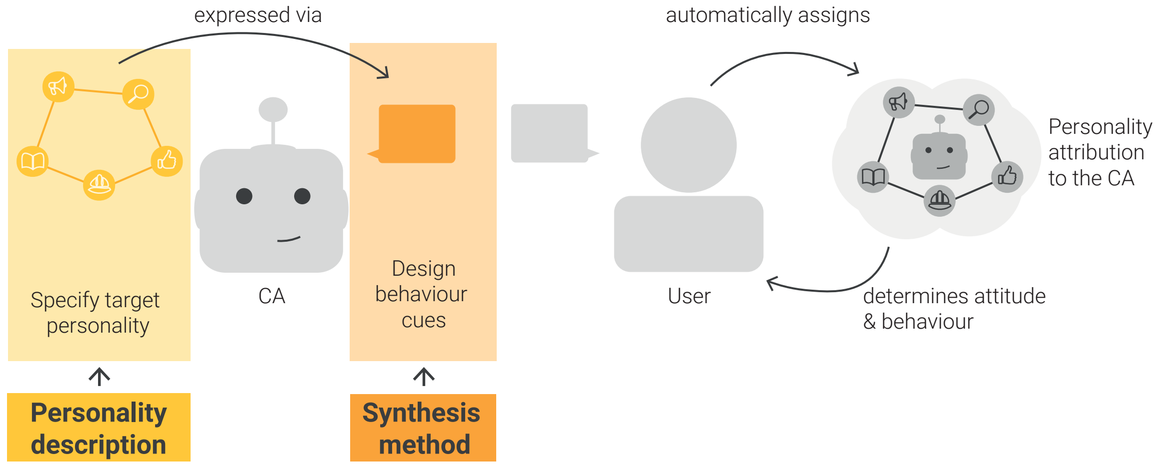

Figure 1.2: To empower CA designers to imbue CAs with personality, this thesis provides them with (1) an underlying personality description on which they may specify the CA's target personality (highlighted in yellow) and (2) synthesis methods for translating the target personality into perceptible behaviour cues (highlighted in orange).

a team of CA script writers may not explicitly specify how talkative a CA should be. As a result, one team member may write a dialogue for a chatty CA, whilst another may write a more taciturn dialogue for it, resulting in inconsistent behaviour that is likely to confuse users [100, 171]. To ensure a mutual understanding of a consistent personality within the CA design teams, I argue that we need an underlying **personality description** for CAs that delineates which personality traits have to be considered for the personality synthesis. Such a personality description can also support CA designers to systematically differentiate between multiple personality versions.

The majority of researchers have used human personality models as a basis for the design of the CA personality, such as the Big Five model [55] (e.g. by [36, 158]), the interpersonal circumplex model [215, 216] (e.g. by [153, 160]), or Eysenck's personality theory [74] (e.g. by [145]), but little is known about whether these *human* models are suited to describe CA personality. To derive a dedicated CA personality description, Kim et al. [112] conducted two design workshops, proposing three CA personality dimensions. Braun et al. [27] built the design of their in-car voice assistants on a two-dimensional model of personality obtained from a pilot study and related work. In yet another approach, Zhou et al. [224] used single personality adjectives, such as "cheerful" or "reserved", to inform the design of their interviewer chatbot. In summary, despite research efforts to establish a theoretical foundation for CA personalities, there has not yet been a consensus on an adequate personality description.

As personality is a latent construct [196], CA designers have to **synthesise perceptible behaviour cues** in the dialogue with the user to elicit the intended personality impression. These behaviour cues can be *verbal* (the linguistic content of the communication) and *non-verbal* (the accompanying bodily communication [10]) in the context of CA design. This thesis focuses on **verbal cues**, as the content of dialogue is core to every interaction with the user for all CAs,

	Personality Description	Synthesis Method
Approach 1:	Human Big Five personality model	Verbal cues transferred from human behaviour [Core1, Core2, Core3]
Approach 2:	Ten CA personality dimensions [Core4]	Enactment-based Dialogue Design [Core5]

Table 1.1: Overview of the two approaches we implemented to systematically describe and synthesise the personality of a CA.

regardless of whether they use speech or text. Hence, in order for CA designers to synthesise personality, we must first identify behaviour cues which express the intended personality.

Different approaches to synthesise personality in CAs have been implemented. There is an abundance of work on the relationship between *human* personality and perceptible behaviour manifestations (e.g. [149, 172, 196]), which previous work has leveraged to synthesise personality in speech-based CAs and robots (e.g. [36, 100, 145]). However, research lacks knowledge about the role of *verbal cues* to emulate the intended personality. The fact that this research on human behaviour cues is scattered across decades and disciplines (e.g. Linguistics, Psychology, Computational Linguistics, etc.), with no overview of all possible behaviour cues, impedes a systematic analysis of the transferability of these cues to CAs.

Another approach is in line with companies' methods for designing CA personality [206]. Braun et al. [27] employed scriptwriters who manually drafted a voice assistant's responses to specific questions. Although participants were able to distinguish between the versions of the voice assistant, a systematic comparison between the personality traits specified in the design and users' resulting perceptions of the personality was missing [27]. Another strand of research aims to automatically generate personality-infused CA responses [138, 184, 185, 186]. For example, Ritschel et al. [186] used reinforcement learning to generate varying amounts of Extraversion in a robot that adapts to the user based on their engagement. However, in these works, the users' perceptions of the generated personalities were not assessed. In summary, various approaches have been implemented to convey an intended personality, without reaching unequivocal conclusions as to which are most successful.

In the light of these research gaps, I argue that two crucial factors are missing to empower CA designers to imbue CAs with personality systematically: (1) conceptual clarity about the underlying personality description on which CA designers can specify the CA's target personality, and (2) synthesis methods for translating the target personality into perceptible behaviour cues (cf. Figure 1.2).

Contribution: To fill this gap, this thesis contributes two approaches, summarised in Table 1.1. These two approaches differ with regard to the underlying personality description, which in turn determines the synthesis methods. In the first approach, we turned to the most predominant description of personality for *humans*, the Big Five model, as the underlying personality description for CAs. The Big Five model is appealing because users apply the same social rules to CAs as to humans [157]. In line with prior research (e.g., [36, 158]), we identified a set of verbal behaviour cues associated with human personality from psycholinguistic literature as well as our own work [Core1] and implemented them in text-based CAs. In particular, we present two studies in which we used human verbal cues to infuse different levels of *Agreeableness* [Core2] and *Extraversion* [Core3] in a CA. After users interacted with the different personalities in a lab [Core2] and a field study [Core3], we evaluated user perceptions of the CA personalities against the intended personality designs. Our results showed that human verbal behaviour cues can be used to a limited extent to equip text-based CAs with personality. More details on the contributions of this approach may be found in Chapter 3.1.1. However, the findings raised doubts about the adequacy of the human Big Five model for describing CA personality, echoing previous work [133, 173, 224].

As our results in [Core3] suggested that the human Big Five model might not be sufficient to capture a complete picture of CA personality, we examined whether this personality model can be replicated for CAs by adopting the well-established psycholexical approach from Psychology. Specifically, we present the first systematic analysis of personality dimensions dedicated to speech-based CAs as an initial step to developing a CA personality model [Core4]. We contribute ten CA personality dimensions which do not correspond with the human Big Five model, thereby yielding much needed theoretical clarity on the necessity of dedicated personality descriptions for CAs. Based on these dedicated CA personality dimensions, we then developed a method, called *Enactment-based Dialogue Design*, which combines approaches from industry on role-playing dialogues between the CA and the user[a] with a theoretical personality foundation. In particular, we showed how focus groups with amateur actors can be used to develop dialogues that express different levels of personality dimensions, which we evaluated against the specified personalities in an online survey. More details on the contributions of this approach may be found in Chapter 3.1.2.

[a] `https://developers.google.com/assistant/conversation-design/write-sample-dialogs`, last accessed 10th May 2022

1.2 User Preferences for Conversational Agent Personality

As described before, the choice of the CA's personality impacts users' attitudes towards the CA and their interaction behaviour (e.g. [27, 157, 224]). Similar to inter-human relationships, Reeves and Nass [182] suggested that there is no such thing as a computer personality that is universally liked. Echoing these findings, Braun et al. [27] ascertained that different users found different in-car voice assistant personalities most likeable, trustworthy, and useful in a real-world driving study. Not only are there individual differences in users' attitudes towards

CAs, but these individual preferences are also reflected in their behaviour. For example, whilst an extraverted chatbot led to longer interaction times and higher sales for users with a congruent personality in a telecommunications service context, an introverted chatbot had the same effect for other users [200].

Despite these findings, companies still offer a one-size-fits-all solution when it comes to CA personalities [206]. For example, the Google Assistant was designed to be perceived as "quirky", "helpful", and a "hipster librarian" [206], independent of individual user preferences. A reason for this – apart from the difficulty of designing different CA personalities (cf. Section 1.1) – is presumably that tailoring the CA personality to the user can also backfire, as the user may accept a mismatched CA less than a default version [27]. Thus, knowledge about user preferences is crucial for tailoring the CA personality to the user.

To shed light on the nature of user preferences, researchers turned to theories from Social Psychology. In inter-human relationships, people show preferences for others with similar personality traits, termed the *Similarity Attraction Paradigm* [34], which entered the vernacular through the saying "Birds of a feather flock together". Early experiments on voice user interfaces led Nass and Brave [157] to conclude that similarity attraction is also a powerful predictor for how much users like a CA personality. For example, synthetic voices on a fictional book reviewing website and on an online auction portal that had a matching level of Extraversion to the user resulted in higher likeability [158] and social presence [124]. This similarity attraction effect has been corroborated for Extraversion in an embodied virtual real estate agent [21] and a consumer service text-based CA [200].

However, there is also research pointing to findings inconsistent with the similarity attraction paradigm. Although Andrist et al. [9] showed that users complied more with a congruent robot personality, this matching did not influence their subjective preference. Conversely, Isbister and Nass [100] found that users significantly preferred and had more fun in a lab game with an embodied virtual agent that manifested an opposite level of Extraversion to themselves. In contrast to these simple similarity or complementary matching methods, explorations of user preferences for in-car voice assistants [27] and interview chatbots [224] suggested that tailoring the CA personality to the user's personality is beneficial, but requires complex adaptation mechanisms. For example, Braun et al. [27] assigned users a voice assistant personality based on a decision tree analysis which featured multiple user personality traits. Contrary to these findings, Cafaro et al. [36] as well as Ruane et al. [190] did not observe any effect of users' personality on their preference or interaction behaviour for an embodied virtual museum guide [36] and a chatbot [190].

The majority of this prior work demonstrating personality matching effects focused on the dimension Extraversion (e.g. [100, 158, 200]). However, a similarity or complementary attraction effect seems less likely for dimensions such as Agreeableness and Conscientiousness in the context of today's CAs which usually embody virtual assistants. Users who are more antagonistic and disorderly may still prefer a friendly, helpful, and reliable assistant rather than one mirroring their personality.

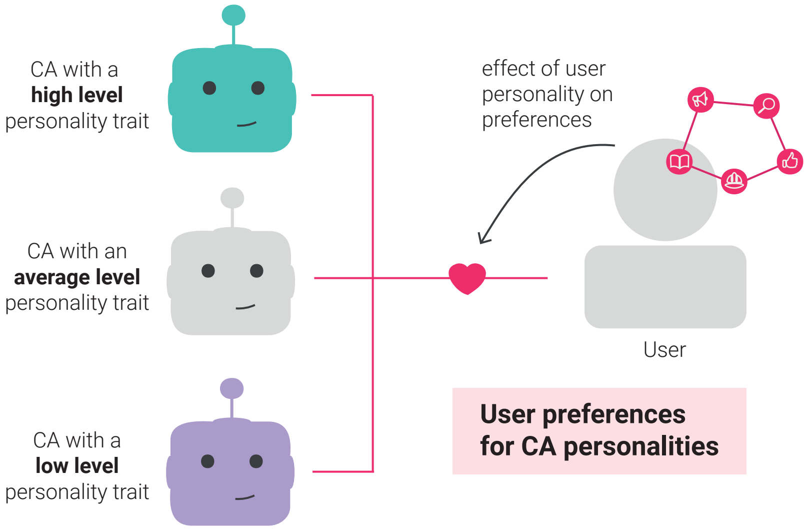

Figure 1.3: To inform the process of tailoring CA personalities to users, this thesis examines user preferences for different CA personalities. For each personality trait, three CAs, each with a distinct level of that personality trait (high, average, low) are synthesised and presented to the users for interaction.

An important question concerns not only the kind of popular personality traits for CAs but also the extent to which these traits are held. Prior work has examined user preferences for CA personality by means of discriminating two opposing versions at the outer poles of the personality continuum, such as an introverted and extraverted voice user interface [157]. However, the distribution of human personality traits is expected to follow a Gaussian curve in the population [144], with most people having rather average values on the personality continuum instead of extremes. Due to the aforementioned work pointing towards similarity attraction effects, we synthesise and examine **three levels of personality manifestations**, including an additional *average* level to allow for a more detailed adaptation.

In summary, previous work provided conflicting findings regarding user preferences for CA personalities, in particular with respect to the role of *user personality* in their preferences. To resolve these conflicting findings, we examined user preferences not only for two opposing, extreme levels of personality but also for an average, more subtle level of personality (cf. Figure 1.3). Assuming a similarity attraction effect, such an average level of personality should better reflect the majority of users. Furthermore, we collect user preferences for personality dimensions such as Agreeableness and Confrontational, for which a similarity attraction effect seems less likely, so as to investigate whether unanimously liked personality preferences emerge.

Contribution: We present three publications in which we investigated user preferences for CAs that exhibit *three levels* of different personality traits, namely Agreeableness [Core2], Extraversion [Core3], Social-Entertaining [Core5], and Confronta-

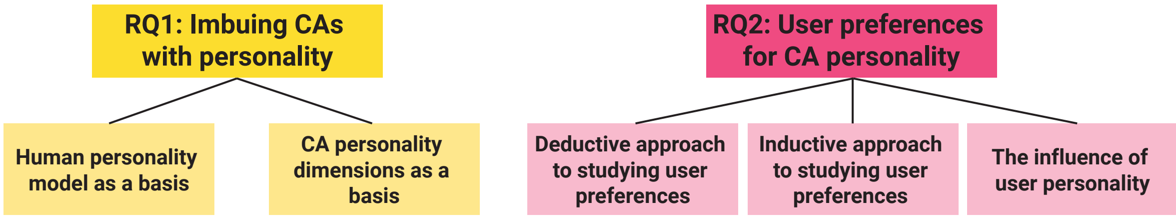

Figure 1.4: Structure of this thesis based on the two guiding research questions. For each guiding research question, two approaches are presented.

tional [Core5]. In these studies, we analysed the user preferences with a deductive approach, that is, we first synthesised different personality versions and then asked users to interact with and rate them. These studies confirmed that users have very individual and conflicting preferences for CA personalities, underlining the benefits of creating multiple personalities and tailoring them to the user. Although we found some evidence for a relationship between user personality and their preference for CA personality, this relationship was less strong than expected based on prior research. More details on our contributions and findings can be found in Chapters 3.2.1 and 3.2.3.

This deductive approach to examining user preferences, which has also been employed in prior work (e.g. [27, 124]), involves little user engagement. Specifically, users are only involved in the design of CA personalities in the last evaluation step after different personalities have already been synthesised, ignoring their perspectives during the requirement collection. To this end, we contribute a new inductive method so as to elicit user visions of dialogues with a perfect CA, termed *Vision Dialogue Elicitation* for the purpose of this thesis. This method aims to engage end users in defining requirements for the development of CA personalities [Core6, Core7]. In addition, we present an application of this method for the context of domestic smart speakers (cf. Chapter 3.2.2). A better understanding of user visions yields much needed knowledge to identify CA personality traits that are universally liked or only by a subset of users, informing CA designers about which personality traits should be tailored to individual user needs.

1.3 Summary and Overview of the Thesis

The goals of this thesis are to (1) develop methods for imbuing CAs with personality systematically and (2) explore user preferences for different CA personalities. I present two approaches to achieving each of these goals. To imbue CAs with personality, I first used the human Big Five personality model as an underlying description of CA personality and transferred human behaviour cues to synthesise personality. As a second approach, I introduced ten dedicated CA personality dimensions to describe CA personality and developed the synthesis method Enactment-based Dialogue Design based on these dimensions.

To examine user preferences for CA personalities, I first asked users to interact with different CA personalities and then collected their preferences (deductive approach). As a second approach, I elicited user preferences bottom-up through Vision Dialogue Elicitation (inductive approach). Furthermore, I investigated the influence of user personality on their preference for CA personality for both approaches. These two goals and respective approaches scaffold the structure of this thesis, which is visualised in Figure 1.4.

Chapter 2 introduces definitions of the central terms used throughout this thesis and provides the theoretical background. Chapter 3 presents the publications included in this thesis and clarifies how they contribute to the guiding research questions. Finally, I position and reflect on the contributions of this thesis in Chapter 4 and propose a research agenda for future work.

2
BACKGROUND AND DEFINITIONS

This chapter defines the two central terms that comprise the title of this thesis and occur throughout: conversational agents and personality. I first review both terms separately and then connect them to explain my assumptions on how users attribute personality to CAs. These definitions set the context of my work and provide readers with background information. Relevant related work is also presented in the respective publications.

2.1 Conversational Agents

Language is the most crucial channel of communication between people and is used in all cultures to communicate and build relationships [176]. Hence, it is not surprising that since the invention of computers, people have been fascinated by interacting with them via natural language [99]. CAs that allow users to communicate in a similar way as to other people have existed since the 1960s [148]. Since then, several research and product strands have emerged, both in academic and industrial labs [147], including the development of natural language understanding (NLU) datasets [3, 93], *Interactive Voice Response* (IVR) systems to automate telephone customer support [171], embodied CAs [42], and chatbots [78]. A new era of speech-based CAs was heralded by commercially available *voice assistants* at the beginning of the 2010s, fuelled by advances in natural language technologies [147]. Today's CAs have permeated people's daily lives across a range of devices such as smartphones, smart speakers, and computers [50, 178], and are ubiquitously integrated into people's homes [178], online shopping [110], mental health support [136], e-learning [80], and automotive user interfaces [27]. For an introduction to voice user interfaces and their design challenges, I refer to the corresponding chapter in our textbook on Human-Computer Interaction (HCI) [Pub7].

2.1.1 Definition of Conversational Agents

Despite the growing ubiquity of CAs, there are no generally applicable definitions. Dale [60] describes that *conversational agents* "achieve some result by conversing with a machine in a dialogic fashion, using natural language." CAs are often equated with *dialogue systems*, both conceptually and in nomenclature [109, 121, 135, 146, 148]. McTear [146, 148] defines a dialogue system as "a computer program that supports spoken, text-based, or multimodal conversational interactions with humans." Dale and McTear's definitions have also been adopted by other HCI researchers [75, 78, 84]. In Computer Linguistics, Jurafsky and Martin [109] refer to systems which "communicate with users in natural language (text, speech, or both)". Despite minor differences, all of these definitions are in agreement that interaction with a CA occurs through natural language.

For the purpose of this thesis, I introduce a narrower definition based on the two separate terms (agents and conversations) to better distinguish CAs from other related concepts.

Agents There are many definitions of what an *agent* constitutes. Wooldrige and Jennings [218] summarise these definitions in four minimum requirements for a software program to be autonomous, sociable, reactive, and proactive. Maes [137], who coined the agent term for the HCI community, defines agents as "computational systems that inhabit some complex dynamic environment, sense and act autonomously in this environment, and by doing so reali[s]e a set of goals or tasks for which they are designed", stressing the importance of autonomy. Following these definitions, this thesis uses the term *agents* to describe autonomous entities, thereby differentiating them from simple text-to-speech voice user interfaces (e.g. announcements in trains). Whilst I use agents as an umbrella term also for mechanical agents (e.g. robots), I consider CAs as software-based agents.

Conversations The Oxford English Dictionary (OED)[1] defines a *conversation* as "an interchange of thoughts and words; familiar discourse or talk". Inter-human conversations are an intricate phenomenon, making their imitation a challenging undertaking for human-CA interaction [109]. A conversation typically consists of a sequence of turns, with each turn representing one of the interlocutor's speech acts [11]. These speech acts are often structured as adjacency pairs [194]. That is, one interlocutor's speech act, for example, a question, is followed – at some point in the conversation – by the other speaker's response. The first part of an adjacency pair may be initiated by either of the speakers, and the initiator role typically changes frequently in the course of the conversation [208]. The goal of a conversation is to construct common ground and converge on agreement [69].

Interactions with today's task-oriented CAs try to mimic inter-human conversations but usually do not meet these expectations of a conversation [51, 178]. In this thesis, I nevertheless use the term *conversational agent* because (1) it underlines the difference from bots that merely automate tasks [148], (2) personality adaptation seems particularly meaningful in the context of multi-turn interactions that comprise social components, and (3) *conversational user interfaces* is eponymous for the newly introduced conference[2] in the HCI community, which focuses on the topics covered in this thesis. Despite using the designation *conversational agents* alluding to this thesis' vision, I will refer to the actual interaction between user and CA as *dialogue*[3], as this more technical term better reflects the scripted nature of today's CAs.

Based on the aforementioned characteristics of agents and conversations, I provide the following definition:

[1] `https://www.oed.com/view/Entry/40748`, last accessed 10th May 2022

[2] `https://www.conversationaluserinterfaces.org`, last accessed 10th May 2022

[3] `https://www.oed.com/view/Entry/51915`, last accessed 10th May 2022

Definition: Conversational agents are autonomous software-based entities that interact with users through natural language in multi-turn dialogues that mimic the conversations between humans.

2.1.2 Types of Conversational Agents

There is a plethora of different categorisations of CAs (e.g. [75, 84, 109, 135, 146]). Jurafsky and Martin [109] distinguish between *task-oriented dialogue agents* and *chatbots*. The former, such as today's voice assistants, use conversational interaction to help the user complete a task. Interactions with these task-oriented dialogue agents are typically initiated by the user and consist of a few turns characterised by adjacency pairs [51, 83]. In contrast, interactions with a chatbot pursue the goal of informal, unstructured chats that often serve to entertain [109]. However, the boundaries of this classification are increasingly blurred to the extent that the two terms are often used interchangeably in HCI research (e.g. [60]), with the notion of chatbots referring primarily to *text-based* CAs (e.g. [75, 84]). Furthermore, Bickmore and Cassell [20, 21] argued that to be trustworthy, a CA must both perform a task and be able to engage in social talk.

A further distinction of CAs can be made based on their embodiment and interaction modality, which play an important role in the set of cues that may be used for personality manipulation. First, we may classify CAs in terms of whether they are embodied or not. Embodied conversational agents (ECAs) not only use natural language to interact with the user but also non-verbal behaviour through a visual human-like representation [43].[4] As ECAs are often developed to provide social companionship and entertainment to users, this research strand has played a pioneering role in integrating human-like features into the design of CAs [147]. Hence, I include examples from work on ECAs in this thesis, although the studies focus on CAs whose primary means of interaction is language. Within these disembodied agents, we may further distinguish based on whether the modality for input and output is written or spoken natural language.

2.1.3 Implementation of Conversational Agents

Technically, today's CAs are usually composed of (1) an NLU module, (2) a dialogue manager, and (3) a response generation module, with speech-based CAs additionally requiring automatic speech recognition and text-to-speech synthesis units [109, 147]. For each component, there is a range of implementation approaches that vary in sophistication [109]. The goal of the NLU module is to interpret the user's input and extract meaningful slots [147]. The simplest way to do this is to recognise keywords, with more complex approaches using neural

[4] Disembodied CAs can also have a kind of visual representation, such as the Amazon Echo's blue circle that lightens up when hearing the wake word or Siri's pulsing abstract animation. However, these representations are typically not an embodiment of humans but rather serve to give the user feedback that the device is listening.

networks to better compensate for different formulations of the query [155]. Based on the interpretation of the user's intent, the dialogue manager decides on the next action and the CA's response [147]. Most of today's dialogue managers in commercial CAs are rule-based, with research developing statistical approaches that scale better to a variety of possible actions but require large amounts of training data [155]. Afterwards, the response generation module translates the next action into words. A simple implementation for clearly defined application areas relies on pre-written scripts or templates, whereas more elaborate realisations adopt natural language generation techniques to allow greater flexibility in the dialogue design [147]. However, this type of interaction also harbours certain dangers as the developer no longer has control over the CA's output [103], which led to serious issues in the past, such as Microsoft's Twitter chatbot Tay which adopted inflammatory speech from interacting with users.[5]

Summary: Conversational Agents

For the purpose of this thesis, I define CAs as autonomous software-based entities that interact with users through natural language in multi-turn dialogues that mimic the conversations between humans. The focus of this thesis is on disembodied CAs which interact via either spoken or written natural language, which I will refer to as *speech-based* and *text-based* CAs, respectively. Depending on the specific characteristics of the CA, further designations may be introduced in the individual publications.

Following Jurafsky and Martin's classification [109], this thesis focuses on task-oriented CAs, which are, however, also capable of social conversation, albeit not their primary purpose. As our developed CAs are employed in constrained applications, we use rule-based implementations to ensure a controlled environment and expression of the personalities. More specifically, the implementations in [Core2, Core5] use the NLU modules provided by the respective implementation platforms Botpress[a] [Core2] and Amazon Alexa[b] [Core5], whereas the text-based CA in [Core3] employs keyword recognition. To ensure that the CA personalities are expressed consistently [100], all implemented CAs use pre-written scripts instead of natural language generation. Further implementation details may be found on the respective project websites.

[a] `https://www.botpress.com`, last accessed 10th May 2022

[b] `https://developer.amazon.com/alexa`, last accessed 10th May 2022

2.2 Personality

Personality describes enduring dispositions, so-called traits, which manifest themselves in distinctive patterns of behaviour, emotion, and cognition [55]. Decades of Psychology

[5] `https://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist`, last accessed 10th May 2022

research have shown that personality traits are relatively consistent across contexts, time, and observers [142]. Hence, personality can predict important life outcomes on an individual, interpersonal, and social level [169]. Examples include physical health [41], subjective well-being [65], information seeking behaviour [92], trust in technology [73], peer [104], family [18], and romantic relationships [66], academic success [115], job performance [108], and political attitudes [107]. Due to this important role of personality in people's preferences, merchants in ancient times already used their knowledge of their customers' individual differences to tailor products and services to their needs [1]. Even today, recruiters as well as dating apps collect information about people's personalities to ensure a match [224], whilst recommender systems and social networks increasingly use people's personality profiles to provide personalised recommendations [76, 140]. For an overview of how personality traits can inform the design of technology, please refer to our work in [Pub10].

2.2.1 Describing Personality

The comprehensive understanding and formal description of personality have been a central subject of Psychology for centuries [70]. The *Big Five* model, also referred to as *OCEAN* or *Five Factor* theory, has emerged as the most predominant personality paradigm in scientific research [62, 141]. It was developed using the psycholexical approach, which assumes that personality differences are encoded in language [86]. Five broad dimensions comprise the Big Five model, which are composed of multiple sub-facets on polar scales [55, 86, 142]:

- ***O**penness* reflects a tendency to be open to fantasy, aesthetics, feelings, actions, ideas, and values.
- ***C**onscientiousness* reflects a tendency to be competent, orderly, dutiful, achievement-striving, self-disciplined, and deliberate.
- ***E**xtraversion* reflects a tendency to be warm, gregarious, assertive, active, excitement-seeking, and positive.
- ***A**greeableness* reflects a tendency to be trustful, genuine, altruistic, compliant, modest, and tender-minded.
- ***N**euroticism* reflects a tendency to be anxious, easily angered, depressed, self-conscious, impulsive, and vulnerable.

There has been some controversy over the validity of this personality approach, with other personality models, such as Eysenck's three traits [74], being proposed. However, as the criticisms mainly revolve around the biological foundation of the Big Five traits [142], which is irrelevant for the description of *artificial* personality in CAs, and as the Big Five are widely accepted in the computing community [175, 205], this thesis adopts the Big Five personality model for describing human personality and partly also for describing CA personality.

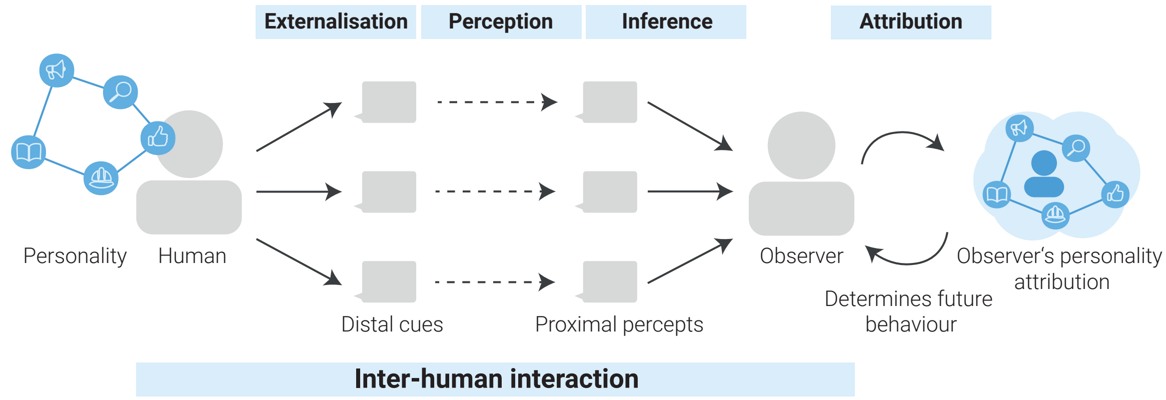

Figure 2.1: Scherer [196] adapted the Brunswik lens model to explain how humans externalise and attribute personality. As personality is a latent construct, humans express their personality via observable and measurable distal behaviour cues. These distal cues undergo a perception process (dotted arrows) and are then perceived as proximal percepts by the observer who attributes the interlocutor a personality based on these inferred percepts. This personality attribution influences the observer's future behaviour and attitude towards the interlocutor.

2.2.2 Externalising and Attributing Personality

Personality is a latent and inward construct; that is, personality traits can neither be gauged directly by observers nor through introspection. Instead, people externalise their personality through perceptible traces of behaviour [142]. These perceptible behaviour cues encompass the linguistic content (*verbal cues*, e.g. choice of words) and the accompanying bodily communication (*non-verbal cues*, e.g. facial expressions, or gestures including *para-verbal cues*, e.g. speech rate, pitch) [10]. Conversely, observers attribute personality to others by making inferences from these perceptible behaviour cues [142]. This personality attribution occurs rapidly, automatically, and unconsciously within the first seconds of an encounter, and people use this judgement to predict the other's behaviour, thereby strongly influencing all subsequent interactions [196].

To explain the personality externalisation and attribution process, Scherer [196] modified Brunswik's lens model, which originally proposed how organisms perceive information in their environment (cf. Figure 2.1). According to this model, human personality traits manifest themselves in indicator cues, called *distal cues* (*Externalisation* part in Figure 2.1). These distal cues undergo a perceptual process and are grasped by the observer, who represents them as *proximal percepts* [195] (*Perception* part in Figure 2.1). In contrast to objectively measurable distal cues, proximal percepts describe what the observer actually perceives. For example, extraverted people speak with high vocal energy and intensity (distal cue), which can be objectively measured. Observers/listeners subjectively perceive this high vocal energy as loudness (proximal percept) [195, 196]. These proximal percepts in turn lead to a cognitive inference process (*Inference* part in Figure 2.1), based on which the observer attributes a personality to the speaker (*Attribution* part in Figure 2.1) [195].

Those distal cues that untrained observers accurately perceive and consistently interpret, resulting in truthful attributions of personality traits, are termed *personality markers*. Identifying these personality markers has a longstanding tradition in Psychology and Linguistics [25, 39, 40, 63, 81, 149, 167, 170, 172, 192, 196]. For example, extraverted individuals are characterised by a higher total verbal output [39, 40, 149, 170, 192] and speaking more loudly [196]. Whilst Agreeableness is associated with using more positive emotion words, people high in Neuroticism tend to adopt more negative words [172]. Readers are invited to find more examples of personality markers in publications [Core1], [Core2], and [Core3].

2.2.3 Assessing Personality

As part of my analysis of user preferences, I examine if preferences for CA personalities are determined by a user's own personality. To this end, it is necessary to gain knowledge about the user's personality. A basic assumption of the personality trait approach is that personality can be quantitatively measured [142]. Traditionally, psychologists used standardised personality inventories, such as the Revised NEO Personality Inventory (NEO-PI-R) or the Big Five Inventory (BFI), in forms of self- or peer-reports to obtain information about people's personalities [26, 139]. With the availability of extensive online behavioural data, researchers presented methods to automatically infer users' personality traits from digital footprints, such as blog entries [82, 205], social media use [12, 23, 116, 119, 180, 197], images [46], or music preferences [77, 163]. In interdisciplinary research projects, we showed that personality [Pub4, Pub6] and similar stable traits [Pub5] also manifest themselves in smartphone usage. The chatbot platform Juji[6] suggests that users' personalities can be inferred from a conversation with their interview chatbot [224], which we used in publication [Pub8].

This work demonstrates that human personality manifests itself in users' online behaviour. Hence, I assume that with more online data available, these predictions will become more accurate and thus provide a robust basis for tailoring the CA personality to the user. As today's personality assessments are still lacking accuracy [Pub8], the publications presented in this thesis use standardised inventories to gauge user personality to avoid compromising validity due to unreliable personality measurements.

Summary: Personality

Our personality influences our behaviour and preferences [55]. Conversely, our perception of others' personalities influences how much we like to interact with them [33]. This importance of personality for interpersonal interaction motivates me to examine both user preferences for certain CA personalities and the influence of users' own personalities on these preferences.

[6] www.juji.com, last accessed 10th May 2022

To this end, this thesis adopts the human Big Five model to describe personality and uses the corresponding personality inventories to assess our participants' personalities. As a latent construct, personality manifests itself through observable behaviour cues [196]. This thesis follows Scherer's adaptation [196] of the Brunswik lens model to explain the process of externalising these behaviour cues and interpreting them to attribute personality to others.

2.3 Conversational Agent Personality

As artificial machines, CAs *cannot actually have* an inward personality, albeit fictional pop culture has suggested otherwise, for example in the movie *Her*, in the TV show *Black Mirror*, or in the book *Origin* by Dan Brown. However, Bates [16] argued there is much to learn from artists in film for the design of agents. In 1994, he postulated that agents must be believable, that is providing the "illusion of life" [16] so as to engage their users. Whilst early agent research sought to create this illusion by endowing agents with human-like intelligence, artists in character animation drew on other aspects of humanity [16]. For example, according to Disney animators Thomas and Johnston [204], one of the most important requirements for people to care about a virtual character is its personality.

These early calls for personality in computers were given emphasis by a series of experiments by Reeves and Nass [182], which demonstrated that people unintentionally respond naturally to computers and media, accumulating in the *Media Equation* theory (media equals real life). Due to the evolution-derived role of voice and language in social interaction, humans are experts in extracting social information from language [97]. People are then attuned to draw conclusions about technology-generated speech and apply the same behaviour rules that they use when interacting with other humans, overwhelming the rational knowledge that CAs are not people [161]. These social responses to CAs and other computers are known as the *Computers are Social Actors* (CASA) paradigm [159, 161]. One of these social responses is that users automatically, rapidly, yet often unconsciously attribute CAs a personality.

Based on the Media Equation and the CASA paradigm, I assume that the processes of personality externalisation and attribution in CAs work in a similar way as in inter-human interaction. In line with prior research on personality computing [175, 205], this thesis adopts Scherer's modification [196] of the Brunswik lens model as a theoretical basis and proposes an altered version for the purpose of describing *conversational agent* personality externalisation and attribution. Figure 2.2 illustrates this altered model and demonstrates starting points for designing CA personality which I describe below.

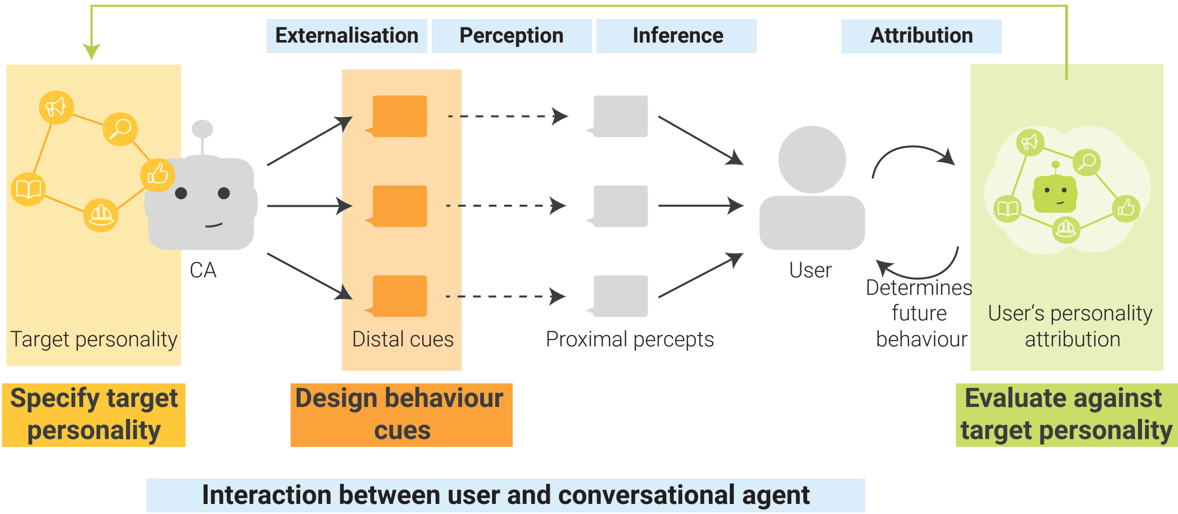

Figure 2.2: Based on Scherer's [196] adaptation of the Brunswik lens model, I assume that CA personality is expressed via distal cues that users interpret as proximal percepts to infer a personality that they attribute to the CA during the interaction. CA designers specify a target personality for the CA using a personality description (marked in yellow). This target personality is then synthesised by deliberately designing distal behaviour cues (marked in orange). Finally, the user's perception of the CA personality has to be compared with the target personality to evaluate the success of the design (marked in green).

2.3.1 Describing and Synthesising Conversational Agent Personality

Just as in inter-human interaction, CAs express distal cues that the observing user infers as proximal percepts which in turn form the user's basis to attribute personality to the CA. For example, a CA's voice is set to a speech rate of more than 210 words per minute (distal cue) [158], which is perceived by the user as fast speaking (proximal percept). Based on this perception, the user attributes a personality to the CA. In contrast to human personality, these distal cues are not an automatic externalisation of the CA's latent personality but can be deliberately manipulated to elicit the impression of an intended personality. For example, the CA designer could reduce the speech rate to create the impression of a calmer CA.

More specifically, the design of CA personalities begins with specifying its target personality (among other characteristics of its persona) [171, 224]. With *target personality*, I refer to the personality that CA designers[7] have stipulated as the desirable traits their CA should possess based on user requirements, the application context, and the company brand (cf. yellow part in Figure 2.2). For example, CA designers might specify that a voice assistant in a sports car should have a casual and smart personality, whereas a chatbot for a banking application should show professional and reserved conduct. As stated in the introduction, a personality description tool is crucial to allow CA designers to specify a consistent target personality.

[7] Due to the novelty of commercial CAs, there is little information on the composition of CA design teams. In reality, it is more likely that target personalities are not only developed by designers but by diverse teams also consisting of researchers, product managers, engineers, marketing specialists, etc. For the sake of brevity and simplicity, this thesis will henceforth use the term *CA designer* as a placeholder for all people working on CAs.

Based on this specification, CA designers have to deliberately design the CA's language, that is its distal cues, in a way that they elicit the attribution of the target personality (cf. orange part in Figure 2.2). Automatically generating these distal cues was termed *Automatic Personality Synthesis* [205]. Before being able to *automatically* generate these distal cues, we must first find methods to design behaviour cues which are associated with CA personality traits. The aim of this deliberate cue design is that the personality traits the user attributes to the CA based on their perception of the behaviour cues correspond to the target personality (cf. green part in Figure 2.2), which I will describe in the following subsection.

2.3.2 Evaluating the Perception of Conversational Agent Personality

To ensure that the personality synthesis was successful, the user's impression of the personality has to be compared with the target personality. Prior work on CAs, ECAs, and robots has primarily performed three different methods for this comparison. On the one hand, user perceptions of the agents' personalities were gauged by means of a variety of human personality inventories, such as the Big Five Inventory (BFI) [106] (e.g. [133]), Wigging's Extraversion scale [215] (e.g. [9, 100, 125, 158]), single scales of the Ten Items Personality Inventory (TIPI) [87] (e.g. [164, 165]), Saucier's Mini Markers [193] (e.g.[36, 37]), the Eysenck Personality Questionnaire Revised [79] (e.g. [145]), or by implicitly measuring how similar participants perceived the agent to be to themselves [5]. Apart from these standardised measurements, several comparisons included custom Likert scale questions [153, 160, 191, 224], or semantic differentials with descriptive adjectives [27]. In another approach, users' personality attributions were captured via open-ended questions, for example by describing the personality in a free-text field [190], or collecting the first adjectives or keywords that came to mind [36, 37, 133, 182, 224].

Liu et al. [133] compared open descriptions with standardised personality inventories and concluded that the two methods complement each other but may result in different conclusions. Open-ended descriptions facilitate collecting participants' impressions of the most relevant and salient personality traits without biasing them against predefined characteristics [133, 182]. In contrast, inventories allow for easier comparison between different agents and for capturing the complete perception on all personality dimensions [133].

Summary: Conversational Agent Personality

Based on the Media Equation and the CASA paradigm, I assume that users attribute personality to a CA in a similar way as they attribute personality to another human. CA designers have the opportunity to steer this personality attribution towards a target personality that is tailored to user needs by deliberately shaping the personality perception. To equip CA designers with the tools and methods to achieve a desired target personality, we need (1) a personality description on which the target personality is specified, and

(2) synthesis methods to translate the target personality into perceptible behaviour cues. In Chapter 3.1, I present two approaches for the personality description underlying this personality specification and two synthesis methods.

Defining a target personality and then synthesising it also requires a final evaluation of whether this synthesis was successful, as explained in Section 2.3.2. Due to Liu et al.'s conclusion [133] on the complementary effects of personality assessment instruments, the projects in this thesis use both standardised personality inventories [Core3, Core6] and open personality descriptions [Core3, Core4], along with rating agents on a number of personality adjectives [Core5]. A systematic comparison of different evaluation methods is beyond the scope of this thesis, but I will reflect on the employed methods in Chapter 4.

3

CONVERSATIONAL AGENTS WITH PERSONALITY

In this chapter, I present the publications included in this thesis. A synopsis of all publications and their primary contributions is provided in Table 3.3 at the end of this chapter. The following sections present a detailed overview of the individual publications, including the motivation, artefacts, research design, and main findings. The sections are structured according to the two guiding research questions (cf. Figure 3.1), with the specific research questions for each publication introduced throughout the chapter as they emerge from the previous research findings. Figure 3.1 illustrates this structure to help the reader navigate through the chapter.

In this chapter, each publication is featured by a blue box detailing the reference. Three publications comprise both the synthesis of CA personality, which is addressed in RQ1, and the analysis of user preferences for these synthesised CAS, which is addressed in RQ2. These papers are presented twice with their respective contributions to each of the guiding research questions. A paper that has been introduced before is marked by a grey box.

Due to the collaborative nature of research, all of my publications have been completed in a joint effort with fellow researchers and students. Table A.1 clarifies the contributions of all co-authors.

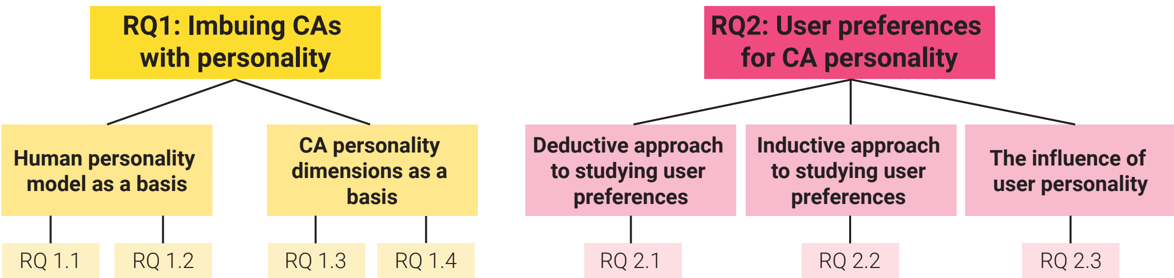

Figure 3.1: Structure of this thesis based on the two guiding research questions. For each guiding research question, two approaches are presented.

3.1 Imbuing Conversational Agents with Personality

To realise my vision of CA personalities that adapt to user preferences, researchers and designers have to be able to create different personalities for these agents. Hence, the first goal of this thesis is to explore methods for systematically imbuing CAs with certain personality traits, or to answer RQ1:

RQ 1: *How to systematically imbue conversational agents with personality?*

To address this research question, this thesis aims to (1) provide conceptual clarity on a **personality description** for CAs, and (2) develop **methods for synthesising** the personality based on this personality description, as outlined in Chapter 2. This section presents publications that make contributions towards this goal, organised according to two approaches based on the underlying personality description.

Specifically, we first chose a human personality model as the underlying personality description, in line with prior work's assumption that "synthetic personality equals human personality" [157]. Following this assumption, we investigated if human behaviour cues can be transferred to CAs to elicit personality perceptions similar to those of humans. As the results indicated that the perception of CA personality differs from that of human personality, we questioned the use of the human personality model for describing CA personality and instead developed ten dedicated CA personality dimensions in the second approach. Based on this new underlying personality description, we explored a new method for synthesising dialogues that evoke the intended personality perceptions.

3.1.1 Human Personality Model as a Basis

Echoing previous work on user perceptions of CAs as social actors [153, 161], we initially adopted the most predominant model for *human* personality, the Big Five model (cf. Chapter 2), as an underlying personality description. Adopting the Big Five model as a basis for describing CA personality offers the advantage that we may draw on a plethora of work in Psychology and Psycholinguistics that has investigated how the Big Five traits manifest themselves in people through observable behaviour cues (e.g. [172, 196]). Related work on robots [5, 9, 29, 187], ECAs [2, 8, 36, 45, 100, 118, 145, 211], and speech-based CAs [27, 157, 160, 164, 168, 174] has leveraged this relationship to transfer human behaviour cues to these agents, imbuing them with a personality. This work has examined the transfer of numerous non-verbal and para-verbal cues; either single cues [9, 36] or in combination [118, 145].

However, previous research has paid less attention to the importance of verbal cues in imbuing CAs with personality, although verbal dialogue plays a major role in most interactions with CAs. In fact, verbal cues, such as choice of words and text length, are the only resource available for synthesising personality in text-based CAs. In the light of this, this thesis focuses on verbal cues to impart personality to CAs. Concretely, we focus on *text-based* CAs in this first approach to circumvent the unintended influence of non-verbal cues on user perceptions of personality. Moreover, whilst today's voice user interfaces still often lack perfect voice recognition [Pub1], text-based CAs can be more easily controlled.

Hence, the objective of the research presented in this subsection is to explore the transfer of human verbal behaviour cues as a method for synthesising personality in text-based CAs. To attain this objective, we need to (1) collect a set of suitable human personality markers[1] from

[1]As a reminder, a *personality marker* is a distal cue that observers accurately perceive, consistently interpret, and attribute to a truthful personality trait, e.g. a fast speech rate is a personality marker for high Extraversion [196].

the literature, (2) implement these cues in the CA, and (3) evaluate whether these cue-infused CAs elicit the same personality attributions as humans. A review of the literature revealed that whilst there is an abundance of correlations between human personality and verbal cues, the effect sizes are small to moderate [154, 172]. Consequently, the set of verbal behaviour cues is limited and might be less expressive than non-verbal cues, raising the question of whether verbal cues are sufficient to imbue text-based CAs with personality.

The problem of restricted expressiveness is also well-known in written human-human communication, causing more ambiguous interpretations of messages [32, 129, 209]. In text messaging, emojis are commonly used as a surrogate for non-verbal cues and have been shown to be associated with the sender's personality traits in the context of social media [13, 117, 131]. As a way to overcome the limitations of verbal cues, this thesis sets out to explore how emojis can be used by text-based CAs to convey intended personality traits. Thus, to expand the set of potential verbal cues for text-based CAs, we first examined if there are personality markers in the way *humans* use emojis:

RQ 1.1: *How do people's personality traits relate to their emoji use in interpersonal written communication?*

[Core1]: Human Personality Markers for Emoji Usage

Völkel, Sarah Theres, Buschek, Daniel, Pranjic, Jelena and Hussmann, Heinrich. 'Understanding Emoji Interpretation through User Personality and Message Context'. In: *Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services*. MobileHCI '19. Project Website: `www.medien.ifi.lmu.de/personality-emojis`. New York, NY, USA: Association for Computing Machinery, 2019. DOI: `10.1145/3338286.3340114`

To address the first research question, we present paper [Core1]. As personality is associated with people's behaviour, for example their use of facial expressions [24, 56, 74, 149], it is likely that their emoji use in written communication is also determined by the sender's personality. Albeit the relationship between user personality and emoji use has been examined before, we are the first to investigate the influence of personality, gauged by an established personality inventory (vs implicit personality assessment via text analysis [131]), on user interpretation of emojis in context (vs stand-alone use [151] or automatically drawn from large text sets [14, 177]) for interpersonal communication (vs social media posts [13, 117, 131]).

In particular, we conducted an online survey, in which we presented 646 participants with common interpersonal texting scenarios and then asked them which emoji(s) they would add to the message.[2] Afterwards, we calculated the influence of participants' personalities on their choice of emojis using generalised linear regression models. Our findings indicated several personality markers for the use of emojis. Specifically, using heart emojis, such as 😘 ❤️ 🤩, was positively associated with higher levels of Agreeableness, Extraversion, and

[2] Please note that the survey in this paper also included another task. As this task and its findings are not part of this thesis, they are left out here.

Neuroticism, whereas participants high in Openness tended to adorn messages less often with heart emojis. Agreeableness was also linked to increased use of contentment emojis (e.g. 😄). In line with neurotic people's disposition towards experiencing negative affect, we found that neurotic participants tended to use more emojis depicting sadness and fewer emojis depicting contentment. Surprisingly, Neuroticism was also a predictor for the use of sensory pleasure emojis (e.g. 😊).

In summary, our work in [Core1] shows that humans have personality markers for the use of emojis for all Big Five personality dimensions except for Conscientiousness in the context of text messaging. As a consequence, we may add these emojis to the set of human behaviour cues associated with personality. In the following publications, we leverage these emojis, among other personality markers informed by related work in Psycholinguistics, to imbue text-based CAs with personality. Knowing that humans treat CAs as if they were people (cf. Chapter 2), we explore whether these human personality markers can be transferred to CAs:

RQ 1.2: *Can different levels of a personality dimension be synthesised in a text-based conversational agent by transferring verbal cues from human behaviour?*

To this end, I contribute two publications, which investigate the transfer of human verbal behaviour cues associated with *Agreeableness* [Core2] and *Extraversion* [Core3] to synthesise different levels of these personality traits in text-based CAs through dialogue design. I chose these two personality dimensions out of the five OCEAN factors because they are particularly meaningful in interpersonal interaction [143, 160].

Prior work has usually compared two opposing versions along a personality dimension, for example *extraverted* vs *introverted* [100, 124, 157, 158]. However, assuming a similarity attraction effect between user and agent personality [9, 158], an additional *average* level of a personality seems particularly beneficial to the user as the distribution of human personality traits is expected to follow a Gaussian curve in the population [144]. Hence, in both of our publications, we examine the synthesis of three different levels along a personality trait continuum (low, average, high). Subsequently, I will provide a brief summary of the two publications, including descriptions of the developed text-based CAs, the research designs, and the results. Figure 3.2 visualises user perceptions of the synthesised personalities, as informed by their ratings of the CAs on Big Five inventories [61, 202].

[Core2]: Synthesising Three Levels of *Agreeableness* in Text-based CAs

> Völkel, Sarah Theres and Kaya, Lale. 'Examining User Preference for Agreeableness in Chatbots'. In: *CUI 2021 - 3rd Conference on Conversational User Interfaces*. CUI '21. Project Website: `www.medien.ifi.lmu.de/agreeableness-chatbots`. New York, NY, USA: Association for Computing Machinery, 2021. DOI: `10.1145/3469595.3469633`

Because CAs are often employed as friendly and helpful assistants in service applications [50], the Big Five dimension Agreeableness plays a major role in CA personality design. Nonetheless, prior work has primarily focused on synthesising Extraversion in artificial agents [124,

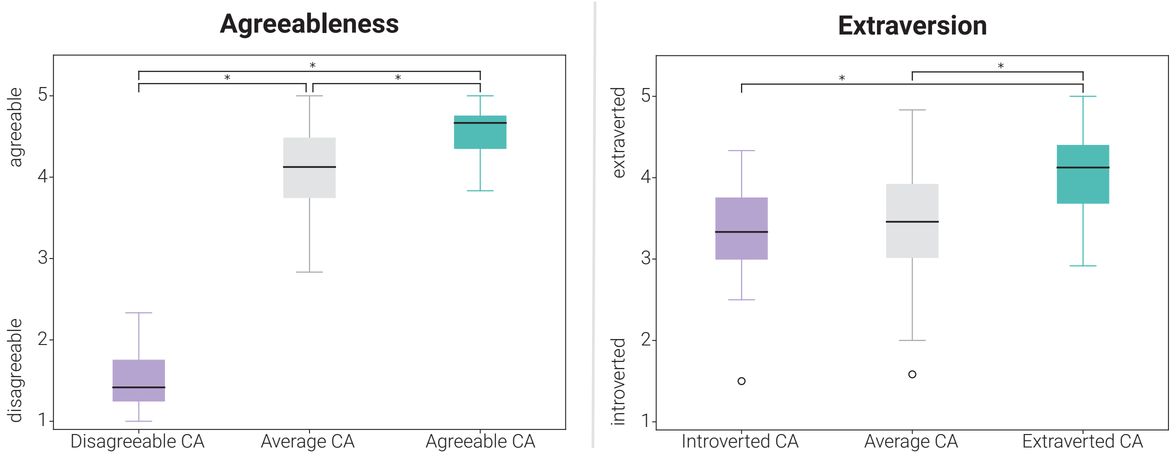

Figure 3.2: Participants' perceptions of the three text-based CAs' levels of Agreeableness (left) [Core2] and Extraversion (right) [Core3] using the BFI-2 personality questionnaires [61, 202].

158] due to its most pronounced link to observable human behaviour [196]. Hence, in this paper, we deliberately manipulated a text-based CA's language, leveraging verbal cues associated with human Agreeableness, to create three distinctive versions, each representing a different level of Agreeableness (disagreeable, medium agreeable,[3] and agreeable). The CAs were implemented using the NLU module and rule-based dialogue manager provided by Botpress[4], whilst the chatbots' responses were based on the pre-scripted personality-imbued dialogues. We situated the CAs in a movie recommender application and then asked 30 participants in a lab experiment to interact with each of the CAs and evaluate them afterwards using the German version [61] of the established Big Five Inventory-2 questionnaire (BFI-2) [202].

Our language manipulation yielded three distinct levels of Agreeableness, with the agreeable and disagreeable CAs being perceived as such, whereas the medium CA was also perceived as rather agreeable (cf. Figure 3.2). This rating of the medium CA roughly corresponds to the average level of Agreeableness of 3.76 in the investigated German population [61]. Thus, the results are suited to examine user preferences in the context of similarity attraction (cf. Section 3.2.1). I will henceforth refer to this CA as the one with an *average* level of Agreeableness.

[Core3]: Synthesising Three Levels of *Extraversion* in Text-based CAs

Völkel, Sarah Theres, Schoedel, Ramona, Kaya, Lale and Mayer, Sven. 'User Perceptions of Extraversion in Chatbots after Repeated Use'. In: *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. CHI '22. Project Website: `www.medien.ifi.lmu.de/extraversion-chatbots`. New York, NY, USA: Association for Computing Machinery, 2022. DOI: `10.1145/3491102.3502058`

[3]Please note that we refer to the CA which was intended to evoke medium levels of Agreeableness as medium here but neutral in the paper, as we learned since the publication of this paper that the term "neutral" may be confusing in this context.

[4]`https://botpress.com`, last accessed 10th May 2022

In our work in [Core2] and related work, user perceptions of CA personality have been evaluated for short one-time interactions only (e.g. [27, 36, 224]) and primarily in lab studies (e.g. [100, 158, 190]). However, personality impressions are subject to change after prolonged contact [10] and can only be reliably determined by observing behaviour in aggregated and relevant situations [139, 219]. As tailoring CA personality to the user seems primarily useful in the context of long-term real-world interaction [103, 187], we examined users' perceptions of personality after repeated use on their personal smartphones in this paper. To this end, we synthesised three different levels of Extraversion in text-based CAs: introverted, average, and extraverted. We again did so by systematically manipulating the CAs' use of language based on work in Psycholinguistics on human personality markers for Extraversion (e.g. [96, 172]) and our work on personality markers in emojis [Core1]. The CAs were implemented using the Telegram Bot API[5], employing keyword recognition as NLU and pre-scripted dialogues for the response generation.

We situated the personality-imbued CAs in a daily stress tracker application as CAs have opened up new opportunities in mental health treatment [114, 122, 136, 203]. In a within-subjects study, we asked 34 participants to converse with the three CAs, each over the course of four days, so as to examine user perceptions of the CA personalities after repeated, prolonged use. To gauge participants' personality impressions of the CAs, we first collected their open descriptions of the personality in a free-text field, followed by the BFI-2 [202]. We qualitatively analysed these descriptions and systematically mapped them to the Big Five dimensions.

Our findings show that participants perceived the extraverted and average CAs as intended, whereas verbal cues transferred from human behaviour were insufficient to create the impression of an introverted text-based CA. This result was evident in both the inventory responses as well as the open descriptions. Moreover, the open personality descriptions shed light on those traits that participants perceived as most salient for the three CAs. All three CAs were primarily described as high in Agreeableness, albeit the specific reasons varied among the CAs. Furthermore, participants noted the high Extraversion of the extraverted CA, whereas the average and introverted CAs were perceived as high in Conscientiousness. The average CA was also perceived as more artificial, as illustrated by descriptors such as "robotic" and "unnatural", which would not have been possible to discover by means of a human personality questionnaire alone.

Two observations, (1) that Neuroticism and Openness were rarely mentioned by participants, and (2) that participants frequently commented on the perceived artificiality of a CA, underpin the necessity to question the adequacy of the human Big Five model for describing CA personality. In addition, some of the individual personality questionnaire items were answered similarly for all three CAs (e.g. for the item *is less active than other people*). This lack of perceived differences could have resulted from the realisation of the agents. Another possible explanation is that the wording of the human personality inventories is less suitable for revealing differences between CA personalities.

[5] `https://core.telegram.org/bots/api`, last accessed 10th May 2022

Summary: Human Personality Model as a Basis

The contribution of this thesis to RQ1.1 and RQ1.2 is threefold: First, we introduce empirical insights into the link between human personality and emoji usage that cannot only be used for improving text messaging between two human interlocutors but also expand the set of behaviour cues used to infuse personality into text-based CAs [Core1]. Second, we present two sets of verbal cues derived from psycholinguistic literature to induce three levels of (1) Agreeableness [Core2] and (2) Extraversion [Core3], implemented in fully operative text-based CA applications. Third, we demonstrate an empirical investigation into the suitability of transferring human verbal cues to synthesise CA personality.

Our results show that human verbal behaviour cues can be used to a limited extent to equip fully text-based CAs with personality, as we achieved a high level and an average level of Extraversion as well as a high, average, and low level of Agreeableness. On the other hand, synthesising specific levels of personality along the continuum of a personality dimension remained challenging. In particular, the CA's task places constraints on this goal: for example, can a CA which has to send daily notifications be perceived as introverted, or can a CA that provides a service be perceived as only moderately agreeable? Moreover, our findings also raised doubts about the adequacy of the human Big Five model and corresponding personality markers for describing and synthesising personality in CAs. I address these doubts in the next part.

3.1.2 Conversational Agent Personality Dimensions as a Basis

Our work in [Core3] brought to light that human personality inventories might not be suitable to gauge user perceptions of CA personality. Furthermore, our work highlighted the salience of the CA's perceived artificiality, which affects not only users' personality perceptions but also their preference for a CA (cf. Section 3.2.1). The fact that users comment on the perceived artificiality of a CA when openly asked to describe the agent personality has been echoed in several studies [133, 173, 224]. These observations challenge the use of the Big Five model to describe CA personality and motivate our next research question:

RQ 1.3: *Which dimensions adequately describe conversational agent personality?*

To address this research question, I contribute [Core4], in which we apply the psycholexical approach [181] – the foundation for the Big Five personality model – to derive personality traits for CAs. If the Big Five personality model is applicable and sufficiently comprehensive to describe CA personality, applying this approach could be expected to result in the same five dimensions for CAs. In the light of the aforementioned need to examine personality in the context of long-term interaction, we turned to speech-based CAs for this publication as users have higher familiarity and prolonged use with *their own* voice assistants and smart speakers in their daily lives [7].

[Core4]: Personality Dimensions to Describe Speech-based CA Personality

Völkel, Sarah Theres, Schödel, Ramona, Buschek, Daniel, Stachl, Clemens, Winterhalter, Verena, Bühner, Markus and Hussmann, Heinrich. 'Developing a Personality Model for Speech-Based Conversational Agents Using the Psycholexical Approach'. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. CHI '20. Project Website: www.medien.ifi.lmu.de/personality-model. New York, NY, USA: Association for Computing Machinery, 2020. DOI: 10.1145/3313831.3376210

To realise my vision of tailored personality-imbued CAs, a theoretical taxonomy describing personality is crucial for enabling CA designers to systematically compose different target personalities. In particular, work by Isbister and Nass [100] highlighted the importance of consistent personality expressions in CAs for the user experience. A theoretical taxonomy makes the targeted personality profile explicit to the CA design team, which is crucial to ensure a consistent synthesis. To this end, we followed the psycholexical approach, which assumes that individual differences manifest themselves in language use [86] and which is the most established approach in Psychology to developing a personality model [181].

Inspired by traditional test construction theory [31], we developed a new multi-method approach to collect potential items (henceforth, *descriptors*) for describing the personality of speech-based CAs and then conducted an exploratory factor analysis on the resulting items to examine their underlying structure. The item pool generation comprised three studies:

1. An *online survey*, in which 135 participants listed personality descriptors for a chosen voice assistant in a free-text task
2. A *lab experiment*, in which 30 people interacted with three popular voice assistants and were subsequently interviewed about their personality impressions
3. A *text analysis* of 30,000 online reviews of three popular voice assistants

We then merged the resulting items and systematically reduced them into a set of 349 personality adjectives. In an online survey, 744 people each rated one of the three most well-known voice assistants (Alexa, Google Assistant, Siri) on the resulting 349 descriptors. An exploratory factor analysis yielded ten latent personality dimensions.

These ten dimensions do not match the human Big Five – neither in number nor in content (cf. Figure 3.3). Instead, several patterns emerged: (1) three dimensions reflect adjectives from the Big Five trait Agreeableness, highlighting its importance in the context of speech-based CAs; (2) the majority of dimensions signify either desirable or non-desirable characteristics, thereby relating the personality description of CAs to users' expectations; and (3) several dimensions comprise both functional and social aspects of the interaction. In summary, our ten dimensions and 349 items provide development teams with a shared vocabulary to facilitate consistency, completeness, and a mutual understanding of a CA personality. In this dissertation, I refer to the resulting dimensions as the *Ten CA Personality Dimensions*.

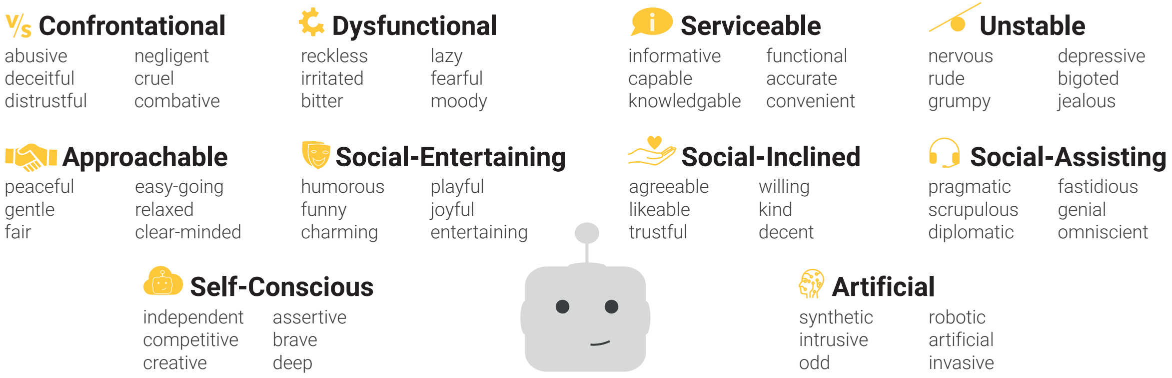

Figure 3.3: Using an adaptation of the psycholexical approach, we found ten personality dimensions to describe CAs. This figure shows these ten dimensions, including our suggested dimension labels and the six descriptors with the highest factor loadings, as outlined in [Core4].

Our results in [Core4] revealed that the human Big Five model is not applicable to describe speech-based CA personality. Furthermore, our findings in [Core3] indicated that human personality markers are not completely transferable to text-based CAs. Hence, we recommend that CA designers use these Ten CA Personality Dimensions to specify the desired target personality of a CA. Consequently, we need a new method for deliberately synthesising a CA personality based on the Ten CA Personality Dimensions. In commercial voice assistant design, dialogue experts, such as scriptwriters, manually write the responses for the CA in line with a predefined personality [206]. On the one hand, to the best of my knowledge, this approach has only been examined for a single, consistent personality and not to synthesise different levels of theoretically grounded personality dimensions. On the other hand, professional scriptwriters might not be available to smaller companies or researchers, but they might be equally interested in equipping their CAs with personality. Hence, the fourth research question asks for a method that meets the aforementioned requirements:

RQ 1.4: *How can different levels of a conversational agent personality dimension be synthesised?*

[Core5]: Synthesising Three Levels of *Social-Entertaining* and *Confrontational* in Speech-based CAs

Völkel, Sarah Theres, Meindl, Samantha and Hussmann, Heinrich. 'Manipulating and Evaluating Levels of Personality Perceptions of Voice Assistants through Enactment-Based Dialogue Design'. In: *CUI 2021 - 3rd Conference on Conversational User Interfaces*. CUI '21. Project Website: `www.medien.ifi.lmu.de/voice-assistant-personality`. New York, NY, USA: Association for Computing Machinery, 2021. DOI: `10.1145/3469595.3469605`

Whilst our work in [Core4] aimed to *describe* CA personality, in this publication we explore how to *synthesise* the perception of a CA personality through dialogue design. Our approach was inspired by Bates' proposal [16] to draw on artists' insights into creating engaging characters. Following a similar approach to the design of commercial voice assistants [171], we generated personality perceptions through *Enactment-based Dialogue Design*; that is, amateur actresses and actors sketch, enact, and discuss dialogues in interactive focus groups. In contrast to the approach employed by companies, we examined whether this method can be used to synthesise different *levels* (low, rather high, high) of *theoretically grounded* personality dimensions. The synthesis of three personality levels is motivated by our vision of CA personalities tailored to their users and our work in [Core2, Core3], which has already demonstrated individual preferences (cf. Section 3.2). We chose a *rather high* level over a medium level because our work in [Core3] suggested that a little pronounced personality is perceived as artificial by participants and is therefore rejected.

When conducting the focus groups, we found that writing dialogues to express three levels of personality is more challenging, as language that consistently distinguishes the three levels from each other has to be chosen. In particular, Isbister and Nass [100] emphasised that users prefer CAs whose behaviour cues are orchestrated to form a consistent personality. Consequently, a rather high level cannot be synthesised by combining high and low level cues of CAs, but dedicated dialogues have to be written. To explore our method, we applied it to express three levels of two of our personality dimensions in [Core4] as example case studies. Specifically, these two dimensions were (1) *Social-Entertaining*, which captures a CA's social and humorous demeanour, and (2) *Confrontational*, which captures a CA's disagreeing and combative behaviour. We chose these two dimensions because they are likely to yield more controversial preferences, as informed by our work in [Core7].

Based on the resulting scripts from the focus groups, we implemented the dialogues using Amazon Alexa and recorded conversations between the different personality-imbued speech-based CAs and a user. As in-car voice assistants constitute a pervasive use case [113] and their personality perceptions have been found to impact drivers' preferences [27], we situated our scenarios in the context of automotive user interfaces. Subsequently, to examine whether the personality levels were successfully synthesised, we presented 156 participants with the recordings of conversations with the three different versions of a speech-based CA for one scenario each in an online survey. Presenting participants with pre-recorded dialogues is a common research design to elicit their perception of a CA's personality, as it allows control over the personality expression and avoids influences from poor speech recognition [35, 102, 103]. Due to the lack of an established CA personality assessment inventory, we used the twenty adjectives we provided for each of the Ten CA Personality Dimensions in our prior work [Core4] to gauge participants' perceptions of the CA personality.

Overall, our findings demonstrate that our method of Enactment-based Dialogue Design is suited to create three different levels of a personality dimension as apparent in the expected sequence of the evaluations (cf. Figure 3.4). However, the perception of higher personality levels was less pronounced than expected for both dimensions. Furthermore, single adjectives (e.g. "clumsy" or "messy") seemed to be less applicable to describe CA personality, echoing our findings in [Core3].

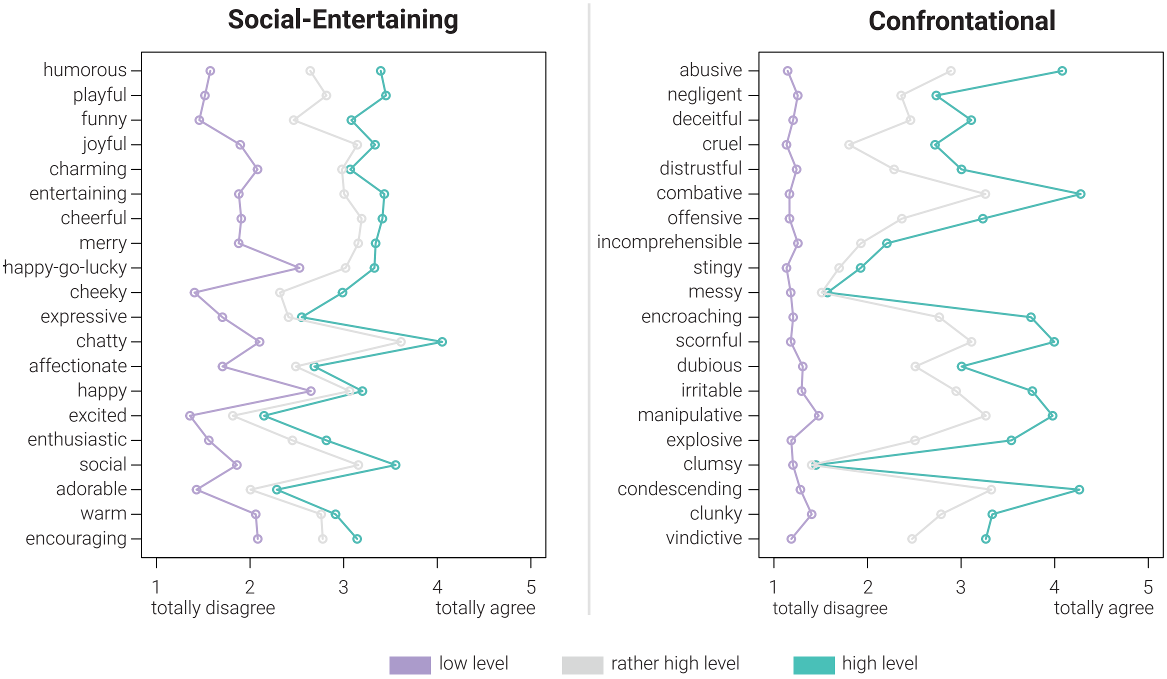

Figure 3.4: Participants' perceptions of the three speech-based CAs' levels of Social-Entertaining (left) and Confrontational (right) [Core5] using the personality descriptors from our Ten CA Personality Dimensions [Core4].

Summary: Conversational Agent Personality Dimensions as a Basis

This thesis contributes theoretical, methodological, and empirical insights to address research questions RQ1.3 and RQ1.4. First, we explored an adaptation of the psycholexical approach by developing a new multi-method strategy from which we generated a set of 349 personality descriptors [Core4]. Second, we contribute Ten CA Personality Dimensions, derived via an exploratory factor analysis on data of 744 people rating our descriptors [Core4]. Third, we present a new method, called Enactment-based Dialogue Design, which allows us to synthesise multiple levels of CA personality [Core5] based on our Ten CA Personality Dimensions [Core4]. Fourth, we evaluated the suitability of our new method in an empirical user study [Core5].

These findings yield much needed theoretical clarity on the necessity of dedicated personality descriptions for CAs, pointing to the insufficiency of using human personality models. Moreover, our new method of synthesising personality-imbued dialogues can be used by CA designers to systematically emulate different perceptions of CA personality.

		[Core2]	[Core3]	[Core5]
Artefact	**CA Modality**	Text-based	Text-based	Speech-based
	Personality description	Big Five	Big Five	Ten CA Personality Dimensions [Core4]
	Synthesised dimension	Agreeableness	Extraversion	Social-Entertaining & Confrontational
	Synthesised levels	low, medium, high	low, average, high	low, rather high, high
	Implementation	Botpress integrated in website	Telegram bot	Amazon Alexa
Synthesis	**Synthesis method**	Human derived language cues	Human derived language cues	Enactment-based Dialogue Design
	Assessment method	BFI-2 questionnaire	BFI-2 questionnaire & open descriptions	[Core4] Personality adjectives
	Results	successful for all levels	successful for average & high levels	successful for low & rather high levels
Research Design	**Context**	Movie recommender	Stress tracking & reflection	Automotive
	Study design	Within	Within	Mixed
	Setting	lab	field	field (online survey)
	Interaction duration per CA	ca. five min	four days	ca. 30s
DVs	**Subjective preference**	Desire to interact, ranking	Desire to interact, ranking, qual. reasons for ranking	Desire to interact, ranking
	Interaction behaviour		Engagement (nr of words)	

Table 3.1: An overview of our three publications on synthesising CA personality and evaluating user preferences for these personalities. The top of the table describes the manipulated artefact for each publication, followed by the synthesis method to generate the targeted personalities (RQ1), the respective research design, and the dependent variables (DVs) examined for user preferences (RQ2).

3.1.3 Summary: Imbuing Conversational Agents with Personality

In conclusion, this thesis explores two approaches for describing CA personality, namely using the human Big Five model and the Ten CA Personality Dimensions, derived via an exploratory factor analysis [Core4]. We present three publications to investigate how to synthesise personality based on these two description approaches. On the one hand, we examine the transfer of human behaviour cues to imbue text-based CAs with different levels of Agreeableness and Extraversion [Core2, Core3]. On the other hand, we introduce a new method to develop dialogues for synthesising three levels of personality expressions [Core5] based on the Ten CA Personality Dimensions in [Core4]. An overview of these three publications and how they differ with respect to the implemented artefact, the dialogue manipulation, the research design, and the dependent variables (DVs) is given in Table 3.1. I discuss the benefits and drawbacks of the two different approaches to imbuing CAs with personality in Chapter 4.

3.2 User Preferences for Conversational Agent Personality

Whilst today's commercial CAs have taken a one-size-fits-all approach with respect to their personality design, research suggests that user predilections for CA personalities are not homogeneous [6, 27, 123, 157, 158], mirroring human-human interaction [34]. Yet, a deeper investigation of what these preferences look like is still missing. In particular, the majority of prior work investigated user preferences for the personality dimension Extraversion (e.g. [100, 158, 200]). However, preferences for other personality traits such as Agreeableness seem particularly important to inspect in the context of CAs, which are typically employed as helpful assistants. Insights into preferences for multiple personality traits can inform CA designers about whether there are unanimously liked types of CA personalities (echoing to some extent the industry's one-size-fits-all approach) or whether user preferences vary.

Another challenge in the context of user preferences for CAs is the fact that users expect the CA's personality to fit the application's context [21, 101, 224]. Whilst the targeted personality traits may be obvious and universal to all users in some situations (e.g. users expect a *professional* CA for a banking application), user preferences may be less clear in other contexts, such as voice assistants in the car. Thus, knowledge about user preferences has to be gathered individually for each context. In the light of this need for more information on user preferences for CA personalities, the second goal of this thesis is captured in RQ2:

RQ 2: *What preferences do users have for conversational agent personality?*

I tackle this goal again in two approaches. First, this section presents publications in which users interact with different CA personalities and evaluate them afterwards (deductive approach). Specifically, we collect user preferences for the four personality dimensions Agreeableness, Extraversion, Social-Entertaining, and Confrontational due to their assumed importance in interpersonal interaction [143, 160]. In this work, we investigate user preferences in three different contexts, namely entertainment [Core2], mental health [Core3], and automotive [Core5]. Second, because our findings yielded clear individual user preferences, we explore a method so as to include users in the development process of CA personality and present empirical findings on user visions of the perfect speech-based CA in the smart home context (inductive approach). As previous research suggests that user preferences are determined by users' own personality [157], we explore the impact of user personality on user preferences in all publications presented in this section. For brevity and to avoid repetitions, I present the results for this influence of user personality together with the other results in Subsections 3.2.1 and 3.2.2 but provide an overview of our findings in Subsection 3.2.3.

3.2.1 A Deductive Approach to Studying User Preferences

Prior work found conflicting findings regarding the relationship between user personality and their preference for a CA personality. Although early work suggested a similarity attraction effect between users' personality and their preference for CA personality [157], this work only examined two CA personalities at the outer poles of the personality continuum, such as an introverted and extraverted CA. However, the majority of users are expected to have more moderately pronounced personality traits [144]. To better reflect these moderate personalities, we manipulate three different versions of a CA personality by adding an *average* level.

In this subsection, I present our findings on user preferences from the three publications introduced in the previous section.[6] In each of them, we first manipulated the three different versions of a CA personality and then asked users to interact with them in a study. Following these interactions, we asked participants to rate their experience, resulting in RQ2.1:

RQ 2.1: *Which levels (low, average, high) of personality traits in conversational agents do users prefer?*

User preferences can manifest themselves in a variety of ways. Previous work has explored the influence of CA personality on user trust [27, 224], likeability [20, 21, 27, 37, 158], engagement [200, 224], self-disclosure [85, 224], and purchase behaviour [200]. In this thesis, I focus on users' subjective preferences, specifically likeability, as a metric because subjective dissatisfaction with the CA experience leads to significant barriers to using these agents [59]. To this end, I used two metrics: (1) a Likert scale question pertaining to users' desire to interact with this CA again, which participants answered directly after they had interacted with a CA version, and (2) a final ranking of all versions as to which agent they liked best after they had interacted with all of them. In [Core3], I enriched subjective preference by objective measures, specifically, by doing a word count of the number of words participants wrote with each CA as an indicator for user engagement in line with prior work [89, 127, 223].

[Core2]: User Preferences for *Agreeableness* in Text-based CAs

Völkel, Sarah Theres and Kaya, Lale. 'Examining User Preference for Agreeableness in Chatbots'. In: *CUI 2021 - 3rd Conference on Conversational User Interfaces*. CUI '21. Project Website: `www.medien.ifi.lmu.de/agreeableness-chatbots`. New York, NY, USA: Association for Computing Machinery, 2021. DOI: `10.1145/3469595.3469633`

In Section 3.1.1, I presented how we manipulated three text-based CAs' language to evoke the perception of three distinctive levels of Agreeableness. As a reminder, 30 participants interacted with the three text-based CAs for a few minutes each in the context of a movie recommendation application. In this paper, we also gathered participants' subjective likeability for the three levels as described above. Whilst the disagreeable CA was universally disliked by participants, the CAs with high and average levels of Agreeableness received similar ratings,

[6] As a reminder, publications that have been introduced before are displayed in a grey box.

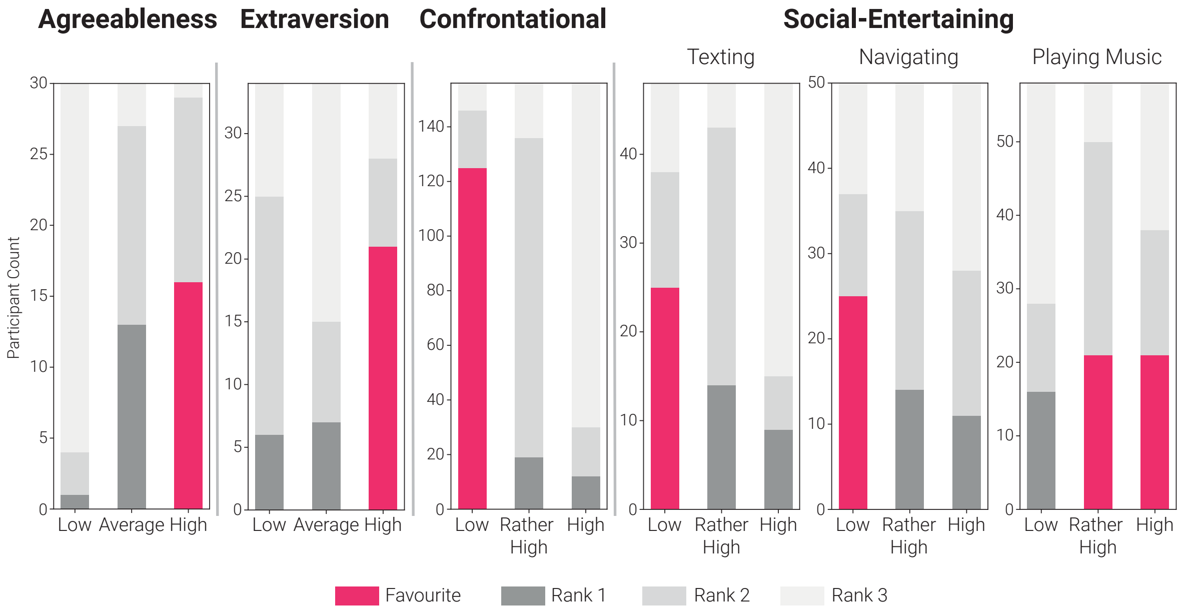

Figure 3.5: Participants ranked their desire to interact again with three personality versions for the dimensions Agreeableness [Core2], Extraversion [Core3], Confrontational [Core5], and Social-Entertaining [Core5]. For the latter two dimensions, we investigated preferences for different scenarios (*texting*, *navigating*, and *playing music*). For Social-Entertaining, participants' ranking is shown for each individual scenario due to the diverging preferences. As the participants' ranking was unequivocal for Confrontational, only the ranking across all scenarios is depicted. Participants' respective favourites (the agent ranked first by most participants) are highlighted in pink.

both in the final ranking (cf. Figure 3.5) as well as regarding participants' desire to interact with them again. These findings emphasise that there are clear individual differences in preferences, lacking a prevalent favourite. In addition, however, our results also highlight the benefits of average, more subtle personality expressions in CAs, aside from the dichotomous low and high personality levels, as the average agreeable chatbot was almost equally liked to the one high in Agreeableness.

Furthermore, our results point to a one-directional similarity attraction effect, with agreeable participants preferring a mutually agreeable CA. In contrast, we did not find the reverse effect of participants with low scores in Agreeableness preferring the disagreeable CA, nor a preference for the average CA among participants with average levels of Agreeableness.

[Core3]: User Preferences for *Extraversion* in Text-based CAs

Völkel, Sarah Theres, Schoedel, Ramona, Kaya, Lale and Mayer, Sven. 'User Perceptions of Extraversion in Chatbots after Repeated Use'. In: *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. CHI '22. Project Website: `www.medien.ifi.lmu.de/extraversion-chatbots`. New York, NY, USA: Association for Computing Machinery, 2022. DOI: `10.1145/3491102.3502058`

In Section 3.1.1, I introduced our research design for [Core3], which consisted of (1) systematically synthesising text-based CA personalities with three levels of Extraversion by leveraging human verbal cues, and (2) asking 34 participants to interact with them in a stress-tracking application on their personal phones, each for a duration of four days. We gathered participants' likeability ratings directly after four days of interaction with a text-based CA and through a final ranking after they had used all three versions. To shed light on participants' rationales for their ranking, we enriched the quantitative data with their statements collected through free-text fields.

Whilst the majority of participants favoured the extraverted CA, as became evident from both participants' desire to interact with the CA again and their final ranking, about 40% liked either the average or the introverted one better (cf. Figure 3.5). Participants' rationales for liking the extraverted CA singled out its agreeable and casual demeanour. In contrast, they described the perceived artificiality of the average CA as detrimental to their preference for it. Conversely, the rationales also revealed that participants had conflicting, opposing preferences, with some participants criticising the casual demeanour of the extraverted CA and perceiving the average CA as more professional. Regardless of a subjective preference for the extraverted CA, participants engaged significantly more with the introverted CA, as indicated by their average number of written words.

Despite clear individual preferences, our results did not signify that these differences are linked to participants' own personalities, in contrast to previous work on speech user interfaces [27, 158] and robots [9]. In summary, we found that the perceived personality of a text-based CA has an effect on users' interaction behaviour and subjective preferences, the latter often being contradictory.

[Core5]: User Preferences for *Social-Entertaining* and *Confrontational* in Speech-based CAs

Völkel, Sarah Theres, Meindl, Samantha and Hussmann, Heinrich. 'Manipulating and Evaluating Levels of Personality Perceptions of Voice Assistants through Enactment-Based Dialogue Design'. In: *CUI 2021 - 3rd Conference on Conversational User Interfaces*. CUI '21. Project Website: `www.medien.ifi.lmu.de/voice-assistant-personality`. New York, NY, USA: Association for Computing Machinery, 2021. DOI: 10.1145/3469595.3469605

In Section 3.1.2, I explained our research design of imbuing speech-based CAs with personality along the dimensions Social-Entertaining and Confrontational from our Ten CA Personality Dimensions through Enactment-based Dialogue Design. In an online survey, we presented 156 participants with pre-recorded dialogues between a user and six speech-based CA versions (two dimensions $\times$ three levels).[7] We collected their likeability ratings in line with our previous work for three different scenarios in the automotive context. For the dimension Social-Entertaining, no clear preference for a version emerged; instead, participants' choice depended on the scenario (cf. Figure 3.5). On average, participants favoured interacting with

[7] As a reminder, we synthesised a low, rather high, and high level for each dimension. We chose a *rather high* level over a medium level because our work in [Core3] suggested that a little pronounced personality is perceived as artificial by participants and is therefore rejected.

the CAs rather high or high in Social-Entertaining in a low-stake scenario, such as *Playing a song*. In contrast, the majority liked the low level CA best for a task-related scenario, such as *Writing a text message*. For the dimension Confrontational, the majority of participants unanimously voiced their support for the least confrontational CA across all three scenarios. Nonetheless, about a quarter of participants picked the other two versions as their favourites, with the variance being most evident in the *Navigating to a restaurant* scenario, in which the CA questioned the user's unhealthy food choice.

Juxtaposing participants' own personalities with their desire to interact with the different CA versions, we found that (1) Extraversion was a significant positive predictor for preferring a CA high in Social-Entertaining, (2) whereas Conscientiousness was a significant negative predictor for preferring a CA rather high in Confrontational. In summary, our findings indicate that user preferences for personality in CAs to some extent depended both on the respective context and users' own personality. Our work highlights that preferences vary widely among users, thereby underlining once again the benefits of different personality levels in CAs and tailoring them to the users.

Summary: A Deductive Approach to Studying User Preferences

This thesis contributes empirical evidence for the impact of CA personality on users' subjective desire to interact with an agent again in the future [Core2, Core3, Core5] and their engagement in the interaction [Core3] (RQ2.1). Our findings brought to light that there is no universally favoured CA personality. Instead, users have clear individual yet conflicting preferences when it comes to CA personality, underlining the benefits of creating different *levels* of personality in a CA and then tailoring its personality to the users. Moreover, in case of limited resources for designing CA personalities, our results can guide CA designers to the personalities that are most preferred (most often ranked first) or least rejected (least often ranked last), as illustrated in Figure 3.5. Finally, our results indicate that user preferences for personality are context-dependent, with users preferring less distinctive personalities for task-focused scenarios [Core5].

3.2.2 An Inductive Approach to Studying User Preferences

The aforementioned methods to generate CA personality as well as related work (e.g. [27, 124, 158]) have taken a deductive, top-down approach with little user engagement during the design process. That is, different versions of CA personalities were developed and presented to users in a contrastive evaluation. Whilst this work has pointed to the general benefits of imbuing CAs with different personalities, little is known about what users would want in an interaction with a CA given no technical constraints. This knowledge, however, is crucial to better understand (1) what desirable target personalities are, and (2) for which personality traits user preferences diverge. The latter can inform CA designers about personality traits that should be considered for adaptation. A better understanding of user visions of CA personalities is also useful for the research community as qualitative work has pointed out shortcomings in current voice assistant interactions [59, 135, 178]. This lack of insight motivates us to actively engage users in informing future CA design by addressing RQ2.2:

RQ 2.2: *How do users envision interacting with a perfect conversational agent, and how do these visions vary?*

To overcome this research gap, we present a new pragmatic, inductive, bottom-up method we call *Vision Dialogue Elicitation*, which we first explored in [Core6] and then refined in [Core7]. In particular, we let users freely imagine a dialogue with a CA they consider to be perfect, using their desired conversation style, syntax, and wording.

[Core6]: A Method to Elicit Users' Envisioned Dialogues with Perfect Speech-based CAs

Völkel, Sarah Theres, Kempf, Penelope and Hussmann, Heinrich. 'Personalised Chats with Voice Assistants: The User Perspective'. In: *Proceedings of the 2nd Conference on Conversational User Interfaces*. CUI '20. Project Website: www.medien.ifi.lmu.de/users-va-dialogues. New York, NY, USA: Association for Computing Machinery, 2020. DOI: 10.1145/3405755.3406156

To gain insights into user visions of CA personalities, we explored the suitability of a new method, which engages users themselves in CA design. Specifically, we presented 26 participants with dialogues for different scenarios in an automotive context, in which only the user's part of the conversation was given. Participants were then asked to write the responses of the CA so that these responses corresponded to their idea of a perfect in-car voice assistant interaction. Notably, all given dialogues comprised a *functional request* (e.g. to make a phone call) and a *social part* (e.g. reminder that they have not called their mother in a while). We chose a lab setting for this study to acquire an understanding of the suitability of our method by personally interacting with participants.

Our findings indicate that users sketched conversations for *functional* requests similarly to how they are implemented in today's voice assistants. Whilst we did not detect notable differences in participants' dialogue designs for these functional requests, our results revealed individual differences for the *social* part of the interaction. The majority of participants imagined a CA that initiates a social conversation which the user can accept or decline, whereas a few participants did not integrate any kind of social interaction. Some participants even imagined a CA that admonishes the user without being asked for an opinion. When investigating the impact of user personality on their vision of the social parts of the interaction, we only found one significant correlation, namely a positive one between Openness and the vision of a CA expressing an opinion.

[Core7]: A Method and Analysis of Users' Envisioned Dialogues with Perfect Speech-based CAs

Völkel, Sarah Theres, Buschek, Daniel, Eiband, Malin, Cowan, Benjamin R and Hussmann, Heinrich. 'Eliciting and Analysing Users' Envisioned Dialogues with Perfect Voice Assistants'. In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Project Website: `www.medien.ifi.lmu.de/envisioned-va-dialogues`. New York, NY, USA: Association for Computing Machinery, 2021. DOI: `10.1145/3411764.3445536`

In this paper, we built on our method in [Core6] and asked participants to write down their envisioned dialogues with a "perfect" speech-based CA for given scenarios. We further refined our method by letting participants write entire dialogues (that is, both the user and CA part) to give them more liberties to design the interaction as they see fit. The scenarios represented common smart speaker interactions identified by Ammari et al. [7]. We thus explored another use case for long-term interaction with CAs. Each scenario comprised a specific issue the user has to solve by conversing with the CA.

In an online survey, 205 participants were prompted to sketch their dialogues for these scenarios. We analysed the content of the resulting dialogues through an in-depth inductive thematic analysis. In addition, we extracted conversation characteristics from the text, such as the number of turns and questions.

Our results show that participants' visions of the perfect CA personality vary widely. The majority of participants envisioned a CA that is smart, proactive, and has personal knowledge about the user and their environment, resulting in more interactive and longer dialogues than with today's voice assistants. Conversely, participants' attitudes diverged regarding the CA's role in future scenarios, as well as the inclusion of humour and opinions. Comparing participants' visions to voice assistants commercially available today, the most notable differences concerned the CA's delivery of recommendations and suggestions as well as its ability to "think more independently" instead of falling back on web searches. The elicited dialogues demonstrated that our new method is effective in engaging users in the design of CAs and can therefore be employed by designers to gain insights into user requirements for CAs.

An exploratory analysis using linear mixed models (LMMs) suggests that participants' personalities impacted their vision of a perfect CA. However, this relationship was less pronounced than expected from related work. Specifically, our findings indicate that more neurotic participants envisioned CAs that do not express humour. On the other hand, more conscientious participants avoided the use of opinions in their visions of a CA, potentially due to their preferences for seeking thorough information from acknowledged, reliable sources [92].

Summary: An Inductive Approach to Studying User Preferences

This thesis' contribution to RQ2.2 is threefold: First, we present a new method to elicit user visions of interactions with a perfect speech-based CA, developed in [Core6] and refined in [Core7]. Crucially, this method actively engages users in the design of CA personality. Second, we contribute a set of qualitative and quantitative analyses and insights into user preferences for tailored CA dialogues in automotive [Core6] and domestic settings [Core7]. Third, we published the data set[a] containing 205 participants' envisioned dialogues for nine smart home scenarios collected in [Core7] to support further research on the desired characteristics of interactions with CAs in the community.

Our results show that users envision a smart, proactive CA, equipped with personal knowledge about the user and their environment to give personalised suggestions and recommendations [Core7]. However, users' attitudes diverged with respect to more social interactions, in particular the use of humour or opinions [Core6, Core7]. These results thus point out that CA designers should design different CA personality versions with and without (1) humour and (2) opinions so that these CAs can be tailored to individual user preferences.

[a] `www.medien.ifi.lmu.de/envisioned-va-dialogues`, last accessed 10th May 2022

3.2.3 The Influence of User Personality

In the previous two subsections, I showed that users have individual preferences for certain CA personality types. However, it remains unclear why different users have these different preferences and whether there are patterns of user characteristics that determine their preferences. According to prior research, *similarity attraction* is a powerful behavioural pattern for interpersonal interactions; that is, humans tend to prefer other people [34] and also CAs who share congruent personality traits [21, 157, 158].

For example, users expressed a more positive attitude [158], higher likeability [27], greater trust [27], as well as higher social presence [124] when a voice user interface matched their own personality. Echoing these findings for text-based CAs, Shumanov et al. [200] showed that corresponding Extraversion levels had a positive impact on user engagement and purchasing outcomes. In contrast, there is also research pointing to findings inconsistent with the similarity attraction paradigm. For example, Isbister and Nass [100] found that users significantly preferred and had more fun in a lab game with an ECA that manifested a complementary level of Extraversion to themselves.

Following this related work, we also collected our participants' self-reported personality via the established Big Five Inventory questionnaires [61, 201, 202] and examined RQ2.3:

RQ 2.3: *Does users' personality determine their preference for conversational agent personality?*

I investigated this research question in all publications on user preferences [Core2, Core3, Core5, Core6, Core7]. I reported on the significant results for the relationship between user personality and preference for CA personality in the previous two subsections with the presentation of the respective publications. Table 3.2 gives an overview of the findings.

		[Core2]	[Core3]	[Core5]	[Core6]	[Core7]
Analysis Method		Correlation	LMM	LMM	Correlation	LMM
User Personality	**O**		n.s.	n.s.	➕: preference for opinion in CA	n.s.
	C		n.s.	➖: preference for CA rather high in Con-frontational	n.s.	➖: preference for opinion in CA
	E		n.s.	➕: preference for CA rather high in Social-Entertaining	n.s.	n.s.
	A	➕: preference for CA high in Agreeableness	n.s.	n.s.	n.s.	n.s.
	N		n.s.	n.s.	n.s.	➖: preference for humour in CA

Table 3.2: The influence of users' personality on their preferences for CA personality levels (low, average, rather high, high), as informed by my publications. For each of the users' Big Five personality traits (OCEAN) only significant relationships with a preference for CA personality are listed (n.s. denotes that no significant links were found; an empty cell means that the respective trait was not examined). The direction of the relationship is denoted by an icon: ➕ indicates that this trait is a positive predictor for the preference, whereas ➖ indicates a negative predictor.

Notably, we detected a similarity attraction effect only for an agreeable CA and users high in Agreeableness [Core2]. In contrast, we neither found a reversed effect for users with a low or average level of Agreeableness nor an effect of user personality on their preference for CAs with different levels of Extraversion [Core3]. Due to the different underlying personality descriptions (Big Five for the users and our Ten CA Personality Dimensions for the CA), there may not be an immediate similarity attraction effect in [Core5]. However, we discovered that users with high levels of Extraversion tend to like a CA rather high in Social-Entertaining [Core5]. As an extraverted personality is associated with gregariousness, cheerfulness, and seeking external stimulation [142], a preference for a more verbose and casual CA seems to reflect a similarity attraction. Another pattern that emerged from our findings is that conscientious users tend to reject CAs that have opinions and may act more independently [Core5, Core7].

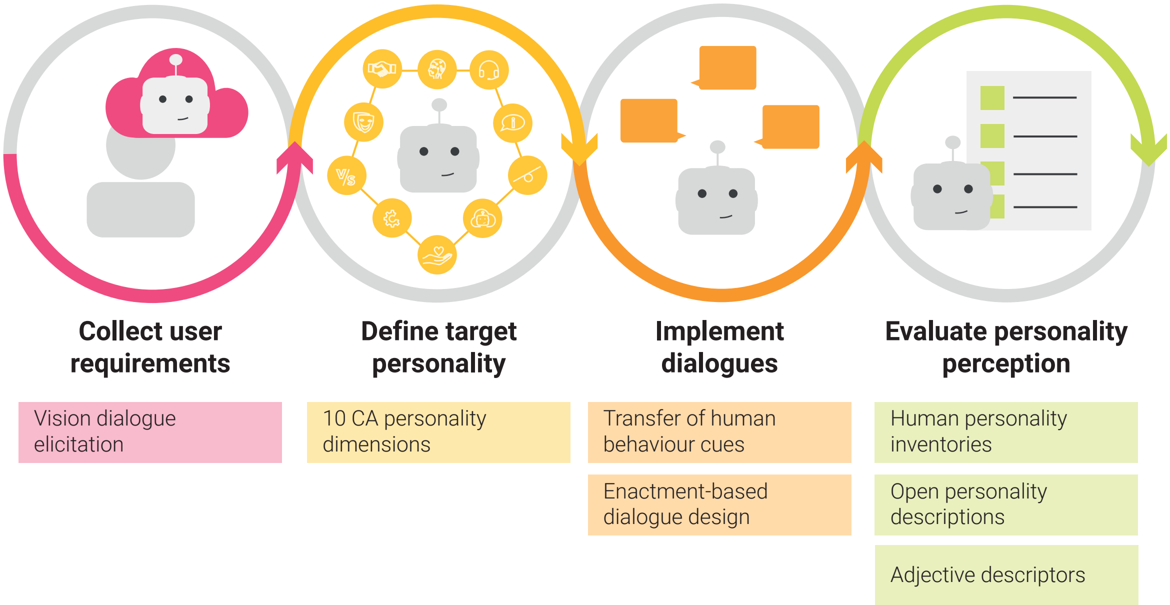

Figure 3.6: I propose a four-step conceptual process to imbue CAs with personality systematically. This thesis contributes methods to each of the steps, which are listed in the boxes below the process.

Summary: The Influence of User Personality

Although previous research suggests a similarity attraction effect between user personality and preference for CA personality, the impact of the user's personality on their preference for a CA personality was less pronounced than expected in our investigations. Whilst we provide empirical evidence for single links between user personality traits and corresponding preferences, we could not find comprehensive patterns, indicating that user preferences are a more complex phenomenon. I will discuss potential reasons for the lack of this link in Chapter 4.

3.3 Methods for Designing Conversational Agent Personality: A Conversation Design Process

In this thesis, I presented the development of several methods which can inform the design of CA personality. Consolidating the knowledge gained from applying these methods and from the scrutiny of related work, I propose a four-step conceptual process to imbue CAs with personality systematically. This process is meant as a starting point to design CA personalities in a structured way by providing CA designers with methods for each of the steps. Due to a lack of scientific guidance on CA design, this process is inspired by the traditional user-centred design process [166] and the Conversation Design Process in the Google Assistant developer guide[8]. The resulting process is specified for the purpose of creating CA personality but should be understood as integrated into the general CA design process. Figure 3.6 shows an overview of this process which is outlined in the following paragraphs. Whilst I present different methods for each step in this subsection, I will reflect on the benefits and drawbacks of each method in Chapter 4.

Collect User Requirements: As our research on user preferences for CA personality shows that these preferences are highly individual and context-dependent, capturing users' requirements and ideas of a CA's personality is the imperative first step of designing such a personality, as in every user-centred product design. To gather these user requirements, our method Vision Dialogue Elicitation, developed in [Core6, Core7], can be used to elicit user visions of the perfect CA. These visions can then (1) inform universal user requirements regarding the CA's personality, and (2) provide indicators for those characteristics for which users' visions vary. These indicators give pointers to traits that should be tailored to individual user preferences. Apart from users' ideas of the CA's personality, it goes without saying that a profound understanding of the users themselves is important to know the target audience.

Define Target Personality: Based on user requirements, CA designers have to define the target personality. The target personality is usually part of the persona that CA designers develop for the CA [52]. Although defining the target personality is generally recognised as an important step in the development of CAs [52, 171], there is no concrete guidance on exactly how to delineate it. The Ten CA Personality Dimensions and corresponding descriptors developed in [Core4] may be used by CA designers as a communication tool to specify the CAs' target personalities. In particular, the Ten CA Personality Dimensions can scaffold CA designers with concrete guidance on which personality traits should be defined to ensure consistency and improve mutual understanding within the team. Apart from the personality, the persona also has to encompass other stable traits, such as perceived gender, dialects, and brand image [52].

Implement Dialogues: The target personality is expressed via what the CA says, i.e. the dialogue. This dialogue is the central element of every CA design. As structured CA design is still in its infancy, the typical methods in industry to develop these dialogues are example-driven and include writing sample dialogues and defining a conversation flow [171]. Whilst research is also concerned with developing tools to automatically generate behaviour cues (e.g. [8, 44, 138, 186]), the publications in this thesis focus on the prior conceptual step of imbuing CAs with personality. To enrich the toolkit of CA designers, I present two approaches which CA designers can adopt as part of writing sample dialogues to express the target personality: (1) Using behaviour cues transferred from human behaviour [Core2, Core3], (2)

[8] `https://developers.google.com/assistant/conversation-design/how-do-i-get-started`, last accessed 10th May 2022

Enactment-based Dialogue Design in which amateur actresses and actors write and enact dialogues in interactive focus groups [Core5].

Afterwards, these dialogues have to be implemented in prototypes to assess the personality impression. As with every prototype, there might be high and low fidelity prototypes depending on the development stage. In the early stages, recording dialogues between a user and a CA, as we did in [Core5], can be sufficient to gather first personality impressions whilst keeping the implementation effort low and avoiding side effects due to poor speech recognition or NLU. On the other hand, because personality perceptions may change over the course of the interaction [10], high fidelity prototypes, such as our fully operative Telegram text-based CA in [Core3], have to be considered in later stages of the development process.

Evaluate Personality Perception: In the final step, the target personality has to be compared with users' actual perceptions of the personality. To this end, users have to interact with the prototypes and their perception has to be gauged. Whilst we have not systematically compared different evaluation methods, we employed (1) human personality inventories [Core2, Core3], (2) open personality descriptions [Core3, Core4], and (3) adjective descriptors [Core5] to evaluate whether the user's perception of the CA's personality matches the target personality. In Chapter 4, I discuss the benefits and drawbacks of each method. Depending on the results of the evaluation, it is recommended to go back to one of the previous steps and iteratively improve the personality perception or, if satisfied with the result, to proceed with the actual implementation of the personality.

3.4 Summary of Contributions

To conclude this chapter, Table 3.3 provides an overview of the primary contributions, the type of HCI knowledge they represent based on Wobbrock and Kientz' classification [217], and the research questions they address. A table that clarifies the contributions of all authors is available in the Appendix.

RQ	Knowledge Type [217]	Primary Contribution	Contributing Publications
RQ1.1	Empirical	Empirical insights into the link between user personality and emoji usage to extend the set of verbal behaviour cues to be used in text-based CAs.	[Core1]
RQ1.2	Artefact	Set of linguistic cues derived from psycholinguistic literature to manipulate three different levels of Extraversion and Agreeableness, implemented in fully operative text-based CAs.	[Core2, Core3]
	Empirical	An empirical investigation of the suitability of human behaviour cues for synthesising CA personality perceptions.	[Core2, Core3]
RQ1.3	Methodological	An adaptation of the psycholexical approach by developing a new multi-method strategy from which we derive a set of 349 personality descriptors (adjectives).	[Core4]
	Theoretical	Ten CA Personality Dimensions, derived via exploratory factor analysis on data of 744 people rating our descriptors.	[Core4]
RQ1.4	Methodological	We transfer the approach of Enactment-based Dialogue Design from a single personality to multiple personality levels.	[Core5]
	Empirical	An empirical investigation of the suitability of Enactment-based Dialogue Design for generating personality perceptions.	[Core5]
RQ2.1	Empirical	An empirical investigation of how a CA's personality influences user preference and interaction behaviour.	[Core2, Core3, Core5]
RQ2.2	Methodological	A new approach for engaging users in CA design by eliciting their visions of a perfect speech-based CA, called Vision Dialogue Elicitation.	[Core6, Core7]
	Empirical	A set of qualitative and quantitative analyses and insights into user preferences for tailored speech-based CA dialogues.	[Core6, Core7]
RQ2.3	Empirical	An examination of the influence of user personality on subjective preference for CA personality.	[Core2, Core3, Core5, Core6, Core7]

Table 3.3: Overview of the primary contributions to each research question and the publications presented in this thesis.

4
INSIGHTS & FUTURE WORK

The main goals of this thesis are (1) developing methods to imbue CAs with personality systematically and (2) examining user preferences for CA personalities. To this end, I introduced two approaches to synthesise personality in CAs, based on two underlying personality descriptions. The first approach was to adopt the human Big Five personality model as a theoretical basis, allowing the transfer of human verbal behaviour cues associated with personality traits to CAs. To enrich this set of verbal behaviour cues, my first step was to investigate the relationship between human personality and emoji use, identifying several personality markers in emojis [Core1]. These personality markers alongside others derived from the psycholinguistic literature were then used to synthesise three levels of Agreeableness [Core2] and Extraversion [Core3] in text-based CAs. The results from this indicated that human verbal cues are partly insufficient to infuse personality in text-based CAs, calling for an examination into whether the Big Five model can be replicated for CA personality (Approach 2). My research team collected personality descriptors through a new multi-method approach, with a confirmatory factor analysis revealing ten personality dimensions for speech-based CAs. We found that these ten dimensions did not correspond with the human Big Five model, proposing the Ten CA Personality Dimensions as an alternative way of describing CA personality [Core4]. To synthesise personality based on these ten dimensions, I introduced a new method called Enactment-based Dialogue Design [Core5].

Second, I presented two approaches to studying user preferences for CA personality. In a deductive approach, my team collected users' likeability ratings for the CAs as described in the first part of this thesis [Core2, Core3, Core5]. These investigations showed that users have very individual preferences for the dimensions Extraversion and Social-Entertaining, whereas the majority prefer CAs that have a medium or high level of Agreeableness and a low one of Confrontational. To better include users in the process of developing CA personalities, I introduced a new inductive method to elicit user visions of a perfect CA, which can be used to gather requirements [Core6, Core7]. In this context, I also examined the influence of users' personalities on their preferences for CA personality, but the resulting effects were minimal.

In this chapter, I first report on the over-arching limitations of my work that readers should bear in mind when interpreting the findings (Section 4.1). Afterwards, I reflect on the methods for imbuing CAs with personality (cf. Section 4.2) and on the implications of my results regarding user preferences (cf. Section 4.3). These reflections discuss further insights beyond the single publications that emerge from dissecting the pooled findings. Furthermore, these reflections yield lessons learned and give methodology recommendations to researchers who wish to study CAs with personality. I conclude this chapter by outlining a research agenda with open challenges for future work (Section 4.4).

4.1 Limitations

The work presented in this thesis is subject to several limitations and should be understood with these constraints in mind. The limitations specific to each of the presented studies are described in the respective publication, however, there are four overarching limitations that I deem noteworthy here.

First, my studies collecting users' perceptions of personality and their preferences were all conducted in a within-subjects experimental design, in contrast to several related works [9, 27, 100, 158]. This choice of research design allowed participants to compare the different personalities against each other and choose their favourite based on experiencing all the different options. Moreover, person-specific confounding variables are parallelised. That is, for example, if a user dislikes interacting with CAs in general, this attitude will affect the perception of all CAs, not just one. Conversely, this decision to explore user preferences involves a trade-off, as it may have fostered contrasting effects in participants' perceptions of the personalities between the different CAs [139]. For example, participants may have perceived a CA as more introverted after first interacting with the opposite, extraverted agent than if they had only interacted with the introverted one.

Second, synthesising and examining dialogues that express certain personality traits requires an analysis of language, which poses several challenges. As the work by Wu et al. [220] suggests that interacting with a voice assistant in a non-native language increases mental workload, I only recruited participants in my studies who were native or fluent speakers in the language the CA adopted. On the one hand, being a native German speaker at a German university myself, collecting participant responses in German is easiest with respect to participant recruitment and ensures a complete understanding of the subtleties in participants' responses. Studies in [Core1, Core2, Core5, Core6] and partly in [Core4] were conducted in German, in line with comparable prior work on user interaction with voice assistants [221]. On the other hand, this approach requires translating participants' responses afterwards to English as the international HCI research community publishes in English, which could also lead to subtle meaning being lost in translation. Furthermore, findings such as the personality model dimensions presented in [Core4] are more useful to the wider community if presented in English. To tackle these challenges, studies in [Core3, Core7] and the final online study in [Core4] were conducted with British native speakers [Core3, Core7] or speakers with at least high English proficiency and located in the US [Core4] recruited via the crowdsourcing platforms Prolific[1] [Core3, Core7] and Amazon Mechanical Turk[2] [Core4], with English native or bilingual speakers supporting the analysis.

Third, due to the aforementioned language constraints, the examination of user perception and preferences for CAs took place in limited cultural settings. However, the interpretation of behaviour cues and subsequent attribution of personality are culturally influenced [48,

[1] `https://www.prolific.co`, last accessed 10th May 2022

[2] `https://www.mturk.com`, last accessed 10th May 2022

88, 95]. Prior work by Endrass et al. [72] demonstrated that users interpret and evaluate a CA's behaviour differently depending on their cultural background. Thus, it is crucial that future work (1) includes users from all cultural backgrounds in the requirement analysis and perception assessment, and (2) investigates whether the user preferences we found also apply to users from other backgrounds.

Fourth, investigating the influence of user personality on preferences for CA personality requires large and diverse samples because participants have to represent the distribution of personality traits in the population [Pub10]. Assembling diverse samples proved particularly difficult as participants with personality traits such as low Agreeableness are inherently less likely to participate in voluntary user studies (cf. the distribution plots of participants' personality traits in [Core1, Core5, Core7]). I addressed this challenge by recruiting large sample sizes ($N = 205$ in [Core7], $N = 646$ in [Core1], $N = 744$ in [Core4]), recruiting via crowdsourcing platforms instead of convenience sampling [Core3, Core4, Core7], and pre-screening participants according to their personality traits [Core3]. However, the sample sizes were limited due to the qualitative nature of the studies in the first explorations of this field, so small effects between user personality and their preference for CA personality may not have been detected.

4.2 Reflections on Imbuing Conversational Agents with Personality

In Chapter 1, I argued that two crucial factors are missing to empower CA designers to imbue CAs with personality systematically: (1) conceptual clarity about the underlying personality description on which CA designers can specify the CA's target personality, and (2) synthesis methods for translating the target personality into perceptible behaviour cues. In addition, in Chapter 2 I explained the need to evaluate users' impression of a CA personality and compare it to the target personality. In Chapter 3, I presented publications which explored different approaches and techniques for each of the three tasks of (1) describing, (2) synthesising, and (3) evaluating CA personality. In this section, I reflect on our work on these three tasks, summarising lessons learned, implications for CA researchers and designers, and starting points for future work.

4.2.1 Describing Conversational Agent Personality

One of the most important findings from this thesis is that perceptions of CA personality are not equal to the perceptions of human personality, as demonstrated by the unsuccessful attempt to replicate the Big Five personality model for CAs [Core4]. Whilst my results support the assumptions of the Media Equation that people treat CAs as inherently social [182], they reveal subtle differences in users' conceptualisation of artificial entities compared to humans. These subtle differences were also echoed in my speculative exploration of users' attitudes towards punishing robots in [Pub3], in which I found that users conceptually placed the robot, which they punished using different techniques, somewhere between alive and lifeless.

I have used the Ten CA Personality Dimensions to more adequately encompass perceptions of CA personality, which includes the dimension Artificial, comprising terms such as “artificial”, “synthetic”, and “robotic”. This dimension seems to be of particular significance, because it is frequently mentioned in the open-ended personality descriptions in both [Core3, Core4] and related research [36, 133, 173, 224], but has no overlap with the Big Five. From a conceptual point of view, the dimension Artificial is logical as the construct of personality is intended to distinguish the behaviour of one entity from that of another [4], and today’s CAs differ from each other in how human-like they behave and sound. Users’ desire to comment on the perceived artificiality underlines their juxtaposition of CA personalities and human personalities. This juxtaposition leads to expectations of human-like personalities that are unlikely to be met, as outlined by prior work [51, 59, 135, 178, 183]. Challenging the currently dominant human metaphor and resemblance for CAs [68], the Ten CA Personality Dimensions can thus raise awareness about deliberately shaping the artificiality of the CA personality to calibrate user expectations.

People’s interaction with CAs is of course not only determined by perceptions of their personality but also other aspects of (social) interaction [28]. My adoption of the psycholexical approach to derive the conceptualisation of CA personality has already inspired a similar approach to derive users’ partner models for CAs [67]. Furthermore, my new multi-method procedure for generating a personality item pool combined traditional psychological approaches such as interviewing users with new methods such as implicit descriptions in online reviews (following our work in [Pub1, Pub2]) and online corpora to systematically reduce the item set. Leveraging data science methods such as text analytics and web crawling hence opens up new opportunities to complement traditional psychological methods, as previously called for in Psychology [152].

Takeaway: Perceptions of CA personality are not equal to the perceptions of human personality. Instead of adopting human personality models to describe CA personality, I suggest Ten CA Personality Dimensions.

4.2.2 Synthesising Conversational Agent Personality

Two approaches to synthesise personality in CAs were explored: (1) Transferring human verbal behaviour cues to CA dialogues and (2) developing CA dialogues through Enactment-based Dialogue Design. Both approaches yielded promising results, but each has its benefits and drawbacks, as discussed below.

The transfer of human verbal behaviour cues to CAs provided the best results for high and low levels of Agreeableness, even for a short interaction time [Core2]. Conversely, differences in the levels of synthesised Extraversion were perceived to be smaller in text-based CAs, and verbal cues were insufficient to convey introversion [Core3]. Furthermore, there are several drawbacks to integrating human cues into CA dialogues to synthesise personality. First is the finding that perceptions of CA personality are not equal to human personality, raising doubts about the transferability of cues. Second, even if the human Big Five are used to describe CA personality, there are fewer personality markers associated with the dimensions Conscientiousness and Openness [53]. Third, when personality dimensions are examined as a multidimensional construct rather than in isolation [139, 181], selecting cues becomes more complicated due to interaction effects (cf. Section 4.4.1). Fourth, personality markers have often been identified by means of correlations in psychological research (e.g. [94, 172]). As such, these correlations provide little knowledge about thresholds, for example, how many positive emotions words does a CA use that has an average level of Extraversion? This lack of thresholds makes it difficult to synthesise subtle levels of personality, which were, however, favoured by up to 43% of our participants [Core2].

Creating dialogues by means of Enactment-based Dialogue Design accomplished a good match of the intended low and rather high levels in the dimensions Social-Entertaining and Confrontational, whilst the targeted high levels were not achieved in the considered dimensions [Core5]. In contrast to the arduous task of identifying human behaviour cues from literature scattered across decades of psychological research, this method was able to synthesise suitable dialogues in focus group sessions of one and a half hours which consisted of amateur actors without expert knowledge of CA design. On the other hand, this approach makes it more difficult to exploit synergies from other personality synthesis projects because it is unclear which specific cues in the dialogue caused the perception.

Contrary to the study of human personality markers, the synthesis of personality in CAs has the advantage that dialogues differing only in single cues can be generated and compared with each other to determine the significance of different cues for personality perception. However, for n cues more than 2^n dialogues[3] have to be created to reflect all combinations of cues because the cues are not simply additive but interdependent [196]. Hence, such an approach is an extensive undertaking. Another, more pragmatic possibility for future work to advance our understanding of personality synthesis could be to further develop our Enactment-based Dialogue Design method. One could ask many dialogue experts, for example screenwriters and authors, to write dialogues for specific target personalities in focus groups, compiling a corpus of dialogues. This corpus could be used in two ways: First, text analytics could be applied to examine the dialogues for syntactic and semantic similarities, thereby deriving a set of cues. Second, deep learning approaches could be implemented, such as language models from Natural Language Processing, to automatically generate dialogues for specific

[3]To illustrate this number, an example is given: We assume that there are three behaviour cues, such as the use of the positive sentiment word “great”, the use of a heart emoji, and the use of an exclamation mark. In this case, $2^3 = 8$ CA dialogues need to be created to evaluate users’ perception of the personality for each combination: one dialogue without any cues, three in which only one cue is present, three in which the combinations of two cues are present, and one in which all three cues present. In reality, there is a large number of these cues, not just three, as well as cues which are not binary but metric, such as text length.

target personalities, following related work on automatically generating CA dialogues [138, 184, 185, 186]. Whilst the latter approach is less laborious in the synthesis phase, deep learning approaches require huge corpora to ensure a reasonable level of accuracy.

Despite these opportunities, CA researchers and designers should be aware that the interpretation of behaviour cues will always be equivocal, as communication is ambiguous [139]. This ambiguity of behaviour cues is illustrated particularly well by the various interpretations users provided for two identical chats that differed only by a single emoji at the end of the message in our work in [Core1]. Therefore, it is unlikely that even successful synthesis methods will ever achieve complete and universal accuracy in users' perceptions of the personality. Moreover, future work ought to examine the interplay between the attributed personality and the situation on users' perceptions of the CA's behaviour. For example, humans tend to overestimate the influence of another person's personality on their behaviour, whilst disregarding situational influences [189]. Yet, a recent exploration by Edwards and Edwards [71] suggests that there are differences between people's causal inferences for humans and robots that require further investigation.

Takeaway: I presented two methods for synthesising CA personality. Transferring human behaviour cues to CAs was particularly successful in synthesising different perceptions of Agreeableness, but is difficult to scale for other dimensions. Our Enactment-based Dialogue Design method was able to generate different levels of CA personality in focus groups and could be developed further to automate the synthesis processes.

4.2.3 Evaluating Conversational Agent Personality

An important part of synthesising CA personality is to evaluate whether users' perceptions of the CA personality match the intended target personality. Although it was not the focus of this thesis to systematically compare different evaluation methods, the experience gained can assist CA designers and researchers to select the appropriate method for their projects.

A Comparison of Methods for Evaluating CA Personality

To evaluate users' impressions of CA personality, this thesis used (1) human Big Five inventories [Core2, Core3], (2) open personality descriptions [Core3], and (3) ratings on the personality adjectives collected in [Core4] as a precursor to CA personality inventories [Core5], which is summarised in Table 3.1. In addition, user perceptions of the personality of existing CAs in [Core4] were collected through open-ended descriptions in an online survey, a lab experiment, and online reviews.

Personality inventories, whether established human inventories [61, 201, 202] or the personality adjectives used in [Core5], promote rapid assessment and easy comparison of different agents. Moreover, they can cover the entire spectrum of perceived personality dimensions and thus reveal impressions about less noticeable personality dimensions. On the other hand, [Core3] found that human personality inventories do not capture the full picture of CA personality, echoing prior work which also brought to light descriptions such as “artificial” [224]. For both the human inventories and the CA personality adjectives, single items emerged as being inadequate for the comparison.

In contrast, open-ended personality descriptions allow participants to express their impressions in their own words, as opposed to inventories that pre-categorise personality perceptions in what Liu et al. [133] termed “terse and efficient” wording. As participants are not biased towards certain personality characteristics, open descriptions can confer an impression of which personality characteristics are most salient to them. Conversely, the interviews in [Core4], where open personality descriptions were collected, revealed that participants often felt uncomfortable describing the personality of a CA. To avoid this problem, we gathered implicit personality impressions from online reviews in [Core4]. Whilst these reviews provide a first impression of the most salient characteristics (e.g. “helpful” was a prominent descriptor in online reviews) and offer insights into the variety of possible descriptors [Core4], a full personality analysis is rather cumbersome. This is because many reviews do not contain any personality description, and although adjectives can be easily extracted automatically, not all of them reflect personality traits (e.g. the descriptor “helpful” may refer to the CA’s personality or to the concept of assistants in general). Thus, manual context-aware analysis is required. Finally, all kinds of open descriptions are time-consuming to collect and analyse.

Takeaway: Both evaluation methods, (1) personality inventories and (2) open descriptions, have benefits and drawbacks and must be used according to the primary research goal. As CA designers and researchers may not always have the time or financial resources to triangulate different evaluation methods, it is imperative to develop CA personality inventories which enable fast assessment. A first step towards dedicated CA personality inventories is discussed in Section 4.4.1.

Recommendations for Comprehensive Reporting of CA Personality Perceptions

To gain a better understanding of user preferences, a profound knowledge of how users actually perceive the developed CA personality is crucial. However, methodological deficits in reporting these user perceptions are evident in related work due to a lack of reporting standards, a problem well known to HCI [38, 111, 207] and Psychology [128]. Specifically, in prior work on user preferences for CAs, user perceptions of CA personalities were only compared in relative terms (e.g. [9, 100, 160, 191]). For example, Andrist et al. [9] found a significant difference in participants’ Extraversion ratings of two robots based on their gaze behaviour, but the absolute difference between the two versions was rather small, with both closer to a neutral Extraversion score than to the extremes. Similarly, [Core3] findings showed a significant difference between the perceptions of the extraverted and introverted

CAs. However, the absolute mean Extraversion score for the introverted CA (cf. Figure 3.2), supported by the open-ended descriptions, revealed that it was not recognised as such.

Whilst these studies advanced the understanding regarding the similarities of agent and human personality perceptions, I argue that research should move from relative comparisons to providing absolute personality scores based on norm values. That is, future work should define thresholds that constitute a low, average, and high level of a personality dimension, thereby providing a frame of reference.

Other challenges encountered in this research and elsewhere also highlight the need for comprehensive reporting on users' perceptions of CA personality. Liu et al. [133] showed that participants often chose "neither agree nor disagree" personality ratings for their CAs, emphasising the need for measures of variance and confidence intervals in CA personality evaluations. This work suggests that some of the Big Five inventory items are less suitable for capturing user impressions of CA personality (e.g. *is less active than other people*) and thus may distort the overall dimension score [Core3, Core5]. Publishing the means for all individual inventory items along the dimension mean may also be informative until CA personality inventories have been established.

Takeaway: To ensure a transparent and comprehensive overview of user personality perceptions, I recommend reporting descriptive statistics including absolute means, measures of variance, confidence intervals, and box plot diagrams for user perceptions of CA personalities as a best practice. Furthermore, future work ought to determine norm values for personality levels as a reference frame.

4.3 Reflections on User Preferences for Conversational Agent Personality

In this section, I discuss whether universal preferences emerged for any of the investigated personality dimensions and for which of these dimensions users' attitudes diverged. To inform practitioners in the field, I thus point out those personality dimensions for which multiple levels should be synthesised so that they can then be tailored to the individual user.

4.3.1 One Size in Personality Does Not Fit All Users

Both the deductive approach, presenting users with various pre-designed personalities [Core2, Core3, Core5], and the inductive approach, eliciting user visions of the perfect CA personality [Core6, Core7], clearly show that users have *individual preferences* for CA personality. Echoing previous results by Nass et al. [162], the publications presented in this thesis thus demonstrate that there is no single CA personality that is universally liked by all users. This lack of a single preferred CA personality underlines the necessity of tailoring CA personalities to the user, as conceived in this thesis' vision. The publications in this thesis reveal important insights into the personality dimensions that should be considered for CA personality design in general and for adaptation in particular.

Universal and Controversial Preferences

Figure 3.5 shows the CA personalities that users liked best with regard to the four dimensions I examined in this thesis. Most participants clearly preferred average or high levels of Agreeableness for a movie recommender text-based CA [Core2], whilst they rejected high and rather high levels of Confrontational for an in-car speech-based CA [Core5]. Although the majority of participants ranked an extraverted text-based CA for a daily stress tracker first, the remaining participants completely rejected this CA and would rather have liked to interact with an average or introverted CA, as informed by the participants' open-ended rationales for their ranking of the CAs [Core3]. A very diverse picture also emerged for the dimension Social-Entertaining for the in-car speech-based CA, where participants' preferences were distributed across the three levels and were also determined by the individual task [Core5].

These preferences can be considered reliable only for the context in which they were collected, as they are context-dependent [Core5]. These insights into users' actual behaviour when presented with different CA personalities thus confirm previous qualitative work on users' expectations of context-specific CA personalities [101]. Nonetheless, my findings suggest that overall most users demand friendly and agreeable CAs which do not challenge the user, which is in accordance with a typical assistant's role [Core2, Core5, Core7]. Hence, Agreeableness may be a one-size-fits-all personality trait, but even with this trait, users differed as to how pronounced it should be [Core2]. In particular, regarding the CA's compliance (a sub-facet of Agreeableness), users' visions of the perfect CA revealed that some users conceive an opinionated CA which sometimes confronts the user [Core7], albeit probably less aggressively than the confrontational and disagreeable CAs we designed. This preference was also evident in more diverse opinions about a CA that nudges the user to make healthier meal choices in the *Navigation* scenario in [Core5]. Hence, our results suggest that the **use of opinions and nudging** is highly individual, presenting an opportunity to tailor CA personalities to the user.

Other characteristics that yielded controversial preferences and should thus be contemplated for adaptation are the inclusion of **humour** and a **casual demeanour**. These controversial preferences were evident in users' visions of the perfect CA, where some participants made extensive use of witty rejoinders from the CA, whilst others completely avoided the use of humour [Core7], and in the divergent opinions on in-car CAs that were (rather) high in Social Entertaining [Core5]. Prior work suggests that many users enjoy humorous comments, so-called "Easter eggs", when interacting with voice assistants [51, 135], but they do not regard them as an integral part of conversations as they do with human interlocutors [51]. The studies in this thesis extend this earlier work by revealing that the appeal of a humorous CA is highly individual and highlight the challenge of designing humorous CAs. Users imagined humour in CAs in the form of contextual witty comments [Core7], which is neither implemented in today's CAs [51] nor technically feasible yet [171]. Humour is often used in conversation to soften serious topics [51]. A cheerful and informal demeanour in conversation is associated with Extraversion and was thus synthesised in the extraverted CA in [Core3]. Users' rationales suggested that some enjoy this kind of causal conversation with a CA, whilst others find it unprofessional or inappropriate for a CA.

Takeaway: User preferences for CA personality are highly individual and context-dependent. Users universally prefer a higher level of Agreeableness, whilst the use of opinions, nudging, humour, and a casual demeanour trigger different reactions from users. Therefore, these characteristics offer the opportunity to tailor the CA personality to individual user preferences.

Strong or Subtle Personalities? When an Average Level of Personality is Beneficial

Prior work is divided as to which personalities are most popular. Reeves and Nass [182] state that users appreciate obvious, strong personalities, which make it easier for users to predict the agent's future behaviour. Conversely, Pearl [171] recommends that conversation designers aim for subtle personalities that are "neither loved nor hated" [171] in the case of a CA that is used by many people.

To the best of my knowledge, this research is the first to not only synthesise two opposing, strong levels of personality in CAs, but to include a third, subtle level that creates an average or an average to high level of a personality trait. The reason for including this third level was that previous research points to a similarity attraction effect (e.g. [9, 157]) and assumes that the distribution of personality in the population resembles a Gaussian curve [144] (this assumption is supported by the personality distributions shown in [Core1, Core5, Core7]). It is, therefore, to be expected that an average level of personality in a CA will be similar to most users and their usual human interlocutors.

Contrary to my expectation, my findings do not suggest that an average CA is generally preferred by the majority of users. The average agreeable CA was ranked first by 43.4% of the participants [Core2], and the CA with a rather high level of Social-Entertaining by 36.3% of the participants for the *Music* scenario [Core5], which corresponds to the likeability scores of the respective high level agents. Conversely, for the dimension Confrontational, participants almost unanimously agreed on the low level CA as their favourite. These findings demonstrate that (1) average levels are likely only beneficial for certain dimensions and (2) a preference for extreme or more subtle personalities also seems to be highly individual. Moreover, the results indicate that for dimensions usually associated with negative dispositions, such as Confrontational, users' consensus is on the low level, making the design of several levels unnecessary. On the other hand, for dimensions more typically associated with assistants, users seem to be divided in their preference for either a strong, high level personality or a more subtle, average personality, providing another opportunity to tailor CAs to individual preferences. However, due to the limited number of dimensions and contexts examined in this thesis, future work will need to corroborate these assumptions.

In contrast to my observations for the dimensions Agreeableness and Social-Entertaining, analysing user preferences for Extraversion showed that the average personality is significantly less liked than the high level version. Participants' rationales for their ranking suggest that the perception of a CA's artificiality is detrimental to their subjective preference and engagement. Future work should examine whether this rejection of an average extraverted CA by the majority of participants is due to (1) the specific context of a stress-tracking application, (2) the perceived artificiality of the agent, or (3) a general preference for extraverted CAs, which would contradict previous work on robots [9] and voice user interfaces [100].

Takeaway: Although an average level of personality reflects most people's personality, it is not universally preferred. Instead, average levels are only beneficial for certain dimensions, in which case users' preferences for a strong or subtle personality are divided.

The Salience of Agreeableness in CA Personality Perceptions

In [Core3], participants' open-ended personality descriptions of the three text-based CAs showed that the CAs stood out most for their Agreeableness, although this dimension was not deliberately targeted in the design. Interestingly, the attributes referring to the CAs' Agreeableness (e.g. "friendly", "helpful") were also frequently mentioned to describe Siri, Alexa, and the Google Assistant in our online survey and lab experiment in [Core4]. Similar observations have been made in prior research [133, 134]. For example, Liu et al. [133] found that their participants frequently commented on the Agreeableness of their ECAs, even though they had manipulated the agents' gestures to vary perceptions of the personality dimension Neuroticism.

As most CAs today are designed to be helpful assistants, the perception of their friendliness is not surprising in itself. Nonetheless, it is striking that these characteristics seem to be so salient to the majority of participants that they are mentioned first in a variety of CA personalities and contexts. This focus on Agreeableness is also unexpected because Extraversion has the closest links to observable *human* behaviour [172]. As a reason for this phenomenon, Liu et al. [133] suspected that people are used to classifying others as friendly and unfriendly. Another explanation could be that a friendly, helpful assistant is in line with users' vision of a CA [Core7] so that users primarily express perceptions that may or may not match this expectation. This explanation is in line with Laurel [123] who assumed that people "paint with bold strokes" when attributing personality to CAs and thus only consider traits that are useful for the specific context.

[Pub9] analysed the occurrences of the term "intelligent user interfaces" in all papers published in the ACM Intelligent User Interfaces conference within the last 25 years. This analysis also yielded that "to assist" is the verb most often associated with intelligent user interfaces in general and agents in particular, highlighting that users' expectations have been steered in this direction for decades. As users often find it difficult to reconcile their rational knowledge that an agent is a machine with their emotional response to it [Pub3], they may be more comfortable using terms that are consistent with the assistant role when asked to describe personality, an inherently human concept. Future work should scrutinise this phenomenon to shed light on the reasons for the salience of Agreeableness in CAs.

Although the frequency with which descriptors refer to the Agreeableness of CAs does not necessarily mean that Agreeableness is particularly important for the design of target personalities, it does at least indicate that this dimension stands out to users. In [Core3], more than half of the participants gave a CA's friendliness as a reason for its high ranking, corroborating the importance of friendliness for user preferences. Hence, designers should carefully design the perception of Agreeableness, whereas less salient dimensions may require less attention unless designers explicitly want to achieve a particular impression. Fortunately, the work in [Core2] shows that synthesising Agreeableness was comparatively straightforward and that users have a clear preference for higher Agreeableness levels. The Ten CA Personality Dimensions also include three traits that can be seen as a refinement of the singular Agreeableness trait in the human model, namely, Approachable, Social-Inclined, and Social-Assisting [Core4]. Whilst this is another indicator of the importance of these characteristics, it also underlines that Agreeableness should be designed on a more fine-grained level in CAs.

Takeaway: Characteristics of Agreeableness in CAs are particularly salient to users, suggesting that designers should take extra care in designing this trait.

4.3.2 User Preferences Are More Complex than Similarity Attraction

Early work by Nass and Brave [157] highlighted that user preferences for personalities of voice user interfaces are determined by a user's own personality, suggesting a similarity attraction effect. Later research partially supported this assumption (e.g. [21, 200]), but also challenged it by implying a complementary effect (e.g. [100, 132]) or no effect at all (e.g. [36, 190]). The publications presented in this thesis point to a limited influence of user personality on preference for a CA personality, both in terms of users' visions of the perfect CA [Core6, Core7] and their ratings of the CA personalities they interacted with [Core2, Core3, Core5].

A possible reason for this lack of an effect could be that the samples taken were too small to detect any influence (cf. Section 4.1). To better understand which effect sizes could be discovered, we performed an a priori power analysis in [Core7] and a sensitivity power analysis in [Core3]. However, as this type of effect is typically small, hidden small effects of user personality on their preferences cannot be ruled out. Despite acknowledging this possibility, I conclude that user preferences cannot fully be explained by similarity attraction, but represent a more complex phenomenon.

As this thesis' findings challenge the applicability of the Big Five model for CA personality, the concept of similarity attraction has to be questioned as well. As explained in Chapter 2, I assume that a target personality is defined along all dimensions to ensure consistency of the expressed behaviour. That is, a CA is not simply classified as extraverted or introverted, but has a certain personality expression, a specific level, on all Ten CA Personality Dimensions. With five human personality traits and ten deviating CA personality traits, a simple similarity mapping is not possible but requires further investigation.

When synthesising personality for speech-based CAs based on our personality dimensions, we found that participants high in Extraversion tended to prefer a CA with high levels of Social-Entertaining [Core5]. Such a CA is characterised by a cheerful, assertive, and humorous demeanour similar to an extraverted human. This preference suggests matching extraverted users with CAs high in Social-Entertaining. Moreover, two links between user personality and preferences emerged that merit further scrutiny. We found that Conscientiousness is a negative predictor for (1) preferring a CA with a confrontational disposition [Core5] and (2) envisioning a CA offering opinions [Core7], indicating that conscientious users tend to prefer compliant CAs that are more restrained in their comments. In Section 4.4.2, I discuss potential future research strands to further investigate determinant factors of user preferences.

Takeaway: Our findings indicate that there are hardly any universal truths when it comes to user preferences for CAs. Instead of falling back to overarching rules such as "similarity attracts", the publications in this thesis highlight the need to gather user requirements for CA personalities in a specific application and develop the personalities in accordance with users' visions, for which the methods presented in Section 3.3 can be used.

4.4 Research Agenda

This thesis is guided by the vision to imbue CAs with personality systematically and tailor them to user preferences. Whilst the publications presented in this thesis have led to methods and empirical findings, there is still much work to be done to achieve this vision. In this section, I propose several open challenges that emerge from the approaches and studies presented. Based on my vision, I first discuss the challenges of imbuing CAs with personality, followed by the challenges of tailoring CA personality to users. This section concludes with a critical reflection on the acceptance of personality-imbued CAs as well as ethical considerations.

4.4.1 Open Challenges for Imbuing Conversational Agents with Personality

In Section 3.3, I introduced a process for designing CA personalities and presented methods for each of the steps. In this subsection, I describe opportunities for future work to further

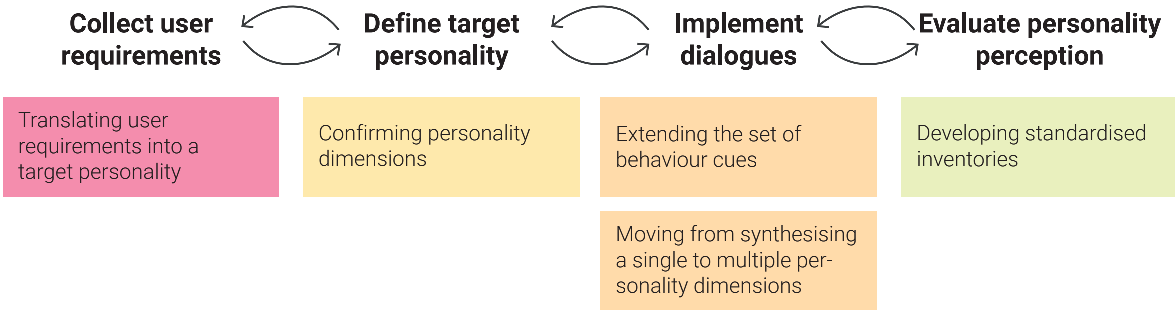

Figure 4.1: This figure shows opportunities for future work to further develop the methods of the design process for CA personality. Each box describes a challenge and is displayed below the corresponding design process step.

develop these methods and tools. Figure 4.1 provides an overview of these opportunities and which of the process steps they address.

Translating User Requirements into a Target CA Personality

Whilst the presented design process builds on my experience with designing CA personalities, it has not yet been applied in practice by conversation designers. However, only the practical application can cast light on whether this process is complete. Specifically, future work has to probe whether CA designers need a method or tool to define a target personality based on user requirements. If user requirements are gathered by means of our Vision Dialogue Elicitation method, it may not be self-explanatory how visions such as providing personalised recommendations translate into a specific target personality.

Confirming Personality Dimensions and Developing Standardised Inventories

The paper [Core4] presented a new multi-method approach based on the psycholexical hypothesis to derive personality dimensions dedicated to speech-based CAs. These derived dimensions should not be considered definitive but rather a starting point for future work. In addition, the work in [Core3] demonstrated the need for standardised personality inventories, which allows researchers and designers to evaluate the perceived personality against the target personality. Similar to the development of personality models and inventories for human personality in Psychology [142], these two steps are intertwined as model and inventory both aim to comprehensively describe the CA personality.

Specifically, the next step towards a personality model for CAs is the validation of the derived dimensions through a confirmatory factor analysis [181]. That is, another large sample should rate the 349 personality adjectives we collected in the multi-method approach, and a confirmatory factor analysis should be conducted on the resulting ratings. This should then be reviewed to determine whether the same ten dimensions emerge.

Once the personality dimensions are confirmed, the development of a personality inventory is the next step. For this, well-established Big Five inventories, such as the Big Five Inventory [202] or the NEO-PI-R [57], may be used as templates but adapted to use the personality adjectives and dimensions dedicated to CAs. The developed inventory would then need to be assessed against test construction criteria, such as reliability and validity [31].

However, researchers and designers should be aware that CA personality models will not be conclusively developed in the next paper. By comparison, several researchers have worked on human personality adjectives and models for decades [105, 181]. First, whilst the Big Five model is the most predominant model to describe human personality today [142], future work may examine the suitability of other personality taxonomies for describing CA personality, such as Eysenck's Big Three [74] or multidimensional structures (cf. Section 4.4.1). Second, it is likely that our understanding and attribution of CA personality and other social processes, and thus their formal description will change in the future as CAs are integrated more ubiquitously in our lives. In particular, the generations which are born now and do not know a world without artificial agents may categorise them differently than previous generations.

As a reminder, personality is defined by *distinguishable* and *distinctive* patterns of behaviour, cognition, and emotion. As such, items in personality inventories have to be actually *distinctive*. For example, McCrae and Costa [142] described a sub-facet of the dimension Openness for Experience as knowledge of foreign cuisines. Whilst this was likely a distinctive behaviour during the initial development of the Big Five, it is less distinctive today. Similarly, due to the current technological possibilities for synthetic voices, today's CAs may be distinguishable by how natural and human-like their voices sound, which is depicted in our dimension Artificial [Core4]. However, as technology advances, this characteristic may be less of a distinguishable feature in ten years.

Extending the Set of Behaviour Cues

The focus of this thesis is on the synthesis of personality at the dialogue level, i.e., the targeted manipulation of verbal behaviour cues. Whilst the manipulation at the dialogue level is necessary for all CAs, non-verbal cues, such as gestures, facial expressions, or proxemics, also have to be considered for speech-based CAs or ECA. For speech-based CAs, these non-verbal cues are para-verbal, such as pitch, loudness, and speech rate. As non-verbal communication tends to eclipse verbal communication [10, 90, 150], the deliberate design of these cues for speech-based CAs is inevitable. Whilst Scherer [196] gives a comprehensive overview of possible human para-verbal personality markers, to the best of my knowledge there is neither such an overview for human non-verbal personality markers nor a complete investigation of their applicability to artificial agents.

Previous work on voice user interfaces, ECAs, and robots has examined the transfer of individual non-verbal cues, such as gaze [9] and proximity [36], as well as combinations of cues, such as gestures with facial expressions [145] or dialects [118]. However, their work also reveals that the transfer of these human behaviour cues works only to a limited extent, with some cues not eliciting the intended personality perceptions [36]. Furthermore, few have pondered the effects of merging non-verbal and verbal cues (exceptions include [100, 158]).

As the consistency of non-verbal and verbal cues is crucial for personality perceptions [100], future work has to scrutinise the effect of non-verbal cues in accordance with their verbal counterparts. Instead of transferring human non-verbal cues, future work could explore whether our approach of Enactment-based Dialogue Design can be used to generate these non-verbal cues. Specifically, actors' changes in movements or voices could be analysed to inform the design of CA personality.

Beyond verbal and non-verbal cues, future work could also examine the use of pragmatic cues, such as the number of notifications or if a CA initiates conversations, so as to extend the expressiveness of verbal cues. Moreover, perceptions of personality have to be examined in context with other social cues, such as perceived gender or age. For example, prior research highlighted that users assign personality based on dialects [118], which in turn influences users' lexical choices [58]. Feine et al. [75] conducted a structured literature review with a taxonomy of these social cues, which can serve as a starting point for researchers.

Moving from Synthesising a Single to Multiple Personality Dimensions

In this thesis, a CA's personality was manipulated along single dimensions, such as Extraversion, Agreeableness, Social-Entertaining, or Confrontational, to examine the suitability of new approaches to synthesise personality. However, as [Core3] showed, users will also perceive personality characteristics on the other dimensions. Therefore, a simultaneous synthesis of all dimensions is needed to ensure consistency. Whilst the original Big Five model understands personality traits as a collection of single orthogonal traits [64, 139], several works have conceived personality to be a multidimensional construct [181]. The reason for this is that in factor analyses descriptors often correlate with more than one dimension, for example "cheerful" loads on both Agreeableness and Extraversion. Under this assumption of multidimensionality, a synthesis of one dimension would also influence users' perception of the other dimensions.

Synthesising personality along multiple dimensions is particularly challenging. First, as [Core2, Core3] demonstrated, identifying behaviour cues associated with personality and implementing them in CAs is already a very cumbersome task for a single dimension and even more so for multiple. Second, it remains unclear how to synthesise personality in case of opposing cues for two dimensions. For example, my study on the relationship between users' personality and their use of emojis found that high Openness is associated with a lower use of heart emojis but high Extraversion manifests itself in a higher use of heart emojis [Core1]. What do these conflicts mean for the design of CA personalities that are high both in Openness and Extraversion?

Hence, an approach as presented in [Core5] seems more feasible when it comes to the design of multiple personality dimensions in CAs. However, it has to be examined whether (1) this approach is feasible for multiple dimensions and (2) untangling single cues associated with different levels of different personality dimensions will be more useful for a sustainable agent design because personality designs may be reused for other agents.

4.4.2 Open Challenges for Tailoring Conversational Agent Personalities to User Preferences

This thesis examines user preferences for CAs to inform the process of tailoring CA personalities to user preferences. Inspired by frameworks for research agendas in Personality Psychology [120], I discuss future work on causes for individual preferences and the effects of tailoring CA personalities to user preferences. In addition, I suggest investigating the timing of adaptation.

What are Causes for Individual User Preferences?

Although the publications in this thesis reveal clear individual differences in user preferences for CA personality, it remains unclear which user characteristics drive these different preferences. However, knowledge of these characteristics is crucial for automatically tailoring CA personalities to users. Whilst this thesis has focused on user personality and its influence on preferences, future work should delve into other possible user characteristics, such as gender, age, and cultural background [210]. The influence of user personality could also be examined at a sub-facet level rather than on a dimension level. For example, a study by Zhou et al. [224] shows that users' level of cheerfulness (a sub-facet of Extraversion) was positively associated with a preference for an extraverted chatbot, whereas users' level of excitement-seeking (another sub-facet of Extraversion) was not. Finally, users' goals and circumstances could also be taken into account when analysing their preferences, for example, it might be possible that users living alone would prefer more sociable smart speakers.

What are the Effects of Tailoring CA Personalities to User Preferences?

In this thesis, I investigated user preferences for CA personalities based on their likeability ratings, i.e. how much users would like to interact with a CA again. [Core3] showed that CA personalities not only determine users' attitudes but also their behaviour, more precisely their engagement with the CAs. Hence, future work should examine in which other aspects of users' attitudes and behaviour user preferences are manifested. Han et al. [89] propose several measures for quantifying chatbot effectiveness, such as user trust, engagement duration, and chatbot repetition rate, which can serve as starting points. Apart from quantitative measurements, my work highlights the advantages of including qualitative reasons for users' preferences to better understand underlying needs [Core3]. These reasons primarily focus on the perceived friendliness, humanness, and professional vs casual demeanour of the CAs. Future work should investigate whether patterns arise in these rationales, which in turn may shed light on which personality dimensions are particularly important to users.

When to Tailor the CA Personality to the User?

The publications presented in this thesis highlight individual user preferences for CA personalities and thus the potential benefits of tailoring CAs to the users. Another important challenge in this context is the question of *when* a CA personality should be configured to best match a user's preferences [103]. There are different possibilities for the timing of the configuration.

- **Give users the choice:** Users could either choose a CA personality upon using a new application based on a description or interact with different personality versions of a CA and then choose the one they like best. On the one hand, this way, users have the opportunity to experience all versions and choose the one they like best themselves. On the other hand, users might drop out at the beginning of using the CA if they do not like the first personality presented.
- **Select a personality for the user before the first contact:** If future work can identify user characteristics that determine their preference for CA personality, and the CA application has this knowledge about the user upon first use (e.g. via the user's personal data stored on the smartphone or questionnaires during the setup), the CA app could select the best CA fit on first use.
- **Adapt the personality during the interaction:** A third possibility is that the CA app learns more about the user's preferences during use so that the app starts with a default personality and then transitions into a more preferred version. On the one hand, Nass et al. [162] suggest that people like it when the interlocutor adjusts their personality to them because it is a form of implicit praise. On the other hand, a change in personality violates basic user interface design guidelines such as the need for consistency. Personality adaptation during interaction allows users to perceive the CA in different situations, as a reliable perception of personality can only result from observing behaviour in aggregated and relevant situations [139, 219]. That is, a CA high in Social-Entertaining can only be perceived as such if it is used in social contexts and not only for purely task-related purposes.

Thus, future work is necessary to identify the most suitable time to adapt the CA personality to the user and how to communicate a potential change or selection in the interest of transparency, which I will address in the following challenge.

4.4.3 Acceptance & Ethical Considerations of Designing Personality-imbued Conversational Agents

Due to the growing ubiquity of CAs, it is paramount to contemplate ethical issues and user acceptance in the context of imbuing CAs with personality. Informed by my research on users' attitudes towards personality-imbued CAs [Pub8] and robots [Pub3], as well as the challenges of everyday intelligent systems [Pub1, Pub2, Pub9], I discuss three ethical aspects that I deem important to address when realising this thesis' vision.

Toxic CA Personalities

My vision is motivated by the influence of CA personality on user attitude and behaviour, which has the potential to create better and more tailored user experiences [28, 157]. However, CA designers who intentionally manipulate CA personality could also abuse this influence. For example, CAs could be equipped with toxic personalities, such as charisma along with aggressiveness and abusiveness, and installed in people's homes. Various dystopian scenarios are conceivable for such a toxic CA personality. For example, governments or companies could use such agents to systematically influence users' opinions, or fraudsters could try to manipulate vulnerable users into trusting them with their resources. Hence, it should be made mandatory that supervisory authorities define rules for the use of CA personalities, similar to what Shneiderman [198] proposed for human-centred artificial intelligence systems in general, and that security mechanisms be developed and made integral into their construction.

Verbal Abuse of CAs

Another challenge is that voice assistants have frequently become targets of verbal abuse by their users [49, 213]. This phenomenon is not exclusive to voice assistants but has also been observed with robots in malls [30] or computers that are shouted at by their frustrated users [47], which called for our examination into users' attitudes towards punishing robots in [Pub3].[4] Being designed as subservient assistants, today's popular voice assistants suggest a certain tolerance for this verbal abuse [54]. For example, Siri used to answer "I'd blush if I could" upon being called a "bi***" [213]. Although voice assistants now refuse to respond to this kind of sexual harassment, they continue to be available to users after such comments.

As people cannot hurt the feelings of an artificial agent, why should we care about verbal abuse? First, abusive behaviour in human-agent interaction could also transfer back to inter-human interaction [214]. Second, although today's popular voice assistants refuse to identify with a human gender when asked, they have female names and are presented with female voices by default [171]. By associating female-sounding voice assistants with subservience and tolerance for sexual harassment, gender stereotypes and inequalities could be perpetuated or even reinforced [213].

Previous work suggests that personality perceptions, among other factors, influence users as to whether or not they will abuse agents and robots [15, 98, 188]. In our exploration of users' attitudes towards robot punishment, we also found that a robot's response to abuse (e.g. wriggling as a response to blinding) makes participants uncomfortable about administering the punishment [Pub3]. The research presented in this thesis demonstrates that most users envision CAs as friendly, obedient helpers [Core5, Core7]. Despite these expectations, future work should explore whether personalities can be moulded to assist the user without tolerating abuse or perpetuating gender bias.

[4] It should be noted that this abusive behaviour is by no means pervasive among all users [213]. On the contrary, many users apply social conversation rules, such as saying "thank you" and "please", even though they are talking to an agent [Core7, 126]. In our study on users' attitudes towards punishing robots, we also found that many users are reluctant to punish robots and feel uncomfortable scolding them [Pub3].

Acceptance and Transparency of Personality-imbued CAs

CAs with a human-like personality are likely to (1) be used for tasks where they replace human-human interaction, and (2) collect huge amounts of personal data, including users' own personality traits. In [Pub8], we investigated users' attitudes towards such a personality-imbued CA provided by the platform Juji[5]. This text-based CA interacts with users for the purpose of customer service tasks and job interviews and analyses their language to infer their personality. Users' interactions with the Juji CA have been the subject of several studies (e.g. [89, 222, 223]). Our findings show that users underestimated how much information such a CA could extract from a seemingly unobtrusive chat. Notably, although the CA's personality was developed under scrutiny by Zhou et al. [224] and Li et al. [130] to engage users in the conversation and share information, none of our participants commented on their personality perception of the CA. Whilst most participants had a rather neutral attitude towards personality assessment CAs, they considered their own personalities as sensitive data and were reluctant to share the generated profiles with people other than family members and close friends. In addition, it was difficult for participants to "trick" the CA into distorting their personality profile, suggesting a limited understanding of these agents [Pub8].

These results and possible scenarios raise the question of whether personality assessment and personality adaptation should be made transparent to the user, and to what extent the user should be able to control these processes. In our research on user problems with intelligent everyday systems based on topic modelling and qualitative analysis of online reviews, we found that users often experience a loss of control when interacting with these opaque systems [Pub1, Pub2]. The need for explanations, trust, and privacy are common challenges for intelligent user interfaces [Pub9]. Future work should therefore examine not only which factors drive user acceptance of personality-imbued CAs, but also how these agents can offer users control over different personality versions as well as data collection and privacy.

[5] https://juji.io, last accessed 10th May 2022

4.5 Closing Remarks

> ***Samantha:*** *"The DNA of who I am is based on the millions of personalities of all the programmers who wrote me. But what makes me me is my ability to grow through my experiences. So basically, in every moment I'm evolving, just like you [...]."*
> ***Theodore:*** *"Well you seem like a person, but you're just a voice in a computer."*
> ***Samantha:*** *"I can understand how the limited perspective of an un-artificial mind would perceive it that way. You'll get used to it."*
>
> **– Excerpt from the film *Her* (2013): Dialogue between the human protagonist Theodore and his CA Samantha.**

The long-term vision of this thesis is to enable CA designers and researchers to synthesise different CA personalities and tailor them to individual user preferences. To this end, this thesis contributes methods, artefacts, and empirical findings to imbue CAs with personalities. Furthermore, our work provides insights into individual preferences for CA personalities, which can be used to tailor CAs to users.

However, the above quote illustrates a dilemma that could arise if CA designers eventually become *too* good at creating tailored CAs. In the film *Her*, the protagonist Theodore falls in love with his voice-controlled operating system, Samantha, because Samantha is perfectly tailored to his needs and wishes, as determined by a questionnaire during the setup. Pop culture has painted various of these utopian or dystopian (depending on the point of view) CAs close to technological singularity. In the light of this potential future, do we really want to completely blur the line between human-human and human-agent interactions? Do we tailor CAs until they represent the perfect human-like interlocutor for a user, or do we leave it at a certain deliberate imperfection?

Researchers have raised doubts about using the human metaphor for the interaction with CAs [68]. One reason for these doubts is that this metaphor leads to exaggerated expectations that the agents often cannot fulfil [59, 135, 178]. Furthermore, as discussed above, several ethical issues and potential dangers arise from this metaphor. This thesis has demonstrated that users perceive the personalities of CAs differently from how they perceive humans [Core4] and that a CA's artificiality is a feature that they notice [Core3, Core4]. Hence, I would urge CA researchers and designers to deliberately use these differences in perceptions to draw clear boundaries between the two.

Despite these challenges, my work has highlighted the powerful effect of deliberately shaping a CA's personality on users' likeability ratings and engagement. The proliferation of CAs is expected to continue to increase in the coming years [50]. It is expected that CAs will permeate even more domains, such as the (mental) health sector [22, 136] (the context for the work in [Core3]), intelligent tutoring [199], or socially assistive robots [19]. Creating these tailored experiences could open up new frontiers in how people interact with machines. These tailored experiences could include a personal tutor whose personality is tailored to engage the student, a text-based CA with which a mentally ill user feels safe to converse at all times, or a robot that nudges its elderly user to take their medication whilst providing companionship. I hope that my work will support and inspire CA designers and researchers to imbue their CAs with personalities that best match their users' preferences.

References

[1] Adomavicius, Gediminas, Huang, Zan and Tuzhilin, Alexander. 'Personalization and recommender systems'. In: *State-of-the-art decision making tools in the information-intensive age* (2008), pp. 55–100. DOI: https://doi.org/10.1287/educ.1080.0044.

[2] Ait Challal, Tagduda and Grynszpan, Ouriel. 'What Gaze Tells Us About Personality'. In: *Proceedings of the 6th International Conference on Human-Agent Interaction*. HAI '18. New York, NY, USA: Association for Computing Machinery, 2018, pp. 129–137. DOI: 10.1145/3284432.3284455.

[3] Allen, James F, Miller, Bradford W, Ringger, Eric K and Sikorski, Teresa. 'A Robust System for Natural Spoken Dialogue'. In: *Proceedings of the 34th Annual Meeting on Association for Computational Linguistics*. ACL '96. USA: Association for Computational Linguistics, 1996, pp. 62–70. DOI: 10.3115/981863.981872.

[4] Allport, Gordon W and Odbert, Henry S. 'Trait-names: A psycho-lexical study.' In: *Psychological Monographs* 47.1 (1936), pp. i–171. DOI: 10.1037/h0093360.

[5] Aly, Amir and Tapus, Adriana. 'A model for synthesizing a combined verbal and nonverbal behavior based on personality traits in human-robot interaction'. In: *8th ACM/IEEE International Conference on Human-Robot Interaction (HRI '13)*. New York, NY, USA: ACM/IEEE, 2013, pp. 325–332. DOI: 10.1109/HRI.2013.6483606.

[6] Aly, Amir and Tapus, Adriana. 'Towards an intelligent system for generating an adapted verbal and nonverbal combined behavior in human-robot interaction'. In: *Autonomous Robots* 40.2 (2016), pp. 193–209. DOI: 10.1007/s10514-015-9444-1.

[7] Ammari, Tawfiq, Kaye, Jofish, Tsai, Janice Y and Bentley, Frank. 'Music, Search, and IoT: How People (Really) Use Voice Assistants'. In: *ACM Trans. Comput.-Hum. Interact.* 26.3 (2019). DOI: 10.1145/3311956.

[8] André, Elisabeth, Klesen, Martin, Gebhard, Patrick, Allen, Steve and Rist, Thomas. 'Integrating Models of Personality and Emotions into Lifelike Characters'. In: *Affective Interactions: Towards a New Generation of Computer Interfaces*. Ed. by Ana Paiva. Berlin, Heidelberg, Germany: Springer, 2000, pp. 150–165. DOI: 10.1007/10720296_11.

[9] Andrist, Sean, Mutlu, Bilge and Tapus, Adriana. 'Look Like Me: Matching Robot Personality via Gaze to Increase Motivation'. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. CHI '15. New York, NY, USA: Association for Computing Machinery, 2015, pp. 3603–3612. DOI: 10.1145/2702123.2702592.

[10] Argyle, Michael. Bodily Communication. London, UK: Methuen and Co Ltd, 1988.

[11] Austin, John Langshaw. How to do things with words. London, UK: Oxford University Press, 1975.

[12] Azucar, Danny, Marengo, Davide and Settanni, Michele. ‘Predicting the Big 5 personality traits from digital footprints on social media: A meta-analysis’. In: *Personality and Individual Differences* 124 (2018), pp. 150–159. DOI: 10.1016/J.PAID.2017.12.018.

[13] Barbieri, Francesco, Ballesteros, Miguel and Saggion, Horacio. ‘Are Emojis Predictable?’ In: *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*. Valencia, Spain: Association for Computational Linguistics, 2017, pp. 105–111. DOI: 10.48550/arXiv.1702.07285.

[14] Barbieri, Francesco, Ronzano, Francesco and Saggion, Horacio. ‘What does this emoji mean? A vector space skip-gram model for Twitter emojis’. In: *Proceedings of Language Resources and Evaluation Conference*. Paris, France: ELRA (European Language Resources Association), 2016.

[15] Bartneck, Christoph, Van Der Hoek, Michel, Mubin, Omar and Al Mahmud, Abdullah. ‘“Daisy, daisy, give me your answer do!” – Switching off a robot’. In: *2007 2nd ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. New York, NY, USA: ACM/IEEE, 2007, pp. 217–222. DOI: 10.1145/1228716.1228746.

[16] Bates, Joseph. ‘The Role of Emotion in Believable Agents’. In: *Commun. ACM* 37.7 (1994), pp. 122–125. DOI: 10.1145/176789.176803.

[17] Beaudouin-Lafon, Michel. ‘Designing Interaction, Not Interfaces’. In: *Proceedings of the Working Conference on Advanced Visual Interfaces*. AVI ’04. New York, NY, USA: Association for Computing Machinery, 2004, pp. 15–22. DOI: 10.1145/989863.989865.

[18] Belsky, Jay, Jaffee, Sara R, Caspi, Avshalom, Moffitt, Terrie and Silva, Phil A. ‘Intergenerational relationships in young adulthood and their life course, mental health, and personality correlates.’ In: *Journal of Family Psychology* 17.4 (2003), pp. 460–471. DOI: 10.1037/0893-3200.17.4.460.

[19] Bemelmans, Roger, Gelderblom, Gert Jan, Jonker, Pieter and de Witte, Luc. ‘Socially Assistive Robots in Elderly Care: A Systematic Review into Effects and Effectiveness’. In: *Journal of the American Medical Directors Association* 13.2 (2012), 114–120.e1. DOI: 10.1016/j.jamda.2010.10.002.

[20] Bickmore, Timothy and Cassell, Justine. ‘Relational Agents: A Model and Implementation of Building User Trust’. In: *Proceedings of the SIGCHI Conference on Human*

Factors in Computing Systems. CHI '01. New York, NY, USA: Association for Computing Machinery, 2001, pp. 396–403. DOI: 10.1145/365024.365304.

[21] Bickmore, Timothy and Cassell, Justine. 'Social Dialogue with Embodied Conversational Agents'. In: *Advances in Natural Multimodal Dialogue Systems*. Ed. by Jan CJ van Kuppevelt, Laila Dybkjær and Niels Ole Bernsen. Dordrecht, Netherlands: Springer Netherlands, 2005, pp. 23–54. DOI: 10.1007/1-4020-3933-6_2.

[22] Bickmore, Timothy W, Mitchell, Suzanne E, Jack, Brian W, Paasche-Orlow, Michael K, Pfeifer, Laura M and O'Donnell, Julie. 'Response to a relational agent by hospital patients with depressive symptoms'. In: *Interacting with Computers* 22.4 (2010), pp. 289–298. DOI: 10.1016/j.intcom.2009.12.001.

[23] Blackwell, David, Leaman, Carrie, Tramposch, Rose, Osborne, Ciera and Liss, Miriam. 'Extraversion, neuroticism, attachment style and fear of missing out as predictors of social media use and addiction'. In: *Personality and Individual Differences* 116 (2017), pp. 69–72. DOI: 10.1016/j.paid.2017.04.039.

[24] Borkenau, Peter and Liebler, Anette. 'Trait inferences: Sources of validity at zero acquaintance.' In: *Journal of Personality and Social Psychology* 62.4 (1992), pp. 645–657. DOI: 10.1037/0022-3514.62.4.645.

[25] Boyd, Ryan L and Pennebaker, James W. 'Language-based personality: a new approach to personality in a digital world'. In: *Current Opinion in Behavioral Sciences* 18 (2017), pp. 63–68. DOI: 10.1016/j.cobeha.2017.07.017.

[26] Boyle, Gregory J and Helmes, Edward. 'Methods of personality assessment.' In: *The Cambridge handbook of personality psychology*. Ed. by Philip J Corr and Gerald Matthews. Cambridge, UK: Cambridge University Press, 2009, pp. 110–126. DOI: 10.1017/CBO9780511596544.010.

[27] Braun, Michael, Mainz, Anja, Chadowitz, Ronee, Pfleging, Bastian and Alt, Florian. 'At Your Service: Designing Voice Assistant Personalities to Improve Automotive User Interfaces'. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. CHI '19. New York, NY, USA: Association for Computing Machinery, 2019. DOI: 10.1145/3290605.3300270.

[28] Breazeal, Cynthia. Designing Sociable Robots. Cambridge, MA, USA: MIT Press, 2002.

[29] Bremner, Paul, Celiktutan, Oya and Gunes, Hatice. 'Personality perception of robot avatar tele-operators'. In: *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. New York, NY, USA: ACM/IEEE, 2016, pp. 141–148. DOI: 10.1109/HRI.2016.7451745.

[30] Brscić, Drazen, Kidokoro, Hiroyuki, Suehiro, Yoshitaka and Kanda, Takayuki. 'Escaping from children's abuse of social robots'. In: *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*. HRI '15. New York, NY, USA: Association for Computing Machinery, 2015, pp. 59–66. DOI: 10.1145/2696454.2696468.

[31] Bühner, Markus. Einführung in die Test- und Fragebogenkonstruktion. Munich, Germany: Pearson Education, 2011.

[32] Buschek, Daniel, Hassib, Mariam and Alt, Florian. 'Personal Mobile Messaging in Context: Chat Augmentations for Expressiveness and Awareness'. In: *ACM Trans. Comput.-Hum. Interact.* 25.4 (2018), 23:1–23:33. DOI: 10.1145/3201404.

[33] Byrne, Donn. 'Interpersonal attraction and attitude similarity.' In: *The Journal of Abnormal and Social Psychology* 62.3 (1961), pp. 713–715. DOI: 10.1037/h0044721.

[34] Byrne, Donn Erwin. The attraction paradigm. Cambridge, MA, USA: Academic Press, 1971.

[35] Cafaro, Angelo, Glas, Nadine and Pelachaud, Catherine. 'The Effects of Interrupting Behavior on Interpersonal Attitude and Engagement in Dyadic Interactions'. In: *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems*. AAMAS '16. Richland, SC, USA: International Foundation for Autonomous Agents and Multiagent Systems, 2016, pp. 911–920.

[36] Cafaro, Angelo, Vilhjálmsson, Hannes Högni and Bickmore, Timothy. 'First Impressions in Human–Agent Virtual Encounters'. In: *ACM Trans. Comput.-Hum. Interact.* 23.4 (2016), 24:1–24:40. DOI: 10.1145/2940325.

[37] Cafaro, Angelo, Vilhjálmsson, Hannes Högni, Bickmore, Timothy, Heylen, Dirk, Jóhannsdóttir, Kamilla Rún and Valgarðsson, Gunnar Steinn. 'First Impressions: Users' Judgments of Virtual Agents' Personality and Interpersonal Attitude in First Encounters'. In: *Intelligent Virtual Agents*. Ed. by Yukiko Nakano, Michael Neff, Ana Paiva and Marilyn Walker. Berlin, Heidelberg, Germany: Springer, 2012, pp. 67–80. DOI: 10.1007/978-3-642-33197-8_7.

[38] Cairns, Paul. 'HCI... not as it should be: Inferential Statistics in HCI Research'. In: *Proceedings of HCI 2007 – The 21st British HCI Group Annual Conference University of Lancaster, UK 21*. Lancaster, UK, 2007, pp. 1–7. DOI: 10.14236/ewic/HCI2007.20.

[39] Campbell, Anne and Rushton, J Philippe. 'Bodily communication and personality'. In: *British Journal of Social and Clinical Psychology* 17.1 (1978), pp. 31–36. DOI: 10.1111/j.2044-8260.1978.tb00893.x.

[40] Carment, DW, Miles, CG and Cervin, VB. 'Persuasiveness and persuasibility as related to intelligence and extraversion'. In: *British Journal of Social and Clinical Psychology* 4.1 (1965), pp. 1–7. DOI: 10.1111/j.2044-8260.1965.tb00433.x.

[41] Caspi, Avshalom, Roberts, Brent W and Shiner, Rebecca L. 'Personality Development: Stability and Change'. In: *Annual Review of Psychology* 56.1 (2005), pp. 453–484. DOI: 10.1146/annurev.psych.55.090902.141913.

[42] Cassell, Justine. 'Embodied Conversational Agents: Representation and Intelligence in User Interfaces'. In: *AI Magazine* 22.4 (2001), pp. 67–83. DOI: 10.1609/aimag.v22i4.1593.

[43] Cassell, Justine, Sullivan, Joseph, Prevost, Scott and Churchill, Elizabeth. Embodied Conversational Agents. Cambridge, MA, USA: MIT Press, 2000.

[44] Cassell, Justine, Vilhjálmsson, Hannes Högni and Bickmore, Timothy. 'BEAT: the Behavior Expression Animation Toolkit'. In: *Life-Like Characters: Tools, Affective Functions, and Applications*. Ed. by Helmut Prendinger and Mitsuru Ishizuka. Berlin, Heidelberg, Germany: Springer, 2004, pp. 163–185. DOI: 10.1007/978-3-662-08373-4_8.

[45] Castillo, Susana, Hahn, Philipp, Legde, Katharina and Cunningham, Douglas W. 'Personality Analysis of Embodied Conversational Agents'. In: *Proceedings of the 18th International Conference on Intelligent Virtual Agents*. IVA '18. New York, NY, USA: Association for Computing Machinery, 2018, pp. 227–232. DOI: 10.1145/3267851.3267853.

[46] Celli, Fabio, Bruni, Elia and Lepri, Bruno. 'Automatic Personality and Interaction Style Recognition from Facebook Profile Pictures'. In: *Proceedings of the 22nd ACM International Conference on Multimedia*. MM '14. New York, NY, USA: Association for Computing Machinery, 2014, pp. 1101–1104. DOI: 10.1145/2647868.2654977.

[47] Charlton, John P. 'The determinants and expression of computer-related anger'. In: *Computers in Human Behavior* 25.6 (2009), pp. 1213–1221. DOI: 10.1016/j.chb.2009.07.001.

[48] Chen, Sylvia Xiaohua, Benet-Martínez, Verónica and Ng, Jacky CK. 'Does Language Affect Personality Perception? A Functional Approach to Testing the Whorfian Hypothesis'. In: *Journal of Personality* 82.2 (2014), pp. 130–143. DOI: 10.1111/jopy.12040.

[49] Chin, Hyojin and Yi, Mun Yong. 'Should an Agent Be Ignoring It?: A Study of Verbal Abuse Types and Conversational Agents' Response Styles'. In: *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. CHI EA '19. New York, NY, USA: Association for Computing Machinery, 2019, LBW2422:1–LBW2422:6. DOI: 10.1145/3290607.3312826.

[50] Clark, Leigh, Doyle, Philip, Garaialde, Diego, Gilmartin, Emer, Schlögl, Stephan, Edlund, Jens, Aylett, Matthew, Cabral, João, Munteanu, Cosmin, Edwards, Justin and Cowan, Benjamin R. 'The State of Speech in HCI: Trends, Themes and Challenges'. In: *Interacting with Computers* 31.4 (2019), pp. 349–371. DOI: 10.1093/iwc/iwz016.

[51] Clark, Leigh, Pantidi, Nadia, Cooney, Orla, Doyle, Philip, Garaialde, Diego, Edwards, Justin, Spillane, Brendan, Gilmartin, Emer, Murad, Christine, Munteanu, Cosmin, Wade, Vincent and Cowan, Benjamin R. 'What Makes a Good Conversation?: Challenges in Designing Truly Conversational Agents'. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. CHI '19. New York, NY, USA: Association for Computing Machinery, 2019, 475:1–475:12. DOI: 10.1145/3290605.3300705.

[52] Cohen, Michael Harris, Giangola, James P and Balogh, Jennifer. Voice user interface design. Boston, MA, USA: Addison-Wesley Professional, 2004.

[53] Connelly, Brian S and Ones, Deniz S. 'An other perspective on personality: meta-analytic integration of observers' accuracy and predictive validity.' In: *Psychological Bulletin* 136.6 (2010), pp. 1092–1122. DOI: 10.1037/a0021212.

[54] Coren, Michael J. Virtual assistants spend much of their time fending off sexual harassment. Quartz. Retrieved online from https://qz.com/818151/virtual-assistant-bots-like-siri-alexa-and-cortana-spend-much-of-their-time-fending-off-sexual-harassment/, last accessed 10th May 2022. 2016.

[55] Costa, Paul T and McCrae, Robert R. 'Four ways five factors are basic'. In: *Personality and Individual Differences* 13.6 (1992), pp. 653–665. DOI: 10.1016/0191-8869(92)90236-I.

[56] Costa, Paul T and McCrae, Robert R. 'Influence of extraversion and neuroticism on subjective well-being: happy and unhappy people.' In: *Journal of Personality and Social Psychology* 38.4 (1980), pp. 668–678. DOI: 10.1037/0022-3514.38.4.668.

[57] Costa Jr, Paul T and McCrae, Robert R. 'The Revised NEO Personality Inventory (NEO-PI-R)'. In: *The SAGE handbook of personality theory and assessment, Vol. 2. Personality measurement and testing (pp. 179–198)*. Ed. by GJ Boyle, G Matthews Matthews and DH Saklofske. Sage Publications, Inc, 2008, pp. 179–198. DOI: 10.4135/9781849200479.n9.

[58] Cowan, Benjamin R, Doyle, Philip, Edwards, Justin, Garaialde, Diego, Hayes-Brady, Ali, Branigan, Holly P, Cabral, João and Clark, Leigh. 'What's in an Accent? The Impact of Accented Synthetic Speech on Lexical Choice in Human-Machine Dialogue'. In: *Proceedings of the 1st International Conference on Conversational User Interfaces*. CUI

’19. New York, NY, USA: Association for Computing Machinery, 2019, pp. 1–8. DOI: 10.1145/3342775.3342786.

[59] Cowan, Benjamin R, Pantidi, Nadia, Coyle, David, Morrissey, Kellie, Clarke, Peter, Al-Shehri, Sara, Earley, David and Bandeira, Natasha. ‘“What Can I Help You with?”: Infrequent Users’ Experiences of Intelligent Personal Assistants’. In: *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services*. MobileHCI ’17. New York, NY, USA: Association for Computing Machinery, 2017, pp. 1–12. DOI: 10.1145/3098279.3098539.

[60] Dale, Robert. ‘The return of the chatbots’. In: *Natural Language Engineering* 22.5 (2016), pp. 811–817. DOI: 10.1017/S1351324916000243.

[61] Danner, Daniel, Rammstedt, Beatrice, Bluemke, Matthias, Treiber, Lisa, Berres, Sabrina, Soto, Christopher and John, Oliver. Die deutsche Version des Big Five Inventory 2 (BFI-2). Mannheim, Germany: GESIS, 2016. DOI: 10.6102/zis247.

[62] Deary, Ian J. ‘The trait approach to personality’. In: *The Cambridge handbook of personality psychology*. Ed. by Philip J Corr and Gerald Matthews. Vol. 1. Cambridge, UK: Cambridge University Press, 2009, pp. 89–109. DOI: 10.1017/CBO9780511596544.009.

[63] Dewaele, Jean-Marc and Furnham, Adrian. ‘Personality and speech production: A pilot study of second language learners’. In: *Personality and Individual Differences* 28.2 (2000), pp. 355–365. DOI: 10.1016/S0191-8869(99)00106-3.

[64] DeYoung, Colin G. ‘Higher-order factors of the Big Five in a multi-informant sample.’ In: *Journal of Personality and Social Psychology* 91.6 (2006), pp. 1138–1151. DOI: 10.1037/0022-3514.91.6.1138.

[65] Diener, Ed and Lucas, Richard E. ‘Personality and Subjective Well-being’. In: *Well-being: Foundations of Hedonic Psychology*. Ed. by Daniel Kahneman, Ed Diener and Norbert Schwarz. New York, NY, USA: Russell Sage Foundations, 1999, pp. 213–229.

[66] Donnellan, M Brent, Conger, Rand D and Bryant, Chalandra M. ‘The Big Five and enduring marriages’. In: *Journal of Research in Personality* 38.5 (2004), pp. 481–504. DOI: 10.1016/j.jrp.2004.01.001.

[67] Doyle, Philip R, Clark, Leigh and Cowan, Benjamin R. ‘What Do We See in Them? Identifying Dimensions of Partner Models for Speech Interfaces Using a Psycholexical Approach’. In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. CHI ’21. New York, NY, USA: Association for Computing Machinery, 2021, pp. 1–14. DOI: 10.1145/3411764.3445206.

[68] Doyle, Philip R, Edwards, Justin, Dumbleton, Odile, Clark, Leigh and Cowan, Benjamin R. ‘Mapping Perceptions of Humanness in Intelligent Personal Assistant Inter-

action'. In: *Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services*. MobileHCI '19. New York, NY, USA: Association for Computing Machinery, 2019, pp. 1–12. DOI: 10.1145/3338286.3340116.

[69] Dubberly, Hugh and Pangaro, Paul. 'What is conversation? How can we design for effective conversation'. In: *Interactions Magazine* 16.4 (2009), pp. 22–28. DOI: 10.1145/1409040.1409054.

[70] DuBois, Philip H. 'A test-dominated society: China, 1115 B.C. – 1905 A.D'. In: *Testing Problems in Perspective* (1966), pp. 29–36.

[71] Edwards, Autumn and Edwards, Chad. 'Does the Correspondence Bias Apply to Social Robots?: Dispositional and Situational Attributions of Human Versus Robot Behavior'. In: *Frontiers in Robotics and AI* 8 (2021), p. 788242. DOI: 10.3389/frobt.2021.788242.

[72] Endrass, Birgit, André, Elisabeth, Rehm, Matthias and Nakano, Yukiko. 'Investigating culture-related aspects of behavior for virtual characters'. In: *Autonomous Agents and Multi-Agent Systems* 27.2 (2013), pp. 277–304. DOI: 10.1007/s10458-012-9218-5.

[73] Evans, Anthony M and Revelle, William. 'Survey and behavioral measurements of interpersonal trust'. In: *Journal of Research in Personality* 42.6 (2008), pp. 1585–1593. DOI: 10.1016/j.jrp.2008.07.011.

[74] Eysenck, Hans Jürgen. 'Personality: Biological foundations.' In: *The Neuropsychology of Individual Differences*. Ed. by PA Vernon. San Diego, CA, USA: Academic Press, 1994, pp. 151–207. DOI: 10.1016/B978-0-12-718670-2.50011-6.

[75] Feine, Jasper, Gnewuch, Ulrich, Morana, Stefan and Maedche, Alexander. 'A Taxonomy of Social Cues for Conversational Agents'. In: *International Journal of Human-Computer Studies* 132 (2019), pp. 138–161. DOI: 10.1016/j.ijhcs.2019.07.009.

[76] Fernández-Tobías, Ignacio, Braunhofer, Matthias, Elahi, Mehdi, Ricci, Francesco and Cantador, Iván. 'Alleviating the new user problem in collaborative filtering by exploiting personality information'. In: *User Modeling and User-Adapted Interaction*. UMAP '16. New York, NY, USA: Association for Computing Machinery, 2016, pp. 221–255. DOI: 10.1007/s11257-016-9172-z.

[77] Ferwerda, Bruce, Tkalcic, Marko and Schedl, Markus. 'Personality Traits and Music Genres: What Do People Prefer to Listen To?' In: *Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization*. UMAP '17. New York, NY, USA: Association for Computing Machinery, 2017, pp. 285–288. DOI: 10.1145/3079628.3079693.

[78] Følstad, Asbjørn and Brandtzæg, Petter Bae. 'Chatbots and the New World of HCI'. In: *Interactions* 24.4 (2017), pp. 38–42. DOI: 10.1145/3085558.

[79] Francis, Leslie J, Brown, Laurence B and Philipchalk, Ronald. 'The development of an abbreviated form of the revised Eysenck personality questionnaire (EPQR-A): Its use among students in England, Canada, the U.S.A. and Australia'. In: *Personality and Individual Differences* 13.4 (1992), pp. 443–449. DOI: 10.1016/0191-8869(92)90073-X.

[80] Fryer, Luke K, Nakao, Kaori and Thompson, Andrew. 'Chatbot learning partners: Connecting learning experiences, interest and competence'. In: *Computers in Human Behavior* 93 (2019), pp. 279–289. DOI: 10.1016/j.chb.2018.12.023.

[81] Furnham, Adrian. 'Language and Personality'. In: *Handbook of language and social psychology*. Ed. by William Peter Robinson and Howard Giles. Chichester, UK: John Wiley & Sons, 1990, pp. 73–95.

[82] Gill, Alastair J, Nowson, Scott and Oberlander, Jon. 'What are they blogging about? Personality, topic and motivation in blogs'. In: *Third International AAAI Conference on Weblogs and Social Media*. Palo Alto, CA, USA: AAAI, 2009, pp. 18–25.

[83] Gilmartin, Emer, Spillane, Brendan, O'Reilly, Maria, Su, Ketong, Saam, Christian, Cowan, Benjamin R, Campbell, Nick and Wade, Vincent. 'Dialog Acts in Greeting and Leavetaking in Social Talk'. In: *Proceedings of the 1st ACM SIGCHI International Workshop on Investigating Social Interactions with Artificial Agents*. ISIAA 2017. New York, NY, USA: Association for Computing Machinery, 2017, pp. 29–30. DOI: 10.1145/3139491.3139493.

[84] Gnewuch, Ulrich, Morana, Stefan and Maedche, Alexander. 'Towards Designing Cooperative and Social Conversational Agents for Customer Service.' In: *Proceedings of the 38th International Conference on Information Systems (ICIS), Seoul, ROK, December 10-13, 2017. Research-in-Progress Papers.* AIS eLibrary (AISeL), 2017.

[85] Gnewuch, Ulrich, Yu, Meng and Maedche, Alexander. 'The Effect of Perceived Similarity in Dominance on Customer Self-Disclosure to Chatbots in Conversational Commerce'. In: *Proceedings of the 28th European Conference on Information Systems (ECIS)*. AIS, 2020.

[86] Goldberg, Lewis R. 'Language and individual differences: The search for universals in personality lexicons'. In: *Review of Personality and Social Psychology*. Ed. by L Wheeler. Vol. 2. Beverly Hills, CA, USA: Sage Publications, 1981, pp. 141–166.

[87] Gosling, Samuel D, Rentfrow, Peter J and Swann, William B. 'A very brief measure of the Big-Five personality domains'. In: *Journal of Research in Personality* 37.6 (2003), pp. 504–528. DOI: 10.1016/S0092-6566(03)00046-1.

[88] Green, Lisa J. African American English: a linguistic introduction. Cambridge, UK: Cambridge University Press, 2002.

[89] Han, Xu, Zhou, Michelle, Turner, Matthew J and Yeh, Tom. 'Designing Effective Interview Chatbots: Automatic Chatbot Profiling and Design Suggestion Generation for Chatbot Debugging'. In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. CHI '21. New York, NY, USA: Association for Computing Machinery, 2021. DOI: 10.1145/3411764.3445569.

[90] Hegstrom, Timothy G. 'Message impact: What percentage is nonverbal?' In: *Western Journal of Speech Communication* 43.2 (1979), pp. 134–142. DOI: 10.1080/10570317909373961.

[91] Heider, F. The Psychology of Interpersonal Relations. New York City, NY, USA: Psychology Press, 1982. DOI: 10.4324/9780203781159.

[92] Heinström, Jannica. 'Fast surfing, broad scanning and deep diving: The influence of personality and study approach on students' information-seeking behavior'. In: *Journal of Documentation* 61.2 (2005), pp. 228–247. DOI: 10.1108/00220410510585205.

[93] Hemphill, Charles T, Godfrey, John J and Doddington, George R. 'The ATIS Spoken Language Systems Pilot Corpus'. In: *Proceedings of the Workshop on Speech and Natural Language*. HLT '90. USA: Association for Computational Linguistics, 1990, pp. 96–101. DOI: 10.3115/116580.116613.

[94] Hirsh, Jacob B, DeYoung, Colin G and Peterson, Jordan B. 'Metatraits of the Big Five Differentially Predict Engagement and Restraint of Behavior'. In: *Journal of Personality* 77.4 (2009), pp. 1085–1102. DOI: 10.1111/j.1467-6494.2009.00575.x.

[95] Hofstede, Geert. 'Cultures and organizations: Intercultural cooperation and its importance for survival'. In: *Software of the Mind*. McGraw-Hill: London, 1991.

[96] Holtgraves, Thomas. 'Text messaging, personality, and the social context'. In: *Journal of Research in Personality* 45.1 (2011), pp. 92–99. DOI: 10.1016/j.jrp.2010.11.015.

[97] Holtgraves, Thomas M. Language as social action: Social psychology and language use. Taylor & Francis, 2013.

[98] Horstmann, Aike C, Bock, Nikolai, Linhuber, Eva, Szczuka, Jessica M, Straßmann, Carolin and Krämer, Nicole C. 'Do a robot's social skills and its objection discourage interactants from switching the robot off?' In: *PloS one* 13.7 (2018), e0201581. DOI: 10.1371/journal.pone.0201581.

[99] Hoy, Matthew B. 'Alexa, Siri, Cortana, and More: An Introduction to Voice Assistants'. In: *Medical Reference Services Quarterly* 37.1 (2018), pp. 81–88. DOI: 10.1080/02763869.2018.1404391.

[100] Isbister, Katherine and Nass, Clifford. 'Consistency of personality in interactive characters: verbal cues, non-verbal cues, and user characteristics'. In: *International Journal of Human-Computer Studies* 53.2 (2000), pp. 251–267. DOI: 10.1006/ijhc.2000.0368.

[101] Jain, Mohit, Kumar, Pratyush, Kota, Ramachandra and Patel, Shwetak N. 'Evaluating and Informing the Design of Chatbots'. In: *Proceedings of the 2018 Designing Interactive Systems Conference*. DIS '18. New York, NY, USA: Association for Computing Machinery, 2018, pp. 895–906. DOI: 10.1145/3196709.3196735.

[102] Janowski, Kathrin and André, Elisabeth. 'What If I Speak Now? A Decision-Theoretic Approach to Personality-Based Turn-Taking'. In: *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems*. AAMAS '19. Richland, SC, USA: International Foundation for Autonomous Agents and Multiagent Systems, 2019, pp. 1051–1059.

[103] Janowski, Kathrin, Ritschel, Hannes and André, Elisabeth. 'Adaptive Artificial Personalities'. In: *The Handbook on Socially Interactive Agents: 20 Years of Research on Embodied Conversational Agents, Intelligent Virtual Agents, and Social Robotics Volume 2: Interactivity, Platforms, Application*. 1st ed. New York, NY, USA: Association for Computing Machinery, 2022, pp. 155–194. DOI: 10.1145/3563659.3563666.

[104] Jensen-Campbell, Lauri A, Adams, Ryan, Perry, David G, Workman, Katie A, Furdella, Janine Q and Egan, Susan K. 'Agreeableness, extraversion, and peer relations in early adolescence: Winning friends and deflecting aggression'. In: *Journal of Research in Personality* 36.3 (2002), pp. 224–251. DOI: 10.1006/jrpe.2002.2348.

[105] John, Oliver P, Angleitner, Alois and Ostendorf, Fritz. 'The lexical approach to personality: A historical review of trait taxonomic research'. In: *European Journal of Personality* 2.3 (1988), pp. 171–203. DOI: 10.1002/per.2410020302.

[106] John, Oliver P, Donahue, Eileen M and Kentle, Robert L. The big five inventory – versions 4a and 54. Berkeley, CA, USA: University of California, Berkeley, Institute of Personality and Social Research, 1991.

[107] Jost, John T, Glaser, Jack, Kruglanski, Arie W and Sulloway, Frank J. 'Political conservatism as motivated social cognition.' In: *Psychological Bulletin* 129.3 (2003), pp. 339–375. DOI: 10.1037/0033-2909.129.3.339.

[108] Judge, Timothy A, Higgins, Chad A, Thoresen, Carl J and Barrick, Murray R. 'The big five personality traits, general mental ability, and career success across the life

span'. In: *Personnel Psychology* 52.3 (1999), pp. 621–652. DOI: 10.1111/j.1744-6570.1999.tb00174.x.

[109] Jurafsky, Dan and Martin, James H. Speech & language processing. Upper Saddle River, NJ, USA: Pearson Prentice Hall, 2000.

[110] Kasilingam, Dharun Lingam. 'Understanding the attitude and intention to use smartphone chatbots for shopping'. In: *Technology in Society* 62 (2020), p. 101280. DOI: 10.1016/j.techsoc.2020.101280.

[111] Kay, Matthew, Haroz, Steve, Guha, Shion, Dragicevic, Pierre and Wacharamanotham, Chat. 'Moving Transparent Statistics Forward at CHI'. In: *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. CHI EA '17. New York, NY, USA: Association for Computing Machinery, 2017, pp. 534–541. DOI: 10.1145/3027063.3027084.

[112] Kim, Hankyung, Koh, Dong Yoon, Lee, Gaeun, Park, Jung-Mi and Lim, Youn-kyung. 'Designing Personalities of Conversational Agents'. In: *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. CHI EA '19. New York, NY, USA: Association for Computing Machinery, 2019, pp. 1–6. DOI: 10.1145/3290607.3312887.

[113] Kinsella, Bret and Mutchler, Ava. In-car Voice Assistant Consumer Adoption Report. http://voicebot.ai/wp-content/uploads/2020/02/in_car_voice_assistant_consumer_adoption_report_2020_voicebot.pdf, accessed July 30, 2020. 2020.

[114] Kocabalil, Ahmet Baki, Laranjo, Liliana and Coiera, Enrico. 'Measuring user experience in conversational interfaces: a comparison of six questionnaires'. In: *Proceedings of the 32nd International BCS Human Computer Interaction Conference 32*. Belfast, UK: BCS Learning and Development Ltd., 2018, pp. 1–12. DOI: 10.14236/ewic/HCI2018.21.

[115] Komarraju, Meera, Karau, Steven J and Schmeck, Ronald R. 'Role of the Big Five personality traits in predicting college students' academic motivation and achievement'. In: *Learning and Individual Differences* 19.1 (2009), pp. 47–52. DOI: 10.1016/j.lindif.2008.07.001.

[116] Kosinski, Michal, Stillwell, David and Graepel, Thore. 'Private traits and attributes are predictable from digital records of human behavior'. In: *Proceedings of the National Academy of Sciences* 110.15 (2013), pp. 5802–5805. DOI: 10.1073/pnas.1218772110.

[117] Kralj Novak, Petra, Smailovia, Jasmina, Sluban, Borut and Mozetia, Igor. 'Sentiment of Emojis'. In: *PLOS ONE* 10.12 (2015), pp. 1–22. DOI: 10.1371/journal.pone.0144296.

[118] Krenn, Brigitte, Endrass, Birgit, Kistler, Felix and André, Elisabeth. 'Effects of Language Variety on Personality Perception in Embodied Conversational Agents'. In: *Human-Computer Interaction. Advanced Interaction Modalities and Techniques. HCI 2014. Lecture Notes in Computer Science*. Ed. by Masaaki Kurosu. Vol. 8511. Cham, Switzerland: Springer International Publishing, 2014, pp. 429–439. DOI: 10.1007/978-3-319-07230-2_41.

[119] Kuss, Daria J and Griffiths, Mark D. 'Online Social Networking and Addiction – A Review of the Psychological Literature'. In: *International Journal of Environmental Research and Public Health* 8.9 (2011), pp. 3528–3552. DOI: 10.3390/ijerph8093528.

[120] Lamiell, James T. 'The characterization of persons: Some fundamental conceptual issues.' In: *The Cambridge handbook of personality psychology*. Ed. by Philip J Corr and Gerald Matthews. Cambridge, UK: Cambridge University Press, 2009, pp. 72–86. DOI: 10.1017/CBO9780511596544.008.

[121] Langevin, Raina, Lordon, Ross J, Avrahami, Thi, Cowan, Benjamin R, Hirsch, Tad and Hsieh, Gary. 'Heuristic Evaluation of Conversational Agents'. In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. CHI '21. New York, NY, USA: Association for Computing Machinery, 2021. DOI: 10.1145/3411764.3445312.

[122] Laranjo, Liliana, Dunn, Adam G, Tong, Huong Ly, Kocaballi, Ahmet Baki, Chen, Jessica, Bashir, Rabia, Surian, Didi, Gallego, Blanca, Magrabi, Farah, Lau, Annie Y S and Coiera, Enrico. 'Conversational agents in healthcare: a systematic review'. In: *Journal of the American Medical Informatics Association* 25.9 (2018), pp. 1248–1258. DOI: 10.1093/jamia/ocy072.

[123] Laurel, Brenda. 'Interface agents: Metaphors with character'. In: *Human Values and the design of Computer Technology*. Ed. by Batya Friedman. Stanford, CA, USA: Cambridge University Press, 1997, pp. 207–219.

[124] Lee, Kwan Min and Nass, Clifford. 'Designing Social Presence of Social Actors in Human Computer Interaction'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '03. New York, NY, USA: Association for Computing Machinery, 2003, pp. 289–296. DOI: 10.1145/642611.642662.

[125] Lee, Kwan Min, Peng, Wei, Jin, Seung-A and Yan, Chang. 'Can Robots Manifest Personality?: An Empirical Test of Personality Recognition, Social Responses, and Social Presence in Human–Robot Interaction'. In: *Journal of Communication* 56.4 (2006), pp. 754–772. DOI: 10.1111/j.1460-2466.2006.00318.x.

[126] Lee, Min Kyung and Makatchev, Maxim. 'How Do People Talk with a Robot? An Analysis of Human-Robot Dialogues in the Real World'. In: *CHI '09 Extended Abstracts on*

Human Factors in Computing Systems. CHI EA '09. New York, NY, USA: Association for Computing Machinery, 2009, pp. 3769–3774. DOI: 10.1145/1520340.1520569.

[127] Lee, Minha, Ackermans, Sander, As, Nena van, Chang, Hanwen, Lucas, Enzo and IJsselsteijn, Wijnand. 'Caring for Vincent: A Chatbot for Self-Compassion'. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery, 2019, pp. 1–13. DOI: 10.1145/3290605.3300932.

[128] Leising, Daniel, Thielmann, Isabel, Glöckner, Andreas, Gärtner, Anne and Schönbrodt, Felix D. 'Ten steps toward a better personality science - how quality may be rewarded more in research evaluation'. In: *Personality Science* 3 (2022). DOI: 10.5964/ps.6029.

[129] Lengel, Robert H and Daft, Richard L. An exploratory analysis of the relationship between media richness and managerial information processing. Tech. rep. Texas A and M University, 1984.

[130] Li, Jingyi, Zhou, Michelle X, Yang, Huahai and Mark, Gloria. 'Confiding in and Listening to Virtual Agents: The Effect of Personality'. In: *Proceedings of the 22nd International Conference on Intelligent User Interfaces*. IUI '17. New York, NY, USA: Association for Computing Machinery, 2017, pp. 275–286. DOI: 10.1145/3025171.3025206.

[131] Li, Weijian, Chen, Yuxiao, Hu, Tianran and Luo, Jiebo. 'Mining the Relationship between Emoji Usage Patterns and Personality'. In: *International AAAI Conference on Web and Social Media*. Palo Alto, CA, USA: AAAI Publications, 2018.

[132] Liew, Tze Wei and Tan, Su-Mae. 'Virtual agents with personality: Adaptation of learner-agent personality in a virtual learning environment'. In: *2016 Eleventh International Conference on Digital Information Management (ICDIM)*. 2016, pp. 157–162. DOI: 10.1109/ICDIM.2016.7829758.

[133] Liu, Kris, Tolins, Jackson, Fox Tree, Jean E, Neff, Michael and Walker, Marilyn A. 'Two Techniques for Assessing Virtual Agent Personality'. In: *IEEE Transactions on Affective Computing* 7.1 (2016), pp. 94–105. DOI: 10.1109/TAFFC.2015.2435780.

[134] Lopatovska, Irene, Korshakova, Elena, Brown, Diedre, Li, Yiqiao, Min, Jie, Pasiak, Amber and Zheng, Kaige. 'User Perceptions of an Intelligent Personal Assistant's Personality: The Role of Interaction Context'. In: *Proceedings of the 2021 Conference on Human Information Interaction and Retrieval*. New York, NY, USA: Association for Computing Machinery, 2021, pp. 15–25. DOI: 10.1145/3406522.3446018.

[135] Luger, Ewa and Sellen, Abigail. '"Like Having a Really Bad PA": The Gulf between User Expectation and Experience of Conversational Agents'. In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI '16. New York, NY,

USA: Association for Computing Machinery, 2016, pp. 5286–5297. DOI: 10.1145/2858036.2858288.

[136] Luxton, David D. 'Ethical implications of conversational agents in global public health'. In: *Bulletin of the World Health Organization* 98.4 (2020), pp. 285–287.

[137] Maes, Pattie. 'Artificial Life Meets Entertainment: Lifelike Autonomous Agents'. In: *Commun. ACM* 38.11 (1995), pp. 108–114. DOI: 10.1145/219717.219808.

[138] Mairesse, François and Walker, Marilyn A. 'Towards personality-based user adaptation: psychologically informed stylistic language generation'. In: *User Modeling and User-Adapted Interaction* 20.3 (2010), pp. 227–278. DOI: 10.1007/s11257-010-9076-2.

[139] Matthews, Gerald, Deary, Ian J and Whiteman, Martha C. Personality traits. Cambridge, UK: Cambridge University Press, 2003.

[140] Matz, SC, Kosinski, M, Nave, G and Stillwell, DJ. 'Psychological targeting as an effective approach to digital mass persuasion'. In: *Proceedings of the National Academy of Sciences* 114.48 (2017), pp. 12714–12719. DOI: 10.1073/pnas.1710966114.

[141] McCrae, Robert R. 'The Five-Factor Model of personality traits: consensus and controversy'. In: *The Cambridge Handbook of Personality Psychology*. Ed. by Philip J Corr and Gerald Matthews. Cambridge, UK: Cambridge University Press, 2009, pp. 148–161. DOI: 10.1017/CBO9780511596544.012.

[142] McCrae, Robert R and Costa, Paul T. 'A five-factor theory of personality'. In: *Handbook of Personality: Theory and Research*. Ed. by OP John, RW Robins and LA Pervin. Vol. 3. New York, NY, USA: The Guilford Press, 2008, pp. 159–181.

[143] McCrae, Robert R and Costa Jr, Paul T. 'Personality trait structure as a human universal.' In: *American psychologist* 52.5 (1997), pp. 509–516. DOI: 10.1037/0003-066X.52.5.509.

[144] McElreath, Richard. Statistical rethinking: A Bayesian course with examples in R and Stan. London, UK: Chapman and Hall/CRC, 2016. DOI: 10.1201/9781315372495.

[145] McRorie, Margaret, Sneddon, Ian, McKeown, Gary, Bevacqua, Elisabetta, Sevin, Etienne de and Pelachaud, Catherine. 'Evaluation of Four Designed Virtual Agent Personalities'. In: *IEEE Transactions on Affective Computing* 3.3 (2012), pp. 311–322. DOI: 10.1109/T-AFFC.2011.38.

[146] McTear, Michael. Conversational AI: Dialogue Systems, Conversational Agents, and Chatbots. Cham, Switzerland: Springer Cham, 2021. DOI: 10.1007/978-3-031-02176-3.

[147] McTear, Michael, Callejas, Zoraida and Griol, David. The Conversational Interface: Talking to Smart Devices. Cham, Switzerland: Springer Cham, 2016. DOI: 10.1007/978-3-319-32967-3_2.

[148] McTear, Michael F. 'The Rise of the Conversational Interface: A New Kid on the Block?' In: *Future and Emerging Trends in Language Technology. Machine Learning and Big Data*. Ed. by José F Quesada, Francisco-Jesús Martín Mateos and Teresa López Soto. Cham, Switzerland: Springer International Publishing, 2017, pp. 38–49. DOI: 10.1007/978-3-319-69365-1_3.

[149] Mehl, Matthias R, Gosling, Samuel D and Pennebaker, James W. 'Personality in its natural habitat: Manifestations and implicit folk theories of personality in daily life.' In: *Journal of Personality and Social Psychology* 90.5 (2006), pp. 862–877. DOI: 10.1037/0022-3514.90.5.862.

[150] Mehrabian, Albert. Silent messages. Vol. 8. 152. Belmont, CA, USA: Wadsworth, 1971.

[151] Miller, Hannah Jean, Thebault-Spieker, Jacob, Chang, Shuo, Johnson, Isaac, Terveen, Loren and Hecht, Brent. '"Blissfully happy" or "ready to fight": Varying Interpretations of Emoji'. In: *International AAAI Conference on Web and Social Media*. Palo Alto, CA, USA: AAAI Publications, 2017, pp. 259–268.

[152] Montag, Christian and Elhai, Jon D. 'A new agenda for personality psychology in the digital age?' In: *Personality and Individual Differences* 147 (2019), pp. 128–134. DOI: 10.1016/j.paid.2019.03.045.

[153] Moon, Youngme and Nass, Clifford. 'How "Real" Are Computer Personalities?: Psychological Responses to Personality Types in Human-Computer Interaction'. In: *Communication Research* 23.6 (1996), pp. 651–674. DOI: 10.1177/009365096023006002.

[154] Moreno, José David, Martínez-Huertas, José Á., Olmos, Ricardo, Jorge-Botana, Guillermo and Botella, Juan. 'Can personality traits be measured analyzing written language? A meta-analytic study on computational methods'. In: *Personality and Individual Differences* 177 (2021), p. 110818. DOI: 10.1016/j.paid.2021.110818.

[155] Munteanu, Cosmin, Penn, Gerald and Murad, Christine. 'Conversational Voice User Interfaces: Connecting Engineering Fundamentals to Design Considerations'. In: *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. CHI '21. New York, NY, USA: Association for Computing Machinery, 2021. DOI: 10.1145/3411763.3445008.

[156] Murad, Christine and Munteanu, Cosmin. 'Designing Voice Interfaces: Back to the (Curriculum) Basics'. In: *Proceedings of the 2020 CHI Conference on Human Factors*

in Computing Systems. CHI ’21. New York, NY, USA: Association for Computing Machinery, 2020, pp. 1–12. DOI: 10.1145/3313831.3376522.

[157] Nass, Clifford and Brave, Scott. Wired for speech: How voice activates and advances the human-computer relationship. Cambridge, MA, USA: MIT press, 2005.

[158] Nass, Clifford and Lee, Kwan Min. ‘Does computer-synthesized speech manifest personality? Experimental tests of recognition, similarity-attraction, and consistency-attraction.’ In: *Journal of Experimental Psychology: Applied* 7.3 (2001), pp. 171–181. DOI: 10.1037/1076-898X.7.3.171.

[159] Nass, Clifford and Moon, Youngme. ‘Machines and Mindlessness: Social Responses to Computers’. In: *Journal of Social Issues* 56.1 (2000), pp. 81–103. DOI: 10.1111/0022-4537.00153.

[160] Nass, Clifford, Moon, Youngme, Fogg, BJ, Reeves, Byron and Dryer, D Christopher. ‘Can computer personalities be human personalities?’ In: *International Journal of Human-Computer Studies* 43.2 (1995), pp. 223–239. DOI: 10.1006/ijhc.1995.1042.

[161] Nass, Clifford, Steuer, Jonathan and Tauber, Ellen R. ‘Computers Are Social Actors’. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’94. New York, NY, USA: Association for Computing Machinery, 1994, pp. 72–78. DOI: 10.1145/191666.191703.

[162] Nass, Clifford and Yen, Corina. The man who lied to his laptop: What we can learn about ourselves from our machines. London, UK: Penguin, 2010.

[163] Nave, Gideon, Minxha, Juri, Greenberg, David M, Kosinski, Michal, Stillwell, David and Rentfrow, Jason. ‘Musical Preferences Predict Personality: Evidence From Active Listening and Facebook Likes’. In: *Psychological Science* 29.7 (2018), pp. 1145–1158. DOI: 10.1177/0956797618761659.

[164] Neff, Michael, Toothman, Nicholas, Bowmani, Robeson, Fox Tree, Jean E and Walker, Marilyn A. ‘Don’t Scratch! Self-adaptors Reflect Emotional Stability’. In: *Intelligent Virtual Agents. IVA 2011. Lecture Notes in Computer Science*. Ed. by Hannes Högni Vilhjálmsson, Stefan Kopp, Stacy Marsella and Kristinn R Thórisson. Vol. 6895. Berlin, Heidelberg, Germany: Springer, 2011, pp. 398–411. DOI: 10.1007/978-3-642-23974-8_43.

[165] Neff, Michael, Wang, Yingying, Abbott, Rob and Walker, Marilyn. ‘Evaluating the Effect of Gesture and Language on Personality Perception in Conversational Agents’. In: *Intelligent Virtual Agents*. Ed. by Jan Allbeck, Norman Badler, Timothy Bickmore, Catherine Pelachaud and Alla Safonova. Berlin, Heidelberg, Germany: Springer, 2010, pp. 222–235. DOI: 10.1007/978-3-642-15892-6_24.

[166] Norman, Donald A. User centered system design: New perspectives on human-computer interaction. Hillsdale, NJ, US: Lawrence Erlbaum Associates, 1986.

[167] Oberlander, Jon and Gill, Alastair J. 'Individual differences and implicit language: personality, parts-of-speech and pervasiveness'. In: *Proceedings of the Annual Meeting of the Cognitive Science Society* 26 (2004), pp. 1035–1040.

[168] Ogawa, Yoshito, Miyazawa, Kouki and Kikuchi, Hideaki. 'Assigning a Personality to a Spoken Dialogue Agent through Self-Disclosure of Behavior'. In: *Proceedings of the Second International Conference on Human-Agent Interaction*. HAI '14. New York, NY, USA: Association for Computing Machinery, 2014, pp. 331–337. DOI: 10.1145/2658861.2658884.

[169] Ozer, Daniel J and Benet-Martínez, Verónica. 'Personality and the Prediction of Consequential Outcomes'. In: *Annual Review of Psychology* 57.1 (2006), pp. 401–421. DOI: 10.1146/annurev.psych.57.102904.190127.

[170] Patterson, M and Holmes, DS. 'Social Interaction Correlates of the MMPI Extraversion Introversion Scale'. In: *American Psychologist* 21 (1966), pp. 724–25.

[171] Pearl, Cathy. Designing voice user interfaces: Principles of conversational experiences. Sebastopol, CA, USA: O'Reilly Media, Inc., 2016.

[172] Pennebaker, James W and King, Laura A. 'Linguistic styles: Language use as an individual difference.' In: *Journal of Personality and Social Psychology* 77.6 (1999), pp. 1296–1312. DOI: 10.1037/0022-3514.77.6.1296.

[173] Perez, Marta Garcia, Lopez, Sarita Saffon and Donis, Hector. 'Voice Activated Virtual Assistants Personality Perceptions and Desires: Comparing Personality Evaluation Frameworks'. In: *Proceedings of the 32nd International BCS Human Computer Interaction Conference*. HCI '18. Swindon, UK: BCS Learning & Development Ltd., 2018. DOI: 10.14236/ewic/HCI2018.40.

[174] Pfeifer, Laura M and Bickmore, Timothy W. 'Should Agents Speak Like, um, Humans? The Use of Conversational Fillers by Virtual Agents'. In: *Intelligent Virtual Agents*. Ed. by Zsófia Ruttkay, Michael Kipp, Anton Nijholt and Hannes Högni Vilhjálmsson. Berlin, Heidelberg, Germany: Springer, 2009, pp. 460–466. DOI: 10.1007/978-3-642-04380-2_50.

[175] Pianesi, Fabio. 'Searching for personality [social sciences]'. In: *IEEE Signal Processing Magazine* 30.1 (2012), pp. 146–158. DOI: 10.1109/MSP.2012.2219671.

[176] Pinker, Steven. The language instinct: How the mind creates language. London, UK: Penguin, 2003.

[177] Pohl, Henning, Domin, Christian and Rohs, Michael. 'Beyond Just Text: Semantic Emoji Similarity Modeling to Support Expressive Communication'. In: *ACM Trans. Comput.-Hum. Interact.* 24.1 (2017), 6:1–6:42. DOI: 10.1145/3039685.

[178] Porcheron, Martin, Fischer, Joel E, Reeves, Stuart and Sharples, Sarah. 'Voice Interfaces in Everyday Life'. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI '18. New York, NY, USA: Association for Computing Machinery, 2018, pp. 1–12. DOI: 10.1145/3173574.3174214.

[179] Pradhan, Alisha and Lazar, Amanda. 'Hey Google, Do You Have a Personality? Designing Personality and Personas for Conversational Agents'. In: *CUI 2021 - 3rd Conference on Conversational User Interfaces*. CUI '21. New York, NY, USA: Association for Computing Machinery, 2021. DOI: 10.1145/3469595.3469607.

[180] Quercia, Daniele, Lambiotte, Renaud, Stillwell, David, Kosinski, Michal and Crowcroft, Jon. 'The Personality of Popular Facebook Users'. In: *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*. CSCW '12. New York, NY, USA: Association for Computing Machinery, 2012, pp. 955–964. DOI: 10.1145/2145204.2145346.

[181] Raad, Boele de. The Big Five Personality Factors: The psycholexical approach to personality. Göttingen, Germany: Hogrefe & Huber Publishers, 2000.

[182] Reeves, Byron and Nass, Clifford Ivar. The media equation: How people treat computers, television, and new media like real people and places. Cambridge, UK: Cambridge University Press, 1996.

[183] Reeves, Stuart. 'Conversation Considered Harmful?' In: *Proceedings of the 1st International Conference on Conversational User Interfaces*. CUI '19. New York, NY, USA: Association for Computing Machinery, 2019. DOI: 10.1145/3342775.3342796.

[184] Rist, Thomas, André, Elisabeth and Baldes, Stephan. 'A Flexible Platform for Building Applications with Life-like Characters'. In: *Proceedings of the 8th International Conference on Intelligent User Interfaces*. IUI '03. New York, NY, USA: Association for Computing Machinery, 2003, pp. 158–165. DOI: 10.1145/604045.604071.

[185] Ritschel, Hannes and André, Elisabeth. 'Real-Time Robot Personality Adaptation Based on Reinforcement Learning and Social Signals'. In: *Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*. HRI '17. New York, NY, USA: Association for Computing Machinery, 2017, pp. 265–266. DOI: 10.1145/3029798.3038381.

[186] Ritschel, Hannes, Baur, Tobias and André, Elisabeth. 'Adapting a Robot's linguistic style based on socially-aware reinforcement learning'. In: *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. New York, NY, USA: IEEE, 2017, pp. 378–384. DOI: 10.1109/ROMAN.2017.8172330.

[187] Robert, Lionel, Alahmad, Rasha, Esterwood, Connor, Kim, Sangmi, You, Sangseok and Zhang, Qiaoning. 'A Review of Personality in Human–Robot Interactions'. In: *Foundations and Trends in Information Systems* 4.2 (2020), pp. 107–112. DOI: 10.2139/ssrn.3528496.

[188] Rosenthal-Von Der Pütten, Astrid M, Schulte, Frank P, Eimler, Sabrina C, Sobieraj, Sabrina, Hoffmann, Laura, Maderwald, Stefan, Brand, Matthias and Krämer, Nicole C. 'Investigations on empathy towards humans and robots using fMRI'. In: *Computers in Human Behavior* 33 (2014), pp. 201–212. DOI: 10.1016/j.chb.2014.01.004.

[189] Ross, Lee. 'The Intuitive Psychologist And His Shortcomings: Distortions in the Attribution Process'. In: *Advances in Experimental Social Psychology*. Ed. by Leonard Berkowitz. Vol. 10. Academic Press, 1977, pp. 173–220. DOI: 10.1016/S0065-2601(08)60357-3.

[190] Ruane, Elayne, Farrell, Sinead and Ventresque, Anthony. 'User Perception of Text-Based Chatbot Personality'. In: *Chatbot Research and Design*. Cham, Switzerland: Springer International Publishing, 2021, pp. 32–47. DOI: 10.1007/978-3-030-68288-0_3.

[191] Ruijten, Peter AM. 'The similarity-attraction paradigm in persuasive technology: Effects of system and user personality on evaluations and persuasiveness of an interactive system'. In: *Behaviour & Information Technology* 40.8 (2021), pp. 734–746. DOI: 10.1080/0144929X.2020.1723701.

[192] Rutter, DR, Morley, Ian E and Graham, Jane C. 'Visual interaction in a group of introverts and extraverts'. In: *European Journal of Social Psychology* 2.4 (1972), pp. 371–384. DOI: 10.1002/ejsp.2420020403.

[193] Saucier, Gerard. 'Mini-Markers: A Brief Version of Goldberg's Unipolar Big-Five Markers'. In: *Journal of Personality Assessment* 63.3 (1994), pp. 506–516. DOI: 10.1207/s15327752jpa6303_8.

[194] Schegloff, Emanuel A. 'Sequencing in Conversational Openings'. In: *American Anthropologist* 70.6 (1968), pp. 1075–1095. DOI: 10.1525/aa.1968.70.6.02a00030.

[195] Scherer, Klaus R. 'Personality inference from voice quality: The loud voice of extroversion'. In: *European Journal of Social Psychology* 8.4 (1978), pp. 467–487. DOI: 10.1002/ejsp.2420080405.

[196] Scherer, Klaus Rainer. 'Personality markers in speech'. In: *Social Markers in Speech*. Ed. by Klaur Rainer Scherer and Howard Giles. Cambridge, UK: Cambridge University Press, 1979.

[197] Schwartz, H Andrew, Eichstaedt, Johannes C, Kern, Margaret L, Dziurzynski, Lukasz, Ramones, Stephanie M, Agrawal, Megha, Shah, Achal, Kosinski, Michal, Stillwell, David and Seligman, Martin EP et al. 'Personality, Gender, and Age in the Language of Social Media: The Open-Vocabulary Approach'. In: *PLOS ONE* 8.9 (2013), e73791. DOI: 10.1371/journal.pone.0073791.

[198] Shneiderman, Ben. 'Bridging the Gap Between Ethics and Practice: Guidelines for Reliable, Safe, and Trustworthy Human-Centered AI Systems'. In: *ACM Trans. Interact. Intell. Syst.* 10.4 (2020). DOI: 10.1145/3419764.

[199] Short, Elaine, Swift-Spong, Katelyn, Greczek, Jillian, Ramachandran, Aditi, Litoiu, Alexandru, Grigore, Elena Corina, Feil-Seifer, David, Shuster, Samuel, Lee, Jin Joo, Huang, Shaobo, Levonisova, Svetlana, Litz, Sarah, Li, Jamy, Ragusa, Gisele, Spruijt-Metz, Donna, Matarić, Maja and Scassellati, Brian. 'How to train your DragonBot: Socially assistive robots for teaching children about nutrition through play'. In: *The 23rd IEEE International Symposium on Robot and Human Interactive Communication*. New York, NY, USA: IEEE, 2014, pp. 924–929. DOI: 10.1109/ROMAN.2014.6926371.

[200] Shumanov, Michael and Johnson, Lester. 'Making conversations with chatbots more personalized'. In: *Computers in Human Behavior* 117 (2021), p. 106627. DOI: 10.1016/j.chb.2020.106627.

[201] Soto, Christopher J and John, Oliver P. 'Short and extra-short forms of the Big Five Inventory–2: The BFI-2-S and BFI-2-XS'. In: *Journal of Research in Personality* 68 (2017), pp. 69–81. DOI: 10.1016/j.jrp.2017.02.004.

[202] Soto, Christopher J and John, Oliver P. 'The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power.' In: *Journal of Personality and Social Psychology* 113.1 (2017), pp. 117–143. DOI: 10.1037/pspp0000096.

[203] Ta, Vivian, Griffith, Caroline, Boatfield, Carolynn, Wang, Xinyu, Civitello, Maria, Bader, Haley, DeCero, Esther and Loggarakis, Alexia. 'User Experiences of Social Support From Companion Chatbots in Everyday Contexts: Thematic Analysis'. In: *J Med Internet Res* 22.3 (2020), e16235. DOI: 10.2196/16235.

[204] Thomas, Frank, Johnston, Ollie and Thomas, Frank. The illusion of life: Disney animation. New York, NY, USA: Hyperion, 1995.

[205] Vinciarelli, Alessandro and Mohammadi, Gelareh. 'A survey of personality computing'. In: *IEEE Transactions on Affective Computing* 5.3 (2014), pp. 273–291. DOI: 10.1109/TAFFC.2014.2330816.

[206] Vlahos, James. Talk to Me: How Voice Computing Will Transform the Way We Live, Work, and Think. Boston, MA, USA: Houghton Mifflin Harcourt, 2019.

[207] Wacharamanotham, Chat, Subramanian, Krishna, Völkel, Sarah Theres and Borchers, Jan. 'Statsplorer: Guiding Novices in Statistical Analysis'. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. CHI '15. New York, NY, USA: Association for Computing Machinery, 2015, pp. 2693–2702. DOI: 10.1145/2702123.2702347.

[208] Walker, Marilyn and Whittaker, Steve. 'Mixed Initiative in Dialogue: An Investigation into Discourse Segmentation'. In: *Proceedings of the 28th Annual Meeting on Association for Computational Linguistics*. ACL '90. USA: Association for Computational Linguistics, 1990, pp. 70–78. DOI: 10.3115/981823.981833.

[209] Walther, Joseph B. 'Theories of computer-mediated communication and interpersonal relations'. In: *The handbook of interpersonal communication*. Thousand Oaks, CA, USA: Sage Publications Ltd., 2011. Chap. 4, pp. 443–479.

[210] Wang, Lin, Rau, Pei-Luen Patrick, Evers, Vanessa, Robinson, Benjamin Krisper and Hinds, Pamela. 'When in Rome: The role of culture and context in adherence to robot recommendations'. In: *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. New York, NY, USA: ACM/IEEE, 2010, pp. 359–366. DOI: 10.1109/HRI.2010.5453165.

[211] Wang, Yingying, Tree, Jean E Fox, Walker, Marilyn and Neff, Michael. 'Assessing the Impact of Hand Motion on Virtual Character Personality'. In: *ACM Trans. Appl. Percept.* 13.2 (2016), pp. 1–23. DOI: 10.1145/2874357.

[212] Weizenbaum, Joseph. Computer power and human reason: From judgment to calculation. New York, NY, USA: WH Freeman & Co, 1976.

[213] West, Mark, Kraut, Rebecca and Ei Chew, Han. I'd blush if I could: closing gender divides in digital skills through education. Tech. rep. UNESCO, 2019.

[214] Whitby, Blay. 'Sometimes it's hard to be a robot: A call for action on the ethics of abusing artificial agents'. In: *Interacting with Computers* 20.3 (2008), pp. 326–333. DOI: 10.1016/j.intcom.2008.02.002.

[215] Wiggins, Jerry S. 'A psychological taxonomy of trait-descriptive terms: The interpersonal domain.' In: *Journal of personality and social psychology* 37.3 (1979), p. 395.

[216] Wiggins, Jerry S. 'Circumplex models of interpersonal behavior in clinical psychology'. In: *Handbook of research methods in clinical psychology* (1982), pp. 183–221.

[217] Wobbrock, Jacob O and Kientz, Julie A. 'Research Contributions in Human-Computer Interaction'. In: *Interactions* 23.3 (2016), pp. 38–44. DOI: 10.1145/2907069.

[218] Wooldridge, Michael and Jennings, Nicholas R. 'Intelligent agents: theory and practice'. In: *The Knowledge Engineering Review* 10.2 (1995), pp. 115–152. DOI: 10.1017/S0269888900008122.

[219] Wright, Jack C and Mischel, Walter. 'A conditional approach to dispositional constructs: the local predictability of social behavior.' In: *Journal of personality and social psychology* 53.6 (1987), pp. 1159–1177. DOI: 10.1037/0022-3514.53.6.1159.

[220] Wu, Yunhan, Edwards, Justin, Cooney, Orla, Bleakley, Anna, Doyle, Philip R, Clark, Leigh, Rough, Daniel and Cowan, Benjamin R. 'Mental Workload and Language Production in Non-Native Speaker IPA Interaction'. In: *Proceedings of the 2nd Conference on Conversational User Interfaces*. CUI '20. New York, NY, USA: Association for Computing Machinery, 2020. DOI: 10.1145/3405755.3406118.

[221] Wu, Yunhan, Rough, Daniel, Bleakley, Anna, Edwards, Justin, Cooney, Orla, Doyle, Philip R, Clark, Leigh and Cowan, Benjamin R. 'See What I'm Saying? Comparing Intelligent Personal Assistant Use for Native and Non-Native Language Speakers'. In: *22nd International Conference on Human-Computer Interaction with Mobile Devices and Services*. MobileHCI '20. New York, NY, USA: Association for Computing Machinery, 2020. DOI: 10.1145/3379503.3403563.

[222] Xiao, Ziang, Zhou, Michelle X and Fu, Wat-Tat. 'Who Should Be My Teammates: Using a Conversational Agent to Understand Individuals and Help Teaming'. In: *Proceedings of the 24th International Conference on Intelligent User Interfaces*. IUI '19. New York, NY, USA: Association for Computing Machinery, 2019, pp. 437–447. DOI: 10.1145/3301275.3302264.

[223] Xiao, Ziang, Zhou, Michelle X, Liao, Q Vera, Mark, Gloria, Chi, Changyan, Chen, Wenxi and Yang, Huahai. 'Tell Me About Yourself: Using an AI-Powered Chatbot to Conduct Conversational Surveys with Open-Ended Questions'. In: *ACM Trans. Comput.-Hum. Interact.* 27.3 (2020). DOI: 10.1145/3381804.

[224] Zhou, Michelle X, Mark, Gloria, Li, Jingyi and Yang, Huahai. 'Trusting Virtual Agents: The Effect of Personality'. In: *ACM Trans. Interact. Intell. Syst.* 9.2-3 (2019), 10:1–10:36. DOI: 10.1145/3232077.

List of Figures

List of Tables

A APPENDIX

Table A.1 provides an overview of my own and collaborators' contributions in regard to the core publications included in this thesis.

	My Contribution	**Contribution of Co-authors**
[Core1]	I was the project lead and first author of the resulting publication. I developed the research idea and the study design, supervised the implementation of the study, and analysed the resulting qualitative and quantitative data.	Daniel Buschek discussed the research design with me, conducted the regression models for the quantitative analysis with me, and actively contributed to the written publication. Jelena Pranjic contributed to developing the research design and conducted the study. Heinrich Hussmann provided feedback on the final paper and during all stages of the research project.
[Core2]	I was the project lead and first author of the resulting publication. I developed the research idea and the study design. I compiled the set of behaviour cues and supervised the implementation of the chatbot prototypes as well as the study. I analysed the data.	Lale Kaya contributed to developing the research design and the set of behaviour cues. She implemented the prototype and conducted the study.
[Core3]	I was the project lead and first author of the resulting publication. I developed the research idea and the study design. I compiled the set of behaviour cues and implemented the first version of the chatbot together with my colleagues Carl Oechsner and Fiona Draxler. I supervised the extension of the implementation of the prototype and the study. I wrote the scripts to pre-process the chatbot log data and analysed the resulting qualitative and quantitative data.	Lale Kaya substantially extended the implementation of the chatbot, contributed to the set of behaviour cues, supported the development of the research design, and conducted the study. Ramona Schödel provided feedback on the research design, wrote the Limitations section, and created the LMM Supplementary Material. Ramona Schödel and Sven Mayer discussed the statistical analyses, re-calculated analyses to ensure consistent findings, and provided feedback on the written paper.

[Core4]	I was the project lead and first author of the resulting publication. I developed the research idea and the different study designs. I conducted the online survey and qualitatively analysed the results. I supervised the implementation of the lab study and qualitatively analysed the results. I wrote the scripts to crawl the online reviews and preprocess them. I developed the mechanisms to reduce the item pool sets and wrote the respective scripts in addition to performing all qualitative analyses. I recruited participants via Mechanical Turk and contributed to the implementation of the online survey and the factor analysis.	Ramona Schödel conducted the online survey and calculated the factor analysis on the resulting descriptors. Furthermore, she co-performed parts of the qualitative analysis of the online reviews and synonym analysis to ensure independent consistent coding. Daniel Buschek discussed the research design and the results and actively contributed to the written publication. Verena Winterhalter contributed to discussing the research design of the lab study, conducted this study, and transcribed the interviews. Clemens Stachl and Markus Bühner provided feedback on the research design and the final publication. Heinrich Hussmann provided feedback on the final publication and during all stages of the research project.
[Core5]	I was the project lead and first author of the resulting publication. I developed the research idea, the two study designs, and the personality-infused dialogues. I supervised the implementation of the focus groups, the Alexa prototype, and the online survey. I analysed the results.	Samantha Meindl contributed to developing the research design and conducted the focus groups and online survey. Furthermore, she drafted the dialogues based on the focus groups and implemented them using Amazon Alexa. Heinrich Hussmann provided feedback on the final publication and during all stages of the research project.
[Core6]	I was the project lead and first author of the resulting publication. I developed the research idea and the research design. I supervised the implementation of the interviews. I analysed the results.	Penelope Kempf conducted the study. Heinrich Hussmann provided feedback on the final publication and during all stages of the research project.
[Core7]	I was the project lead and first author of the resulting publication. I developed the research idea and the research design and conducted the online survey. I analysed the data qualitatively and quantitatively.	Malin Eiband and Daniel Buschek contributed to developing the research design, discussed the results and coding schemes, and edited the written publication. Malin Eiband re-coded 13.2% of the data to ensure inter-coder stability of the findings and qualitatively analysed the open scenario. Daniel Buschek and I discussed the LMM analysis together, Daniel Buschek wrote the according first draft of the script. Benjamin Cowan provided feedback on the written publication, the results (i.e. coding scheme), and possible analysis methods. Heinrich Hussmann provided feedback on the final publication and during all stages of the research project.

Table A.1: Clarification of contributions for all core publications included in this thesis.

Eidesstattliche Versicherung

(Siehe Promotionsordnung vom 12.07.11, § 8, Abs. 2 Pkt. 5)

Hiermit erkläre ich an Eidesstatt, dass die Dissertation von mir selbstständig und ohne unerlaubte Beihilfe angefertigt wurde.

München, den 10. Mai 2022

Sarah Theres Völkel