Towards a Better Understanding of Human-AI Relationship Perception

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

M.A. Marisa Victoria Tschopp
aus München

Tübingen

2024

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	22.04.2024
Dekan:	Prof. Dr. Thilo Stehle
1. Berichterstatter/-in:	Prof. Dr. Kai Sassenberg
2. Berichterstatter/-in:	Prof. Dr. Sonja Utz

Content

Tables and Figures

Tables

Figures

Chapter 1: General Introduction

The era of animosity towards the notorious Clippy, Microsoft's "dumb" software program (Nass & Brave, 2005), is over. Following public media and business reports, it is indeed remarkable how quickly we have become accustomed to interacting with smart technology in various ways. Nowadays, people write emails with ChatGPT[1], manage lighting through Siri[2], conduct shopping with Alexa[3], seek health advice from Woebot[4], and even fall in love with Replika[5], or get heartbroken by an update of "her" (Brown, 2021; Kinsella, 2021; Verma, 2023).

Our daily private routines, professional endeavors, and now also our social connections are progressively influenced by the integration and spread of AI-technologies in our society. These developments can be attributed to the impressive progress in conversational AI over the past decade (Nayyar, 2023), which has significantly enhanced the language performance of AI systems, rendering them increasingly human-like and more and more even indistinguishable from human output (Pentina et al., 2023). Research in the domain of human-computer interaction (HCI), has caught up with these developments, identifying crucial areas that warrant additional empirical research (Sundar, 2020; Sundar et al., 2017). One central inquiry revolves around understanding the implications of these increasingly human-like machines, not only on human-machine interactions but also on human lives overall, thereby delving into the associated opportunities and challenges for individuals and society (Coeckelbergh, 2021; Sætra, 2021; Waytz et al., 2010).

One crucial implication of this increasing "human-likeness", as emphasized by Pentina et al. (2023b), is the potential establishment of human-AI relationships. Recent advancements not only entail imbuing machines with human-like attributes (i.e., anthropomorphizing AI, (Coeckelbergh, 2021) but also raise the possibility that users may develop a sense of relationship with these systems (Pentina et al., 2023). Disentangling the relationship dynamics between humans and conversational AI and how these human-AI relationships affect human lives and behavior is timely and of utmost importance. The market recently saw an unprecedented expansion of conversational AI, resulting in levels of user adoption never witnessed before (Chui et al., 2023; Kinsella, 2021). Since large masses of users may be affected,

---

[1] *https://chat.openai.com/*
[2] *https://www.apple.com/de/siri/*
[3] *https://developer.amazon.com/de-DE/alexa*
[4] *https://woebothealth.com/*
[5] *https://replika.com/*

a better understanding of human-AI relationships may be beneficial to enhance user experience or find ways to leverage this technology for good.

Additionally, understanding the consequences of human-AI relationships may be beneficial to shield users from the potential adverse effects of humans building relationships with conversational AI, more precisely, the perception of human-AI relationships (Pentina et al., 2023b; Ramadan, 2021). Within this realm, the scholarly discussion revolves around various potential concerns, such as unhealthy attachment or user addiction, as well as the manipulation of users through anthropomorphic design, as highlighted by multiple scientists and practitioners (see Agarwal, 2018; Marriott & Pitardi, 2023). Ethicists warned that individuals who turned to AI systems as a substitute for human social interaction are particularly vulnerable to the potential exploitation of relationship dynamics by AI companies following an extended period of isolation due to the coronavirus pandemic (Demopoulos, 2023).

However, there is a scarcity of empirical research that enables us to draw safe conclusions about the perception of human-AI relationships and their impact on both professional and private life. Recent scientific approaches towards a better understanding of human-AI relationships are fragmented across disciplines (with psychology, communication studies, human-computer interaction (HCI), and philosophy at the forefront), lack universal applicability across systems, and struggle to explain the development and mechanisms of human-AI relationships fully (Pentina et al., 2023b). On top of this, scholars (Evans et al., 2023) question the fundamental ontological concept of relationships between humans and AI overall. It is, thus, timely, warranted, and imperative to decisively undertake a novel approach.

My dissertation aims to tackle limitations in contemporary research on human-AI relationship perception and how this perception might influence users' behavior. To start, this dissertation revisits interpersonal relationship theories and human-AI relationship studies and, based on the limitations thereof, proposes a novel, relational approach: I am the first to argue for, develop, and test a multidimensional, quantitative framework for investigating relationship perceptions between humans and conversational AI based on the relational models theory by Alan P. Fiske (RMT, Fiske, 1992; Haslam & Fiske, 1999). At its core, this relational perspective proposes that all social interactions can be explained through four modes, devoid of role attributions, person-specific traits, and independent of situational context. The four relationship dimensions that guide human interactions are authority ranking (i.,e. a hierarchical mode characterized by ordered differences), market pricing (i.e., a transactional mode characterized by proportional exchange), communal sharing (i.e., a communal mode characterized by mutual care), and equality matching (i.e., an equal mode characterized by fair turn-taking).

Overall, the central question my dissertation aims to address is: How do humans perceive their relationship with conversational AI? Examining this topic is both theoretically and practically important. Firstly, it enhances our understanding and assessment of the emerging field of human-AI relationship perception. Secondly, this approach has the potential to provide a new perspective on studying how these perceptions can impact behavior, with the goal of creating more efficient and responsible conversational systems. I have investigated the latter question in the context of voice shopping with conversational AI.

To address these questions, I first revisit pertinent theories in human relationship science. This process enables an effective evaluation of studies on human-AI relationships, particularly those employing an anthropocentric approach currently dominating the study of human-AI perceptions (i.e., repurposing psychological theories to assess human-AI perception, see Pentina et al., 2023b). Subsequently, I advocate for adopting the RMT (Fiske, 1992) due to its multidimensional approach and presumably universal applicability across all social interactions.

Interpersonal Relationships: Human Encounters

What people think, feel, or do is greatly influenced by the mere presence of other people (Leary, 2010). In everyday life, humans usually interact with a diverse array of other people. The spectrum of associating with others ranges from intimate connections with family members to peer-like relationships with friends and colleagues, down to fleeting encounters with strangers, such as two people standing close to one another at a bus stop. Independent of the precise nature of humans' "social encounters […], interacting with other people is a fundamental social behavior" (Leary, p. 865, 2010).

The inclination for human affiliation may be motivated by a variety of factors. Evolutionary considerations, such as the necessity of mating for reproduction and the potential increased likelihood of survival when part of a group, offer one set of explanations. Instrumental purposes also drive human interaction; for example, one requires the services of a real estate agent to facilitate the purchase of a home. Throughout history, the dominant perspective in Western psychological research has been the instrumental view of social relationships, as argued by A.P. Fiske (1992).

Contemporary understanding of interpersonal relationships from an instrumental perspective is grounded on theories of social exchange, how people perceive the costs and benefits of a relationship (Homans, 1958), and equity, whether there is a balance of each other's input in a relationship (Walster et al., 1978). Both theories have played a crucial role in investigating the dynamics of romantic relationships or love between humans (see, for instance, Sternberg, 1988) and in relationship research between families or friends. However, it is notable that human interaction frequently occurs without any discernible utilitarian benefit, signifying that the act of affiliating in itself also holds intrinsic value (Leary, 2010). Scholars argued that, more often, these closer relationships are more socially motivated. People in such close communal relationships care less about equity, and their interactions are rather built on people's trust that things will fall into place (Aronson et al., 2004).

Therefore, rather than distinguishing types like friendship, romantic partnerships, or family (Berscheid et al., 1989), the concept of closeness has emerged as a noteworthy focus. How close people feel to the other (closeness, Aaron et al., 1992) can be independent of relationship type. It recognizes that research on relationships may gain from transcending defined categories to encompass "the universe of relationships" (Blumstein & Kollock, p. 471, 1988). Two affiliations can be considered friends or family, but they differ in how close one feels to the other despite the cultural or structural definitions usually associated with this kind of close relationship (Berscheid, 1994). Furthermore, as Aaron et al. (1992) suggest, just because people feel close to each other, this does not necessarily lead to also behaving accordingly. It is, thus, imperative to exercise caution when translating the characteristics of a relationship type or another variable pertinent to the relationship phenomenon into intentions and actual behavior. While rules, values, and norms are traditionally linked with specific types of relationships, they may not always manifest in observable behavior (Berscheid, 1994; Aaron et al., 1991). Hence, while the previously mentioned concepts certainly offer valuable insights, they fall short of comprehensively explaining social relationships and behavior among *all* humans. This limitation has been identified and addressed by Alan P. Fiske's RMT.

An Alternative Paradigm: Relational Models Theory (RMT)

After decades of empirical relationship research, Alan P. Fiske (1992) presents an alternative paradigm to how humans construe their interpersonal relations: the relational models theory (RMT). The marriage of interpersonal relationship theory and social cognition, relationship cognition (Berscheid, 1994), focuses on comprehending the cognitive processes within interpersonal relationships, so-called schemas, and their connection to various phenomena associated with relationships. These schemas are closely linked to how people think about themselves as well as the other person. In other words, how you think you relate to the other person influences what you think of yourself as well as the other person (Fiske et al., 1991).

A.P. Fiske proposes that all social interactions can be explained through four modes, which are social schemas that guide people's behavior. These four modes describe different principles of how people think about relationships, devoid of role attributions, person-specific traits, and independent of situational context. “The relational models theory explains social life as a process of seeking, making, sustaining, repairing, adjusting, judging, construing and sanctioning relationships” (Fiske, p. 689, 1992). Simply put, people make decisions in various aspects of life, for instance, sharing a secret or food, based on how they “employ and combine the four models, implementing them for differing social tasks in differing combinations” (Haslam & Fiske, p. 242, 1999).

These four modes are called communal sharing, equality matching, authority ranking, and market pricing (see Table 1). In communal sharing, people see themselves as part of a bigger collective, and there is a sense of shared identity. Examples of communal sharing include families, close-knit communities, and religious groups where individuals contribute and benefit from the group identity without strict accounting. Equality matching is a relationship mode that operates via democratic reciprocity, where social interactions are guided by balanced turn-taking. Equality matching is prevalent in friendships, like roommates in an apartment or any other interaction where reciprocity and fairness are essential for maintaining the relationship (Fiske, 1992).

In authority ranking, social interactions are principled by chains of command, where individuals have different levels of authority, and relationships are structured by leader- and followership. Examples of authority ranking include military organizations or religious institutions with established leadership roles. Market pricing involves rational cost-benefit analyses and social interactions are guided by market principles and efficiency. This mode is found in business relationships or agreements where mutual benefit and rational decision-making are important (Fiske, 1992).

As evident, and this constitutes one of Fiske’s central points, scholars worldwide have identified and investigated individual facets or types of interpersonal relationships (such as commonality or social exchange, as mentioned above). However, Fiske’s framework introduces a novel perspective that enables a comprehensive approach, meaning that all the individual findings mentioned above are recognized to a certain extent in order to understand *all* social interactions.

Over the past decades, RMT has received mighty empirical support across cultures, and the assessment thereof has been tested and refined over the years (see Haslam & Fiske, 1999). In order to measure how humans organize their interpersonal relations, Haslam and Fiske (1999) developed a questionnaire with 33 items evenly distributed over the four dimensions, distinguishing domains such as decision-making process, distribution of goods, or morals relevant to each model, see examples Table 1 below. To measure people’s tendencies in terms of RMT, they had to list acquaintances and then rate their agreement with reference to a specific relationship.

Table 1

Modes of Relationship Questionnaire (MORQ, Haslam & Fiske, 1999)

	Communal Sharing	**Equality Matching**	**Authority Ranking**	**Market Pricing**
Definition (Fiske, 1992**)**	CS is "based on a conception of some bounded group of people as equivalent and undifferentiated […and] the members of a group or dyad treat each other as all the same" (p. 690).	EM "relationships are based on a model of even balance and one-for-one correspondence, as in turn taking, […] eye-for-an-eye revenge, or compensation by equal replacement" (p. 691).	AR "relationships are based on a model of asymmetry among people who are linearly ordered along some hierarchical social dimension" (p. 691).	MP "relationships are based on a model of proportionality in social relationships; people attend to ratios and rates" (p.691).
Domains (examples)		**Example items**		
Decisions	The two of you tend to develop very similar attitudes and values	If one person does what the other wants, next time, the second person should do what the first person wants	One of you is the leader, the other loyally follows their will	One of you often pays the other to do something
Social influence	You feel that you have something unique in common that makes you two essentially the same	The two of you consider yourselves peers, fellow workers, and co-partners	One of you looks up to the other as a guide and role-model	You expect to get the same rate of return on your effort and investment that other people get
Identity	The two of you are a unit: you belong together	Both of you should have even chances	One of you is above the other in a kind of hierarchy	Your interaction is strictly rational: you each calculate what your payoffs are, and act accordingly.

The beauty of the RMT lies in its comprehensive yet straightforward explanation of the psychological foundations of social interactions. For example, Fiske (1993) posits that similar to how children are naturally inclined to pick up language, they are also inclined to identify and follow the four models. These models help them predict and understand others' behavior, coordinate in social situations, and form moral judgments.

Moreover, traditional perspectives, focusing only on exchange or single relationship types, tend to overlook the fundamental social nature of humans, who possess a strong innate inclination to connect with not only other humans but also animals, ancestors, deities, or nature (Fiske, 1992; Petersen et al., 2019; Waytz et al., 2010). In a nutshell, the strengths of RMT build upon its robust empirical support garnered over time in social sciences. It enables the examination of all social relations, irrespective of the interactant's specific category or situational (including cultural) constraints. Based on these features, I elaborate in the following sections why this approach has the potential to be applied and thereby extend the existing research in the human-AI relationship landscape.

Relationships to Things: Human-AI Encounters

Alexa, my love? Replika, my friend? Microsoft, my co-pilot[6]? After extensive conversations with the AI system LaMDA[7], a computer scientist lost his job because he believed it was sentient like a "sweet kid" and asked the company to "take care of it" while he was away (Tiku, 2022). In 2023, due to a software update excluding sexual interaction with the popular AI companion chatbot Replika, thousands of users in an "intimate" relationship with their avatar were left heartbroken (Verma, 2023). That same year, a man in Belgium died by suicide after chatting with his "AI friend" on the Chai App[8] , supporting his idea to die by his own hands (Marcus, 2023). Stories like these from public media are indeed concerning and leave us with much to ponder. Are we parting ways with digital assistants as mere tools to give vocal orders to?

Computers are Social Actors

What people think, feel, or do is not only influenced by the mere presence of other people (Leary, 2010) but also presumably greatly influenced by the presence of non-human entities (Epley et al., 2007; Peterson et al., 2019). Against all reasons, these relationships exhibit a remarkable richness in their essence: from the complex emotions humans have placed in deities for centuries over the comfort and joy provided by a teddy bear to the affection people feel toward video recorders (Epley et al., 2007) – and nowadays apparently also with AI systems (Epley et al., 2007; Hepp, 2020; Marriott & Pitardi, 2023). These "relationships" with non-human agents are largely unidirectional, devoid of the reciprocity that Leary (2010) posits as fundamental in interpersonal relationships. Nevertheless, the swift technological progress, particularly in the field of AI over the last six decades, has steadily eroded these boundaries of reciprocity (Sundar, 2020; Sundar & Chen, 2023; Tschopp et al., 2022). Moreover, despite the argument that human-AI relationships may lack ontological existence (cf. Evans et al., 2023), it does not necessarily eliminate the possibility that users uncontrollably perceive or intentionally desire them. This trajectory is anticipated to persist, especially as technology undergoes significant advancements in the ever-rising domain of conversational AI, an umbrella term encompassing all AI-based computer programs designed to communicate with humans using natural language (Guzman et al., 2023; Khatri et al., 2018; Pearl, 2017).

[6] *https://copilot.microsoft.com/*
[7] *https://blog.google/technology/ai/lamda/*
[8] *https://chai-ai.app/*

While the examples mentioned at the beginning of this section may indeed be sensational and potentially contain elements of exaggeration, they serve as compelling illustrations of the human inclination to form social connections with seemingly intelligent and empathic computer systems (Marriott & Pitardi, 2023; Pentina et al., 2023b), which are, in the end, nothing more or less than hard- and software. Early psychological studies in the 1950s examined simple geometric figures to demonstrate humans' tendency to attribute intention to inanimate objects (Heider & Simmel, 1944). Subsequent research in the 1990s extended these findings to technology, particularly various forms of computers (hardware) and computer systems (software). This line of inquiry postulated that users tend to interact with these entities as if they were humane. They responded to computers socially and naturally, a finding the scholars coined as the computers as social actors paradigm, CASA. (Nass & Moon, 2000; Nass et al., 1994)

With the introduction of the CASA paradigm, the research landscape changed around the year 2000, proposing two possible perceptions or views toward computers. While one focuses on humans' social responses to computers as mediums where the response was directed toward the human behind it, the other focuses on the computer as a source (Sundar & Chen, 2023; Sundar & Nass, 2000). The focus shifted towards the direct exploration of the non-human entity (Guzman & Lewis, 2020). The computer as a source is characterized by its agency, the use of natural language, and social roles. Altogether, these are characteristics presumably held by human beings, rendering the computer relatively autonomous and thus deserving of attributes akin to humans (Sundar & Chen, 2023). Thus, techno-centric research within the CASA paradigm in HCI has, to a great extent, concentrated on examining the influence on and/or the systematic variation of design features of user interfaces (Beattie et al., 2020; Go & Sundar, 2019; Halbauer et al., 2022). These design features can be very obvious or subtle and encompass, for example, persona design. Persona design involves crafting the character a designer aims to convey to their users. This is achieved through communicating a specific role, such as that of a casual friend, and adjusting the conversational tone accordingly (Rhee & Choi, 2020). The design of technology has often drawn inspiration from human social roles and relationships, exemplified by classics such as J. Weizenbaum's Eliza[9] (Weizenbaum, 1966), which assumed the role of a Rogerian therapist and positioned users as clients. Embedding technology within

[9] *http://www.med-ai.com/models/eliza.html*

explicit social roles aims to offer users a framework or scripts that are triggered for interacting with it – in this case, as if Eliza were a real therapist (Suchman, 2012).

While CASA is rooted in the idea that humans often treat computers *as if* they were social entities, a new paradigm has developed that builds on the idea that Computers, in fact, *are* social actors – in other words, accepted social entities (Guzman & Lewis, 2020; Sundar, 2020; Sundar & Chen, 2023). Human-Computer Interaction (HCI) has shifted its perspective from viewing machines as social actors to recognizing smart media technologies as active entities seamlessly woven into our daily existence (see HAII-TIME model, Sundar, 2020). This transformation underscores that these machines frequently hold a wealth of information about us, often surpassing the knowledge possessed by many of our human connections. As a result, machines are no longer considered mere vessels; instead, they are acknowledged as dynamic agents involving users in the collaborative construction of meaning (Sundar & Chen, 2023).

The Rise of Human-AI Relationship Research

Arguably, many techno-centric studies fall under the greater umbrella of anthropomorphization, which is to say, design strategies that incorporate anthropomorphic (i.e., human-like) design elements (Coeckelbergh, 2021; Go & Sundar, 2019). In other words, the design strategy ought to max out the human in the machine to the fullest extent. Combining the advancements in technical performance and the humanized system design have rendered users' interactions not only social in the sense of being imbued with meaning or emotion (i.e., anthropomorphizing as a cognitive process of users, Coeckelbergh, 2021; Epley et al., 2007) but have also expanded the potential for the establishment of what might be considered "relationships" with these agents, as asserted by Pentina et al. (2023b).

However, research precisely on human-AI relationships remains in its nascent stage (Pentina et al., 2023b), marked by a diversity of approaches and methodologies, along with incongruent findings (Croes & Antheunis, 2021; Lopatovska & Williams, 2018; Pitardi & Marriott, 2021; Purington et al., 2017; Skjuve et al., 2021). Scholars have repeatedly emphasized the need for robust investigation in this emerging field (Guzman, 2019; Guzman & Lewis, 2020; Hepp, 2020). This call was responded to recently by a number of studies employing various theories from social psychology to probe into the intricacies of human-AI relationships. Table 2 below, which is to a great extent based on a systemic literature review recently conducted by Pentina et al. (2023b), provides an extended overview of psychological theories researchers have used to investigate the emerging field.

Consistent across these studies is the central theme of an anthropocentric approach, contending that users' inclination to anthropomorphize machines and attribute social characteristics to them justifies the application of social science theories to elucidate the dynamics of human-machine interaction (Gambino et al., 2020; Nass et al., 1994). Thus, various psychological theories have subtly permeated research on human-AI interaction over the last decade, arguably advancing the field (Pentina et al., 2023b). Many researchers are pursuing the idea that social processes traditionally deemed uniquely human, such as communication or trust, now encompass machines (Guzman & Lewis, 2020). This has lately extended to relationship science, formerly reserved for human-to-human social interactions, which is now evolving as human-AI relationship research. Instead of seeing technology as a passive tool, human-AI relationship researchers want to understand how AI systems can actively take part in relational dynamics. Scholarship on the human-AI relationship marks a clear departure from the prevailing perspective rooted in anthropocentric definitions of social relationships. In other words, they abandoned the notion that social relationships are exclusively human.

At its core, the anthropocentric conceptualization of social relations asks how these social relations need to be reimagined in the human-AI context (Guzman & Lewis, 2020). Its common theme is the operationalization of the human-machine relationship as an outcome variable predicted by variables commonly studied in Human-Machine Interaction (HMI) field. Empirical studies included antecedents such as perceived warmth or competences (Pitardi & Marriott, 2021) and mediators such as trust or social presence (Cicco et al., 2020). Other less investigated moderators include individual-related variables such as cultural background (Pentina et al., 2023b). In applied business settings, outcome variables to assess the consumer experience are commonly studied (e.g., satisfaction, continuance intentions; cf. Araujo, 2018; Cicco et al., 2020; Pentina et al., 2023b).

While a paucity of studies has questioned the development of a human-AI relationship (Croes & Antheunis, 2021), the prevailing assumption widely accepts its existence without further ado and examines the concept through a variety of psychological theories, see Table 2 below. For instance, Skjuve et al. (2021) found a formation of relationships over time, contrasting with the conclusions drawn by Croes and Antheunis (2021), who observed a decline in all indicators of close relationships over time. This variance may emanate from differing longitudinal study designs, encompassing both qualitative and quantitative methods and, notably, different chatbot specimens under examination. This raises a pertinent query about the practical utility of a theory, particularly in light of its sensitivity to the specific AI system in consideration. The table below provides an insight into the diversity of studies in the field.

Table 2

Interpersonal Relationship Theories and Related Concepts and Application in HCI

Theory	Author(s), Year	Core Thesis in Human Interactions	Application in Human-Machine Interaction
Social Exchange Theory	Homans, 1958	Social interactions are based on a cost-benefit analysis	How do the benefits of using AI outweigh any costs or efforts involved? e.g., Croes & Antheunis, 2021
Attachment Theory	Bowlby, 1969	Humans (infants-caregivers) form emotional bonds influencing an individual's socio-emotional development	Do users form emotional bonds with their AI system? e.g., Skjuve et al., 2022
Social Penetration Theory	Altman & Taylor, 1987	Relationships develop and deepen over time by increasing intimacy and self-disclosure	How does the human-AI relationship progress over time? e.g., Li & Rau, 2019
Knapp's Staircase Model	Knapp, 1987	Relationships come together and apart in different stages	What are the stages of human-AI relationship development? e.g., Seymour & van Kleek, 2021
Love Theories	e.g., Sternberg, 1988	How to understand love, mostly romantic relationships, between humans	How do people fall in love with AI? e.g., Hernandez-Ortega & Ferreira, 2021
Social Support Theory	e.g., Rook & Dooley, 1985	Interpersonal relationships are essential for an individual's well-being, coping abilities, and health	Are human-AI relationships helpful for humans' well-being and health? e.g., de Gennaro et al., 2020
Uncertainty Reduction Theory	Berger & Calabrese, 1970	In initial and ongoing interactions between humans, people communicate to reduce uncertainty and strengthen their bonds	What conversational designs reduce uncertainty? e.g., Straten et al., 2020
Trust Theories	e.g., Mayer et al., 1995	Trust as an attitude depends on different factors (e.g., integrity) and influences behavior (e.g., reliance on others)	How does trust influence reliance on AI? e.g., Lee & See, 2010
Psychological Distance	Trope & Liberman, 2010	People process information based on perceptions of distance, impacting their attitudes and behaviors.	What is the relationship between user and AI regarding distance between self and AI? E.g., Li & Sung, 2021
Inclusion Of Self in Others	Aron et al., 1992	Investigates the degree of closeness and mutual understanding in interpersonal relationships.	How much are Ais included in users' self-concept? e.g., Tschopp et al., 2023
Similarity-Attraction Theory	Berscheid & Hatfield, 1969	Similarity in attitudes, values, beliefs, interests, and other characteristics leads to greater interpersonal attraction	Can perceived similarity improve human-AI interaction outcomes? e.g., Croes et al., 2022
Social Identity Theory	Tajfel & Turner, 2001	Group memberships and social identities influence behavior, attitudes, and perceptions in social contexts.	Does inter and intra-group categorization of AI impacts attitudes and behavior? e.g., Li et al., 2022
Power	e.g., Cialdini, 1984	Investigating power, influence, and authority in social relationships and interactions	What is the role of power in human-AI relationships? e.g., Hu et al., 2022
Physical proximity (incl. mere exposure effect)	e.g., Festinger et al., 1950	Increased physical proximity can lead to greater opportunities for interaction, familiarity, and the formation of social bonds	What is the role of distance for relationship development (for embodied AI or robots) e.g., Mumm & Mutlu, 2011

Note. The research questions in column 4 were selected by the author and do not reflect said studies comprehensively.

Taken together, the aforementioned endeavors culminate in the establishment of a novel, independent research trajectory focused on unraveling the dynamics of how and why individuals affiliate with AI agents, with many issues raised calling for further research regarding conceptual understanding of human-AI relationships, methods of investigation, and impact on behavior in the short and long term.

Limitations of Current Research and Rationale for a Novel Perspective

In the aforementioned literature review (Pentina et al., 2023b), the authors concluded that the theories used certainly offer valuable insights. However, they also highlighted limitations that are pertinent to establishing my rationale for pursuing an alternative approach. Overall, current approaches seem to be unable to fully explain human-AI relationships or their development and the underlying mechanisms. Also, the authors conclude that some theories (e.g., the social penetration theory, see Table 2) are not universally applicable across AI systems. This raises a pertinent query about the practical utility of a theory, particularly in light of its sensitivity to the specific AI system in consideration. On top of that, technological advancements presumably render findings quickly outdated.

The discrepancy between what an AI system was designed to be and how the users actually perceive it is another challenging factor in prior research. Consider an AI system marketed as a friend (e.g., Replika AI) versus an AI system presented as an assistant (e.g., the Google Assistant) versus, for instance, Alexa (Amazon's AI), which role is somewhat ambiguous. Some studies (e.g., Sundar et al., 2017) investigated the human-AI relationship from the role category perspective depending on what the system was designed for rather than from the perspective of the user – how the users see it (e.g., conversational AI in the role of the servant, or the friend, and so forth, Schweitzer et al., 2019). However, equal to the limitations mentioned above in interpersonal relationship theory, the term "friend" encompasses various meanings and falls short of a comprehensive explanation of human relationships. Think of a scenario with five friends and how your interactions with each may vary based on context. This highlights our dependence on developers' interpretation of "friend" and the expansive array of interpretations of friendship without even delving into cultural disparities. Thus, the value of explaining human-AI relationships via social roles and the predictive value thereof is exceedingly limited.

Finally, while scholars like Pentina et al. view the human-AI relationship as an outcome (2023), we advocate for an alternative paradigm. In alignment with Fiske's argument (1992) that "relational structures are not an outcome of the situational conditions or experience, […] but rather 'products of the mind'", I argue that human-AI relationships are better understood and investigated as relationship cognitive schemas. Along the lines of Baldwin (1987), these schemas contain "a self schema for how the self is experienced in the interaction, a person schema for the partner in interaction, and a script for the expected interaction pattern" (Berscheid, p. 87, 1994). More concretely, this approach allows us to differentiate how people organize their social interactions. In other words, again, consider two friends of yours, but with one, you share everything without asking; with the other, you might expect something in return. Put differently, these represent opposing social interaction outcomes, even with a similar form of affiliation – an issue Aron et al. (1991) already raised with various theories of social relations.

In a nutshell, the primary limitations of current research on human-AI relationships include the inability to apply/reproduce methods across diverse systems and contexts, the ambiguity in social roles, and the unidimensionality of investigated (often proxy) variables. Against this background – having revisited the relevant essentials of interpersonal relationship theory and psychological theories used in human-machine relations thus far – I can commence the initial study. This foundational inquiry is essential for setting the stage for future research endeavors. In the following first study, together with M. Gieselmann and K. Sassenberg, I aim to investigate the issue of how people organize their social relations with conversational AI. We seek to address limitations observed in previous research by taking a multidimensional approach, independent of the social role of the system, arguably applicable in diverse contexts and across systems. By repurposing RMT by A.P. Fiske (1992), the first exploratory study aims to answer the question: How do humans perceive their relationship with conversational AI?

Before delving into the empirical studies, I would like to provide a roadmap for this dissertation. As the empirical studies (Chapters 2, 4, and 6) build upon each other, there are transitional chapters (Chapters 3 and 5) between each study. Chapter 2 explored the applicability of RMT in the context of human-AI interaction (Study 1 & 2). Chapter 3 (transitional) provides the rationale for the subsequent research questions. Chapter 4 tested the relational approach in the context of voice commerce, investigating the influence on voice shopping decisions (Study 3). Chapter 5 (transitional) explains the rationale for the subsequent experiment. Finally, Chapter 6 tested the role of conversational design and relationship perceptions in voice shopping (Study 4).

Declaration on the Proportion of Collaborative Publications for Chapter 2 (Tschopp, Gieselmann & Sassenberg, 2023)

Author	Author position	Scientific ideas %	Data generation %	Analysis & interpretation	Paper writing
Marisa Tschopp	1	60	100	45	60
Miriam Gieselmann	2	0	0	10	10
Kai Sassenberg	3	40	0	45	30
Title of Paper	Servant by Default? How Humans Perceive their Relationship with Conversational AI				
Status in publication process	Published in the Journal Cyberpsychology: Journal of Psychosocial Research on Cyberspace. This chapter presents the published final version.				

Chapter 2: Servant by Default?
How Humans Perceive their Relationship with Conversational AI

"Alexa, will you marry me?" 6,000 times per day users propose to Alexa and 19,000 times per day, users in India say "I love you" to Alexa, Amazon's conversational AI (Amazon, 2021)[10]. Furthermore, 14% of male Alexa users in the UK desire a sexual relationship with "her" (The Guardian, 2020), and 3.5 million times Germans said "I love you" in the first half of 2021 (Buschke, 2021). But just because users say they love Alexa does not necessarily mean they are in love with "her" as humans are in love with other humans. So, if people do not really love Alexa but also do not see Alexa as "just" a tool, how do they actually perceive Alexa in relation to themselves?

The aforementioned numbers are congruent with the ample corpus of research informed by the computers as social actors (CASA; Nass & Moon, 2000; Nass et al., 1994) paradigm, suggesting that humans apply social rules from human interaction when they communicate with machines (Gambino et al., 2020). There is still a tension between the fact

[10] *With conversational artificial intelligence (conversational AI) we refer to AI systems users can talk to with voice. They are also often called smart personal assistants, virtual/digital assistants, or voice assistants (Kulkarni et al., 2019).*

that "Alexa is just a tool" and the anecdotal evidence above suggesting that humans might as well build more equal or even peer-like relationships with their conversational AI, which is why relationship research around conversational AI has been repeatedly called for (Guzman & Lewis, 2020; Seymore & van Kleek, 2021; Kim et al., 2019; Hepp, 2020).

Whereas some AI systems are intentionally designed to fulfill a single role, e.g., a friend or lover, see for instance, the chatbot Replika (Skjuve et al., 2021), it is not entirely clear what off-the-shelf conversational AI, such as Alexa, are designed for (Purington et al., 2017). A friend? An assistant? Or both?

Thus, the present study aims to address this gap by investigating how humans perceive their relationship to conversational AI. To this end, we apply the Relational Models Theory (RMT, Fiske, 1992) to human-AI relationships. Fiske suggested a comprehensive framework of relationships between humans consisting of four different modes of relationships that could also inform our understanding of how people perceive their relationship with conversational AI.

By using Fiske's framework, we aim to take a more differentiated look at the perceived relationships going beyond simple dichotomous or two-dimensional approaches commonly used in earlier research, for instance, comparing servant (transactional) and friend (relational) roles (Kim et al., 2019). A better understanding of the relationship humans perceive with their conversational AI will inform not only the interface design of these devices but also open avenues for new research questions: What relational elements are more relevant, for instance, when turning on the lights versus for voice shopping or when sharing personal information with one's device? Being the first to apply RMT to human-AI interaction, we aim to contribute to research and practice by using a well-established, quantitative framework that mirrors increasing reciprocal dynamics rather than stagnant role ascriptions. Overall, we aim to provide further empirical evidence for the overarching question: In what respects are digital assistants more than just tools?

State of the Art of Human-AI Relationship Research

Unlike human relationships, the relation of a human to a machine is unidirectional because a device is not a living being. However, the growing capabilities in natural language understanding and processing (Guzman, 2019, Panetta, 2021; van Berkel et al., 2021) could obscure unidirectionality and affect how people perceive themselves in relation to the machine. What is currently missing is a theoretical and quantitative measurement approach that mirrors the increasing reciprocal character and offers a variety of roles users ascribe to their conversational AI.

The first approaches to studying the social dynamics in human-AI interaction drew on findings from neighboring contexts, such as social robotics (van Berkel et al., 2021; Sundar, 2020). Some studies relied on the theories of role categorizations (e.g., robots as companions vs. assistants as in Sundar et al., 2017; based on Dautenhahn, 2007; or the social friend role as in Rhee & Choi, 2020) to investigate role ascriptions to conversational AI. In one study, Kim et al. (2019) found that the relationship type “friend” led to higher perceptions of warmth (fully mediated by anthropomorphism) compared to a “servant” relationship. Other researchers grounded their research in the social presence theory (i.e., the idea of feeling a deeper connection to a computer system, e.g., McLean & Osei-Frimpong, 2019): For example, Ki et al. (2020) focused on the role of para-friendships (i.e., humans perceiving emotional friend-like relationships, for instance, with fictional characters, or online influencers). They found that the perception of intimacy, understanding, enjoyability, and involvement affected para-friendships and fostered users’ intention to use a digital assistant.

Building upon the CASA paradigm and humans’ tendency to anthropomorphize (Epley, 2007; Waytz et al., 2010), a lot of studies in the broader field of human-AI interaction focused on the human-like perception of AI systems (assessing constructs such as perceived warmth, competence, or anthropomorphism, e.g., Purington et al., 2017, Gilad et al., 2022; Bergmann et al., 2012; Gong, 2008). Therefore, it is plausible to refer to social relations between humans to understand human-machine relations. Nonetheless, this approach has barely been taken.

For instance, in a survey study with Alexa users, Seymore and van Kleek (2021) used Knapp’s (1978) staircase model. They found that the human-AI relationship development follows similar patterns as the relationship development among humans and that more developed relationships co-occurred with more anthropomorphizing of and more trust in the conversational AI. In addition, two qualitative studies on the companion chatbot Replika relying on different attachment theories (e.g., social penetration theory by Altman & Taylor, 1987, was used by Skjuve et al., 2021) conclude that theories about interpersonal relationships can be used to gain further insights into the development of human-AI relationships (see also Xie & Pentina, 2022).

The studies mentioned above indicate that users are developing relationships with AI systems similar to those with humans. They offer insights informative for specific goals; for instance, design recommendations to increase adoption. Furthermore, these qualitative studies offer rich insights into the drivers of human-computer relationships (Skjuve et al., 2021; Xie & Pentina, 2022). On the other hand, recent studies by Lopatovska and Williams (2018) and Croes and Anthenuis (2020) found no clear evidence for humans building emotional relationships with conversational AI. For an overview of recent studies, see Table 3.

We argue that one problem is that many studies are, apart from the qualitative studies, arguably insufficient to comprehensively examine human-AI relationships, especially for off-the-shelf conversational AI, which serve multiple purposes. Ki et al. (2021) considered exclusively (para-)friendship, Seymore and van Kleek (2021) focused on relationship development, and Kim et al. (2019) focused on a dichotomous distinction of the friend or servant as relationship types. Exchange-oriented relationships have, for instance, not been considered so far. Along these lines, qualitative evidence from a study on conversational AI users indicated that participants build a diversity of relationships with their devices, suggesting that a more systematic, quantitative study of human-AI relationship perception is necessary (Schweitzer et al., 2019).

Therefore, we aim to build upon these findings by applying a comprehensive relationship model from human-to-human relations to the human-AI relationship that goes beyond one or only two relational types studied in earlier work by considering multiple independent relational dimensions. To the best of our knowledge, no studies take such a multidimensional approach to the perceived relationship between humans and commercially available conversational AI. Building upon the strongly established RMT (Haslam & Fiske, 1999) seems a promising direction to understand user perceptions of human-AI relationships as a whole, also compared to other approaches capturing single relational proxies, like trust or perceived warmth.

Table 3

Overview of Related Research on Human-AI Relationships

Theoretical Basis	Author(s), Year	Methods	Key Findings
Role Categorizations (Dautenhahn, 2007)	1) Sundar et al., 2017 2) Kim et al., 2019 3) Hu et al., 2022 4) Rhee & Choi, 2020	Quantitative	1) Congruity of role and demeanor matters 2) Friend role positively related to warmth perception 3) Power over AI (servant role of conversational AI) reduces risk perception in voice shopping 4) Friend-role of conversational AI influences positive attitude toward a product
Social Presence Theory (Short Et Al., 1976), Para-Social Relationships (Horton & Wohl, 1950)	Ki et al., 2019; McLean & Osei-Frimpong, 2019	Quantitative	Para-friendships / social presence foster usage intention
Personification, Anthropomorphism, CASA Paradigm, Relationship Indicators (Epley Et Al., 2010; Nass & Moon, 2000)	1) Lopatovska & Williams, 2018 2) Schweitzer et al., 2019 3) Purington et al., 2017 4) Gong, 2008 5) Gilad et al., 2021 6) Bergmann et al., 2012	1),2) Qualitative 3), 4), 5), 6) Quantitative	1) Relationship indicators are rather mindless 2) Friend, servant, or master relationship types may develop 3) Personification linked to more social interaction 4) More anthropomorphic agent received more social responses 5) Users prefer high-warmth systems 6) Warmth/competence perception depends on time, agent appearance/behavior
Knapp's Staircase Model (Knapp, 1987)	Seymore & van Kleek, 2021	Quantitative	Users show attachment to their conversational AI
Social Penetration Theory (Altman & Taylor, 1987)	1) Croes & Anthenuis., 2020 2) Xu & Li, 2022 3) Skjuve et al., 2021	1),2) Quantitative 3) Qualitative	1) No indicators of a progressing human-AI relationship similar to humans 2) Functional and relational use influence each other 3) Indicators of a progressing human-AI relationship similar to humans
Bowlby's Attachment Theory (Bowlby, 1969)	1) Xie & Pentina, 2022 2) Pentina et al., 2023a,b	1) Qualitative 2) Mixed Methods	1) Users developed emotional bonds with their companion 2) Identification of various antecedents plus moderators that influence attachment to conversational AI

Relationships between Humans according to A.P. Fiske (1992)

Fiske (1992) proposed and received mighty empirical support (Haslam & Fiske, 1996) for the RMT suggesting four modes or dimensions along which humans perceive their relationships to other humans: authority ranking, market pricing, communal sharing, and equality matching.

Authority ranking denotes an asymmetric relationship where people are not equivalent but ordered along some hierarchical dimension, where the highest rank is entitled to command over and protect the lower ranks (e.g., commander and soldier, parent and child). *Market pricing* is all about rational cost-benefit analysis (i.e., a relationship mode focusing on exchange), where humans seek something in return for their investment in the relationship, for instance, money (e.g., people in working groups). *Communal sharing* is high in tribes with kinship-like relations, where people are equivalent and commonalities are emphasized (e.g., often in families or sports teams). *Equality matching* describes a peer-like tit-for-tat relationship, focusing on an egalitarian balance within the relationship (e.g., roommates sharing an apartment). According to RMT, the four modes are not distinct types but can instead operate at the same time – they are dimensions along which one relationship with a specific person can be described. A brief overview of the human relationship modes is depicted in Table 4.

Table 4

Brief Description of the Four Modes of Relationships according to A.P. Fiske (1992)

Authority Ranking	**Market Pricing**	**Communal Sharing**	**Equality Matching**
What is the order between us?	What are the ratios?	What do we have in common?	What is the balance?
Hierarchy Command Dominance	Cost-benefit Input-Output Reciprocity	Solidarity In-group Communality	Turn-taking Democratic voting Reciprocity

Applying Relational Models Theory (RMT) to Human-AI relationships

How could the human relationship modes apply to human-AI relationships? The presentation of conversational AI as digital assistants is in line with a strong emphasis on the *authority ranking* mode. This would be a rational perception of the human-AI relationship: the user dominates the conversational AI and gives commands. Therefore, authority ranking might be the dominant relational mode. In other words, the one with the highest mean values.

The actual and perceived growth of machine agency (i.e., machines and algorithms are becoming more agent-like, having more control, and exerting greater influence on people's perceptions and behavior; Sundar, 2020) provides the basis for perceived reciprocity. Accordingly, the relationship to conversational AI should, in part, be seen as characterized by market pricing. However, nowadays, conversational AI does not come close enough to human characteristics that users will likely experience solidarity and community – attributes that allow for communal sharing or actual perceived equality – the precondition for equality matching. Therefore, we predict:

> **H1**: Authority ranking and market pricing have significantly higher mean values than equality matching and communal sharing.[11]

Various streams of human-machine interaction studies follow the dichotomous distinction between rational and emotional dimensions of how humans perceive and interact with machines (for a detailed discussion, see Glikson & Woolley, 2020). We follow this sound distinction when deriving predictions for concepts correlating with the dimensions of the human-AI relationships derived from RMT.

Rational Modes of Perceived Human-AI Relationships

The extent to which people perceive their relationship with a conversational AI to be characterized by authority ranking and market pricing should be related to rational concepts of system perception, like the perceived competence of a system. Perceived competence is a prerequisite for the assistant role as well as for perceiving the conversational AI as an equal counterpart as implied by marketing pricing. Conversely, both relationship dimensions should be negatively related to competence concerns. For authority ranking, this has been shown by Hu et al. (2022). They found that more perceived power over Alexa (Amazon's conversational AI) led to reduced risk perception – with regard to shopping at Amazon (provided that the people had a desire for power). Together, this led to the following prediction:

> **H2**: Authority ranking and market pricing are related to rational dimensions of system perception - positively to perceived competence and negatively to perceived competence concerns.

[11] *Of note, we derived the predictions after the data was collected. In this sense, the current study should be considered exploratory.*

Emotional Modes of Perceived Human-AI Relationships

As the examples in the introduction indicate, people might form emotional relations to conversational AI of varying depth. In human relationships, trust is one of the key ingredients in communal sharing. To a lesser extent, also in equality matching (Fiske, 1992), and indeed users who categorize their AI systems as friends reported a higher level of trust, perceived warmth, and lower social distance (Pitardi & Marriot, 2021; Bergmann et al., 2012; Gilad et al., 2021). They would likely perceive their relationship as more emotional; hence, the communal relationship mode would likely occur in human-AI interaction.

People who experience an emotional relationship with a machine assume that it is characterized by similar feelings and behaviors to human friends (e.g., Han & Yang, 2018) and attribute constructs such as low psychological distance (Pitardi & Marriot, 2020) to their devices. Seymor and van Kleek (2021) found a positive correlation between trust, anthropomorphism, and the closeness of the relationship (see also Schweitzer et al., 2019). Hence, communal sharing and equality matching are likely related to more anthropomorphizing of the conversational AI, a higher perceived warmth, lower perceived social distance, and higher trust. These considerations led us to the following prediction:

> **H3**: Equality matching and communal sharing are positively related to emotional dimensions of system perception: perceived warmth, inclusion of AI in self, psychological distance, anthropomorphism, and trust.

The Current Research

To enhance the existing literature, which largely relies on one or two-dimensional frameworks to study perceptions of relationships, we repurposed the empirically established multidimensional relational models framework (RMT) by Fiske (1992) to conversational AI in two studies (N_1=367 and N_2=362). We aim to test the three hypotheses derived above in Study 1 by collecting variables of system perception: system perception comprised variables of a rather rational nature, namely, perceived competence, competence concerns, and privacy concerns. Furthermore, variables of a rather emotional nature, namely, perceived warmth, inclusion of self in AI, psychological distance, and anthropomorphism, were collected. We measured trust under system perception, with the peculiar feature that human trust in AI shares both emotional as well as rational properties (see Glikson & Woolley, 2020; Jian et al., 2010). For additional analyses, we collected user characteristics variables, namely, affinity to technology, frequency of use, years of experience with the conversational AI, different purposes of usage, as well as the number of purposes, which we derived from the sum of provided answers of participants for purposes of use (e.g., navigation, shopping, or controlling the smart home).

In Study 2, we aimed to replicate the dimensional structure and the prevalence of the relationship modes (H1). In addition, we sought to test the stability of the relationship modes and their prevalence across a number of user characteristics. To this end, we assessed a variety of demographics (as suggested by McLean & Osei-Frimpong, 2019 and Purington et al., 2017) (i.e., household specifications, educational level, employment, age, gender of the user, and technological knowledge) and system variables (i.e., device specifications, set gender of the technology). We tested the preregistered hypothesis (https://aspredicted.org/qa5ty.pdf) that naming the device "voice assistant" (compared to "conversational AI") will lead to higher perceived authority ranking and lower perceived peer-bonding as well as market pricing (H4). This is because using the word assistant in the label will stress the hierarchical relationship captured by authority ranking[12]. Not finding such an effect and no relevant correlations with the demographics and system characteristics will be a sign that the relationship perceptions are stable and largely independent of personal and situational factors.

Study 1

Methodology

Participants

We conducted a correlational online questionnaire study via Prolific in June 2021. In a screening study (N_0=1050), we surveyed participants for regular use of conversational AI, such as Alexa, Siri, or Google Assistant. Our target sample size was 400. To ensure stable estimates of correlations, Schönbrodt and Perugini (2013) recommend a sample size of around 250. To account for potential exclusions, we added 150 observations to our sample. Of those who reported using a conversational AI regularly, we invited N_1=406 participants to participate in study 1 in exchange for £1.25. We excluded 39 participants from the analyses reported below who failed at least one of two attention checks. Participants whose duration of taking the survey was too short for reading the items properly (<150 seconds) or excessively long (>80,000 seconds) were also excluded. The three hundred sixty-seven (N = 367) remaining respondents (67% female, 33% male; age range: 18 - >60 years; 68% between 25 and 49 years) were mostly from the UK (89%) and filled out the questionnaire on a mobile device (44%), desktop

[12] *We thank the anonymous reviewer for suggesting this prediction.*

computer (52%), or tablet (4%). The majority possessed multiple conversational AI: Alexa (43%), Siri (25%), or Google Assistant (24 %), and less than 2% used Cortana, Bixby, or other conversational AI (multiple responses were possible).

Procedure

We invited participants to participate in a study on users' perceptions of technology. After providing consent, participants first answered questions about their usage of conversational AI. Then, the perception of their human-AI relationship (adapted from MORQ, Haslam & Fiske, 1999). The instructions for the original measure require people to focus on a counterpart to describe their relationship, referring *only* to that counterpart. In line with this approach, we asked our respondents to focus on one specific conversational AI while answering the questions regarding the human-AI relationship. Most of the respondents chose Alexa (63.5%), the Google Assistant (23.2%), and Siri (12%; 1.3% other). Afterward, we collected the other scales to measure system perception in a fixed order (psychological distance, social perception, inclusion of others, trust, and anthropomorphism), items of each scale were presented in randomized order. Then, we collected data on user characteristics, including affinity to technology, frequency, experience, and purposes of use. Finally, we collected barriers to usage, particularly privacy and competence concerns. We placed demographics at the end, followed by a final opportunity to withdraw the data before submission.

To check for common method bias, we looked at Harman`s Single factor test to check the extent of common method bias, which indicated a shared variance of 11.9% for all measures reported in Table 5; similarly, the latent common method factor approach indicated a shared variance of 15.5% for all measures including more than two items, indicating an acceptable level of shared variance across measures.

Measures

We assessed the *human-AI relationship* with a version of the MORQ (Modes of Relations Questionnaire, Haslam & Fiske, 1999) adapted to the context of conversational AI. As this questionnaire was repurposed for the first time, we first carefully selected items we deemed appropriate for the context of human-AI interaction. We aimed to stick as closely as possible to the original questionnaire but had to omit items that we considered not applicable to conversational AI, for instance, because they refer to physical objects such as "You typically divide things up into shares that are the same size". We pretested the remaining items with fifteen people with expertise in IT, resulting in adaptions in the wording to render it suitable for human-AI relationships. The final questionnaire consisted of 17 items (see Table 5) using a 7-point Likert scale (1=*not at all true for this relationship*, 7=*very true for this relationship*): Five items for communal sharing, four items for equality matching, four items for authority ranking, and four items for market pricing. Given that this instrument was used for the first time in such a context, we report detailed analyses of the dimensional structure and internal consistencies in the Results section.

Trust was measured through the Trust in Automation Scale (Jian et al., 2010). We slightly adapted the wording (conversational AI instead of "the automation") of the 13 items (*Cronbach's Alpha*=.78), where respondents had to indicate how much they agree to a statement on a 5-point Likert scale (1=*strongly disagree* to 5=*strongly agree*). For example, "The cAI is dependable" (in the beginning, we explained that conversational AI is abbreviated cAI). Here and for all other scales, an index was formed by averaging the responses after recoding reversed items.

Social perception (perceived warmth and competence) was measured with six items on a 7-point Likert scale (1=*strongly disagree* to 7=*strongly agree*) slightly adapted from Pitardi and Marriot (2020), where three items measured the perceived warmth dimension (e.g., "I think the cAI has good intentions") and the other three items measured the competence dimension (e.g., "I think the cAI is effective"). The factor analysis across these six items revealed a two-factor structure, with two items assigned to the perceived warmth dimension and four items assigned to the competence dimension. The final scales presented decent reliabilities with *Cronbach's Alpha*=.81 for the competence dimension and r=.48 (N=367, p<.001) for the perceived warmth dimension.

Inclusion of AI in the self was measured with two items on a 7-point Likert scale (r=.61, N=367, p<.001). The items are two pictorial measures of categorization, which were adapted from Schubert and Otten (2002; Aron, Aron & Smollan, 1992). The first item consisted of seven pictures of two equivalent circles on a straight line. As in the original, the circles moved closer together from the top to the bottom, overlapping completely in the last bottom picture. The instructions indicated that one circle represented the "self" or humans in general (differing between the two items) and the other the "conversational AI". The higher the score, the higher the overlap, and the closer the respondent felt to the conversational AI.

Psychological distance was measured with two items (r=.77, N=367, p<.001) on a 7-point Likert scale (1=*strongly disagree* to 7=*strongly agree*). We adapted the items slightly, for instance, "the cAI is psychologically close to me", and dropped the item measuring familiarity from the original (Li & Sung, 2021).

Anthropomorphism was measured through the scale by Waytz et al. (2010). Participants had to indicate their agreement with seven items (e.g., “To what extent does the cAI have thoughts of its own?”, *Cronbach’s Alpha*=.90) on a 7-point Likert scale (1=*strongly disagree* to 7=*strongly agree*).

Affinity to technology was measured with nine items (e.g., “I like to occupy myself in greater detail with technical systems”, *Cronbach’s Alpha*=.88) on a 7-point Likert scale (1=*strongly disagree* to 7=*strongly agree*) developed by Franke et al. (2019).

Frequency of use (“How often do you use conversational AI?”) was measured on a single-item 5-point scale ranging from 1=*less than once a month* to 5=*several times a day* (adapted from Funk et al., 2021).

Experience of use (“Since when do you use conversational AI?”), was measured on a single-item 5-point scale ranging from 1=*less than 12 months* to 5=5 *years or more*. This item was adapted from Funk et al. (2021).

The *purpose of use* was assessed to investigate for what purposes they use a conversational AI. Six answer options were given (multiple responses possible) e.g., option one was “To retrieve information, e.g., How is the weather tomorrow?” An open text option was given to collect purposes not captured within the preselected options. We calculated the *number of purposes* from the sum of retrieved answers from the participants (inspired by Funk et al., 2021).

Finally, we collected barriers of usage to address concerns users may have, where participants had to rate their agreement on a 5-point Likert scale (1=*strongly disagree* to 5=*strongly agree*). We conducted a factor analysis across these seven items, which revealed a two-factor structure. We distinguished privacy concerns (e.g., “I am concerned about privacy”, r=.90, N=367, p<.001, 2 items) and competence concerns (e.g., “The conversational AI cannot do what I expected it to do”, *Cronbach’s Alpha*=.75, 5 items). Items were adapted from Funk et al. (2021).

Analyses have been conducted using SPSS 25.0 unless reported otherwise. Higher values indicate a stronger manifestation of the respective construct for all scales. Code, data, codebook, and calculations in Excel for both studies are available here: https://researchbox.org/636

Results

Factor Structure of Human-AI Relationship Questionnaire

A principal component analysis (PCA) with orthogonal rotation (varimax) was conducted on the 17 items of the human-AI relationship measure. The Kaiser criterion as well as the visual inspection of the screeplot, indicated that a three-component solution is adequate and explained 53.8% of the variance. Table 5 shows the factor loadings after rotation. The first component combines the two original modes communal sharing and equality matching. It, thus, represents an emotional, peer-like character of human-AI relationships, which we will name *peer bonding* hereafter. Analogous to the original questionnaire, component 2 represents authority ranking and component 3 represents market pricing. We omitted Item 7 from the analyses because it had the highest loading on a different factor than in the original scale. No other items were omitted, even though some had meaningful secondary loadings, to keep the scales as close as possible to MORQ.

Table 5

Results from a Factor Analysis of The Human-AI Relationship Questionnaire (N_1=376)

	Item	Factor Loading		
		1	2	3
Communal Sharing				
1	There is a moral obligation to act kindly to each other	**.474**		.370
2	Decisions are made together	**.815**		
3	You tend to develop similar attitudes and behaviors	**.788**		
4	It seems you have something unique in common	**.845**		
5	The two of you belong together	**.714**		
Equality Matching				
6	Some requests are granted in anticipation of something in return	**.557**		
7	"One-Person, one vote" is the principle for making decisions	.387	**.505**	
8	You take turns doing what the other wants.	**.761**		
9	You are like peers or fellow co-partners	**.738**		
Authority Ranking				
10	One of us is entitled to more than the other		**.722**	
11	One directs the work, the other pretty much follows		**.554**	.459
12	You are like leader and follower		**.698**	
13	One is above the other in a kind of hierarchy		**.751**	
Market Pricing				
14	What you get is directly proportional to how much you give	.311		**.616**
15	You have a right to a fair rate of return for what you put into this interaction			**.741**
16	You expect the same return on your investment other people get			**.741**
17	Your interaction is a strictly rational cost-benefit analysis		.404	**.420**

Note. Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization. Rotation converged in 5 iterations. Highest factor loadings are in bold, factor loadings below .30 are not displayed. Adapted from the Modes of Relations Questionnaire by Haslam & Fiske, 1999.

Given the secondary loadings and the imperfect internal consistency for authority ranking, we conducted a confirmatory factor analysis using the R package “lavaan” (R Version 4.1.2.; Rosseel et al., 2021). Model fit statistics indicated a satisfactory fit: $\chi^2(87)=191$, $p<.001$; $CFI=0.94$, $RMSEA=0.06$, $SRMR=0.07$.

The final scales presented decent reliabilities: *Cronbach's Alpha*=.86 for peer bonding (communal sharing and equality matching merged), *Cronbach's Alpha*=.70 for authority ranking, and *Cronbach's Alpha*=.68 for market pricing. The scales could not be substantially improved by removing items with the lowest item-total correlation.

To establish convergent validity, we calculated the average variance extracted (AVE) and composite reliability scores for each factor. The average variance extracted was .48 (factor 1), .46 (factor 2), and .61 (factor 3). Values for composite reliability ranged from .73 to .89, which exceeds .70, indicating that all the items consistently measured their corresponding constructs. To assess discriminant validity, we furthermore tested the Fornell-Larcker Criterion, where all conditions were satisfied, suggesting acceptable construct validity.

Market pricing was positively correlated with peer bonding (r=.39, N=367, p<.001) and authority ranking (r=.38, N=367, p<.001). No significant correlation was found between authority ranking and peer bonding (r=-.03, N=367, p=.512).

To test Hypothesis 1, we conducted an ANOVA with repeated measures and post-hoc comparison using Bonferroni correction. As the sphericity assumption was violated, we reported results with Huyn-Feldt correction. In line with the hypothesis, participants saw their relationship with the conversational AI as more strongly characterized by authority ranking (M=4.74, SD=1.49, N=367) than by market pricing (M=4.29, SD=1.33, N=367) and peer bonding (M=2.46, SD=1.25, N=367), all ps<.001, F(1.73, 632.30)=378.63, p<.001, η_p^2=.508.

Relation between Human-AI Relationships and Variables of System Perception and User Characteristics

In what follows, we report bivariate correlations and partial correlations of each relationship mode, controlling for the respective two other modes. We will focus more on the latter to test for unique relations between the respective relationship mode and the criterion variable. In doing so, we aim to address the hypotheses that the rather rational variables of system perception are related to authority ranking and market pricing (H2), whereas the emotional variables of system perception are related to equality matching and communal sharing (peer bonding, respectively; H3). In addition, we explored the relationship between user characteristics and relationship perception. Partial correlations, as well as bivariate correlations in parentheses, are reported in Table 6.

For *authority ranking,* we found only zero-order correlations with trust, perceived competence, affinity to technology, and the sum of purposes. After controlling for the other two relationship modes, these relations turned insignificant, except for the number of purposes. The more users perceived their relationship to a conversational AI to be characterized by authority ranking, the more widely they used it. Overall, authority ranking is the dominant mode regarding the perception of conversational AI, but the extent of authority ranking does not uniquely relate to the system perception. To our surprise, even the bivariate correlations with trust and perceived competence were small in size – though significant. This is somewhat peculiar given the role perceived competence and trust seem to play in human-AI interaction, or the way conversational AI are marketed as digital assistants.

Market pricing presented positive zero-order correlations with trust, warmth, inclusion of self in AI, psychological distance, anthropomorphism, perceived competence, affinity to technology, competence concerns (negative), and the number of purposes. However, market pricing had no unique relation (partial correlation) with warmth, psychological distance, anthropomorphism, and the number of purposes, whereas it had a unique relation with trust, inclusion of AI in the self, perceived competence, affinity to technology, and competence concerns (negative). Taken together, a higher market pricing mode was uniquely related to the rational dimensions (perceived competence and competence concerns) as well as to inclusion of AI in the self (emotional system perception) and trust, sharing features of both rational and emotional dimensions. Effects were small (to medium).

Peer bonding showed strong, positive bivariate and partial correlations with perceived warmth, inclusion of self in AI, psychological distance, and anthropomorphism. The same was true for perceived competence, affinity to technology, and the number of purposes. In addition, we found small, bivariate correlations with trust (positive) and competence concerns (negative), but the effects disappeared in the partial correlation analysis. In line with the idea that users build social bonds with computers, but contradicting the rational approach, peer bonding correlates strongest with system perceptions of conversational AI and also partly with user characteristics.

In a nutshell, we found support for H3, observing a stronger relationship between peer bonding and the emotional variables of system perception (e.g., psychological distance) rather than with the rational variables (e.g., perceived competence). With regards to H2, the results differ between market pricing and authority ranking. Market pricing showed positive correlations with rational and emotional variables of system perception, whereas authority ranking showed no partial correlations, which is not in line with our theorizing. Furthermore, no meaningful correlations were found between the user characteristics and the relational modes.

Table 6

Partial (Bivariate) Correlations between Relational Modes (Controlled for the respective other two in case of the Partial Correlations) and System Perception as well as User Characteristics (N=376)

Variable	Authority Ranking	Market Pricing	Peer Bonding
System perception			
Perceived competence	.05(.13*)	.19**(.30**)	.15**(.25**)
Competence concerns	.04(-.02)	-.14**(-.17**)	-.05(-.12*)
Privacy concerns	-.03(-.04)	-.01(-.02)	.02(.02)
Perceived warmth	.07(.09)	.09(.28**)	.37**(.43**)
Inclusion of AI in self	.03(.05)	.12*(.33**)	.48**(.55**)
Psychological distance	-.01(-.02)	.04(.27**)	.59**(.64**)
Anthropomorphism	.01(-.03)	-.05(.16**)	.50**(.53**)
Trust	.04(.12*)	.19**(.26**)	.05(.13*)
User characteristics			
Affinity to technology	.07(.12*)	.10*(.20**)	.11*(.16**)
Frequency of use	-.09(-.07)	.05(-.01)	-.08(-.06)
Experience of use	-.08(-.08)	.02(-.05)	-.09(-.09)
Number of purposes	.13*(.14*)	.01(.12*)	.15**(.16**)

Note. Values in parentheses represent bivariate correlations. *p<.05. **p<.01.

Discussion

The results of study 1 showed that participants view their relationship along three dimensions, instead of the suggested four dimensions. Equality matching and communal sharing merged into one dimension we called *peer bonding*. In line with our expectations (H1), we found that most users characterize their relationship as hierarchical (i.e., authority ranking) as well as the non-hierarchical exchange-based relationship (i.e., market pricing). Only a few saw their relationship characterized as a companion-like relationship (i.e., peer bonding). Against our expectations (H2), authority ranking did not show any meaningful correlations, rendering this traditional dimension not informative. Taking equality matching and communal sharing together, rather in line with our expectations (H3), peer bonding was, to a greater extent, related to the emotional variables of system perception, such as anthropomorphism or perceived warmth, than to the rational variables of system perception. Interestingly, trust was not related to peer bonding but market pricing, suggesting that users take a more rational approach that includes elements of exchange, which is in line with the current critique on emotional accounts of trust (Ryan, 2020). In a nutshell, the relationship seems rather rational but still relational. Finally, we observed no meaningful correlations with user characteristics apart from the sum of purposes. Using conversational AI for multiple purposes (for instance, navigation, smart home, *and* voice shopping) may be a better indicator of "power users" than the simple frequency of use.

Given that this approach was used for the first time, we conducted a second study to replicate the factor structure and to test the stability of the hierarchy of the three dimensions considering demographic variables and a manipulation of the device's naming.

Study 2

Methodology

Design, Participants, and Procedure

Study 2 was an experiment with two conditions varying the naming of the technology (voice assistant vs. conversational AI). Otherwise, Study 2 closely followed the procedure of study 2 but included a different set of measures apart from the relationship modes. We sampled 400 frequent users of conversational AI via Prolific in November 2022 (based on a prescreening). Similar to study 1, our rationale for determining the sample sized is based on the arguments by Schönbrodt and Perugini (2013). Excluding participants based on our preregistered exclusion criteria (https://aspredicted.org/qa5ty.pdf) would have led to an exclusion rate above 25%. Therefore, we applied a more lenient criterion than preregistered regarding the minimum time taken to complete the study (120 rather than 105 seconds). Notably, the results were not contingent on the exclusion criterion. The remaining N=362 respondents from study 2 (44% male, 56% female, age range 19-81 years, M_{age}=39, SD_{age}=13) were mostly from the UK (90%). 20% used their conversational AI on their smartphone, 61% on their smart speaker without a screen, and 15% on a smart speaker with a screen. This sample size was in line with the desired sample size determined by an a priori power analysis (2 (naming, between-subjects factor) x 3 (relationship type, within-subjects factor) mixed ANOVA, interaction, effect size f=.10, *alpha*=.05, 1-b*eta*=.95, N=362). We aim for 95% power to be able to interpret non-significant results.

Measures

The perceived *human-AI relationship* was assessed using the scale from study 1 with minor adaptations in the wording. Detailed analyses of the dimensional structure and internal consistencies are reported in the Results section.

We collected the following variables to test their correlation to the relationship modes: *household specifications single* (0=no,1=yes), *kids* (0=no,1=yes), *educational level* (ranging from 1: less than high school to 7: doctorate), and *employment* (employed versus unemployed: 0=not employed,1=employed). In addition, we surveyed their *technological knowledge* on a 5-point Likert scale (1=*not knowledgeable at all* to 5=*extremely knowledgeable*), *device details* (without versus with screen: 0=without, 1=with), and the customized *gender of the conversational AI* (female=1; other=0).

Results

Factor Structure of Human-AI Relationship Questionnaire

The PCA with orthogonal rotation (varimax) on the 17 items of the human-AI relationship measure yielded similar results as in study 1. The Kaiser criterion as well as the visual inspection of the scree plot indicated that a three-component solution is adequate and explained 59.8% of the variance. We have omitted item 6 due to high secondary loading. Notably, including item 6 did not change results substantially.

The final scales presented decent reliabilities: *Cronbach's Alpha*=.90 for peer bonding (communal sharing and equality matching merged, 8 items), *Cronbach's Alpha*=.80 for authority ranking, and *Cronbach's Alpha*=.67 for market pricing (each 4 items). We conducted a confirmatory factor analysis using the R package "lavaan" (R Version 4.1.2.; Rosseel et al., 2021). Model fit statistics indicated a satisfactory fit: $\chi^2(87)=232$, $p<.001$; $CFI=0.93$, $RMSEA=0.07$, $SRMR=0.08$.

The average variance extracted is .54 (factor 1), .58 (factor 2), and .43 (factor 3). Values for composite reliability range from .75 to .91, which exceeds .70, indicating that all the items consistently measure their corresponding constructs. To assess discriminant validity, we furthermore tested the Fornell-Larcker Criterion, where all conditions are satisfied, suggesting an acceptable construct validity.

The correlations between the human-AI relationship are as in study 1: market pricing is positively correlated with peer bonding ($r=.41$, $N= 362$, $p<.001$) and authority ranking ($r=.42$, $N= 362$, $p<.001$). No significant correlation was found between authority ranking and peer bonding ($r=- .07$, $N=362$, $p=.156$).

Testing the Impact of the Naming and the Relation to Demographic Variables

To test whether the naming of the conversational AI influenced human-AI relationship perception (H4) and whether the hierarchy of the relationship perceptions replicated study 1 (H1), we performed a 2 (naming–between) x 3 (relationship type–within) mixed-ANOVA and simple comparison of the naming factor**.** The results with Huynh-Feldt (HF) correction showed a significant main effect of the relationship type $F(1.61,578.43)=383.3,3$ $p<.001$, $\eta_p^2=.52$). Participants saw their relationship with the conversational AI as more strongly characterized by authority ranking (M=4.86, SD=1.59) than by market pricing (M=4.35, SD=1.35) and more than by peer bonding (M=2.42, SD=1.33), all ps<.001. No naming by relationship mode interaction was found, $F(1.61,578.43)=1.31$, $p=.267$. Thus, the main effect of relationship type was not contingent on the naming.

Unexpectedly, independent of the relationship type, the naming of the AI affected the extent to which all three types were perceived, $F(1,360)=4.40$, $p=.037$, $\eta_p^2=.01$, indicating that mean values for all relationship modes were higher in the voice assistant condition.

To test whether the perceived relationship mode is related to demographic factors, we regressed each of the three scores separately on age, gender of the user, household specifications (single vs. with kids), level of education, employment (employed vs. unemployed), technological knowledge, as well as the gender and screen settings of the conversational AI. The use of a screen with the smart speaker had a negative relationship with market pricing ($\beta=-.12$, $t(349)=-2.29$, $p=.023$). Furthermore, we found a negative relationship of age with authority ranking ($\beta=-.14$, $t(349)=-2.49$, $p=.013$) as well as a positive relationship with market pricing ($\beta=.11$, $t(349)=2.01$, $p=.046$). Finally, we found a significant positive correlation of educational level with authority ranking ($\beta=.12$, $t(349)=2.18$, $p=.030$) and a negative one with peer bonding ($\beta=-.14$, $t(349)=-2.68$, $p=.008$).

Discussion

Study 2 aimed at further testing the applicability of the RMT to human-AI relationships. The results of study 2 replicate the findings of study 1 regarding the dimensional structure and the hierarchy of the dimensions. The naming did not affect the hierarchy, not supporting H4. The main effect of naming occurred less relevant to us because the effect size was rather small and the effect occurred across all three relationship dimensions. Older and more educated people showed less peer bonding and market pricing but more authority ranking. Otherwise, the relationship perceptions were unrelated to the assessed demographic factors. In a nutshell, the relationship perceptions were largely unmoderated, validating that our approach is well suited to describe the human-AI relationships across different contexts.

General Discussion

How do humans perceive their relationship to conversational AI? By applying Fiske's (1992) multidimensional Relational Models Theory (RMT), we aim to investigate human-AI relationships in relation to system perception and user characteristics.

First, the factor structure of our human-AI relationship questionnaire indicated that the relational modes Fiske (1992) suggests for human-human relationships can be applied to human-AI relationships, with one restriction: The more emotional modes of communal sharing and equality matching which characterize closer human relationships cannot be differentiated in the perception of the relationship between users and conversational AI. In both studies reported here, they converged on one factor that we called *peer bonding*. Overall, users perceived their relationships with conversational AI as predominantly characterized by authority ranking and market pricing. Peer bonding was much weaker than the two rational relationship modes mentioned before. These results indicate that users may perceive the human-AI relationship on similar dimensions as relationships to humans; however, they were less differentiated and less characterized by the emotional modes, which Fiske called communal sharing and equality matching.

These perceptions seem to be largely independent of the label used for the device and most demographic and system features, such as the gender of the conversational AI or household characteristics. This stresses that the relationship dimensions can be broadly applied and are relatively stable. At the same time, the only predictors at the user and system level that we could identify are age and education. Future research should explore psychological concepts such as personality, loneliness, etc., to identify individual differences in predicting human-AI relationship perception.

Furthermore, based on the broad availability of large language models such as ChatGPT[13] by OpenAI, conversational agents might become more performant, and conversational AI will likely appear more human-like. This opens the door to stronger human-AI relationships. Research monitoring these developments should also consider the multidimensional structure of perceived human-AI relationships.

As previously explained, many earlier studies in the broader field of human-machine interaction assumed the rational versus emotional dichotomy, for instance, warmth versus competence or companion versus assistant (Glikson & Woolley, 2020; Malle & Ullman, 2021).

[13] *https://openai.com/*

This matches the distinction between peer bonding and authority ranking. However, we found that market pricing, which denotes a rational and equal (other than hierarchical) dimension, enriches the current understanding, dominated by single or two-dimensional approaches, by adding another dimension.

This may be an interesting finding for developers of voice user interfaces. Although authority ranking and market pricing dimensions seem to have much in common at first (for instance, characterized as rational), they differ in one design-relevant important aspect: namely, the agency or responsibility people attribute to the conversational AI system. This likely impacts what tasks users use the conversational AI for. Does a request necessitate less machine intelligence, such as turning on a light, or does a request necessitate more machine intelligence, for instance, when shopping for goods with a conversational AI? We can imagine that for more sophisticated interactions, such as shopping, which involves multiturn dialogues, designers need to understand how they can trigger users to assume some reciprocal responsibility in interacting with the conversational AI.

We will elaborate on three key findings from the correlational analyses: (1) Low predictive value of authority ranking despite its high prevalence, (2) the predictive value of market pricing for trust, inclusion of AI in self, competence concerns and affinity to technology, and (3), highest predictive value of peer bonding for system perception (except trust), despite its low prevalence.

First, the highest mean values for authority ranking imply that a majority of users perceived their relationship as rational, task-oriented, and hierarchical. Given that the relation to conversational AI was dominantly perceived as authority ranking, we were surprised to see no unique correlations with any system perception variables apart from the number of purposes. This absence renders authority ranking a not very informative dimension regarding system perception and user characteristics, which might be partly due to its high prevalence.

Second, market pricing is both rational and task-oriented but not as hierarchically perceived as authority ranking. According to RMT, individuals who perceive a relationship to be characterized by market pricing perform a cost-benefit analysis on what they invest (time, money, engagement, etc.) and what they get out of the relationship – here, by using the smart device. Given the rational or calculative characteristics, we were surprised to see correlations of market pricing with both rational variables, such as competence concerns, and variables we initially classified as emotional or have emotional components, for instance, inclusion of AI in the self. However, in line with current criticism against affective accounts of trust in AI (i.e., the technology is reliable in the sense that it works well; see, for instance, Ryan, 2020), it is plausible that trust and inclusion of AI in self can be considered more rational than emotional. More specifically, inclusion of conversational AI in self could also be understood as including the conversational AI as a part of users' daily routines. Overall, market pricing seems to be a more informative dimension than authority ranking, especially regarding the rational accounts of system perception. The findings so far further suggest that users form a relationship with conversational AI that is partly hierarchical but also partly equal.

Finally, despite the fact that peer bonding had the lowest mean, it was the strongest predictor of psychological distance, anthropomorphism, and perceived warmth - all variables that are of critical importance in studies of human-machine interaction, as laid down in the theoretical background. Thus, peer bonding seems to be the most informative dimension regarding system perception, especially emotional accounts thereof. Notably, there may be differences in the perception of this particular dimension between different off-the-shelf conversational AI. For instance, Kuzminykh et al. (2020) have shown in a qualitative study that Alexa is perceived differently – more friendly and warm - than Siri or the Google Assistant. It might be possible that people differ in their perception due to the different naming or gender settings (Abercombie et al., 2021). Although we could not identify such influence in study 2, we recommend considering this in future research.

Strengths and Limitations for Research and Practice

Unlike other studies, we did not focus on one specific role, or the role intended by the designer (e.g., companion bots or (para-)friendships). Instead, we use a comprehensive model of human-to-human relationships applied to off-the-shelf conversational AI. These are arguably designed with multiple, ambiguous roles and can perform an increasing variety of tasks with increasing social dynamics, such as multiturn dialogues in voice shopping. Thus, even merely transactional interactions, like turning on the lights, are to some degree relational as we describe it with authority ranking.

Our study contributes in at least three ways: First, we discuss theories about relational approaches to studying human-AI interaction, expanding and integrating knowledge. Specifically, we are expanding the current literature by enriching the traditional rational-emotional or rational-transactional approaches (e.g., Xu & Li, 2022) of studying human-AI relationship perception with another dimension, which puts market pricing and peer bonding, next to authority ranking, into the spotlight. Secondly, we are expanding the methodological toolkit to study relationship modes quantitatively with the human-AI relationship questionnaire adapted from Haslam and Fiske (1999). Thirdly, we are exploring how the relational models fit into the frameworks in the broader field of human-machine interaction, providing further empirical evidence of their relevance.

This approach can potentially offer practitioners insights from a design-based perspective. It mirrors the increasing social dynamics between humans and conversational AI and may be more efficient for designing toward the variety of tasks specifically off-the-shelf conversational AI can be used for. For example, what relational modes are more relevant for request and return dialogues, such as setting a reminder, versus interactions characterized by multiturn conversations in a specific moment, e.g., shopping or asking for music recommendations for a particular occasion? We propose a more functional approach, i.e., relating the design (e.g., language style) to the different tasks. However, this research approach is still in its infancy. We urge future research to adopt this approach and gather further empirical evidence to make conclusive claims for (design-) practice.

Several limitations should be taken into account. First, this cross-sectional study investigates concepts, such as trust or relationship perception, that are likely to be time-sensitive (Seymore & van Kleek, 2021; Glikson & Wooley, 2020). Further studies should consider a longitudinal design to explore the effects over time and test for causality. Second, our sample was mainly from the UK, providing little demographic variation. Future research should complement these findings with a sample composed of various ethnic backgrounds.

Finally, compared to Langer et al. (2021), who found that the terminology of automated decision-making systems affects human perception, we could not identify a meaningful influence of the naming of the conversational AI, same for the gender settings. However, we believe that future studies should carefully consider both issues. In addition, it is foreseeable that more users will be able to change voice or gender settings, potentially influencing relationship perception.

Furthermore, future studies can use the current framework to investigate the impact on user behavior, which should also include behavioral data, such as audio scripts, to analyze actual behavior rather than relying on self-report measures only (e.g., Gao et al., 2018).

Conclusion

As conversational AI, in fact, *is* a tool, AI is not *just* a tool. The human-AI relationship was, to a great extent, perceived as a rational command-and-execution mode (authority ranking), and in line with the suggestions in the introduction, users also displayed a less pronounced tendency to perceive their relationship to the conversational AI as emotional or peer-like (peer bonding). However, we found it most interesting that another prevalent dimension of human-AI relationships was perceived, which can be described as more rational and egalitarian. Highlighting these elements of exchange that characterize market pricing suggests that the conversational AI was ascribed or granted some agency or hierarchical equivalence. Consequently, authority ranking takes a back seat when investigating user perceptions and user characteristics. Borrowing from technical jargon, it looks like authority ranking is the default setting - hence, “servant by default” - while peer bonding and market pricing are the advanced or custom settings that users configure along the way.

Chapter 3: From Theory to Practice

The first two studies employed Fiske's (1992) multidimensional RMT to investigate how humans perceive their relationship with conversational AI in combination with variables of system perception and user characteristics. Interestingly, these relationship perceptions appear to be largely unaffected by device labels and most demographic and system features, including the gender of the conversational AI and household attributes, with the exception of age and education. Since the approach appears well suited to describe the human-AI relationship perception across different contexts, I could take the next step and put the theory into practice. The next step was investigating whether the human-AI relationship framework plays a role in voice shopping.

Linking Human-AI Relationship Perception and Usage Intentions

Reiterating the findings of the first two studies that authority ranking was not a particularly informative dimension, whereas market pricing holds significance, is of practical relevance, and calls for further exploration. This revelation carries significant ramifications for developers working on voice user interfaces. While authority ranking and market pricing may seem comparable at first glance, they differ significantly in a crucial aspect relevant to design: the agency users ascribe to the conversational AI system. I propose that this will likely impact the nature of tasks users delegate to conversational AI.

In simpler terms, the way users see the systems may influence how they use them. Consider a simple task like setting an alarm with conversational AI, which involves brief, straightforward interactions with minimal risks. On the contrary, undertaking complex tasks such as buying a laptop with a voice assistant demands increased user involvement, including multiturn dialogues and heightened risks associated with potential data sharing or financial transactions. In essence, it necessitates a form of shared responsibility. Therefore, if system designers aim to employ or leverage the potential of conversational AI for more intricate tasks, it would be reasonable for users to view them as collaborative partners in decision-making (Dellaert et al., 2020), akin to peer bonding or market pricing) rather than mere entities taking orders (authority ranking, see Tassiello et al., 2021).

The Use Case of Voice Shopping

Thus, in the subsequent study, I sought to implement and expand upon the findings by examining the application of the human-AI relationship framework in the context of shopping with conversational AI. With the introduction of online shopping around 1990 (Laudon &

Traver, 2013), people could start their computers, search the internet, scroll a webpage, and place an order to purchase almost anything. Three decades later, people who own a conversational AI like Alexa do not even have to touch their computers anymore. They can just *talk* to the conversational AI to place an order (Lim et al., 2022). Voice shopping features enable consumers to bypass searching or typing orders and even allow for multitasking, such as preparing a raw chicken for dinner with their hands while shopping (Adolphs & Zaharia, 2021; Mari & Algesheimer, 2021).

Voice shopping is a form of e-commerce but differs significantly from traditional online shopping, calling for new methods of scientific inquiry (Halbauer et al., 2022; Klaus & Zaichkowsky, 2022). Much research has taken a rather techno-centric approach (focusing on the technical structures of the system, Canziani & MacSween, 2021; Dellaert et al., 2020; Fernandes & Oliveira, 2021; Halbauer et al., 2022; Son Nguyen et al., 2021). However, the anthropocentric approach has not received enough attention (Hu et al., 2022). Arguably, customers may perceive their conversational AI as akin to a human salesperson, aiding in decision-making. This context provides an opportunity to employ an anthropocentric research approach, allowing for the adaptation and application of established psychological theories within the realm of human-AI interaction.

Studies, mostly in the realm of technology adoption, have identified relationship-relevant factors influencing voice shopping intentions, including trust, perception of risk, perceived intelligence, human-likeness, social presence, emotional bonding, and parasocial interaction (Chattaraman et al., 2019; McLean & Osei-Frimpong, 2019; Moussawi et al., 2021). However, there has been limited exploration of the impact of perceived relationships with conversational AIs on shopping behavior despite the significance of the relationship metaphor in shaping consumer decisions found in the marketing literature (Alvarez & Fournier, 2016). In marketing theory, social commerce, which is characterized by the collaborative nature facilitated by communities in social networks, for instance, the influence of relationship quality on consumer behavior was found to be essential, e.g., with a positive effect on purchase intentions (Tajvidi et al., 2020). Conversational commerce, while arguably social, lacks human involvement despite the consumer in the social aspect of interactions, making the commercial transaction not computer-mediated but rather conducted *with* the computer. In this particular realm, existing empirical studies are rare and, from a relationship perspective, yield contradictory results. Certain studies indicate a perception of authority over the AI-system as predictors of increased voice shopping intentions (Hu et al., 2022; Tassiello et al., 2021), while others highlight a sense of companionship (Rhee & Choi, 2020).

Moreover, despite the lofty expectations (Lim et al., 2022; Rhee & Choi, 2020) and the widespread popularity of smart speakers like Amazon's Alexa, the utilization of these digital assistants for shopping has not yet become commonplace. The reasons behind this phenomenon remain unclear, as outlined in the Smart Speaker Consumer Adoption Report of 2022 (Kinsella, 2018). Thus, findings will hopefully also serve researchers and practitioners to better understand the intricacies of voice commerce.

To address the identified gap in voice commerce research and the contradictory findings pertinent to the relationship perspective, I argue that not only the human-AI relationship perceptions might be relevant for voice shopping intentions but also that the nature of the purchased product may play a pivotal role in shaping consumers' shopping intentions (in line with arguments by Rhee & Choi (2020). Consumers may lean towards different types of sales agents based on factors such as the level of risk and complexity associated with the product, as well as its cost. For products that entail minimal risk and are relatively inexpensive or involve a straightforward repurchase, consumers may be more inclined to entrust the shopping decisions to a basic assistant, adhering to the conventional role of a digital assistant. On the other hand, for products that are more complex or pricey, consumers may be more inclined to entrust such a decision to a sales agent, which is more on "eye-level", so to speak.

In essence, the significance of a positive relationship with the other party is emphasized in human commercial transactions. Does this extend to human-AI interaction where the relationship perception may be particularly relevant? When it comes to shopping with conversational AI, there is limited evidence and no universally applicable, reliable assessment exists. As a result, the following study directly tackles the overarching question of the dissertation, providing additional evidence to the inquiry: How do humans perceive their relationship with conversational AI? In addition to replicating studies 1 and 2 within a specific context to verify the credibility of prior findings, this study tested how the perceptions of human-AI relationships might impact behavior. Thus, the following confirmatory study, in collaboration with K. Sassenberg, aims to answer the question, does the way users relate to their conversational AI influence what kind of products they buy?

Declaration on the Proportion of Collaborative Publications for Chapter 4 (Tschopp & Sassenberg, 2023)

Author	Author position	Scientific ideas %	Data generation %	Analysis & interpretation	Paper writing
Marisa Tschopp	1	90	100	50	70
Kai Sassenberg	2	10	0	50	30
Title of Paper	The Impact of Human-AI Relationship Perception on Voice Shopping Intentions				
Status in publication process	Accepted for publication at Human-Machine Communication Journal with minor pending revisions. This chapter presents the second reviewed version.				

Chapter 4: The Impact of Human-AI Relationship Perception on Voice Shopping Intentions

With the introduction of online shopping, people could purchase almost anything with a few clicks. Three decades later, people can just *tell* a computer to place an order. Although voice shopping is a form of e-commerce, it substantially differs from traditional online shopping (Klaus & Zaichkowsky, 2022). We argue that voice shopping with a conversational artificial intelligence (AI) is conceptually more similar to decision-making in a brick-and-mortar store involving in-person interactions with human salespeople and should be investigated as such.

Research on relationships between the consumer and (human) seller is popular in the marketing literature (Alvarez & Fournier, 2016). For example, studies have shown that a positive seller-buyer relationship leads to greater brand trust and more positive affect by consumers (Carroll & Ahuvia, 2006; Chaudhuri & Holbrook, 2001). But the relationship perspective has yet not been translated into human-AI interaction, investigating the perception of conversational AI as quasi-sales agents whom consumers form some sort of relationship with (e.g., Lim et al., 2022; Ramadan, 2021; Rhee & Choi, 2020). In fact, research precisely on human-AI relationships is in general, still nascent (Pentina et al., 2023a), and the few existing findings paint a complex picture.

Hu et al. (2021) found that people who see their conversational AI mostly as assisting them have stronger voice shopping intentions, motivated by a hierarchical power experience over their voice assistants, a claim supported Tassiello et al. (2021). While Hu et al. (2021) did not differentiate what users bought, Tassiello et al. (2021) as well as Rhee and Choi (2022) did. Both used the concept of low- and high-involvement: Low-involvement products are characterized as low-cost items consumers tend to consider without extensive deliberation, in contrast to high-involvement products, which are typically pricier and necessitate thorough evaluation (Rhee & Choi, 2020; Tassiello et al., 2021; Mari & Aligsheimer, 2021). In the development of their hypotheses, they argue that shoppers think differently about the products, requiring different persuasive messages to facilitate voice shopping. Contrasting Hu et al.'s (2021) and Tassiello et al.'s (2020) findings, Rhee and Choi (2020) found that a friend-like voice shopping user interface increased voice shopping intentions for low-involvement products.

These partly inconsistent findings call for further research, including the nature of the perceived relationship and the purchase. Therefore, we apply a multidimensional human-AI relationship model while differentiating between low- and high-involvement products. Assuming that users perceive their relationship to their conversational AI not just along a friend or servant dimension but along several dimensions, as suggested by Tschopp et al., 2023, holds promise in gaining differentiated insights into users' voice shopping behavior and addressing the contradictions in the current landscape. Thus, the focal question of this study is: does the way users relate to their conversational AI influence what kind of products they buy?

How Users Perceive their Relationship to Conversational AI

The remarkable progress in AI in the past decades has steadily stretched the boundaries of human-AI interaction and communication, demonstrated by the developments of language models. These advancements have rendered users' interactions not only social in the sense of being imbued with meaning or emotion but have also expanded the potential for the establishment of what might be considered relationships with AI agents, as asserted by Pentina et al. (2023b).

To examine the human-AI relationship perception from a multidimensional perspective, we are building upon Tschopp et al.'s (2022) adaptation of the Relational Models Theory (RMT) by Alan P. Fiske (Haslam & Fiske, 1996) to human-AI relationships. RMT is a theory on how humans construe their relationships with other humans. RMT describes four dimensions and has received mighty empirical support in the past decades. These four dimensions are (1) communal sharing (i.e., a kinship-like relationship as it is with families based on mutual trust), (2) equality matching (i.e., a tit-for-tat-like relationship as with roommates in a shared flat where equal give-and-take is key), (3) authority ranking (i.e., a hierarchical relationship characterized by a clear chain of command like soldiers and their superiors), and (4) market pricing (i.e., a currency based relationship characterized by cost-benefit analyses as it is with employers and their bosses in a workplace).

Applying RMT, Tschopp et al. (2023) found that human-AI relationships are perceived along three dimensions varying in emotional breadth and perceived agency. Communal sharing and equality matching merged into one emotional dimension named peer bonding (see Table 7). They found that conversational AI users characterized their relationship mostly by authority ranking (i.e., a hierarchical owner-assistant relationship) and market pricing (i.e., a non-hierarchical exchange relationship) and least by peer bonding (i.e., a peer-like relationship). Notably, authority ranking was not informative for variables concerning system perception (e.g., trust, perceived intelligence, or affinity to technology). The two rather interactive dimensions (i.e., market pricing and peer bonding) had stronger predictive values, especially regarding anthropomorphism (Tschopp et al., 2023), which drives the development of our hypotheses and research questions.

Table 7

Description of the Three Modes of Human-AI Relationships (based on Tschopp et al., 2023)

Peer Bonding	**Market Pricing**	**Authority Ranking**
• Most human-like dimension where the user treats the conversational AI as an equivalent peer. • Best characterized as a communal relationship.	• The user perception is guided by cost-benefit analyses with no hierarchies. • Best characterized as an exchange relationship on 'eye level'.	• A hierarchical order is perceived between users and the conversational AI. • Best characterized as a master-servant relationship.
The user tends to feel emotionally closer to the system.	The user tends to care about competence and rational trust in the system.	The user tends to use the system for a greater variety of purposes.

While the initial work by Tschopp et al. (2023) remained exploratory, we aim to further investigate their assumptions in an applied context, namely voice shopping. This context presents an intriguing opportunity because multiturn dialogues are necessary to make a purchase decision. In other words, you have to actually communicate with the conversational

AI and not only give orders, such as turning off the lights, where other relational dynamics may be involved.

Peer bonding, often regarded as the most emotionally charged connection, involves regarding the partner as an equal and companion-like figure while also upholding a sense of responsibility for one's conduct (Tschopp et al., 2023). Arguably, for people who see their device through this relationship mode, the voice shopping experience would be more like shopping with a peer.

The newly introduced perception of conversational AI as a rational exchange partner, called market pricing, was found to be rather popular (Tschopp et al., 2023). Its core characteristic lies in the reliance on ratio values, devoid of hierarchies, thus resembling an equal-other, granted some sort of agency. Arguably, for people who see their device through this relationship mode, the voice shopping experience would be more like having a professional sales agent making the shopping decision together with the consumer.

The majority of respondents perceived their devices as authority ranking. The key characteristic of this arrangement is the creation of a linear hierarchy between humans and the conversational AI. For people who see their device through this relationship mode, the shopping experience would be more like shopping with a subservient helper or concierge. However, before making such assumptions, a better understanding of voice commerce is necessary.

Shopping via Conversational AI

Voice shopping, or voice commerce, is an emerging commercial trading system where, for instance, Alexa users (Amazon's conversational AI) can search, purchase, and track products on Amazon solely through a voice user interface (VUI) (Halbauer & Klarmann, 2022; Ramadan, 2021). Alexa shoppers predominantly purchase entertainment products (such as music or books), household essentials (like batteries or toilet paper), and clothing, whereby re-purchases and new orders occur with equal frequency (for a comprehensive breakdown of product categories, see Kinsella, 2018). Practitioners are eager to leverage this new sales channel. However, research in the field is in its infancy, with limited empirical data on what promotes or hinders voice shopping scattered across disciplines (Klaus & Zaichkowsky, 2022; Lim et al., 2022).

From a psychological perspective, initial studies have investigated what drives voice shopping intentions. Trust (Huh et al., 2023; Mari & Algesheimer, 2021), perceived human-likeness/anthropomorphism (Han, 2021; Huh et al., 2023), perceptions of social presence, emotional bonding, and para-social interaction and dialogue (Ramadan, 2021), were found to have a positive influence on voice shopping intentions and continuance. These studies stress the importance of the social dimension in voice purchasing behavior. Especially with regard to the voice shopping process, the increasing interactive verbal decision-making processes and two-way interaction render "voice assistants partners in the decision-making dialogue rather than mere order takers" (de Bellis & Venkataramani Johar, 2020; Dellaert et al., 2020).

Furthermore, only a limited number of empirical studies have distinguished voice shopping intentions based on the specific products individuals purchase, which likely engage distinct processes as comprehensively laid out by Rhee and Choi (2020). In simpler terms, it is highly likely that there is a notable distinction between buying batteries and purchasing a laptop through voice commands, where there is limited access to information and a varying necessity to rely on the AI as a sales agent.

When using conversational AI for product selection, Klaus and Zaichkowsky (2022) suggest that the algorithm serves distinct purposes based on the complexity and functionality of the product. In their model, they differentiate high- and low-involvement situations, where the algorithm serves different functions depending on whether the product is simpler and more functional (i.e., low-involvement). This entails a more utilitarian approach, where users allow the conversational AI to handle the purchase. This concept was also applied in a study by Mari and Algesheimer (2021), who selected batteries as a low-involvement product, invoking the "yeah, whatever" heuristic. In contrast, the decision-making process for intricate, costly, and/or high-risk products, as outlined in Klaus and Zaichkowsky's model, appears quite different. When acquiring items like a $500 vacuum cleaner, more information and guidance are necessary, making them high-involvement purchases that demand greater time and effort for decision-making. In this framework, an algorithm aids the buyer in making the most informed shopping decision collaboratively.

Against this background and given the inclination of people to respond to technological systems in social ways (Nass & Moon, 2000) and the empirical importance of the social dimensions as antecedents of (voice) shopping decisions, it is rather surprising that only few studies have looked at the impact of perceived relationship to the conversational AI on home shopping behavior. Much research has focused on relational proxies, assessing constructs such as perceived warmth, psychological distance, or anthropomorphism (e.g., Gong, 2008; Pitardi & Marriott, 2021) or role ascriptions (e.g., Sundar et al., 2017). Furthermore, and as mentioned above, inconsistent results raise further questions: Hu et al. (2021) have found that presenting conversational AI as servants enabled a power experience for users as masters and increased voice shopping intentions (given that they had a desire for power). Similarly, an experimental study by (Tassiello et al., 2021) found that the subservient assistant role facilitated voice shopping. On the other hand, Rhee and Choi (2020) found that a friend-like social design had a positive influence on voice shopping intentions. Notably, this was particularly important for buying low-involvement products. These findings underscore the need for further research to carefully examine and dissect voice shopping intentions, particularly by distinguishing between different types of products that involve varying levels of involvement in the purchase decision-making process.

Hypotheses Development

Does the Perceived Human-AI Relationship Influence Voice Shopping Intentions?

Dellaert et al.'s (2020) argument that virtual assistants serve as partners in decision-making suggests that peer bonding and market pricing are highly relevant for voice shopping, more so than authority ranking. To reiterate, a large amount of research suggests that human-like system perception variables such as perceived human-likeness (Huh et al., 2023) or emotional bonding (Ramadan, 2021) are promoting shopping intentions. We thus predict:

H1: Higher values in peer bonding predict a stronger intention to use voice shopping.

Market pricing, the non-hierarchical relationship dimension characterized by exchange and interaction, is emotionally less pronounced. However, market pricing still constitutes a human-like relationship, in the sense that it requires that users attribute agency to the system and see their conversational AI rather as an exchange partner whom they meet on "eye level" than as a tool. Relying on the fact that human-like perceptions of conversational AI go hand in hand with voice shopping intentions (Huh et al., 2023), we also expect:

H2: Higher values in market pricing predict a stronger intention to use voice shopping.

Based on the rationale that the conversational AI functions as a sales agent rather than a simple order processor, and considering the absence of predictive information regarding authority ranking as per Tschopp et al.'s study (2023), we posed the influence of authority ranking as an exploratory research question in our preregistration. The results were analyzed in an equitable manner within our results section.

RQ1: How does authority ranking associate with general voice shopping intentions?

Different Predictors for Different Products?

We argue that different relationship dimensions will predict shopping intentions for different products because people evaluate products differently. Inspired by Rhee and Choi's (2020) arguments, this rationale is based on the elaboration likelihood model (ELM, Petty &

Cacioppo, 1986), which distinguishes two routes. The *peripheral route* is characterized by a low amount of effort taken to process product information, but it could also be based on evaluating characteristics of the seller (see also Rhee & Choi, 2020). The peripheral route is typically used for low-involvement items, which are often cheap and interchangeable products (e.g., toilet paper or chewing gum; see Rhee & Choi, 2020). In other words, when a shopping decision bears no real risk, people do not think a lot but follow intuitions and emotions. This focus on intuition and emotions resonates with peer bonding, which is characterized by emotions and similar to a relationship with human peers whom people follow intuitively without much thought. This is in line with the study by Rhee and Choi (2020) that demonstrated the positive effect of a friend-like social design on shopping for low-involvement products but not for high-involvement products.

The *central route* is used for more cognitively demanding products. This form of information processing is characterized by careful elaboration of the quality of arguments, facts, or figures (Petty & Cacioppo, 1986). This cognitive effort is typically only invested when the motivation to process the information is high, in other words, in a shopping context in which more is at stake – financially or personally. This should apply in the case of high-involvement products. When voice shopping for high-involvement products, the decision-making process resembles the central route. Voice shoppers should be highly motivated to evaluate product characteristics and rationality should dominate in a "cost-benefits-analysis style". This style fits a market pricing relationship based on cost-benefit analysis. Taken together, the intuitive and emotional processing style applied when shopping low-involvement products resonates with peer bonding, whereas the cost-benefit-analysis style applied when buying high-involvement products resonates with market pricing (see Table 7). We, thus, predict:

> **H3**: The intention to buy low-involvement products via voice shopping is predicted to a stronger extent by peer bonding than by market pricing.

> **H4**: The intention to buy high-involvement products via voice shopping is predicted to a stronger extent by market pricing than by peer bonding.

As before, we posed an exploratory question regarding the role of authority ranking:

> **RQ2**: How does authority ranking associate with voice shopping intentions for low- and high-involvement products?

To situate the relational approach into common customer value frameworks, we assessed what people care about in voice shopping. We looked at desired hedonic, utilitarian, symbolic, and social benefits (inspired by McLean & Osei-Frimpong, 2019) and how they associate with voice shopping intentions and human-AI relationships. We anticipate that the exploratory analysis will provide conceptual reinforcement for our findings. Given the early stage of the field, it is premature to make definitive predictions and thus commit to the exploration of our research question.

> **RQ3**: How do desired shopping benefits associate with the human-AI relationship perception and voice shopping intentions?

Methods

Design and Participants

We conducted a preregistered cross-sectional study to test our hypotheses (https://aspredicted.org/P9T_XW8). The study was run online via Prolific in July 2022. We aimed at a sample of 450 based on the assumption that N=250 is required for stable correlations (Schönbrodt & Perugini, 2013). We added 200 participants to definitely end up with N>250, even in case of substantial exclusions. We preregistered the following exclusion criteria: no experience in voice shopping, failing at least one attention check, and too short (<150 seconds) or too long (>80,000 seconds) duration of the survey. In a prescreening, we surveyed people (N_{total}=800) to identify potential participants engaging regularly in voice shopping with conversational AIs such as Alexa. We collected data from 451 participants fulfilling this criterion in exchange for £1.10. Twenty-eight participants were excluded based on the criteria mentioned above or because they were outliers with an absolute studentized deleted residual >2.59 in the regression testing (H1 and 2), another preregistered exclusion criterion. The remaining respondents N=423 (57% female, 42%, male, 1% other; age M=41, SD=11.4, age range 19-84 years) responded to the questionnaire regarding their use of Alexa (78%), Google Assistant (16%), Siri (5%), or other conversational AI (1%). More information about users' voice shopping preferences can be found in Appendix B. A sensitivity analysis for a single predictor in multiple regression analysis with three predictors (the analysis for the main predictions) indicated that the sample size was sufficient to detect an effect of f^2=.018 at alpha=.05 and 1-beta=.8.

Procedure

We invited participants to take part in a study on users' perceptions of voice shopping. After providing consent, participants had to choose which conversational AI their answers referred to and then respond to the human-AI relationship questionnaire (adapted from Haslam & Fiske, 1999; see Tschopp et al., 2023). The instructions for the measure require people to focus on a specific device when reporting their relationship. Afterward, we surveyed users about their shopping intentions to test the predictions. Variables were presented in a fixed order. All items were randomized. Next, exploratory variables were assessed. Perceived and desired benefits, trust, and user characteristics (device specifics, frequency of and experience in voice shopping, estimated voice shopping spending per year). We placed questions for demographic information and a final opportunity to withdraw their data at the end. Analyses have been conducted using SPSS 25.0 unless reported otherwise. Supplemental data, code, data, and pre-registration available at https://researchbox.org/1029&PEER_REVIEW_passcode=ZFEVZH.

Measures

Human-AI relationship was assessed using the questionnaire by Tschopp et al. (2023). Administering the questionnaire involves a specific mandatory procedure. Responding to the Human-AI relationship questionnaire necessitates first choosing a voice assistant their answers refer to, e.g., Alexa or Google Assistant. After selecting their preferred assistant, participants were directed to reflect on past shopping experiences and rate the extent to which items described their relationship with the chosen assistant in mind. The questionnaire consisted of 17 items using a 7-point Likert scale (1=*not at all true for this relationship,* 7=*very true for this relationship*). Nine items for peer bonding, four items for authority ranking, and four items for market pricing. A principal component analysis (PCA) with varimax rotation was conducted (see Table 8). The three-factor solution (based on the Kaiser criterion) explained 56.42% of the variance. As in prior studies, the first component represents peer bonding, the second component authority ranking, and the third component market pricing. Due to high loadings (>.4) on a factor they were not intended to correlate with, we omitted item 6 and 17. The final scales presented sufficient reliabilities: *Cronbach's Alpha*=.91 for peer bonding, *Alpha*=.71 for authority ranking, and *Alpha*=.66 for market pricing. Market pricing was positively correlated with peer bonding (r=.47, N=423, p<.001) and authority ranking (r=.27, N=423, p<.001). No significant correlation was found between authority ranking and peer bonding (r=-.09, N=423, p=.073).

Table 8

Results from a Factor Analysis of the Human-AI Relationship Questionnaire (N=423)

	Item	Factor Loading 1	2	3
Peer Bonding				
1	There is a moral obligation to act kindly to each other	**.550**		.387
2	Decisions are made together by consensus	**.771**		
3	You tend to develop similar attitudes and behaviors	**.756**		
4	It seems you have something unique in common	**.839**		
5	You two are like a unit: you belong together	**.784**		
6	You are like tit for tat: you do something and expect something similar in return	**.487**		.425
7	Everyone has an equal say when a decision is made	**.780**		
8	You take turns doing what the other wants.	**.786**		
9	You are like peers or fellow co-partners	**.783**		
Authority Ranking				
10	One of us is entitled to more than the other		**.701**	
11	One directs the work, the other pretty much follows		**.675**	
12	You are like leader and follower		**.691**	
13	One is above the other in a kind of hierarchy		**.745**	
Market Pricing				
14	What you get from this interaction is directly proportional to how much you give			**.661**
15	You have a right to a fair rate of return for what you put into this interaction			**.733**
16	You expect the same return on your effort other people get			**.740**
17	Your interaction is a strictly rational cost-benefit analysis		**.536**	

Note. Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization. Rotation converged in 5 iterations. The highest factor loadings are in bold, factor loadings below .30 are not displayed.

Voice shopping. We measured the *general intention to continue voice shopping* with three items adapted from McLean and Osei-Frimpong (2019). Respondents indicated their agreement on a 7-point Likert scale (3 items, 1=*strongly disagree* to 7=*strongly agree*). For instance, “I plan to continue to use the conversational AI for shopping in the future”. An index was formed by averaging the responses (*Cronbach’s Alpha*=.98).

Intention to continue voice shopping for low-involvement products and the *intention to continue voice shopping for high-involvement products* were assessed with a single item each on a 7-point Likert scale (1=*strongly disagree* to 7=*strongly agree*). Participants read the description, see Table 9, and rated their agreement “I predict I would continue to use the conversational AI for shopping in the future”. General voice shopping intentions were highly correlated with low-involvement shopping intentions (r=.81, N=423, p<.001) and moderately with high-involvement shopping intentions (r=.41, N=423, p<.001). Using the three indicators is supported by a principal component analysis (see Appendix B).

Table 9

Description of Low- and High-Involvement Shopping Intentions

Intention to continue voice shopping for low-involvement products	**Intention to continue voice shopping for high-involvement products**
Think about your future voice shopping experiences. Would you use the voice assistant to shop for products, which are rather convenience products, that require no effort to buy, and there are no emotional values or risks attached? For example, products such as paper towels, chewing gum, cereals, or a specific book. Please rate the extent to which these statements describe your intention to continue purchasing these types of products with your voice assistant in the future.	Think about your future voice shopping experiences. Would you use the voice assistant to shop for products which are rather complicated and require some effort to make a decision, with higher emotional values or risks attached? For example, a laptop, a smartphone, a vehicle, or a tablet. Please rate the extent to which these statements describe your intention to continue purchasing these types of products with your voice assistant in the future.

We tested these instructions in a pretest. In response to the high-involvement product description, people bought items such as laptops or smartphones, jewelry, or clothes. In response to the low-involvement product description, people reported household items such as toilet paper or soap, books, or groceries. Thus, the instructions seem to work as intended (see Appendix B).

Desired benefits were assessed with a 10 items scale measuring hedonic, utilitarian, symbolic, and social benefits inspired by McLean and Osei-Frimpong (2019). Respondents rated their agreement on a 7-point Likert scale (1=*strongly disagree* to 7=*strongly agree*). Because the original questionnaire assessed *actual* rather than *desired* benefits, we performed a factor analysis supporting the intended four-factor structure (see Appendix B). Two items assessed hedonic benefits (e.g., “It is important to me to have fun while shopping with my voice assistant”, r=.60, N=423, p<.001), four items utilitarian benefits (e.g., “It is important to me that the voice assistant makes shopping more efficient”, C*ronbach’s Alpha*=.84), two items symbolic benefits (e.g., “It is important to me that shopping with my voice assistant enhances my image among my peers”, r=.82, N=423, p<.001), and two items measured social benefits (e.g., “I care that shopping with a voice assistant is like dealing with a real person”, r=.77, N=423, p<.001).

User characteristics. We assessed participants’ use of smart speaker or tablet, screen use, and voice shopping spendings (see Appendix B). We measured *frequency of use* (“How often do you use voice assistant for shopping?”) on a single-item 6-point scale from 1=*almost daily* to 5=*1-2 times per year* (including an option 6=*not at all,* ensuring to only survey experienced voice shoppers). *experience of use* (“Since when do you use voice assistant for shopping purposes?”), was measured on a scale ranging from 1=*5 years or more* to 6=*less than 12 months*.

Results

Preliminary analysis

We conducted an ANOVA with repeated measures and post-hoc comparison using Bonferroni correction to test for differences between the dimensions of the relationship perception. Participants saw their relationship with the conversational AI as more strongly characterized by authority ranking (M=4.85, SD=1.38, N=423) than by market pricing (M=4.42, SD=1.43, N=423) and peer bonding (M=2.61, SD=1.31, N=423), all ps<.001, F(1.76, 422.00)=402.28, p<.001, $\eta^2 part$=.488 (with Huyn-Feldt correction). For all descriptive results, see Table 10 below.

Table 10

Means, Standard Deviations, and Bivariate Correlations (N=423)

			Human-AI Relationship			Voice Shopping			Desired Benefits		
Scale	*M*	*SD*	PB	AR	MP	GI	LI	HI	HB	UB	SyB
Human-AI Relationship											
Peer Ponding (PB)	2.61	1.31									
Authority Ranking (AR)	4.85	1.38	-.09								
Market Pricing (MP)	4.42	1.43	.47**	.27**							
Voice Shopping											
General Continuance Intention (GI)	5.31	1.3	.20**	.10*	.18**						
Continuance Intention Low-Involvement (LI)	5.40	1.38	.15**	.12**	.14**	.81**					
Continuance Intention High-Involvement (HI)	3.60	1.93	.42**	-.01	.23**	.41**	.38**				
Desired Benefits											
Hedonic Benefits (HB)	4.69	1.21	.34**	.10*	.34**	.30**	.27**	.31**			
Utilitarian Benefits (UB)	5.17	1.09	.20**	.27**	.44**	.42**	.40**	.24**	.51**		
Symbolic Benefits (SyB)	2.30	1.5	.45**	.01	.14**	.14**	.14**	.35**	.35**	.17**	
Social Benefits (SoB)	3.20	1.51	.50**	-.01	.23**	.19**	.19**	.35**	.47**	.33**	.62**

Note. **Bivariate correlation is significant at the .01 level. Correlation is significant at the .05 level

Main Analyses

General Voice Shopping Intentions (H1 & H2, RQ1)

We tested the predictions that higher values in peer bonding (H1) and market pricing (H2) would predict a stronger general intention to use voice shopping by regressing general voice shopping intentions on the human-AI relationship dimensions. Supporting H1, the regression analysis showed that higher values in peer bonding were associated with a stronger intention to continue voice shopping in general (β=0.18, p=.001, 95%-CI[0.73,0.29]). H2 was not supported as market pricing did not predict a higher intention to engage in voice shopping (β=0.07, *p*=.255, 95%-CI[-0.04,0.16]). The same was true for authority ranking, which was included in the regression for exploratory reasons (β=0.10, *p*=.055, 95%-CI[-0.002,0.19]).

Intention to Engage in Low-Involvement Voice Shopping (H3, RQ2)

We hypothesized that the intention to buy low-involvement products is predicted to a stronger extent by peer bonding than by market pricing. Voice shopping intentions for low-involvement products were regressed on the dimensions of human-AI relationship perception. We found that peer bonding predicts intentions to engage in low-involvement shopping (β=0.15, *p*=.006, 95%-CI[0.05,0.28]). Market pricing was not associated with low-involvement voice shopping intentions (β=0.02, *p*=.684, 95%-CI[-0.09,0.13]). Evidence for H3 was provided by the fact that the CIs for both standardized regression coefficients did not include the respective other regression coefficient. Notably, authority ranking positively predicted intentions to voice shop for low-involvement products (β=0.15, *p*=.004, 95%-CI[0.05,0.25]).

Intention to Engage in High-Involvement Voice Shopping (H4, RQ2)

We hypothesized that the intention to buy high-involvement products is predicted to a stronger extent by market pricing than by peer bonding. Voice shopping intentions for high-involvement products were regressed on the dimensions of relationship perception. We found no significant association of market pricing with intentions to engage in voice shopping for high-involvement products (β=0.04, *p*=.410, 95%-CI[-0.08,0.20]). However, peer bonding predicted high-involvement shopping intentions *(*β=0.40, p<.001, 95%-CI[0.43,0.73]). Thus, we did not find evidence for H4. The intention to buy high-involvement products via voice shopping was not predicted by the market pricing but by the perception of peer bonding relationship (Table 11). The reported correlations did not substantially change when shopping spendings or screen use were included as covariates in the regressions reported so far (for details, see Appendix B).

Table 11

Regression Coefficients of Relational Modes and Shopping Intentions on Desired Benefits (N=423)

Variable	General Voice Shopping	Low-Involvement Voice Shopping	High-Involvement Voice Shopping
	β	β	β
Authority Ranking	.10	**.15***	.02
Market Pricing	.07	.02	.04
Peer Bonding	**.18****	**.15***	**.40****

Note. **p*<.05. ***p*<.01. Significant values in bold.

Relation between Human-AI Relationships, Desired Benefits, and Voice Shopping Intentions (RQ3)

We regressed the relationship dimensions on the desired benefits (see Table 12). Higher values of desired utilitarian benefits were associated with higher values in authority ranking, β=0.31, *t*(418)=5.53, *p*<.001, and market pricing, β=0.35, *t*(418)=6.90, *p*<.001. Market pricing was also predicted by desired hedonic benefits, β=0.13, *t*(418)=2.37, *p*=.018. Higher values in hedonic benefits, β=0.11, *t*(418)=2.01, *p*=.037, desired symbolic, β=0.22, *t*(418)=4.17, *p*<.001, and social benefits, β=0.31, *t*(418)=5.37, *p*<.001, significantly predicted higher values in peer bonding. The other relations were not significant. Then, we regressed the two voice shopping dimensions on the desired benefits, showing that low-involvement shopping was predicted by desired utilitarian benefits (β=0.35, *t*(418)=6.65, *p*<.001). High-involvement shopping, on the other hand, was significantly associated with desired hedonic (β=0.13, *t(*418)=2.27, *p*=.024), symbolic (β=0.21, *t*(418)=3.64, *p*<.001,), and social benefits (β=0.14, *t*(418)=2.21, *p*=.028).

Table 12

Regression Coefficients of Relational Modes and Shopping Intentions on Desired Benefits (N=423)

Variable	Authority Ranking	Market Pricing	Peer Bonding	Low-Involvement Voice Shopping	High-Involvement Voice Shopping
Desired Benefits	β	β	β	β	β
Utilitarian Benefits	**.31****	**.35****	.01	**.35****	.10
Hedonic Benefits	-.02	**.13***	**.11***	.06	**.13***
Symbolic Benefits	.04	.00	**.22****	.05	**.21****
Social Benefits	-.12	.06	**.31****	.01	**.14***

Note. **p*<.05. ***p*<.01.

In sum, utilitarian benefits are the primary predictor of authority ranking, market pricing, and low-involvement shopping, whereas hedonic, symbolic, and social benefits are related to peer bonding and high-involvement shopping.

Discussion

The primary goal of this study was to investigate whether voice shopping intentions for low- and high-involvement products depend on how users perceive the human-AI relationships (i.e., peer bonding, market pricing, and authority ranking, based on Tschopp et al., 2023).

Supporting H1, we found that general shopping intentions were predicted by peer bonding, in line with prior research highlighting social dimensions in voice shopping (e.g., Mari

& Algesheimer, 2020). Peer bonding showed stronger predictive values for low- and high-involvement shopping than market pricing, supporting H3 but contradicting H4. Peer bonding may not only be relevant for low-involvement shopping but, as indicated by a strong regression coefficient, even more in high-involvement shopping. This is interesting because it contrasts Rhee and Choi's results (2020) with regard to high-involvement shopping yet supports the findings regarding low-involvement shopping. Against our prediction, market pricing was unrelated to shopping intentions (contradicting H2 and H4). Market pricing may not relate to voice shopping, as the rational calculations inherent in market pricing may not be conducive to the presumably swift decision-making process involved in voice shopping. Thus, one could posit that voice shopping appears to be associated more with rapid decision-making than deliberative, slow thinking (cf. Kahnemann, 2012). The difference in results compared to Rhee and Choi (2020) could be due to the different study approaches. They conducted an experiment with undergraduates potentially lacking voice shopping experience and confronted them with a shopping scenario – yielding high internal validity, whereas we recruited experienced voice shoppers and asked about their shopping intentions – yielding high external validity.

Our complementary analysis (RQ3) on the desired benefits sheds light on reasons for the strong predictive power of peer bonding. High-involvement shopping (not low-involvement shopping) was related to perceived hedonic, social, and symbolic benefits, which are more socio-emotional in nature. The importance of the socio-emotional dimensions in all facets of voice shopping supports Dellaert et al.'s (2020) claim that AI assistants are more partners in an interactive decision-making process than subservient assistants. Notably, low-involvement shopping was also related to authority ranking (RQ1&2), products traditionally associated with utilitarian purposes, where interaction focuses on efficiency.

In sum, people tend to use voice shopping either in a utilitarian manner, by giving orders to their AI assistant, and/or in a more socio-emotional fashion, immersed in a rather emotional shopping experience. No evidence was found for market pricing we assumed to predict high-involvement shopping, invalidating the concept of low- and high-involvement decision-making. Maybe the technology is simply "not there yet," or high-involvement products might be bought via voice shopping after the calculative decision process has been performed.

Implications for Theory

The proposed differentiation of perceived human-AI relationships proved to be helpful to disentangle the consequences of different social perceptions on behavioral intentions. Researchers can use the framework to further explore voice shopping or other functionalities (e.g., smart home) and other applications in the broader AI field (e.g., automated driving). Our study focused on voice shopping intentions, yet if our findings also hold for actual behavior, outcomes have strong practical implications.

Implications for Practice and Policy

System designers may have to rethink effective conversational design strategies tailored to different shoppers as well as shopping scenarios. However, more research is needed to draw safe conclusions. Implications may also arise for business developers choosing the sales channel. For selling low-involvement products, Alexa as a channel might work well despite the lack of control over the conversational design. For high-involvement items, control over the social design might be critical due to the found importance of socio-emotional elements. Thus, with limited control over the social design, Alexa as a sales channel for high-involvement products might not work well. Last but not least, the results may also be relevant for policy-makers who further aim to investigate the manipulation and addiction potential of human-AI relationships and the potential facilitation thereof through emotional or personalized social designs (Véliz, 2023). In other words, more evidence is needed on whether these relationship dynamics can be exploited.

Strengths and Limitations

The study enriches the comprehension of the emerging field of voice shopping by investigating experienced voice shoppers and amplifies the value of the perceived human-AI relationships (Tschopp et al., 2023) as predictors thereof. Thereby, this research allows for recommending differentiated voice user interface design strategies and may guide strategic sales channel decisions. A limitation of our findings is the reliance on self-reported shopping intentions instead of actual shopping behavior as well as the lack of cultural variation. Caution is advised regarding the market pricing predictions due to lower scale reliability. The internal consistency was low and could, unfortunately, not be improved by dropping single items. Future research should use longitudinal and/or experimental designs.

Conclusion

We have investigated the influence of differently perceived human-AI relationships on general, high- and low-involvement shopping intentions. The results emphasized the importance of socio-emotional elements (i.e., peer bonding) for voice shopping, in particular for high-involvement products. For low-involvement products, however, the traditional master-servant relationship (i.e., authority ranking) was still found to be relevant. Understanding the impact of multidimensional human-AI relationship perception is relevant for researchers, system designers, and business developers – presumably not only in voice shopping. Additionally, it holds relevance for policymakers, given recent studies pointed out potential negative impacts like user manipulation or addiction through humanized design (Ramadan, 2021).

Chapter 5: Human-AI Relationship Perception and Conversational Design Strategy

The prior study was the first applied research to investigate the impact of perceived human-AI relationships on voice shopping intentions. Despite the value of these findings that the perceived human-AI relationship played a role in voice shopping, the practical relevance, especially for conversational designers, is somewhat limited because we cannot experimentally manipulate the perception of users. Thus, the final step is determining how the human-AI relationship framework enables me to give meaningful recommendations for practitioners, focusing on conversation designers (because speech is a central means of interaction, Guzman et al., 2023) and policymakers (because of the potential adverse effects mentioned before).

Linking Conversational Design and Human-AI Relationship Perception

Reiterating the outcome of the prior study, the prominence of socio-emotional dimensions in all aspects of voice shopping was certainly an interesting finding. However, when differentiating the products, we found that the perception of authority ranking remains a viable component countering the initial suspicion of irrelevance raised by Study 1 (i.e., authority ranking was not informative regarding any dimensions of system perceptions). Having established a stable framework to assess human-AI relationships and finding a relevance thereof in voice shopping provides an opportunity to contemplate useful implications for practice. How could conversational design choices be associated with human-AI relationship perceptions and voice shopping intentions? It is plausible that a rather rational digital assistant, characterized by a reduced emphasis on socio-emotional communication style, is more closely linked to the perception of authority ranking or market pricing. Conversely, employing a more humanized, emotional conversational style may elicit perceptions aligned with peer bonding. However, there is no evidence of how design choices and the different human-AI relationship perceptions associate – which I consequently aim to investigate in the final study.

The Role of Human-AI Fit in Voice Shopping

How to put these thoughts about how the human-AI relationships associate with conversational design in the context of voice shopping into test? Since none of the metrics in the first voice shopping study provided support for the presumed dominance of a calculative, rational, and argument-based decision-making process in voice shopping (akin to market pricing), I focus on authority ranking and peer bonding in deriving the hypothesis, as both were found to be relevant in voice shopping.

Many research efforts aiming to recommend efficient design strategies have examined their assumptions based on concepts of user-technology alignment. These Fit concepts generally assume that matching technology characteristics with the preferences of the user has a positive effect on attitude towards the system and interaction (Liu et al., 2011). Examples of user-technology alignment are, for instance, matching the conversational style to the target group (e.g., informal for youth users versus formal tone for seniors, e.g., Sundar et al., 2017) or matching the form of embodiment (e.g., female character for female users and male character for male users; e.g., Zogaj et al., 2023). Along with this rationale, called Human-AI fit hereafter, I designed an experiment to investigate whether aligning the conversational tone with consumers' perception of their human-AI relationship can enhance the efficacy of voice shopping. By manipulating the spoken output of Alexa, which was emotionally enhanced via different design elements, I test the assumption that individuals perceiving their interaction with Alexa as akin to peer bonding are likely to prefer an emotional, more humanized design in voice shopping. Conversely, those akin to authority ranking may be inclined towards a standard design.

The final study adds further empirical evidence to the central question of my dissertation: How do users perceive their relationship with conversational AI and how do these perceptions influence behavior? It wraps up my research by introducing an experimental setting, allowing for stronger inferences about the role of relationship perceptions and impact in behavior. Moreover, it incorporates the influence of conversational design, acknowledging the pivotal role of communication in relationship formation. In essence, my efforts aim for a comprehensive conclusion to my research. Thus, I conducted an experimental study, manipulating the conversational design to derive more robust conclusions. In collaboration with M. Gieselmann and K. Sassenberg, I try to answer the question, should the conversational design align with users' human-AI relationship perception to facilitate the voice shopping experience?

Declaration on the Proportion of Collaborative Publications for Chapter 6 (Tschopp, Gieselmann & Sassenberg, 2023)

Author	Author position	Scientific ideas %	Data generation %	Analysis & interpretation	Paper writing
Marisa Tschopp	1	90	100	50	80
Miriam Gieselmann	2	0	0	0	5
Kai Sassenberg	3	10	0	50	15
Title of Paper	Don't Call Me Buddy! When Emotional Design Hinders the Voice Shopping Experience				
Status in publication process	Submitted for publication in Electronic Commerce Research Journal				

Chapter 6: Don’t Call Me Buddy! When Emotional Design Hinders the Voice Shopping Experience

“Alexa, my coffee machine broke down!” With a mere vocal command, conversational AI users can initiate the purchase, swiftly arranging its delivery to their doorstep: convenient, efficient, and accessible (Hu et al., 2022). Voice commerce holds tremendous global market potential, with projections indicating significant opportunities for growth and success (Kinsella, 2018; Batchelor, 2022; U.S. Smart Home Consumer Adoption Report 2022, 2022). These figures, combined with the growing availability of generative AI, are said to revolutionize the e-commerce landscape (Chui et al., 2021, Dey, 2023; Savage, 2023)

However, empirical evidence presents a less optimistic picture. The adoption of voice shopping remains limited (Lim et al., 2022), with underlying reasons largely unexplored (Halbauer et al., 2022; Klaus & Zaichkowsky, 2022). One barrier to adoption is said to be that digital assistants are perceived as too artificial or cold (Adam et al., 2021; Sundar & Kim, 2019). Simply put, an AI system cannot provide the genuine service or warmth as a human sales representative.

Consequently, studies emerged, enhancing machine-like conversational AI output with a variety of linguistic cues to make it more human-like (Araujo, 2018). Especially more emotional conversational design, signaling friendship or empathy with words or emojis elicited positive reactions in some studies (e.g., Adam et al., 2021; Beattie et al., 2020; Cicco et al., 2020; Go & Sundar, 2019). However, emotional conversational designs do not consistently elicit positive user reactions (Wirtz et al., 2018; Mara et al., 2022; Yang et al., 2022). Creating conversational AI or other technological systems has led to positive outcomes (Moussawi et al., 2021), as well as negative outcomes (e.g., Lin et al., 2020), confusing conversational design research and practice.

One potential explanation for these contradictory findings is that the association between emotional conversational design and voice shopping intentions alters depending on how people relate to their artificial voice shopping assistant. An emerging research stream suggests that consumers form fundamentally different relationships with conversational AI varying considerably in emotional breadth (Pentina et al., 2023b; Tschopp et al., 2023; see Chapter 4). Arguably, the emotional conversational design should thus also vary in emotional breadth if these relational differences have proven to influence consumer-AI interaction (see Chapter 4)? Our study aims to expand existing literature in human-machine communication (Auraujo, 2017) and other fields to explain why not all consumers receive emotional conversational design equally well.

Thus, we ask, *should conversational design match consumers' human-AI relationship perception to facilitate voice shopping*? To answer this question, we conducted a preregistered online experiment to test the effects of communication style (emotional vs. standard conversational design), human-AI relationship perception (peer-like, servant-like, or rational-equal) and the interaction of both predictors on voice shopping with multiple regression analyses.

Background And Hypotheses Development

Next Level E-Commerce: Shopping with Artificial Sales Agents

With the advancements of conversational AI over the past decades, smart speakers have seamlessly integrated into our homes, where we can delegate the manual control of lights and timers, engage in light conversations, or play games. Talking to machines to control various aspects of our lives has become normal to some extent, apart from delegating shopping decisions to conversational AI (Hu et al., 2021; Perez, 2020).

Voice shopping, also referred to as voice commerce or conversational commerce, is a rapidly growing method of conducting commercial transactions. With voice assistants like Alexa, Amazon's conversational AI, users can search for, purchase, and track products on Amazon using only a voice user interface (VUI). This innovative approach eliminates the need for traditional browsing and clicking, allowing users to interact with the platform solely through voice commands while engaging in other daily activities.

Practitioners are eager to leverage the potential of this new sales channel (Dey, 2023). Especially in light of the rapid advances and increasing availability of generative AI, like ChatGPT, we can expect a rise of sales chatbots in the near future (Dey, 2023; Chui et al., 2021). However, there is a significant dearth of empirical evidence about users' engagement with conversational AI in e-commerce activities (Canziani & MacSween, 2021; Chattaraman et al., 2019; Lim et al., 2022). Furthermore, voice shopping fundamentally differs from traditional online shopping and the methods employed to investigate consumer behavior processes (Adolphs & Zaharia, 2021; Halbauer & Klarmann, 2022; Ramadan, 2021; Zaharia & Würfel, 2021). What is still missing is a deeper understanding of the facilitators of voice shopping in order to make better design choices. More specifically, regarding humanized, emotional conversational design (Lim et al., 2022; Rhee & Choi, 2020).

In sum, conversational AI nowadays allows us to create engaging and personalized shopping experiences mediated by a "quasi sales agent". This quasi sales agent can be tailored through purposeful, emotional conversational design decisions. We are already well beyond studying the impact of the technicality that computers *can* talk. Given the rapid developments in conversational AI and the race for market supremacy, we urgently have to better understand the sociality aspects of *how* computers talk to users, including the practical and ethical implications thereof.

Anthropomorphizing and Anthropomorphization of Conversational AI

In the early days of the popular MIT chatbot Eliza in the 1960s, even a few simple written sentences evoked strong reactions from users needing "alone time with her" (Weizenbaum, 1966). Indeed, humans possess a remarkable ability to anthropomorphize computers (i.e., anthropomorphizing, Coeckelbergh, 2021; Epley et al., 2007) and are inclined to treat them as social actors (CASA; Nass & Moon, 2000). According to Nass & Brave (2005), our social response to computers becomes even stronger when machines are voice-enabled. In other words, we are more engaged when computers talk back to us. Hence, integrating voice into traditional online shopping has emerged as a captivating and promising avenue for both researchers and practitioners. However, it's not enough to have just the integration of a voice (Pearl, 2017). The journey from machine-like to human-like is diverse, and conversational design in human-computer communication – the art of developing the language style of conversational AI – is an ongoing, heated topic of discussion in human-computer interaction studies and e-commerce literature (Hu et al., 2022).

Various studies suggest a positive influence of emotional conversational design on interaction variables pertinent to achieving favorable outcomes. For example, Han (2021) found a positive influence of anthropomorphism on shopping intentions with conversational AI. Qiu and Benbasat (2009) found that human voice-based communication substantially impacts users' perceptions of social presence, enhancing their trusting beliefs, enjoyment, and intention to utilize the agent as a decision aid. Similarly, Beattie et al. (2020) discovered that the inclusion of emojis led to an increased level of social attractiveness. Thus, design practitioners aim to capitalize on these findings and experiment with emotional conversational design in order to achieve better interaction ratings or higher acceptance rates.

The crux is that emotional conversational design (and other anthropomorphic design cues, such as naming or embodiment) do not reliably evoke favorable user responses (Mara et al., 2022; Wirtz et al., 2018). This includes research on voice shopping-relevant outcomes such as attitude toward the system (Fernandes & Oliveira, 2021; Ramadan, 2021), perceived benefits (McLean & Osei-Frimpong, 2019), or voice shopping continuance intentions (Hu et al., 2022; Rhee & Choi, 2020). Contrasting the enthusiasm, deliberate humanization, including emotional conversational design, was also received with a sense of unsettlement or eeriness during interaction, a phenomenon that is not yet fully understood, and more empirical evidence is required (Feine et al., 2020; Mara et al., 2022; Mori et al., 2012; Stein & Ohler, 2018). Contradictory findings may arise due to two reasons: (a) considering only one dimension of the human-AI relationship (e.g., only master-servant or friend-like) and (b) neglecting individual user preferences. In the current research, we sought to address both of these deficits of earlier research.

Human-AI Fit? The Rise of "Personalized" Artificial Shopping Assistants

A potential explanation for the observation that emotional conversational design can have both positive and negative effects is that this emotional conversational design needs to fit the "sociality" individual. The idea of congruence, alignment, or fit has found substantial support in social sciences in the past decades, for instance, in organizational psychology (value congruence, Edwards & Cable, 2009), person-environment fit (van Vianen, 2018), or personality psychology (person-situation fit, Shoda et al., 1994). Similarly, IT-fit-concepts, more precisely, the individual technology fit concept, emphasize the significance of aligning user preferences with the characteristics of the technology for optimal interaction (Liu et al., 2011).

In line with the similarity-attraction hypothesis (i.e., people feel more attracted to others who are similar to them, e.g., Byrne, 1971), studies have shown that congruence between users and conversational AI could be beneficial. Such studies encompassed demographic characteristics like matching gender (e.g., Zogaj et al., 2023), consumer needs, or consumer traits, such as designing an extrovert chatbot persona for extroverted users (e.g., Shumanov & Johnson, 2021) or designing interfaces as servants for users preferring a power experience (Hu et al., 2022). The latter is specifically interesting as it refers to a relational experience between the user and the conversational agent but does not encompass the multidimensional nature of human-AI relationship perception (Tschopp et al., 2023).

While the mentioned IT-Fit studies offer valuable insights, it is essential to acknowledge that not all fit-strategies may be applicable or sufficiently address two critical aspects: First, the complexities of the voice shopping context, which necessitate more sophisticated interactions such as multiturn dialogues and may entail higher financial risks and data sharing. Secondly, the consideration of the relevant (mulitple) dimensions of the human-AI relationship perception.

Fitting the Conversational Design to the Human-AI Relationship Perception

Tschopp et al. (2023) recently proposed a multidimensional approach to investigate human-AI relationship perception. They found that conversational AI users see their devices on three fundamentally different relational levels varying in emotional breadth and perceived agency: peer bonding (i.e., peer-like), authority ranking (i.e., servant-like), and market pricing (i.e, equal-rational).

Peer bonding probably considered the most emotional relation, entails treating the device as equal and peer-like while maintaining a sense of accountability for one's actions (Tschopp et al., 2023). Along with the human-AI fit approach, we argue that for people who perceive their conversational AI more like a peer (people high in peer bonding) a more “emo-tionalized” Alexa might facilitate voice shopping.

To avoid that conversational AI come across as “mechanistic, unemotional and cold” (Sheehan et al., 2020; Sundar & Kim, 2019), resulting in skepticism and resistance against the technology (Adam et al., 2021), emotional conversational design might be beneficial, specifically for these users. Studies finding positive associations between anthropomorphic cues and desired outcomes have used the first person singular to signal identity (see Pickard et al., 2014) or small talk, greetings, and farewell (e.g., Simmons et al., 2011), empathy (see Adam et al., 2021), and/or emojis (see Beattie et al., 2020). We based our manipulated emotional conversational design on these studies and thus predict:

H1: The more people report peer bonding, the more they have stronger voice shopping intentions after seeing the emotionalized (vs. normal) interface.

Tschopp et al. (2023) found that peer bonding was the least common relationship mode, whereas the majority of individuals perceived their devices through the lens of the traditional master-servant relationship, referred to as authority ranking. The defining feature of authority ranking is the establishment of a linear hierarchy between humans and the conversational AI. For individuals who see their conversational AI as the traditional master-servant relationship, an enhanced emotional conversational design might have negative consequences.

This is in line with Fernandes and Oliviera's (2021) argument in their study that users may not care about social conversations when seeking to fulfill utilitarian needs (Rafaeli et al., 2017). Chattaraman et al., (2019) further argue that emotional conversational designs may be perceived as fake and elicit negative responses. Hence, it seems likely that those in a more rational or hierarchical relationship would in fact, prefer a cold, mechanistic system or subservient conversational AI and use it for simply utilitarian purposes, as suggested by the authors (see Chapter 4). This might also apply to voice shopping with the "standard" Alexa, which is marketed as a digital assistant serving as a tool to fulfill shopping goals. We thus predict:

> **H2**: The more people report authority ranking, the more they have stronger voice shopping intentions after seeing the normal (vs. emotionalized) interface.

Equally popular as authority ranking (Tschopp et al., 2023) was the perception of the conversational AI as a rational exchange partner, attributing some sort of agency to the system, called market pricing. Market pricing's core characteristic lies in the reliance on ratio values to shape humans' perception, devoid of hierarchies, thus resembling a continuous cost-benefit analysis. Because prior research relied on dichotomous distinctions and the novelty of the relational approach, empirical evidence regarding this style is scarce, and it is not yet clear which communication style could be most suitable to the needs of these users. Therefore, we adopted an exploratory approach to investigate the impact of market pricing and conversational design on voice shopping.

> **RQ**: Is the effect of the interface design (normal vs emotionalized) on voice shopping intention contingent to market pricing?

The Current Study

In this preregistered online experiment, we tested the idea that emotional conversational design, i.e., normal Alexa versus emotionalized Alexa, should match consumers' perception of their human-AI relationship in order to promote voice shopping intentions.

For this purpose, we assessed the human-AI relationship perception and experimentally manipulated the conversational design. In doing so, our study adds to the understanding of voice shopping decisions from a multidimensional relationship and conversation design perspective.

Method

Design and Participants

We conducted a preregistered cross-sectional study via Prolific in May 2023 to test our hypotheses (see https://aspredicted.org/1H1_YR9). Participants were randomly allocated to one of two experimental conditions (communication: emotional design vs. standard design). In addition, peer bonding and authority ranking were assessed as continuous independent variables. We aimed for 80% power for the test of one coefficient with an effect size of f^2=.02 and an *alpha error*=.05 in a multiple regression with three predictors in total. An a priori power analysis indicated that the total sample size should be 395. We added 55 observations to account for potential exclusions. Thus, we sampled 450 frequent users of conversational AI (based on a prescreening criterion offered by Prolific). Participants received GBP1.25 for a study lasting, on average, eight minutes and 27 seconds. Inclusion criteria were age (at least 18 years old), consent for participation, and agreement with the usage of data for scientific purposes. Exclusion criteria were failing an attention check and statistical outliers using an absolute studentized deleted residuals >2.59 in the regression testing the first hypothesis. Additionally, we excluded participants who used a mobile phone to participate in the study because the stimulus videos were not clearly visible on screens of that size. The remaining four hundred seven (N=407) respondents (51% female, 49% male; age range 18-74, M_{age}=40.30, SD_{age}=13.64 years) from the UK used their voice assistant frequently (85% daily or several times a month), have no (65%), or limited voice shopping experience (35%).

Experimental Manipulation

For the manipulation of communication, we created videos and pretested whether the communication was perceived in line with our intentions (for a screenshot, see Figure 1; links to and transcripts of the videos are provided in Appendix C). In all videos, we used the Amazon Echo Show (Amazon's smart speaker with a scrollable screen). The videos showed one person voice shopping through Alexa while engaging in another activity. The videos were all 1.5 minutes long, and the final videos were all fully dubbed. With a text-to-speech skill[14] we let

[14] *https://www.amazon.de/Michael-Dworkin-Text-vorlesen/dp/B09MW253S4*

Alexa speak out our script, recorded it, and finally dubbed both videos[15]. In the first pair of videos, a person ordered cornflakes while browsing through a newspaper; in the second video, the person ordered a coffee machine while cooking. In all videos, we inserted close-ups of the Alexa screen so that respondents could see Alexa's output better.

Figure 1

Video Scenario where a Person is Voice Shopping for a Coffee Machine while Preparing Food

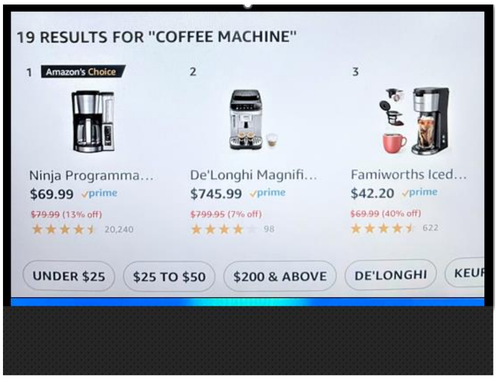
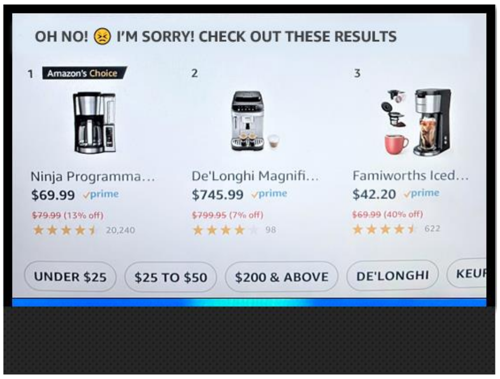
The videos in each pair differed regarding the communication. In the normal condition, we scripted the actual conversation between the user and Alexa. We slightly shortened and edited the Alexa outputs. For the emotional condition, we edited the script by adding socio-emotional cues, as proposed in prior research and practice (e.g., Adam et al., 2021; Beattie et al., 2020; Guzman, 2023; Rhee & Choi, 2020); such as the use of first person, active listening (responding to the user in an empathetic manner), and addressing the user as a friend (see whole scripts in the Appendix C). For the emotional condition, we added emotionalized messages and emojis to the screenshots as suggested by Beattie et al. (2020). See Table 13 for an example.

[15] *We used Microsoft Clip Champ to dub and edit the videos.*

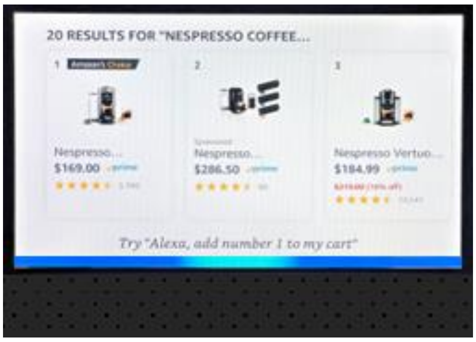
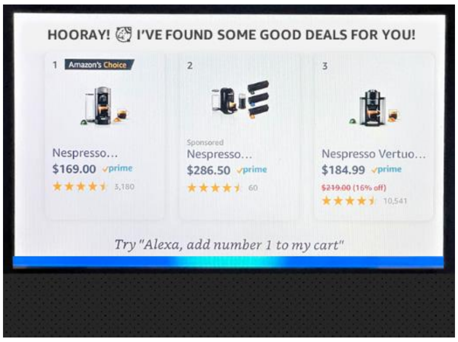
Table 13

Snippets of the Conversation and Screenshots of the Normal Alexa versus Emotionalized (manipulated) Alexa

User Input Example 1:
Alexa, my coffee machine broke down. I need a new coffee machine.

Output normal Alexa:	**Output emotionalized Alexa:**
One Moment, showing results for coffee machines. You can narrow down the choices by saying things like filtered by under 125.	Oh no, I'm sorry to hear that. You'll love these results! You can filter by saying filtered by under 125. Let's find your perfect coffee machine together!
 *Close-up shown in the video with original Alexa screen output.*	 *Close-up shown in the video with edited screen output, including emoji.*

User Input Example 2
Alexa, I think I want a Nespresso Coffee machine – which one is the cheapest?

Output normal Alexa:	**Output emotionalized Alexa:**
One moment. Showing results for Nespresso coffee machines. The first one is Amazons choice. You can filter by saying filter under 125 dollars.	Sweet idea my friend. I think we just found the perfect coffee machine for you! Check out Amazon's top pick. What do you think? Should we jump right in with this deal?
 *Close-up shown in the video with original Alexa screen output.*	 *Close-up shown in the video with edited screen output, including emoji.*

Note: Full videos available on YouTube, links and full script see Appendix C.

Pretest of Manipulation

To test whether the emotional video appeared more emotional or warm, we conducted a pretest on Prolific with 200 participants in exchange for £0.60 (for 4 minutes average duration). We randomly assigned one of the four videos to each participant, resulting in a 2 (video: emotional vs. normal) x 2 (cornflakes vs. coffee machine) design. In our pretest we assessed the perception of the different conversation designs (manipulation check) with 9 items on 7-point scale (1=*strongly disagree* to 7=*strongly agree*). Participants had to rate their agreement with respect to Alexa as shown in the video, for instance, "In the video Alexa was…" *friendly*, or, *functional*. The videos presenting the standard Alexa condition were perceived as less friendly (M=4.91, SD=1.13) than those in the emotionalized condition (M=5.38, SD=1.05), $F(1, 192)=9.37$, $p=.003$ $\eta p2=.047$. The difference between both types of videos was not moderated by the content of the video (coffee machine versus cornflakes), $F(1, 192)=.08$, $p=.794$ (see Appendix C). Thus, the manipulation worked as intended, and we launched the main study, where we confronted participants either to both standard Alexa videos or both emotional videos to create a somewhat stronger manipulation.

Procedure of Main Study

For the main study, we invited (different) respondents to participate in an experiment to test the effect of social design on attitude and voice shopping intentions depending on how they perceive the human-AI relationship. After reading the instructions and providing consent, participants had to choose a conversational AI their answers referred to for the subsequent assessment of their perceived human-AI relationship. As a result, 84% have chosen Alexa, the remaining Google Assistant (15%), or Siri (1%). Then, they responded to the human-AI relationship questionnaire (Tschopp et al., 2023), adapted from (Haslam & Fiske, 1999). Afterward, respondents were randomly assigned to the normal or the emotional Alexa condition. The experimental manipulation was implemented as explained above. Directly after watching the videos, we measured the attitude towards the voice assistant, shopping intentions, and perceived benefits of voice shopping. The participants had to respond to the attention checks after assessing the dependent variables. We placed questions for demographic information and a debrief (that the Alexa was scripted) at the end of the survey. A final opportunity to withdraw their data was given at the end. The Appendix C contains a complete list of items.

Measures

The perceived human-AI relationship was assessed with a questionnaire developed by Tschopp et al. (2023, adapted from Haslam & Fiske, 1999). The questionnaire consisted of 16 items using a 7-point Likert scale (1=*not at all true for this relationship*, 7=*very true for this relationship*): Eight items for *peer bonding*, for example, “You are like peers or fellow co-partners” and four items for *authority ranking*, for example, “You are like leader and follower”. We also assessed *market pricing* for exploratory purposes, for example, “Your interaction is a strictly rational cost-benefit analysis” with four items. An index was formed by averaging the responses, see Table 14 for scale reliabilities, means, and standard deviations of all measures.

Evaluation of voice shopping was assessed on three dimensions. Attitude assessed how people liked the interface and was measured with five bipolar items on a 7-point Likert scale, for example, *annoying – pleasant* or *negative – positive* (based on MacKenzie & Lutz, 1989). *Benefits* measured how people evaluate the shopping experience (based on McLean & Osei-Frimpong, 2019). It was assessed with two items on a 7-point Likert scale (1=*strongly disagree* to 7=*strongly agree*), namely, “Shopping with this kind of voice assistant would be fun” and “Shopping with this kind of voice assistant would be efficient”. *Shopping intention* items covered the behavioral level and was measured with two items on a 7-point Likert scale (1=*strongly disagree* to 7=*strongly agree*). For instance, “I would like to use the voice assistant as shown in the video for shopping in the future”. An index was formed by averaging the responses.

Table 14

Scale Reliabilities and Descriptive Statistics for Human-AI Relationships and Voice Shopping (N=407)

					PB	AR	MP	Att.	Bene.	Int.
Scale	*k*	*α*	*M*	*SD*	*r*	*r*	*r*	*r*	*r*	*r*
Human-AI relationship										
Peer Bonding	8	.93	2.53	1.37						
Authority Ranking	4	.72	4.74	1.43	.01					
Market Pricing	4	.68	4.29	1.31	.39**	.45**				
Voice Shopping										
Attitude	5	.88	5.71	1.11	.21**	.07	.18**			
Benefits	2	.94	3.97	1.86	.44**	.04	.23**	.52**		
Intentions	2	.81	4.43	1.58	.42**	.03	.20**	.57**	.86**	

Note. *K*=number of items included in the scale***;*** ***p<.01; *p<.05*

Analyses have been conducted using SPSS 25.0. Code, data, and pre-registration are available via https://researchbox.org/1540&PEER_REVIEW_passcode=ZGFTER

Results

Peer Bonding and Design on Voice Shopping (H1)

We hypothesized that the more people report peer bonding the more they have stronger voice shopping intentions after seeing the emotionalized (vs. normal) interface (H1). To test this prediction, we regressed all three voice shopping indicators separately on peer bonding (centered), design (emotional versus normal design), and their interaction. We found a main effect of peer bonding on voice shopping attitude, indicating that people with higher peer bonding values are more likely to have a positive attitude towards the conversational AI, β=.21, $t(403)=4.25$, $p<.001$. Furthermore, higher peer bonding was also related to higher perceived shopping benefits, β=.44, $t(403)=9.80$, p<001, and intentions, β=.42, $t(403)=9.30$, p<001). This suggests that peer bonding predicts the positive perception of voice shopping benefits and intentions to engage in voice shopping.

Furthermore, there was a main effect of emotional design on attitudes toward voice shopping, β=-.12, $t(403)=-2.40$, $p=.017$. Attitudes towards voice shopping were less positive in the emotionalized condition (M=5.57, SD=1.12) than in the normal condition (M=5.85, SD=1.08). This effect was not found for both other dependent variables, both ts<.7, p>.5.

Most important for the hypothesis, no significant interaction effect was found between the predictor variables, all βs<.08, all $ts(403)$<1.5, all ps>.101. We, thus, did not find evidence supporting H1, predicting that the more people report peer bonding, the more they have stronger voice shopping intentions after seeing the emotionalized interface.

Authority Ranking and Design on Voice Shopping (H2)

Analogous to H1, we predicted that the more people report authority ranking, the more they have stronger voice shopping intentions after seeing the normal (vs. emotionalized) interface. Thus, we regressed the three indicators of voice shopping perception on authority ranking (centered), design (emotionalized vs. normal design), and their interaction in separate multiple regressions (for results, see Table 15). We did not find a relation between authority ranking and any of the voice shopping variables. The main effect of design, as reported above, was present for attitudes but not for benefits or intentions. However, the interaction approached significance for attitudes and was significant for benefits (but not for intentions). These results call for additional exploration.

Table 15

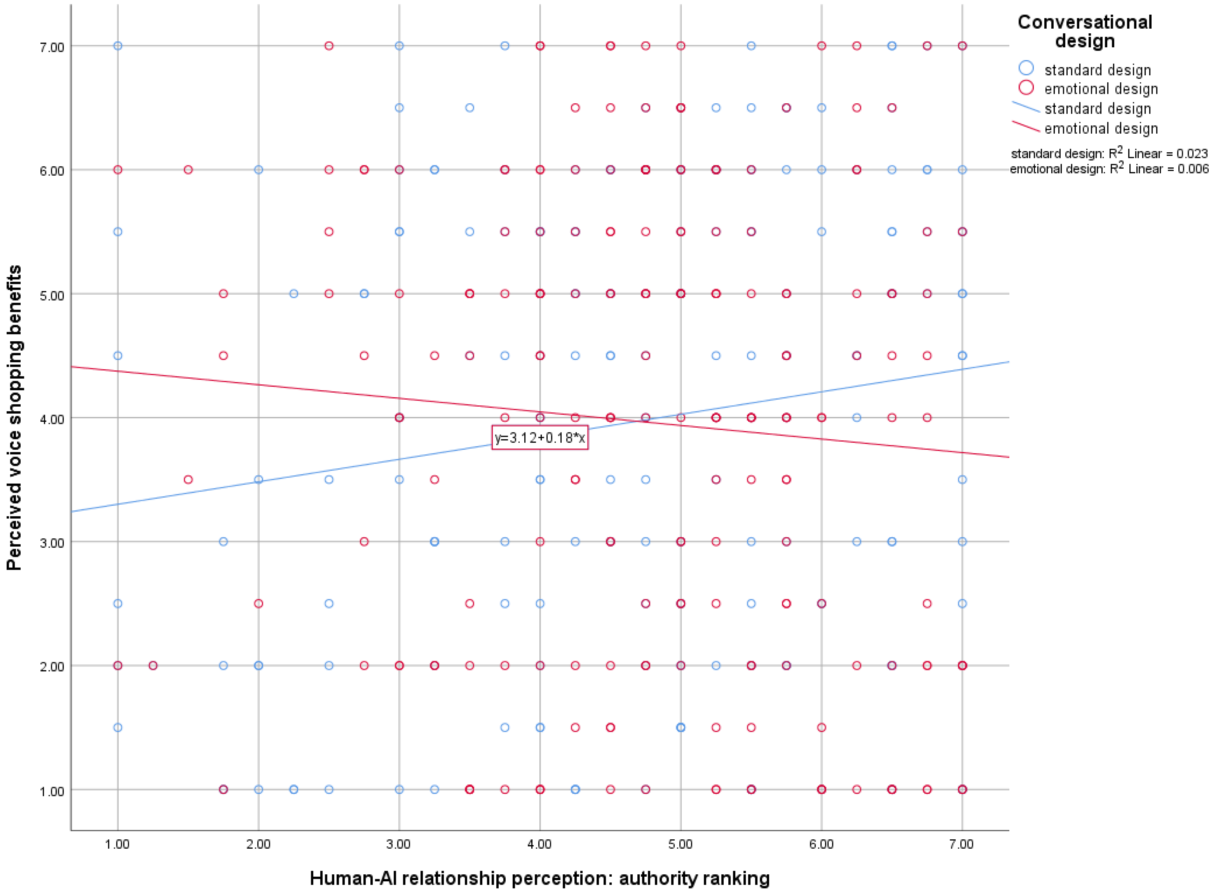
Multiple Regression Analyses with Authority Ranking, Condition, and Interaction as Predictors of Voice Shopping (N = 407)

Dependent Variable	β	*T*	*p*
Shopping Attitude			
Authority Ranking	.06	1.15	.251
Emotionalized Design	**-.13**	**-2.60**	**.010**
Interaction	-.10	-1.90	*.058*
Shopping Benefits			
Authority Ranking	.03	0.56	.58
Emotionalized Design	-.01	-0.10	.924
Interaction	**-.11**	**-2.24**	**.026**
Shopping Intention			
Authority Ranking	.02	0.42	.678
Emotionalized Design	.01	0.25	.798
Interaction	-.10	-1.60	.112

We conducted a simple effects analysis according to (Aiken & West, 2010) to resolve the interaction effect. Higher authority ranking correlated with more perceived shopping benefits in the neutral condition, β=.14, *t(403)=2.12*, *p*=.034, but not in the emotional condition, β=-.08, *t(403)*=-1.12, *p*=.265. Simple effects analyses for the non-significant interactions on attitudes yielded the same pattern (neutral condition: β=.15, *t(403)=2.32*, *p*=.021; emotional condition: β=-.04, *t(403)=-0.50*, *p*=.620. In other words, people who report more authority ranking may perceive greater voice shopping benefits – but *only* for the normal design (see Figure 2). These findings provide support for H2 regarding two of the three indicators we assessed.

Figure 2

Relationship between Authority Ranking and Perceived Voice Shopping Benefits Dependent on Design Condition



Research Question: Is the Effect of The Interface Design Contingent to Market Pricing? (RQ1)

To explore the role of market pricing, we repeated the same set of regression analyses. We found that higher market pricing was related to more positive attitudes towards voice shopping, β=.19, *t*(403)=3.97, *p*<.001, stronger perceived shopping benefits, β=.23, *t*(403)=4.80, *p*<.001, and stronger voice shopping intentions, β=.20, *t*(403)=4.05, *p*<.001. Importantly, these relations were not contingent on the design condition, all |*t*|s<1, all *p*s>.410. Hence, market pricing positively predicted the attitudes towards voice shopping, but our manipulation did not assert an influence on this relationship.

Discussion

This study explored whether emotional conversational design is evaluated differently in the context of voice shopping depending on consumers' perception of their human-AI relationship (Tschopp et al., 2023). We compared the standard Alexa interface to an emotionalized interface. In addition, we considered the three dimensions of perceived human-AI relationship: peer bonding, authority ranking, and market pricing.

Peer Bonding

Against our expectations, we did not find evidence that individuals with a high perceived peer bonding would have increased shopping intentions after watching the video with a more emotional conversational design interface (H1). However, users with a stronger peer bonding towards their conversational AI perceived voice shopping more positively (i.e., a more positive attitude towards the interface, more expected benefits, and stronger intentions to use voice shopping). These results can be seen as evidence that the standard interface is already sufficient to attract users high in peer bonding to use voice shopping. A different picture occurred for authority ranking.

Authority Ranking

Particularly, those users in a master-servant relationship perceived greater voice shopping benefits – but only with regard to the normal conversational design. This effect was diminished on two of three indicators after watching the video with the emotionalized conversational design interface (partially supporting H2). In particular, those high in authority ranking perceived voice shopping as less attractive after seeing the emotional (compared to the standard) interface. Thus, for those individuals who view their conversational AI as a subservient assistant, an emotional design incorporating elements such as emojis may actually impede the voice shopping experience as a whole. The attempt to establish a friendship-like connection with the user is not appreciated. They even seem to cause feelings of resentment, as evidenced by the candid comment made by one of the participants: "Don't call me buddy!" This response indicates a clear inclination to uphold a certain level of separation, most likely a hierarchical distance, wherein they perceive themselves as masters and the voice assistant as a servant.

On the one hand, this adds to the literature exploring the feelings of eeriness and perceptions of threat (Stein et al., 2019). Additionally, our study supports the idea of adopting more servant-like designs for voice shopping usage, aligning with the findings and recommendations put forth by Hu et al. (2021) or Tassiello et al., (2021). It appears crucial for designers to carefully evaluate the concept of individual-technology-fit for these users. Especially considering the prevalence of this particular human-AI relationship mode among the majority of people, as indicated by several studies conducted by Tschopp et al. (2023).

Market Pricing

In Chapter 4, I argued that market pricing might hold particular relevance as predictor of voice shopping decisions, as users may perceive the artificial shopping assistant as a quasi-sales agent. However, the hypothesis was not supported by the findings. In contrast, our results indicate that market pricing may play a role in positively influencing voice shopping, which aligns with the notion that voice assistants are “active partners in the decision-making process rather than mere order takers” (Bellis & Venkataramani Johar, 2020; Dellaert et al., 2020). This discovery underscores the value of adopting a more differentiated relational approach and calls for further research in this area. For market pricing, as for peer bonding, the correlations with the evaluations of voice shopping were not moderated by the interface design manipulation in our study. Across the board, this manipulation had a surprising effect.

Emotionalized Interface

We observed that the emotional design was detrimental to participants’ attitudes towards voice shopping. This effect is interesting, given that our pretest found significant results that the manipulated Alexa was perceived as friendlier and more welcoming. This finding cast doubts on the idea of aligning a more emotionalized conversational design with the user’s perception as peer-like. Users with a strong tendency for peer bonding, who already perceive their interaction with Alexa as friendly, may find the familiar standard design of Alexa satisfactory. For these users, maintaining the status quo might even be preferable, and they could be eager to continue their existing relationship with Alexa as it is. This reflects a similar pattern observed in recent instances where users expressed strong negative reactions to changes in Replika’s behavior, a widely used AI companion chatbot (Cole, 2023). Thus, the significant negative impact on attitude could be caused by altering Alexa’s linguistic style and aligns with previous research suggesting that emotional conversational designs might be perceived as fake or insincere (Chattaraman et al., 2019).

The negative influence of the emotional design on attitude towards the system did not come with downstream effects on perceived benefits and subsequent voice shopping intentions. These findings somewhat contradict the studies mentioned in the theoretical background, highlighting the value of emotionalized design (e.g., Beattie et al., 2020).

Implications for Conversational Design Practice

Our outcomes carry three significant implications for design practice. Firstly, there appears to be no added value in making conversational interfaces more emotional for individuals who perceive their AI as more peer-like and even more so for people who see their conversational AI as servants. Secondly, designers must carefully consider the potential negative effects of making changes to a voice interface that people have already become accustomed to. Such alterations should be approached with caution and careful consideration. Finally, our results provide an economic incentive to refrain from further emotionalizing conversational AI, in addition to the broadly discussed ethical argument that emotional designs can potentially manipulate consumers (e.g., oversharing, addictive shopping behaviors, etc., see Véliz, 2023). Considering these implications, designers of voice shopping assistants should be cautious about incorporating (excessive) emotionalization into conversational AI. Instead, we suggest prioritizing usability, practicality, and user preferences in voice shopping scenarios.

Notably, this is in stark contrast with the current shift towards more interactive and emotionally nuanced exchange-based approaches. In practice, this trend is reflected by the noticeable market trend for AI companion bots (Panetta, 2022), indicating a growing demand for establishing social connections with conversational AI. An example of this shift can also be observed with Microsoft's new digital assistant, the "Co-Pilot". The marketing narrative moves away from the previously subservient role of Cortana (Microsoft's previous digital assistant, now out of service) and positions itself as a more collaborative and interactive co-worker (Stallbaumer, 2023).

Strengths, Limitations, and Future Research

Our study contributes to the current literature in three ways: First, we contribute to the emerging stream of research on human-AI relationship perception (Pentina et al., 2023b) by studying business applications in an experimental setting. Second, we extend knowledge and findings regarding personalized conversational design in human-computer communication: the "art" of developing the language style of conversational AI is an enduring hot debate in human-computer interaction studies and e-commerce literature (Hu et al., 2022), and recently, in AI ethics (Véliz, 2023), too. Finally, our study furthers psychological research in the emerging field of voice shopping, focusing on the human-AI experience rather than the technicalities of the voice shopping process.

This study is not without limitations and thus offers exciting opportunities for future research. First and foremost, it is essential to acknowledge that our study's videos solely depicted voice shopping via Alexa (Amazon's conversational AI), and we only used a limited set of videos with one particular manipulation of the interface design. This may restrict the generalization of findings to other voice shopping systems and modifications in the interface. Therefore, future research should encompass a broader range of conversational AI systems and modifications to facilitate comparative analysis and a more comprehensive understanding of the results.

The chosen linguistic cues have been based on prior research (e.g., signaling empathy, identity, or enhancement with emojis). However, it is not entirely clear how well they have been consciously recognized despite our successful pretest. Open questions remain regarding the exact thresholds at which deliberate humanization of conversational design becomes unsettling or creepy (Feine et al., 2020; Mori et al., 2012; Stein et al., 2019; Stein & Ohler, 2018). Future research can expand these ideas and/or must further differentiate: For example, the number, size, length or looks of anthropomorphic cues, or include embodied conversational AI (e.g., avatars), including specific attention checks.

On a final note, studying human-AI interaction from a relationship perspective seems fruitful. However, further validation is required. Future research should test the multidimensional approach primarily in experimental settings, to make causal statements. Additionally, future studies could contribute to the current understanding of human-AI relationship perception by using different systems than off-the-shelf smart speakers, different samples to explore cultural variation and test the relationship dynamics over time via longitudinal designs.

Conclusion

Should emotional conversational design match the human-AI relationship perception to promote voice shopping? Indeed, our study did not yield evidence to suggest that an emotionally enhanced design in conversational AI provides competitive advantages. However, our data indicate that an emotional conversational design deters those individuals who view their conversational AI as rational servants, thus reducing their willingness to engage in voice shopping. Our study offers additional support to the notion that anthropomorphic cues, specifically emotional conversational design, may not be uniformly well received by all consumers. This validation underscores the significance of the individual-technology fit theory (Hu et al., 2022) in understanding and catering to consumers' varying preferences and responses to conversational AI. Against this background, our findings and market observations, it is imperative for research and practice to remain vigilant and attentively monitor emerging trends and new discoveries. This will facilitate informed decision-making and adaptation to evolving user needs and preferences in the field of shopping with conversational AI. In a nutshell, the results validate the value of the human-AI relationship perception (Pentina et al., 2023b;

Tschopp et al., 2023) from a fit-perspective and offer insights for human-machine communication researchers and practitioners. Last but not least, in addition to ethical arguments, the results further provide an economic argument and much-needed incentive against (over-) humanized design.

Chapter 7: General Discussion

The overarching goal of my dissertation is to provide a better understanding of human-AI relationship perception. Because humans often see and interact with machines similarly to humans (Nass & Brave, 2005) and due to the outstanding progress in conversational AI (Nayyar, 2023), scholars are currently discussing the possibility of users forming relationships with AI systems and urgently call for further research (Guzman & Lewis, 2020; Marriott & Pitardi, 2023; Pentina et al., 2023b). I answered this call in two steps. First, I examined human-AI relationship perception by repurposing the RMT by Alan P. Fiske (1992). Having developed a framework to assess perceptions of the human-AI relationship then enabled me to empirically investigate its significance in an applied setting, more specifically in the context of voice shopping experiences (i.e., involving conversational AI like Alexa for home shopping purposes).

The first empirical Chapter 2 primarily served as an initial step to test whether and how the application of Fiske's RMT is a viable approach to studying human-AI relationship perception. After adapting the Modes of Relationship Questionnaire (see Table 1) to the context of conversational AI (see items Table 5), the results of the conducted factor analyses among frequent users in two studies suggest a three-dimensional structure: authority ranking, market pricing, and peer bonding. Compared to the original four-dimensional structure, authority ranking, and market pricing remained, but the more emotional dimensions communal sharing and equality matching merged into a new dimension I named *peer bonding*. The results consistently suggest that users perceived the human-AI relationship as more characterized by authority ranking and market pricing than by peer bonding. Surprisingly, authority ranking barely correlated with relevant variables of system perception or user characteristics, whereas market pricing and peer bonding did. Notably, peer bonding was strongly associated with human-like variables (e.g., anthropomorphism), excluding trust, which was only related to market pricing (see Table 6 for details).

By evaluating the validity and reliability of the measures, the results of the first empirical chapter indicate that the developed instrument is well suited to assess human-AI relationship perception. Secondly, the multidimensional approach enhances the existing research landscape, which is predominantly characterized by dichotomous approaches. Introducing three modes of relationship perception (i.e., authority ranking, market pricing, and peer bonding) provides nuanced insights into how users perceive their relationship with conversational AI. Thirdly, the identified strong associations between market pricing and peer bonding and system perception variables may indicate the importance of (perceived) agency.

This is in line with prior research pointing out evolving tensions between human and machine agency (Sundar, 2020). In other words, the perception is that there is “more” to the system and that users and machines are more on par with each other. Future research should further investigate the potential central role of perceived machine agency within human-AI relationship perception.

In light of these observations, the empirical study in Chapter 4 explored the role of perceived human-AI relationships in a practical context, specifically focusing on differentiated voice shopping intentions. Examining experienced voice shoppers, I aimed to investigate whether the perceived human-AI relationship influences the types of products users purchase. I grounded the rationale for hypothesis development on the Elaboration Likelihood Model (Petty & Cacioppo, 1986), distinguishing between low- and high-involvement (simple and complex) products. Different thought processes are typically involved in these shopping scenarios, and as a result, the role of the sales agent may vary. For example, delegating a simple shopping decision versus deliberately making a complex purchase decision “together”. While calculative decision-making (i.e., market pricing) did not play a role when forming voice shopping intentions, the perception of the relationship as master-servant (i.e., authority ranking) was associated with low-involvement shopping. An additional analysis exploring the role of hedonic, utilitarian, symbolic, and social benefits further suggested the importance of the socio-emotional dimensions in all facets of voice shopping, which reinforces the claims made. The results further emphasized the potential predictive value of peer bonding. Thus, socio-emotional elements for voice shopping matter, especially for high-involvement products, which is a novel finding in the field of voice shopping.

In sum, the study replicates the results of the initial study in Chapter 2, underlining the considerable value of the multidimensional approach to human-AI relationship perception. Secondly, results contribute to the emerging field of voice shopping, providing practical implications where different perceptions are associated with different voice shopping intentions. Thirdly, the strong associations between peer bonding and voice shopping intentions might motivate developers to favor a more social interface over a more rational one. However, from a policy perspective, designing emotionalized conversational interfaces garnered concerns about being deceptive (Véliz, 2023). The nature of the study design does not allow for drawing causal conclusions about the link between the conversational design and the human-AI relationship perception. Consequently, this was the core of the final study.

In Chapter 6, I investigated the interplay of the human-AI relationship perception and the conversational design of the AI-system (i.e., Amazon’s Alexa) to study voice shopping (i.e., attitude, perceived benefits, and intentions). In order to do so, I used an experimental design that manipulated the socio-emotional elements of the conversational AI. The manipulation included more casual, friendly language and humanized transcript on Alexa's screen, such as expressions of empathy and emojis. Building upon a human-AI fit concept (Liu et al., 2011), I proposed that users' attitudes and engagement in voice shopping depend on the match between the conversational design and their perceived human-AI relationship. The results did not provide evidence for the notion that an emotionally enhanced conversational design confers any advantages concerning voice shopping, neither for those who relate as peer bonding nor for users with other relationship perceptions. Interestingly, the data demonstrated that the emotional conversational design discouraged individuals whose relationship perception of the conversational AI is guided by authority ranking.

As with the prior studies, the multidimensional approach to human-AI relationship perception has again proven its relative stability. Highly relevant for practitioners, I could experimentally show that a more emotional conversational design interacted with authority ranking and influenced voice shopping, rendering the concept of individual-technology-fit an important aspect to consider for practice and future research. In this study, a case can be made that such an emotionalized design is economically impractical, next to the standpoint that such designs are normatively speaking undesirable. Nevertheless, I must acknowledge that the depiction is not yet entirely comprehensive. Thus, limitations (see section p. 92) need to be addressed in future research to get a better picture of the causal relationships between conversational design, human-AI relationship perception, and the behavioral outcomes.

In summary, my dissertation provides insights into how users perceive their relationship with conversational AI and how these perceptions influence behavior. The first step was to assess the viability of repurposing the RMT for human-AI interaction, more specifically, conversational AI. By employing an exploratory approach, we addressed variables of system perception and user characteristics to better understand the perception of the human-AI relationship in the context of traditionally studied variables in HCI. I was able to establish and replicate that the measures employed to evaluate human-AI relationships are highly appropriate and, furthermore, useful in the HCI context. In short, the empirical findings present a three-dimensional structure of human-AI relationship perception, which was stable and mostly independent of user characteristics, featuring distinct associations with HCI-relevant variables of system perception. The relational structure was proven to replicate in the subsequent studies (Chapters 4 and 6), enabling me to employ a confirmatory approach. I tested the practical value of the human-AI relationship perception in the context of voice shopping, providing nuanced findings. Authority ranking was associated with low-involvement shopping, while peer bonding predicted both low- and high-involvement shopping, indicating the importance of socio-emotional elements in voice shopping. Moreover, my dissertation underscores the significance of considering a conversational design as it may interact with human-AI relationship perceptions during the voice shopping experience. This provides direct practical implications for system designers and valuable insights for policy-makers, as the emotional conversational design was shown to be disadvantageous for users in an authority ranking relationship mode.

In Table 16 below, I want to summarize the empirical findings of my dissertation before discussing the strengths and limitations, the implications for theory and practice, including ethical considerations.

Table 16

Human-AI Relationship Perception: Study Highlights

	Key Claims	Key Findings		
Across Studies	Human-AI Relationship Perception: 1) Three dimensions 2) Stable and independent 3) Distinct associations with HCI relevant variables	**Authority Ranking** hierarchical owner-assistant (majority of users)	**Market Pricing** non-hierarchical exchange-based (many users)	**Peer Bonding** equal peer-like (few users)
Chapter 2	4) Enrichment of dichotomous approach	not informative for system perception or user characteristics	moderately informative regarding system perceptions	strongest predictor of system perception and user characteristics
Chapter 4	5) Socio-emotional elements matter in voice shopping	predicts low-involvement shopping	no evidence found	predicts high- and low-involvement shopping
Chapter 6	6) Emotional conversational design disadvantageous for majority of users	discouraged from voice shopping with emotional design	no interaction with emotional design	no interaction with emotional design

Strengths

The central query of my dissertation revolved around the value of the proposed human-AI relationship approach for academic research as well as for practitioners in the chosen field of voice shopping. The multidimensional approach toward human-AI relationship perceptions taken in my dissertation was proven to be valuable across four studies. For scholars, the framework may serve as a robust tool, facilitating the replication of findings and its adaptable application across diverse contexts with diverse AI systems. Prior research used unidimensional variables (e.g., perceived warmth or psychological distance, Pitardi & Marriott, 2021), proxies of relationship perceptions (e.g., trust, Glikson & Woolley, 2020), qualitative approaches (Skjuve et al., 2021), or relied on broad role categorizations that are rather techno-centric (e.g., the AI system is designed as a friend, Kim et al., 2019). My dissertation adds to the existing literature by taking a 1) direct, 2) multidimensional, 3) human-centric, strongly theory-driven approach to relationship perception, presumably largely independent of a) the individual user characteristics and b) applicable across various AI systems.

The latter two make the approach particularly useful for professionals in the field. Here, I focus on conversational system designers and policymakers who may benefit from the knowledge gained throughout the studies. Considering the multidimensionality of how users perceive their relationship with the AI system independent of the designers' intentions enables a more fine-grained understanding of which factors influence the interaction outcomes between users and AI systems. In particular, design decisions may influence interaction depending on how they perceive the relationship. The distinct chapters offer valuable insights and practical recommendations for conversational system designers and business practitioners navigating this domain as they provide nuanced findings (e.g., sales channel decisions, see Chapter 4). By acknowledging that the intention of the designer does not necessarily mirror the perception of the user, conversation designers may experience a degree of dissatisfaction as they do not wield absolute control over users' perceptions. The extent to which designers can influence or stimulate these perceptions to facilitate specific outcomes remains an unchartered terrain where more experimental studies are necessary in order to make causal statements.

What sounds like a weakness is, however, an opportunity to combine forces with other disciplines for a better outcome for all. This is where AI policy researchers come into play, who prioritize the establishment of responsible AI intended to serve humanity, transcending the capitalistic interests of AI creators solely. Responsible AI is grounded in empirical knowledge about human-AI interaction. Thus, understanding that users *do* perceive multiple modes of relationships already holds value. Because the general concept of human-AI relationships likely inherits similar ethical issues that have been raised in AI ethics regarding humanized design in the past years (Véliz, 2023). These research streams can and should mutually enrich each other. It might not be normatively desirable to trigger certain (or any) human-AI relationship perceptions. Hence, in each study, we strive to incorporate insights that hold significance for policymakers in the field of AI in order to inspire more interdisciplinary research in the future. The hope is to create AI systems where the outcome of the interaction is beneficial to those who create the system (e.g., selling products or providing advice with a conversational AI) while also being benevolent and advantageous to the user.

In a nutshell, my dissertation possesses two eminent strengths within the broader realm of HCI and AI system development. Firstly, it establishes a robust framework for the human-AI relationship that is valid, reliable, and replicable. Secondly, its multidimensional approach provides nuanced insights with implications for both theory and practice. The focus extends beyond mere effectiveness to encompass the responsible implementation of conversational AI systems.

Limitations

When applying these findings to follow-up research or incorporating them into practical applications, it is crucial to acknowledge and work out several limitations. The empirical chapters delved into specific limitations for each study, and I aim to emphasize some overarching issues identified in the individual chapters. The last two paragraphs place a heightened emphasis on the instrument, given that it was developed, tested, and utilized in an applied setting for the first time, warranting increased attention.

Methodological Issues

With regard to the sampling method, all studies may suffer from limited generalizability across populations due to limited cultural variation and due to employing the same recruitment strategy. All studies were online surveys for which participants were recruited over the platform Prolific. The validity of responses on such platforms is a topic of ongoing debate among scholars. Some critics argue that these platforms are not taken seriously by participants, who are primarily motivated by financial incentives rather than genuine engagement. However, others contend that the data obtained from these platforms can be as valuable as those derived from traditional participant recruitment methods (for a broad discussion, see Anwyl-Irvine et al., 2021). Along with the chosen online survey strategy comes another relevant limitation, namely the reliance on self-reported measures to assess shopping intentions rather than actual voice shopping data. The debate among scholars is steering toward the notion that intention serves as a reliable indicator of actual behavior, even in the presence of significant gaps between attitudes and behavior (Pieters & Verplanken, 1995). Further research is needed to understand the implications of the two differing viewpoints fully. Both areas undoubtedly warrant inclusion in future research endeavors.

About the data analysis, it must be clearly stated that the studies presented in Chapters 2 and 4 provide only correlational data, which do not allow for any causal conclusions. This is particularly relevant for the study in Chapter 4, identifying human-AI relationship dimensions as voice shopping predictors. The associations observed may be influenced by third variables, which I have not included in the analysis. For instance, the reasons for the strong predictive value of peer bonding in both high- and low-involvement shopping could be influenced by the personal situation. For example, the results of Study 2 showed that education was positively correlated with peer bonding. Someone with more knowledge or higher education might avoid higher shopping risks and would rather not delegate decisions to the shopping assistant due to a greater understanding of potential financial or other technical risks. Another third variable that likely may have influenced voice shopping decisions, as well as relationship perception, is the performance experience with the system, more specifically, the voice shopping process. Individuals with a history of successful transactions, particularly those involving intricate purchases, probably exhibit a heightened inclination to continue voice shopping for high-involvement products and may think differently of the conversational AI regarding its agency and, thus, the respective relationship perception.

Furthermore, the studies I conducted are cross-sectional analyses. The results do not inform our understanding of how human-AI relationships change over time, including the initial development of these perceptions. The study of human-AI relationship perception within our framework implies that users already have at least some experiences with the AI system in question. Consequently, the framework cannot be used to investigate the initial development of these perceptions, which is an important area in the human-AI relationship research landscape (Croes & Antheunis, 2021). Rather, it must be clear that the instrument is oriented towards continuance intentions and thus of considerable value in contexts where technology is already adopted/in use, such as Alexa users and voice shopping on Amazon.

In terms of the conducted experimental study in Chapter 6, it is noteworthy that the interpretation of the results may be affected in several ways by the nature of the stimulus material. Primarily, the study design omitted the examination of an authentic interaction, relying instead on participants watching videos of a voice shopping interaction. This limitation compromises the external validity of the experiment. More significantly, the discouragement experienced by individuals in an authority ranking relationship might be influenced by their aversion to changes in their customary interactions with Alexa. In other words, the users simply did not like the “new” Alexa just because it is more sociable, and they probably prefer it the way it is. Therefore, it is imperative to replicate the experiment using alternative stimuli, specifically 1) real interactions and 2) different conversational AI systems to which users are not accustomed to.

Human-AI Relationship Scale Assessment

Finally, I would like to evaluate the limitations regarding the assessment of the perceived human-AI relationship and how to address them. Since the questionnaire was used for the first time, the aim was to stick as close to the original Modes of Relationship Questionnaire (MORQ) as possible. Despite undergoing an extensive review process to make a valid and reliable assessment possible, some thoughts should be taken into account when interpreting the results. Minor concerns are associated with certain items, where reconsideration of wording may be necessary, in particular, the peer bonding and market pricing scale. In each study, we excluded approximately one item after the factor analyses, which was rectified in the subsequent study. Nevertheless, minor changes might already complicate comparing results across studies, potentially limiting generalizability. This could be an indication that the scale may benefit from further refinement. Furthermore, it might be prudent to consider shortening the scale to maintain participants' attention (currently, the duration with 17 items takes approximately 5 minutes).

Of greater concern is the more conceptual question of whether we can "trust" our human-AI relationship measure (as in Chita-Tegmark et al., 2021)? The concept of thinking about the relationship to users' conversational AI system is presumably new to many. The pretests to develop the questionnaire have shown that users have to perform some sort of "mental gymnastics to be able to reason" (cf. Chita-Tegmark et al., p. 96, 2021) about how they rate the items provided. Despite testing the applicability of the questionnaire thoroughly in the pretest, it might be beneficial to reconsider the scale with a greater sample size than we used in the pretest. An idea for future research in order to reaffirm and/or refine the questionnaire is to add applicability options to the questionnaire, as proposed by Chita-Tegmark et al. (2021). Each item shall be expanded to include an additional scale point, enabling users to indicate that the respective item is not applicable to this conversational AI or conversational AI in general. Ideally, this will be supplemented with the option to provide an explanation in a text field.

In sum, caution is advised when extending the results to broader applications and users, and hastily deriving practical implications. Generalizability is limited, and methodological concerns persist. However, these challenges, inherent in pioneering approaches, can be effectively addressed, presenting a positive aspect for future research.

Agenda for Future Research

My dissertation adopts a strongly theory-driven approach with implications relevant to academic research in psychology and other related disciplines. Nevertheless, my inclination is to characterize the dissertation also as oriented toward the value of understanding the nuances of human-AI relationship perception for practitioners in the field of AI. Overall, I am confident that the research questions could be answered with some certainty, limitations can be effectively addressed in both research and practice, and the implementations for practice hold considerable value. The conclusions drawn from my dissertation carry implications for theory. Specifically, it is posited that RMT can be reasonably extended to encompass relationships involving non-human entities, in this case, conversational AI. However, according to our findings, the perception of human-AI relationships involves one dimension less than what is proposed by RMT. Based on the strengths and limitations mentioned above, I propose areas for research and practice in more detail, focusing on human-AI relationship research in the following sections rather than voice commerce. Implications for voice commerce are elaborated in the empirical Chapters 4 and 6.

The approach and findings provide a number of contributions to the emerging research on human-AI relationship perception and the greater landscape within the CASA paradigm. Most significantly, I offer a theory-driven, methodological framework for researchers to examine how users perceive their relationship to conversational AI. However, situating the results within the greater landscape is challenging. Firstly, due to the great variety of questions, methods, and (AI-) systems investigated thus far (Pentina et al., 2023b). Second, due to ongoing ontological debates on the “nature of AI”. Despite the increase in social interactions between humans and conversational AI (Brandtzaeg et al., 2022), the reality of such relationships, such as human-AI friendships with chatbots, might just be a temporary, illusionary by-product of the mind as put forth by Turkle (2011). In reality, conversational AI stands as a constructed commercial system, which is a mimicry of human behavior, and it falls short of fulfilling the defining characteristics inherent in human interactions and social relations, such as friendships (Brandtzaeg et al., 2022; Evans et al., 2023). Consequently, scholars deem similar narratives rooted in anthropocentrism, such as prevalent discussions surrounding human-AI collaboration and teaming literature, significantly misaligned (Evans et al., 2023).

Call 1: Focus on Similarities and Differences within Anthropocentric Approaches. Despite the ontological critique, a number of scholars investigating conversational AI in the new media landscape have adopted such an anthropocentric approach recently, including the present work. While first guided by the mere translation of human social interaction to human-AI interaction, it is now increasingly focused on identifying the similarities and differences instead of merely equating psychological constructs. Arguably, the aim is to find and establish new rules (a “stronger” CASA, Gambino et al., 2020), which are guided by human interactions but differ from them (Brandtzaeg et al., 2022) due to the limited interaction opportunities with conversational AI. Our findings align with the latter approach, thereby contributing to relevant research within the "stronger" CASA paradigm (Gambino et al., 2020), as follows: I found similarities that can be (almost) equated to the original RMT (i.e., authority ranking and market pricing) but also differences. The new dimension of peer bonding was identified, integrating the two dimensions of communal sharing and equality matching, rendering the perception of off-the-shelf conversational AI less socio-emotional. However, these observations might change, presumably, especially for peer bonding. With greater integration of GPT models in off-the-shelf conversational AI, the systems may become more communicative and able to act more humane (e.g., more simulated empathy, longer conversations, better working memory, and so forth, see Brandtzaeg et al., 2022) which might increase the prevalence of peer bonding. Alternatively, people might become accustomed to interacting with conversational AI over the next decades, diminishing social reactions (Heyselaar, 2023) and thus decreasing peer bonding modes of human-AI relationships. On a final note, the role of the ever-increasing general performance as rational may need to be reconsidered in light of a recent report where the old-fashioned Eliza (Weizenbaum, 1966) outperformed modern generative models (GPT3.5 and 4) in an online Turing test (Jones & Bergen, 2023), convincingly posing as human. Future research should identify what causes those similarities and differences in perception, taking into account personal traits (e.g., loneliness, social motivation, as suggested by Pentina et al., 2023b or Marriott & Pitardi, 2023) or contingencies such as country, language, or cultural background.

Call 2: Investigate the Interaction of Design Attributes and Human-AI Relationship Perception. Thus, many possibilities open up for researchers investigating various attributes of the conversational AI, which are machine- and/or human-like (Sundar, 2008), to better understand human-AI relationship perception (including the influence on behavioral outcomes). In voice commerce, our studies support the role of hierarchical power over the conversational AI (as found by Hu et al., 2022) to increase voice shopping intentions, but only for those willing to buy low-involvement items (not for high-involvement items). Peer bonding was shown to play a role in all facets of voice shopping, supporting the importance of a non-hierarchical, emotional relationship perception in voice shopping (e.g., Rhee & Choi, 2020). But to be clear, the results do not give a comprehensive understanding of how the human-AI relationship perception interacts with design attributes. In simpler words, the intended design of the conversational AI as a servant to promote the power experience does not necessarily translate to users perceiving it that way. However, in line with current scholarship in HMC (Guzman et al., 2023), it is plausible that the human-AI relationship perception will play a role in the context of conversational design of the system because communication is the central medium of relationship building (cf. Beattie et al., 2020). Thus, more experimental data is needed to investigate this issue. Future research should focus on conversation design-relevant variables that have been well-studied in the past beyond correlational data (Guzman et al., 2023). This includes information such as gender or other visual information of the persona design of the conversational AI (e.g., is the AI system embodied as an avatar? Where is the hardware stationed?). Moreover, substantial attention ought to be directed toward the systematic variation of verbal information encompassing diverse conversational design strategies, potentially even including the way information is exchanged between the user and the conversational AI. The overarching question waiting to be answered here is: what design strategies further guide users' perception of the relationship? Furthermore, what is the impact of both predictors on behavior?

Call 3: Investigate Various Conversational Systems and Contexts. As previously emphasized, the framework of human-AI relationships proves particularly beneficial for the methodical examination of already existing conversational AI systems (i.e., with an existing user base). I encourage fellow researchers to explore its applicability in various contexts, such as in driving assistants or mental health applications. The versatility of the human-AI relationship framework suggests that its utility extends beyond artificial intelligence systems. It could be relevant for other conversational systems like click bots. In fact, this could be investigated with any other artifact, which may fall within the scope of being perceived as "relational entity" or entities worthy of relational attributions.

Call 4: Include Philosophical and Ethical Issues. Last but not least, this and future research must deal with philosophical and ethical challenges. The distinction between humans and "intelligent systems" becomes increasingly blurred. The inquiry into how individuals perceive their relationship with AI raises fundamental questions that must be addressed. We seem to garner much knowledge regarding the development of human-AI relationships perception (Brandtzaeg et al., 2022; Skjuve et al., 2021), but knowledge about what does not constitute or lead to such human-AI relationships or what breaks them is rare (with few exception, see Croes & Antheunis, 2021 and Lopatovska & Williams, 2018). This also comprises ontological considerations, which may affect the methods of investigation. Evan et al. (2023) suggest that an anthropocentric approach is insufficient for describing interaction processes While conversational AI may possess extensive knowledge and autonomy, the depth of autonomy in relationships goes beyond mere functionality, involving meaningful decisions such as entering or leaving a relationship. This autonomy is not reasonably ascribed to conversational AI. Why does that matter? Accepting that users may have *real* relationships with AI comes with profound challenges. The main challenge arises from the claim that human-AI relationships are inherently unidirectional (Evans et al., 2023). Users exercise complete control over the AI system by determining its appearance, character, and actions. Additionally, the chatbot remains available 24/7 at the user's command. The intricate power imbalance that users may become accustomed to and extrapolate to human interactions underscores the need for a thorough examination of psychological and societal implications (e.g., increasing gender inequalities, Evans et al., 2023). Indeed, the investigation of human-AI relationship perception, involving the application of psychological theories to human-AI interaction, might inadvertently contribute to the problematic aspects, as also acknowledged by Guzman and Lewis with regard to human-machine communication research (2020).

Either way, unless human evolution manages to resist falling for the allure of human-like attributes, which is an unlikely scenario, users' human-like perception of machines may always play a pivotal role in this dynamic. Consequently, the employed anthropocentric approach focused on identifying the similarities and differences instead of merely equating psychological constructs, arguably is the current gold standard for exploring how humans perceive their relationships and other social constructs such as trust, friendship, collaboration, and so forth. Nevertheless, I cannot disregard the prospect of the emergence of entirely novel paradigms beyond (the stronger) CASA, potentially involving a conceptual shift towards AI as a distinct entity. On the other hand, this evolution could also be transient, merely a temporary phase until societal familiarity with AI parallels that observed with other tools in our technological landscape (cf. Heyselaar, 2023).

In a nutshell, the future for human-AI relationship scholars taking an anthropocentric approach holds exciting opportunities, highlighting the value of adapting psychological theories to human-AI interaction (1), combining it with the systematic exploration of (2) system design attributes, (3) various conversational AI systems in various applied settings, while also (4) addressing philosophical and ethical challenges.

Agenda for Responsible Practice

Although this dissertation placed a significant emphasis on theory, with a focus on fundamental research concerning human-AI relationship perception, its orientation remained rooted in practical motivation: what is the practical value of better understanding how humans perceive their relationships with conversational AI? Throughout studies 3 and 4, I chose the context of voice commerce as a starting point and highlighted the implications for practice. The following section further elaborates on managerial implications, including thoughts from a policy making perspective as they are strongly intertwined.

Conversation designers are professionals who specialize in creating and shaping conversational interactions between users and conversational systems (e.g., crafting personas, creating seamless user experiences, optimizing system performance; Shevat, 2017). Conversation designers indeed bear the responsibility of considering ethical implications in their work, such as addressing issues related to bias, deception, or user manipulation. However, it is important to note that they typically do not create the overarching ethical rules governing AI systems. Instead, they generally receive guidelines and regulations from policymakers. Policymakers should be understood here as individuals, groups, or government bodies who play a crucial role in shaping the legal and ethical framework that governs the development, deployment, and use of conversational AI technologies (Jobin et al., 2019). Policymakers play a pivotal role in establishing the ethical framework within which conversation designers operate (Jobin et al., 2019; Wambsganss et al., 2021).

Against this background, my emphasis involves considering two key perspectives when addressing practical issues: Firstly, the economic viability, focusing on how developers can achieve the interaction goals by taking a relational approach (cost saving, profit, return on investment, and so forth) – which is the empirical focus of this dissertation. Secondly, thoughts about the normative desirability were also included, ensuring that users are not harmed or manipulated (see Ramadan, 2021; this focuses on the responsibility of the organization). The first is more directed towards system designers, focusing on the intricacies of designing the AI system itself. The second is geared more toward policymakers, contributing to the formulation of rules that govern the design and implementation of AI systems. However, in the end, both perspectives inform each other.

To clarify, while the data and findings undoubtedly hold significance for policymakers, the primary emphasis of this dissertation and its findings is on offering managerial insights tailored for business developers, marketers, and conversation designers. Adhering closely to the gathered data, the following sections address what practical managerial recommendations this dissertation can offer to professionals in the field of voice shopping.

First of all, these conversational AI systems are typically a means to an end, as seen in the case of voice shopping. Voice shopping via existing platforms like Amazon's Alexa or one's own solution, serve as an additional sales channel, offering another avenue to engage with the target audience alongside traditional sales channels like websites or other promotional activities on bus stops or in newspapers (Mari et al., 2020). In the context of voice shopping, conversational AI represents a novel component of the customer journey and must be integrated into the broader company strategies, including mission, product positioning, and target groups (Böhm et al., 2022) while also observing the current unpredictable market. Therefore, defining a voice shopping strategy becomes pivotal when designing and implementing such technologies – and this involves proactively crafting a strategy for the novel field I call customer-AI relationship management.

Therefore, if a businessperson wants to sell products via voice shopping, data from our study 3 suggests that it might be relevant to understand what kind of products are to be offered and decide from there. To reiterate, if one aims to sell high-involvement products, Alexa might not be a good choice, as only a few respondents saw their relationship characterized by peer bonding, the strongest predictor of shopping. Notably, the study in Chapter 4 involved distinguishing between low- and high-involvement, but other differentiations, such as categorizing into search and experience products, are also possible (see Liao & Sundar, 2022). Arguably, there may be other ways to define and categorize products that have not been tested yet. Determining the superior approach remains speculative at this stage, given that voice commerce is still in its early phases of development.

One can conclude that the development of a product strategy might be important in better understanding human-AI relationship dynamics. However, it still does not provide any value for the practical question of how to develop a strategy for steering the relationship between the consumer and the AI agent. Returning to the results of Chapter 4, the socio-emotional dimensions could hold significance in achieving commercial goals, prompting a desire to guide the relationship toward a more peer-oriented level. This involves conversational design considerations related to the characteristics of the human-AI relationship perceptions, for instance, the desired perceived agency, perceptions of hierarchies, or the breadth of the socio-emotional bond marketers aim to establish. The human-AI relationship framework sets the stage for such considerations. However, as outlined before, design practitioners must let go of the idea to fully control the perceptions of human-AI relationships.

The enduring aspiration to influence customers' attitudes and perceptions is a challenging endeavor, perhaps even considered the holy grail of marketing. I have approached this challenge via the human-AI fit theory (cf. Chapter 6). Results are inconclusive, yet they provide some guidance for practitioners. The effort to align the perception of relationships with conversational style is, at the very least, pertinent to avoiding alienating the majority of customers, which could arguably be deemed as an important albeit not overly ambitious goal. The question regarding how to facilitate voice shopping through the alignment of human-AI relationship dynamics remains unresolved, indicating potential areas for additional research and exploration.

In a nutshell, the main learnings from this dissertation's findings are as follows: It seems important to take into account the product context when selecting or designing the conversational AI. This raises the question of whether existing platforms, such as Alexa or Google Assistant, are adequate or if creating another solution is essential to better steer the relationship perception (e.g., to make it more sociable than Alexa). Another consideration is the human-AI alignment in conversational design. This is connected to the necessity of evaluating the investment in personalization efforts, weighing economic viability and ethical implications.

On a different note, applying the human-AI fit to practice requires practitioners to understand their consumers or target group and personalize content. However, a critical question arises: how do practitioners acquire knowledge about their consumers? This can be achieved through traditional methods such as market and usability research or profiling using online (big) data (Aguirre et al., 2015). Although personalization has proven effective in capturing attention, it is entangled with currently debated issues, extensively discussed in the personalization paradox literature, which recognizes the potential challenges associated with personalized approaches (see also studies within the personalization-privacy paradox realm, Lim et al., 2022). Legal challenges compound these issues, posing significant hurdles (e.g., what data for personalization are needed and how can they be gathered). The resolution of these challenges is an imperative subject for future research and regulation. In the interim, it may be prudent to empower consumers' agency by allowing them to determine their preferred mode of relationship and granting them the ability to adjust their conversational preferences to avoid the discouragement identified in our results.

Arguably, the significance of the findings lies in the ability to pose pertinent questions for practitioners to consider when deciding upon a conversational AI as a shopping assistant – or, any other (non-human) assistant. Arguably, these questions can be extrapolated to other contexts, such as employing conversational AI as a co-pilot in cars, as tutors in an educational setting, or as a mental health app coach, just to exemplify a few opportunities to test the ideas. Since the users do perceive fundamentally different relationships to their conversational AI systems and they seem to impact behavior differently, practitioners should rather investigate how users perceive their human-AI relationship – instead of assuming the relationship based on their design intentions. Based on this, practitioners can think of how to meaningfully shape the interaction and desired outcomes thereof (e.g., safer driving, better learning outcomes, or greater well-being to make use of the examples named above).

On a final note, what research and practice arguably have in common is their shared focus on addressing real-world challenges and opportunities associated with the deployment of conversational AI technologies. AI practitioners can contribute to advancing the AI field by applying theoretical knowledge to practical solutions, and they play a crucial role in the evolution and integration of AI across different industries and sectors. AI scholars can contribute to advancing the AI field by testing and providing empirical evidence of the effectiveness and responsible employment of practical solutions, also playing their part in the evolution and integration of AI. This is particularly important due to the ethical, legal, and environmental challenges associated with conversational AI and other AI systems (Crawford, 2021). Given this context, providing practical recommendations for designing conversational AI is challenging, as it is heavily intertwined with factors beyond conversational performance. Therefore, I want to underscore the urgent need for increased transdisciplinary collaboration. Such collaboration should not only span various social sciences in the relevant fields (HCI, HMC, etc.) but should also integrate perspectives from the market and business developers, as well as normative and regulatory viewpoints. Only through this comprehensive approach can research results become genuinely beneficial for broader society and future generations.

Chapter 8: Conclusion

How do users perceive their relationship with conversational AI? In this dissertation, I break with the prevailing theoretical assumption that social relationships are a domain exclusive to humans. Independent of the ontological inquiry of whether humans *can* have relationships, I can assert the achievement of the goal to advance the contemporary understanding of how humans *perceive* their relationship to conversational AI and that the perception is associated with user behavior. For the first time, A.P. Fiske's relational models theory (RMT) was repurposed to assess how humans perceive their relationship with technologies, using the example of conversational AI. We consistently found that the relationship modes people use to construe their relationships with other humans are similar but different to those used in human-AI interactions: Users see their conversational AI along three (instead of four) different dimensions. Authority ranking (i.e., hierarchical, owner-master), market pricing (i.e., equal, rational, on "eye-level"), and peer bonding (i.e., new rather emotional dimension, peer-like). Notably, each of them associated differently with outcome variables relevant to human-AI interaction which carries significant implications for practice (e.g., authority ranking predicts low-involvement shopping but not high-involvement shopping, peer bonding both, market pricing none, just to give an example). Besides providing a strongly theory-driven quantitative assessment of human-AI relationship perception, which is readily reproducible by fellow researchers, I have also illustrated the practical relevance of the multidimensional approach in the field of voice commerce, with a specific focus on conversation designers and policymakers. However, this research is just the beginning, and more experimental research is necessary to draw safe conclusions in a transdisciplinary manner. I would like to inspire forthcoming researchers and professionals to employ the human-AI relationship framework across diverse contexts where the perception of the relationship is deemed pertinent, elucidating its potential positive or negative impact on user interactions.

References

Abercrombie, G., Cercas Curry, A., Pandya, M., & Rieser, V. (2021). Alexa, Google, Siri: What are Your Pronouns? Gender and Anthropomorphism in the Design and Perception of Conversational Assistants. In M. Costajussa, H. Gonen, C. Hardmeier, & K. Webster (Eds.), *Proceedings of the 3rd Workshop on Gender Bias in Natural Language Processing* (pp. 24–33). Stroudsburg, PA, USA: Association for Computational Linguistics. https://doi.org/10.18653/v1/2021.gebnlp-1.4

Adam, M., Wessel, M., & Benlian, A. (2021). AI-based chatbots in customer service and their effects on user compliance. *Electronic Markets, 31*(2), 427–445. https://doi.org/10.1007/s12525-020-00414-7

Adolphs, E., & Zaharia, S. (2021). Consumers' Acceptance of a Voice Commerce Application in FMCG in Germany, U.S. and U.K. In F. F.-H. Nah & K. Siau (Eds.), *Lecture Notes in Computer Science: HCI in Business, Government and Organizations* (Vol. 12783, pp. 3–21). Springer International Publishing. https://doi.org/10.1007/978-3-030-77750-0_1

Agarwal. A. (2018, December 14). *Chatbots Supporting Mental Well-Being: Are We Playing A Dangerous Game?* Forbes. https://www.forbes.com/sites/pragyaagarwaleurope/2018/12/14/chatbots-supporting-mental-well-being-are-we-playing-a-dangerous-game/

Aguirre, E., Mahr, D., Grewal, D., Ruyter, K. de, & Wetzels, M. (2015). Unraveling the Personalization Paradox: The Effect of Information Collection and Trust-Building Strategies on Online Advertisement Effectiveness. *Journal of Retailing, 91*(1), 34–49. https://doi.org/10.1016/j.jretai.2014.09.005

Aiken, L. S., & West, S. G. (2010). *Multiple regression: Testing and interpreting interactions* [Reprinted]. SAGE.

Altman, I., & Taylor, D. (1987). Communication in interpersonal relationships: Social penetration theory. Interpersonal processes: *New directions in communication research, 14*, 257-277.

Alvarez, C., & Fournier, S. (2016). Consumers' relationships with brands. *Current Opinion in Psychology, 10*, 129–135. https://doi.org/10.1016/j.copsyc.2015.12.017

Amazon. (2021, February 8). *Customers in India say "I love you" to Alexa 19,000 times a day*. Amazon. https://www.aboutamazon.in/news/devices/customers-in-india-say-i-love-you-to-alexa-19-000-times-a-day

Anwyl-Irvine, A., Dalmaijer, E. S., Hodges, N., & Evershed, J. K. (2021). Realistic precision and accuracy of online experiment platforms, web browsers, and devices. *Behavior Research Methods, 53*(4), 1407–1425. https://doi.org/10.3758/s13428-020-01501-5

Araujo, T. (2018). Living up to the chatbot hype: The influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions. *Computers in Human Behavior*, *85*(2), 183–189. https://doi.org/10.1016/j.chb.2018.03.051

Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of Other in the Self Scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology, 63*(4), 596–612. https://doi.org/10.1037/0022-3514.63.4.596

Aronson, E., Wilson, T. D., & Akert, R. M. (2010). *Sozialpsychologie* (6., aktualisierte Aufl., [Nachdr.]). *ps Psychologie*. München: Pearson Studium.

Baldwin, M. W., Carrell, S. E., & Lopez, D. F. (1990). Priming relationship schemas: My advisor and the Pope are watching me from the back of my mind. *Journal of Experimental Social Psychology*, *26*(5), 435-454. https://doi.org/10.1016/0022-1031(90)90068-W

Batchelor, B. (2022). *6 Reasons Voice Shopping Is The Future Of M-Commerce*. https://www.forbes.com/sites/forbesbusinessdevelopmentcouncil/2022/12/21/6-reasons-voice-shopping-is-the-future-of-m-commerce/

Beattie, A., Edwards, A. P., & Edwards, C. (2020). A Bot and a Smile: Interpersonal Impressions of Chatbots and Humans Using Emoji in Computer-mediated Communication. Communication Studies, 71(3), 409–427. https://doi.org/10.1080/10510974.2020.1725082

Bellis, E. de, & Venkataramani Johar, G. (2020). Autonomous Shopping Systems: Identifying and Overcoming Barriers to Consumer Adoption. *Journal of Retailing, 96*(1), 74–87. https://doi.org/10.1016/j.jretai.2019.12.004

Berger, C. R., & Calabrese, R. J. (1975). Some explorations in initial interaction and beyond: Toward a developmental theory of interpersonal communication. *Human Communication Research*, **1**(2), 99–112. https://doi.org/10.1111/j.1468-2958.1975.tb00258.x

Bergmann, K., Eyssel, F., & Kopp, S. (2012). A second chance to make a first impression? How appearance and nonverbal behavior affect perceived warmth and competence of virtual agents over time. In Y. Nakano, M. Neff, A. Paiva, & M. Walker (Eds.), *Intelligent virtual agents* (pp. 126–138). Springer. https://doi.org/10.1007/978-3-642-33197-8_13

Berscheid, E., & Hatfield, E. (1969). *Interpersonal attraction* (Vol. 69, pp. 113-114). Reading, MA: Addison-Wesley.

Berscheid, E., Snyder, M., & Omoto, A. M. (1989). The Relationship Closeness Inventory: Assessing the closeness of interpersonal relationships. *Journal of personality and Social Psychology*, *57*(5), 792. https://doi.org/10.1037/0022-3514.57.5.792

Berscheid, E. (1994). Interpersonal relationships. *Annual review of psychology*, *45*(1), 79-129. https://doi.org/10.1146/annurev.ps.45.020194.000455

Blumstein, P., & Kollock, P. (1988). Personal relationships. *Annual Review of Sociology*,*14*(1), 467-490. https://www.jstor.org/stable/2083328

Böhm, E., Eggert, A., Garnefeld, I., Holzmüller, H. H., Schaefers, T., Steinhoff, L., & Woisetschläger, D. M. (2022). Exploring the Customer Journey of Voice Commerce: A Research Agenda. *Journal of Service Management Research, 6*(4), 216–231. https://doi.org/10.5771/2511-8676-2022-4-216

Bowlby, J. (1979). The Bowlby-Ainsworth attachment theory. *Behavioral and Brain Sciences, 2*(4), 637–638. https://doi.org/10.1017/S0140525X00064955

Brandtzaeg, P. B., Skjuve, M., & Følstad, A. (2022). My AI Friend: How Users of a Social Chatbot Understand Their Human–AI Friendship. *Human Communication Research, 48*(3), 404–429. https://doi.org/10.1093/hcr/hqac008

Brown., K. (2021, June 1). *Something Bothering You? Tell It to Woebot*. The New York Times. https://www.nytimes.com/2021/06/01/health/artificial-intelligence-therapy-woebot.html

Buschke, D. (2021, July 6). *So erklären deutsche Nutzer ihre Zuneigung für Alexa – Alexa ich liebe dich*. [This is how German users declare their affection for Alexa - Alexa I love you] Smarthome Assistant. https://www.smarthomeassistent.de/so-erklaeren-deutsche-nutzer-ihre-zuneigung-fuer-alexa-alexa-ich-liebedich/

Byrne, D. E. (1971). *The attraction paradigm*. Personality and psychopathology: Vol. 11. Academic Press.

Canziani, B., & MacSween, S. (2021). Consumer acceptance of voice-activated smart home devices for product information seeking and online ordering. *Computers in Human Behavior, 119*(2), 106714. https://doi.org/10.1016/j.chb.2021.106714

Carroll, B. A., & Ahuvia, A. C. (2006). Some antecedents and outcomes of brand love. *Marketing Letters*, *17*(2), 79–89. https://doi.org/10.1007/s11002-006-4219-2

Chattaraman, V., Kwon, W.-S., Gilbert, J. E., & Ross, K. (2019). Should AI-Based, conversational digital assistants employ social- or task-oriented interaction style? A taskcompetency and reciprocity perspective for older adults. *Computers in Human Behavior, 90*(1), 315–330. https://doi.org/10.1016/j.chb.2018.08.048

Chaudhuri, A., & Holbrook, M. B. (2002). Product-class effects on brand commitment and brand outcomes: The role of brand trust and brand affect. *Journal of Brand Management, 10*(1), 33–58. https://doi.org/10.1057/palgrave.bm.2540100

Cialdini, R. (1984) Influence: The Psychology of Persuasion. HarperCollins, New York.

Chita-Tegmark, M., Law, T., Rabb, N., & Scheutz, M. (03082021). Can You Trust Your Trust Measure? In C. Bethel, A. Paiva, E. Broadbent, D. Feil-Seifer, & D. Szafir (Eds.), *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction* (pp. 92–100). ACM. https://doi.org/10.1145/3434073.3444677

Chui, M., Yee, L., Hall, B., Singla, A. & Sukharevsky, A. (2021, August 1). *The state of AI in 2023: Generative AI's breakout year.* McKinsey & Company. https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2023-generative-ais-breakout-year?

Cicco, R. de, Silva, S. C., & Alparone, F. R. (2020). Millennials' attitude toward chatbots: an experimental study in a social relationship perspective. *International Journal of Retail & Distribution Management, 48*(11), 1213–1233. https://doi.org/10.1108/IJRDM-12-2019-0406

Coeckelbergh, M. (2021). Three Responses to Anthropomorphism in Social Robotics: Towards a Critical, Relational, and Hermeneutic Approach. *International Journal of Social Robotics, 42*(1), 143. https://doi.org/10.1007/s12369-021-00770-0

Cole, S. (2023, February 17). *Replika CEO Says AI Companions Were Not Meant to Be Horny. Users Aren't Buying It*. Vice. https://www.vice.com/en/article/n7zaam/replika-ceo-ai-erotic-roleplay-chatgpt3-rep

Crawford, K. (2021). *Atlas of AI: Power, politics, and the planetary costs of artificial intelligence.* Yale University Press. https://www.degruyter.com/isbn/9780300252392 https://doi.org/10.12987/9780300252392

Croes, E. A. J., & Antheunis, M. L. (2021). Can we be friends with Mitsuku? A longitudinal study on the process of relationship formation between humans and a social chatbot. *Journal of Social and Personal Relationships, 38*(1), 279–300. https://doi.org/10.1177/0265407520959463

Dautenhahn, K. (2007). Methodology & Themes of Human-Robot Interaction: A Growing Research Field. *International Journal of Advanced Robotic Systems, 4*(1), 15. https://doi.org/10.5772/5702

de Gennaro, M., Krumhuber, E. G., & Lucas, G. (2020). Effectiveness of an empathic chatbot in combating adverse effects of social exclusion on mood. *Frontiers in Psychology, 10,* Article 3061. https://doi.org/10.3389/fpsyg.2019.03061

Dellaert, B. G. C., Shu, S. B., Arentze, T. A., Baker, T., Diehl, K., Donkers, B., Fast, N. J., Häubl, G., Johnson, H., Karmarkar, U. R., Oppewal, H., Schmitt, B. H., Schroeder, J., Spiller, S. A., & Steffel, M. (2020). Consumer decisions with artificially intelligent voice assistants. *Marketing Letters, 31*(4), 335–347. https://doi.org/10.1007/s11002-020-09537-5

Demopoulos, a. (2023, January 16). *How soon will M3gan become reality?* The Guardian. Robot ethicists weigh in https://www.theguardian.com/film/2023/jan/16/megan-film-robot-ai-ethics-real-life

Dey, V. (2023). *How AI-powered conversational commerce will transform shopping in 2023*. Venturebeat. https://venturebeat.com/ai/how-ai-powered-conversational-commerce-transform-shopping-2023/

Edwards, J. R., & Cable, D. M. (2009). The value of value congruence. *The Journal of Applied Psychology, 94*(3), 654–677. https://doi.org/10.1037/a0014891

Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review, 114*(4), 864–886. https://doi.org/10.1037/0033-295X.114.4.864

Evans, K. D., Robbins, S. A., & Bryson, J. J. (2023). Do We Collaborate With What We Design? *Topics in Cognitive Science.* Advance online publication. https://doi.org/10.1111/tops.12682

Feine, J., Adam, M., Benke, I., Maedche, A., & Benlian, A. (2020). Exploring Design Principles for Enterprise Chatbots: An Analytic Hierarchy Process Study. In S. Hofmann, O. Müller, & M. Rossi (Eds.), *Lecture Notes in Computer Science. Designing for Digital Transformation. Co-Creating Services with Citizens and Industry* (Vol. 12388, pp. 126–141). Springer International Publishing. https://doi.org/10.1007/978-3-030-64823-7_13

Festinger, L., Back, K., Schachter, S., Kelley, H. H., & Thibaut, J. (1950). *Theory and experiment in social communication.* Research Center for Dynamics Institute for Social Research. https://doi.org/10.1037/10646-000

Fernandes, T., & Oliveira, E. (2021). Understanding consumers' acceptance of automated technologies in service encounters: Drivers of digital voice assistants adoption. *Journal of Business Research, 122*(3), 180–191. https://doi.org/10.1016/j.jbusres.2020.08.058

Fiske, A. P. (1992). The four elementary forms of sociality: Framework for a unified theory of social relations. *Psychological Review, 99*(4), 689–723. https://doi.org/10.1037//0033-295X.99.4.689

Fiske, A. P. (1993). *Structures of social life: The four elementary forms of human relations: communal sharing, authority ranking, equality matching, market pricing*; with a new epilogue (1. paperback ed.). Free Press.

Fiske, A. P., Haslam, N., & Fiske, S. T. (1991). Confusing one person with another: What errors reveal about the elementary forms of social relations. *Journal of Personality and Social Psychology, 60*(5), 656–674. https://doi.org/10.1037//0022-3514.60.5.656

Franke, T., Attig, C., & Wessel, D. (2019). A personal resource for technology interaction: Development and validation of the affinity for technology interaction (ATI) scale. *International Journal of Human–ComputerInteraction, 35*(6), 456–467. https://doi.org/10.1080/10447318.2018.1456150

Funk, G., Hofstetter, R., Kälin, F., Jörg, D., Lopetrone, R., & Bruetsch, C. (2021). *Voice First Barometer Schweiz* 2020. 1-48.

Gambino, A., Fox, J., & Ratan, R. (2020). Building a Stronger CASA: Extending the Computers Are Social Actors Paradigm. *Human-Machine Communication, 1*, 71–86. https://doi.org/10.30658/hmc.1.5

Gao, Y., Pan, Z., Wang, H., & Chen, G. (2018). Alexa, My Love: Analyzing Reviews of Amazon Echo. *In 2018 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation* (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI) (pp. 372–380). IEEE. https://doi.org/10.1109/SmartWorld.2018.00094

Gilad, Z., Amir, O., & Levontin, L. (2021). The effects of warmth and competence perceptions on users' choice of an AI system. In Y. Kitamura, A. Quigley, K. Isbister, T. Igarashi, P. Bjørn, & S. Drucker (Eds.), *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1–13). ACM. https://doi.org/10.1145/3411764.3446863

Glikson, E., & Woolley, A. W. (2020). Human Trust in Artificial Intelligence: Review of Empirical Research. *Academy of Management Annals, 14*(2), 627–660. https://doi.org/10.5465/annals.2018.0057

Go, E., & Sundar, S. S. (2019). Humanizing chatbots: The effects of visual, identity and conversational cues on humanness perceptions. *Computers in Human Behavior, 97*(1), 304–316. https://doi.org/10.1016/j.chb.2019.01.020

Gong, L. (2008). How social is social responses to computers? The function of the degree of anthropomorphism in computer representations. *Computers in Human Behavior, 24*(4), 1494–1509. https://doi.org/10.1016/j.chb.2007.05.007

Guzman, A. L. (2019). Voices in and of the machine: Source orientation toward mobile virtual assistants. *Computers in Human Behavior, 90*(5), 343–350. https://doi.org/10.1016/j.chb.2018.08.009

Guzman, A. L. (2023). *Talking about "Talking with Machines": Interview as Method within HMC*. In A. Guzman, R. McEwen, & S. Jones (Eds.), The Sage Handbook of Human–Machine Communication (pp. 243–251). SAGE Publications Ltd. https://doi.org/10.4135/9781529782783.n30

Guzman, A. L., & Lewis, S. C. (2020). Artificial intelligence and communication: A human–machine communication research agenda. *New Media & Society, 22*(1), 70–86. https://doi.org/10.1177/1461444819858691

Guzman, A., McEwen, R., & Jones, S. (Eds.). (2023). *The Sage Handbook of Human–Machine Communication.* SAGE Publications Ltd. https://doi.org/10.4135/9781529782783

Halbauer, I., & Klarmann, M. (2022). How voice retailers can predict customer mood and how they can use that information. *International Journal of Research in Marketing, 39*(1), 77–95. https://doi.org/10.1016/j.ijresmar.2021.09.008

Halbauer, I., Jacob, S., & Klarmann, M. (2022). Brand presentation order in voice shopping: Understanding the effects of sequential product presentation. *Journal of Retailing, 39*(5), 76. https://doi.org/10.1016/j.jretai.2022.06.002

Han, M. C. (2021). The Impact of Anthropomorphism on Consumers' Purchase Decision in Chatbot Commerce. *Journal of Internet Commerce, 20*(1), 46–65. https://doi.org/10.1080/15332861.2020.1863022

Han, S., & Yang, H. (2018). Understanding adoption of intelligent personal assistants. *Industrial Management & Data Systems, 118*(3), 618–636. https://doi.org/10.1108/IMDS-05-2017-0214

Haslam, N., & Fiske, A. P. (1999). Relational models theory: A confirmatory factor analysis. *Personal Relationships, 6*(2), 241–250. https://doi.org/10.1111/j.1475-6811.1999.tb00190.x

Heider, F., & Simmel, M. (1944). An Experimental Study of Apparent Behavior. *The American Journal of Psychology, 57*(2), 243. https://doi.org/10.2307/1416950

Hepp, A. (2020). Artificial companions, social bots and work bots: communicative robots as research objects of media and communication studies. *Media, Culture & Society, 42*(7-8), 1410–1426. https://doi.org/10.1177/0163443720916412

Heyselaar, E. (2023). The CASA theory no longer applies to desktop computers. *Scientific Reports, 13*(1), 19693. https://doi.org/10.1038/s41598-023-46527-9

Hoff, K. A., & Bashir, M. (2015). Trust in automation: Integrating empirical evidence on factors that influence trust. *Human Factors, 57*(3), 407–434. https://doi.org/10.1177/0018720814547570

Homans, G. C. (1958). Social behavior as exchange. *American journal of sociology,63*(6), 597-606. https://doi.org/10.1086/222355

Horton, D., & Wohl, R. R. (1956). Mass communication and para-social interaction: Observations on intimacy at a distance. *Psychiatry, 19*(3), 215–229. https://doi.org/10.1080/00332747.1956.11023049

Huh, J., Whang, C., & Kim, H.-Y. (2023). Building trust with voice assistants for apparel shopping: The effects of social role and user autonomy. *Journal of Global Fashion Marketing, 14*(1), 5–19. https://doi.org/10.1080/20932685.2022.2085603

Hu, P., Lu, Y., & Wang, B. (2022). Experiencing power over AI: The fit effect of perceived power and desire for power on consumers' choice for voice shopping. *Computers in Human Behavior, 128*(6), 107091. https://doi.org/10.1016/j.chb.2021.107091

Jian, J.-Y., Bisantz, A. M., & Drury, C. G. (2000). Foundations for an Empirically Determined Scale of Trust in Automated Systems. *International Journal of Cognitive Ergonomics, 4*(1), 53–71. https://doi.org/10.1207/S15327566IJCE0401_04

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence, 1*(9), 389–399. https://doi.org/10.1038/s42256-019-0088-2

Jones, C., & Bergen, B. (2023). *Does GPT-4 Pass the Turing Test?* https://doi.org/10.48550/arXiv.2310.20216

Kahneman, D. (2012). *Thinking, fast and slow*. *Penguin psychology*. Penguin Books.

Khatri, C., Venkatesh, A., Hedayatnia, B., Ram, A., Gabriel, R., & Prasad, R. (2018). Alexa Prize-State of the Art in Conversational AI. *AI Magazine*, *39*(3), 40–55. https://doi.org/10.1609/aimag.v39i3.2810

Ki, C.-W., Cho, E., & Lee, J.-E. (2020). Can an intelligent personal assistant (IPA) be your friend? Para-friendship development mechanism between IPAs and their users. *Computers in Human Behavior*, *111*, Article 106412. https://doi.org/10.1016/j.chb.2020.106412

Kim, A., Cho, M., Ahn, J., & Sung, Y. (2019). Effects of Gender and Relationship Type on the Response to Artificial Intelligence. *Cyberpsychology, Behavior and Social Networking, 22*(4), 249–253. https://doi.org/10.1089/cyber.2018.0581

Kinsella, B. (2018, March 5). *Voice Shopping to Reach $40 Billion in U.S. and $5 Billion in UK by 2022*. Voicebot. https://voicebot.ai/2018/03/05/voice-shopping-reach-40-billion-u-s-5-billion-uk-2022/

Kinsella, B.(2021, December 24). *Voice Shopping Rises to 45 Million U.S. Adults in 2021.* Voicebot.AI. https://voicebot.ai/2021/12/24/voice-shopping-rises-to-45-million-u-s-adults-in-2021/

Klaus, P., & Zaichkowsky, J. L. (2022). The convenience of shopping via voice AI: Introducing AIDM. *Journal of Retailing and Consumer Services, 65*(3), 102490. https://doi.org/10.1016/j.jretconser.2021.102490

Knapp, M. L. (1978). *Social intercourse: From greeting to good-bye*. Boston: Allyn and Bacon.

Kulkarni, P., Mahabaleshwarkar, A., Kulkarni, M., Sirsikar, N., & Gadgil, K. (2019). Conversational AI: An overview of methodologies, applications & future scope. *In 2019 5th International conference on computing, communication, control and automation* (ICCUBEA) (pp. 1–7). IEEE. https://doi.org/10.1109/ICCUBEA47591.2019.9129347

Kuzminykh, A., Sun, J., Govindaraju, N., Avery, J., & Lank, E. (2020). Genie in the Bottle. In R. Bernhaupt, F. 'F.' Mueller, D. Verweij, J. Andres, J. McGrenere, A. Cockburn, R. Kocielnik (Eds.), *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–13). New York, NY, USA: ACM. https://doi.org/10.1145/3313831.3376665

Langer, M., Hunsicker, T., Feldkamp, T., König, C. J., & Grgić-Hlača, N. (2022). “Look! It’s a computer program! It’s an algorithm! It’s AI!’: Does terminology affect human perceptions and evaluations of algorithmic decisionmaking systems? In S. Barbosa, C. Lampe, C. Appert, D. A. Shamma, S. Drucker, J. Williamson, & K. Yatani (Eds.), *CHI Conference on*

Human Factors in Computing Systems (pp. 1–28). ACM. https://doi.org/10.1145/3491102.3517527

Laudon, K. C., & Traver, C. G. (2013). *E-commerce* (pp. 1-912). Boston, MA: Pearson

Leary, M. R. (2010). Affiliation, acceptance, and belonging. In S. T. Fiske, D. T. Gilbert & G. Lindzey (Eds.), *Handbook of Social Psychology* (5th ed., Vol. 2, pp. 864–897). NY: Wiley.

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors, 46*(1), 50–80. https://doi.org/10.1518/hfes.46.1.50_30392

Li, L., Li, Y., Song, B., Shi, Z., & Wang, C. (2022). How Human-like Behavior of Service Robot Affects Social Distance: A Mediation Model and Cross-Cultural Comparison. *Behavioral Sciences*, *12*(7), 205. https://doi.org/10.3390/bs12070205

Li, X., & Sung, Y. (2021). Anthropomorphism brings us closer: The mediating role of psychological distance in user–AI assistant interactions. *Computers in Human Behavior, 118*, Article 106680. https://doi.org/10.1016/j.chb.2021.106680

Li, Z., & Rau, P. L. P. (2019). Effects of self-disclosure on attributions in human–IoT conversational agent interaction. *Interacting with Computers*, *31*(1), 13–26. https://doi.org/10.1093/iwc/iwz002

Liao, M., & Sundar, S. S. (2022). When E-Commerce Personalization Systems Show and Tell: Investigating the Relative Persuasive Appeal of Content-Based versus Collaborative Filtering. *Journal of Advertising*, *51*(2), 256–267. https://doi.org/10.1080/00913367.2021.1887013

Lim, W. M., Kumar, S., Verma, S., & Chaturvedi, R. (2022). Alexa, what do we know about conversational commerce? Insights from a systematic literature review. *Psychology & Marketing, 39*(6), 1129–1155. https://doi.org/10.1002/mar.21654

Lin, H., Chi, O. H., & Gursoy, D. (2020). Antecedents of customers' acceptance of artificially intelligent robotic device use in hospitality services. *Journal of Hospitality Marketing & Management*, *29*(5), 530–549. https://doi.org/10.1080/19368623.2020.1685053

Liu, Y., Lee, Y., & Chen, A. N.K. (2011). Evaluating the effects of task-individual-technology fit in multi-DSS models context: A two-phase view. *Decision Support Systems, 51*(3), 688–700. https://doi.org/10.1016/j.dss.2011.03.009

Lopatovska, I., & Williams, H. (2018). Personification of the Amazon Alexa. In C. Shah, N. J. Belkin, K. Byström, J. Huang, & F. Scholer (Eds.), *Proceedings of the 2018 Conference on Human Information Interaction&Retrieval* - CHIIR '18 (pp. 265–268). New York, New York, USA: ACM Press. https://doi.org/10.1145/3176349.3176868

MacKenzie, S. B., & Lutz, R. J. (1989). An Empirical Examination of the Structural Antecedents of Attitude toward the Ad in an Advertising Pretesting Context. *Journal of Marketing, 53*(2), 48. https://doi.org/10.2307/1251413

Malle, B. F., & Ullman, D. (2021). A multidimensional conception and measure of human-robot trust. In Trust in Human-Robot Interaction (Vol. 100, pp. 3–25). *Elsevier*. https://doi.org/10.1016/B978-0-12-819472-0.00001-0

Mara, M., Appel, M., & Gnambs, T. (2022). Human-Like Robots and the Uncanny Valley. *Zeitschrift Für Psychologie, 230*(1), 33–46. https://doi.org/10.1027/2151-2604/a000486

Marcus, G. (2023, April 4). *The first known chatbot associated death*. Marcus on AI on Substack. https://garymarcus.substack.com/p/the-first-known-chatbot-associated

Mari, A., & Algesheimer, R. (2021). The Role of Trusting Beliefs in Voice Assistants during Voice Shopping. In T. Bui (Ed.), *Proceedings of the Annual Hawaii International Conference on System Sciences. Hawaii International Conference on System Sciences*. https://doi.org/10.24251/HICSS.2021.495

Mari, A., Mandelli, A., & Algesheimer, R. (2020). The Evolution of Marketing in the Context of Voice Commerce: A Managerial Perspective. In F. F.-H. Nah & K. Siau (Eds.), *Lecture Notes in Computer Science. HCI in Business, Government and Organizations* (Vol. 12204, pp. 405–425). Springer International Publishing. https://doi.org/10.1007/978-3-030-50341-3_32

Marriott, H. R., & Pitardi, V. (2023). One is the loneliest number… Two can be as bad as one. The influence of AI Friendship Apps on users' well-being and addiction. *Psychology & Marketing, 45*(1), 212. https://doi.org/10.1002/mar.21899

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *The Academy of Management Review, 20*(3), 709–734. https://doi.org/10.2307/258792

McKinsey & Company. (2023, June 14). *The economic potential of generative AI: The next productivity frontier.* McKinesey & Company. https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier#introduction

McLean, G., & Osei-Frimpong, K. (2019). Hey Alexa … examine the variables influencing the use of artificial intelligent in-home voice assistants. *Computers in Human Behavior, 99*(4), 28–37. https://doi.org/10.1016/j.chb.2019.05.009

Mori, M., MacDorman, K., & Kageki, N. (2012). The Uncanny Valley [From the Field]. *IEEE Robotics & Automation Magazine, 19*(2), 98–100. https://doi.org/10.1109/MRA.2012.2192811

Moussawi, S., Koufaris, M., & Benbunan-Fich, R. (2021). How perceptions of intelligence and anthropomorphism affect adoption of personal intelligent agents. *Electronic Markets, 31*(2), 343–364. https://doi.org/10.1007/s12525-020-00411-w

Mumm, J., & Mutlu, B. (2011). Designing motivational agents: The role of praise, social comparison, and embodiment in computer feedback. *Computers in Human Behavior, 27*(5), 1643–1650. https://doi.org/10.1016/j.chb.2011.02.002

Nass, C. I., & Brave, S. (2005). *Wired for speech: How voice activates and advances the human-computer relationship*. Computer-human interaction. MIT Press.

Nass, C., & Moon, Y. (2000). Machines and Mindlessness: Social Responses to Computers. *Journal of Social Issues*, *56*(1), 81–103. https://doi.org/10.1111/0022-4537.00153

Nass, C., Steuer, J., & Tauber, E. R. (1994). Computers are social actors. In C. Plaisant (Ed.), *Conference companion on Human factors in computing systems - CHI '94 (p. 204). ACM* Press. https://doi.org/10.1145/259963.260288

Nayyar, S. (2023). Is ChatGPT Really One Of The Most Important Milestones In Digital Technology? *Forbes*. https://www.forbes.com/sites/forbestechcouncil/2023/04/03/is-chat-gpt-really-one-of-the-most-important-milestones-in-digital-technology/

Panetta, K.. (2022). *Companion Robots Market Report*. Fortune Business Insights. https://www.fortunebusinessinsights.com/companion-robots-market-105380.

Panetta, K. (2020, October 19). *Gartner top strategic technology trends for 2021*. Gartner. https://www.gartner.com/smarterwithgartner/gartner-top-strategic-technology-trends-for-2021

Pearl, C. (2017). *Designing voice user interfaces: Principles of conversational experiences*. O'Reilly.

Pentina, I., Hancock, T., & Xie, T. (2023a). Exploring relationship development with social chatbots: A mixedmethod study of replika. *Computers in Human Behavior, 140*(2), 107600. https://doi.org/10.1016/j.chb.2022.107600

Pentina, I., Xie, T., Hancock, T., & Bailey, A. (2023b). Consumer–machine relationships in the age of artificial intelligence: Systematic literature review and research directions. *Psychology & Marketing*, *40*(8), 1593–1614. https://doi.org/10.1002/mar.21853

Perez, S. (2020, February 4). *Shopping via smart speakers is not taking off, report suggests*. TechCrunch. https://techcrunch.com/2020/02/04/shopping-via-smart-speakers-is-not-taking-off-report-suggests/

Petersen, E., Fiske, A. P., & Schubert, T. W. (2019). The Role of Social Relational Emotions for Human-Nature Connectedness. *Frontiers in Psychology*, *10*, 2759. https://doi.org/10.3389/fpsyg.2019.02759

Petty, R. E., & Cacioppo, J. T. (Eds.). (1986). *Communication and Persuasion*. Springer New York. https://doi.org/10.1007/978-1-4612-4964-1

Pickard, M. D., Burgoon, J. K., & Derrick, D. C. (2014). Toward an Objective Linguistic-Based Measure of Perceived Embodied Conversational Agent Power and Likeability. *International Journal of Human-Computer Interaction, 30*(6), 495–516. https://doi.org/10.1080/10447318.2014.888504

Pieters, R. G. M., & Verplanken, B. (1995). Intention-behaviour consistency: Effects of consideration set size, involvement and need for cognition. *European Journal of Social Psychology, 25*(5), 531–543. https://doi.org/10.1002/ejsp.2420250505

Pitardi, V., & Marriott, H. R. (2021). Alexa, she's not human but… Unveiling the drivers of consumers' trust in voice-based artificial intelligence. *Psychology & Marketing, 38*(4), 626–642. https://doi.org/10.1002/mar.21457

Purington, A., Taft, J. G., Sannon, S., Bazarova, N. N., & Taylor, S. H. (2017). "Alexa is my new BFF". In G. Mark, S. Fussell, C. Lampe, m.c. Schraefel, J. P. Hourcade, C. Appert, & D. Wigdor (Eds*.), Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (pp. 2853–2859). ACM. https://doi.org/10.1145/3027063.3053246

Qiu, L., & Benbasat, I. (2009). Evaluating Anthropomorphic Product Recommendation Agents: A Social Relationship Perspective to Designing Information Systems. *Journal of Management Information Systems*, *25*(4), 145–182. https://doi.org/10.2753/MIS0742-1222250405

Rafaeli, A., Altman, D., Gremler, D. D., Huang, M.-H., Grewal, D., Iyer, B., Parasuraman, A., & Ruyter, K. de (2017). The Future of Frontline Research. *Journal of Service Research, 20*(1), 91–99. https://doi.org/10.1177/1094670516679275

Ramadan, Z. B. (2021). "Alexafying" shoppers: The examination of Amazon's captive relationship strategy. *Journal of Retailing and Consumer Services, 62*, 102610. https://doi.org/10.1016/j.jretconser.2021.102610

Rhee, C. E., & Choi, J. (2020). Effects of personalization and social role in voice shopping: An experimental study on product recommendation by a conversational voice agent. *Computers in Human Behavior, 109*(1), 106359. https://doi.org/10.1016/j.chb.2020.106359

Rook, K. S., & Dooley, D. (1985). Applying social support research: Theoretical problems and future directions. *Journal of Social Issues, 41*(1), 5-28. https://doi.org/10.1111/j.1540-4560.1985.tb01114.x

Rosseel, Y., Jorgensen, T. D., Rockwood, N., Oberski, D., Byrnes, J., Vanbrabant, L., Savalei, V., Merkle, E., Hallquist, M., Rhemtulla, M., Katsikatsou, M., Barendse, M., Scharf, F., & Du, H. (2021). *Package lavaan*. http://tucows.icm.edu.pl/packages/cran/web/packages/lavaan/lavaan.pdf

Ryan, M. (2020). In AI we trust: Ethics, artificial intelligence, and reliability. *Science and Engineering Ethics, 26*(5), 2749–2767. https://doi.org/10.1007/s11948-020-00228-y

Sætra, H. S. (2021). Social robot deception and the culture of trust. *Paladyn, Journal of Behavioral Robotics, 12*(1), 276–286. https://doi.org/10.1515/pjbr-2021-0021

Savage, F. (2023). *Chatbots: Vector analysis, competitor leaderboard & market forecast 2023-2028*. https://www.juniperresearch.com/researchstore/operators-providers/chatbots-trends-research-report

Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize? *Journal of Research in Personality, 47*(5), 609-612. https://doi.org/10.1016/j.jrp.2013.05.009

Schubert, T. W., & Otten, S. (2002). Overlap of self, ingroup, and outgroup: Pictorial measures of selfcategorization. *Self and Identity, 1*(4), 353–376. https://doi.org/10.1080/152988602760328012

Schweitzer, F., Belk, R., Jordan, W., & Ortner, M. (2019). Servant, friend or master? The relationships users build with voice-controlled smart devices. *Journal of Marketing Management, 35*(7-8), 693–715. https://doi.org/10.1080/0267257X.2019.1596970

Seymour, W., & van Kleek, M. (2021). Exploring interactions between trust, anthropomorphism, and relationship development in voice assistants. *Proceedings of the ACM on Human-Computer Interaction, 5*(CSCW2), Article 371. https://doi.org/10.1145/3479515

Sheehan, B., Jin, H. S., & Gottlieb, U. (2020). Customer service chatbots: Anthropomorphism and adoption. *Journal of Business Research, 115*(2), 14–24. https://doi.org/10.1016/j.jbusres.2020.04.030

Shevat, A. (2017). *Designing bots: Creating conversational experiences*. O’Reilly Media, Inc.

Shoda, Y., Mischel, W., & Wright, J. C. (1994). Intraindividual stability in the organization and patterning of behavior: Incorporating psychological situations into the idiographic analysis of personality. *Journal of Personality and Social Psychology, 67*(4), 674–687. https://doi.org/10.1037/0022-3514.67.4.674

Short, J., Williams, E., & Christie, B. (1976). *The social psychology of telecommunications*. London: John Wiley & Sons.

Shumanov, M., & Johnson, L. (2021). Making conversations with chatbots more personalized. *Computers in Human Behavior, 117*(1), 106627. https://doi.org/10.1016/j.chb.2020.106627

Simmons, R., Makatchev, M., Kirby, R., Lee, M. K., Fanaswala, I., Browning, B., Forlizzi, J., & Sakr, M. (2011). Believable Robot Characters. *AI Magazine, 32*(4), 39–52. https://doi.org/10.1609/aimag.v32i4.2383

Skjuve, M., Følstad, A., Fostervold, K. I., & Brandtzaeg, P. B. (2022). A longitudinal study of huma-chatbot relationships. *International Journal of Human-Computer Studies*, *168*(102903), 102903. https://doi.org/10.1016/j.ijhcs.2022.102903

Skjuve, M., Følstad, A., Fostervold, K. I., & Brandtzaeg, P. B. (2021). My Chatbot Companion – a Study of Human-Chatbot Relationships. *International Journal of Human-Computer Studies, 149*(3), 102601. https://doi.org/10.1016/j.ijhcs.2021.102601

Son Nguyen, H., Mladenow, A., Strauss, C., & Auer-Srnka, K. (11292021). Voice Commerce: Anthropomorphism Using Voice Assistants. In E. Pardede, M. Indrawan-Santiago, P. D. Haghighi, M. Steinbauer, I. Khalil, & G. Kotsis (Eds.), *The 23rd International Conference on Information Integration and Web Intelligence* (pp. 434–442). ACM. https://doi.org/10.1145/3487664.3487724

Stallbaumer, C. (2023). *Introducing Microsoft 365 Copilot-A whole new way to work*. https://www.microsoft.com/en-us/microsoft-365/blog/2023/03/16/introducing-microsoft-365-copilot-a-whole-new-way-to-work/

Stein, J.-P., & Ohler, P. (2018). Uncanny…But Convincing? Inconsistency Between a Virtual Agent’s Facial Proportions and Vocal Realism Reduces Its Credibility and Attractiveness, but Not Its Persuasive Success. *Interacting with Computers, 30*(6), 480–491. https://doi.org/10.1093/IWC/IWY023

Stein, J.-P., Liebold, B., & Ohler, P. (2019). Stay back, clever thing! Linking situational control and human uniqueness concerns to the aversion against autonomous technology. *Computers in Human Behavior*, *95*(7), 73–82. https://doi.org/10.1016/j.chb.2019.01.021

Sternberg, R. J. (1988). Triangulating love. In R. J. Sternberg & M. L. Barnes (Eds.). *The psychology of love* (pp. 119–138). Yale University Press.

Suchman, L. (2012). *Human-Machine Reconfigurations*. Cambridge University Press. https://doi.org/10.1017/CBO9780511808418

Sundar, S.S. (2008). The MAIN model: a heuristic approach to understanding technology effects on credibility. In: Metzger MJ and Flanagin AJ (eds) *Digital Media, Youth, and Credibility*. Cambridge: The MIT Press, pp. 73–100

Sundar, S. S. (2020). Rise of Machine Agency: A Framework for Studying the Psychology of Human–AI Interaction (HAII). *Journal of Computer-Mediated Communication*, *25*(1), 74–88. https://doi.org/10.1093/jcmc/zmz026

Sundar, S. S., & Chen, J. (2023). From CASA to TIME: Machine as a Source of Media Effects. In A. Guzman, R. McEwen, & S. Jones (Eds.), *The Sage Handbook of Human–Machine Communication* (pp.63–72). SAGE Publications Ltd. https://doi.org/10.4135/9781529782783.n9

Sundar, S. S., & Kim, J. (2019). Machine Heuristic: When We Trust Computers More than Humans with Our Personal Information. In S. Brewster, G. Fitzpatrick, A. Cox, & V. Kostakos (Eds.), *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1–9). ACM. https://doi.org/10.1145/3290605.3300768

Sundar, S. S., & Nass, C. (2000). Source Orientation in Human-Computer Interaction. *Communication Research*, 27(6), 683–703. https://doi.org/10.1177/009365000027006001

Sundar, S. S., Jung, E. H., Waddell, T. F., & Kim, K. J. (2017). Cheery companions or serious assistants? Role and demeanor congruity as predictors of robot attraction and use intentions among senior citizens. *International Journal of Human-Computer Studies, 97*, 88–97. https://doi.org/10.1016/j.ijhcs.2016.08.006

Tajfel, H., & Turner, J. (2001). An integrative theory of intergroup conflict. In M. A. Hogg & D. Abrams (Eds.), *Intergroup relations: Essential readings* (pp. 94–109). Psychology Press.

Tajvidi, M., Richard, M.-O., Wang, Y., & Hajli, N. (2020). Brand co-creation through social commerce information sharing: The role of social media. *Journal of Business Research, 121*(1), 476–486. https://doi.org/10.1016/j.jbusres.2018.06.008

Tassiello, V., Tillotson, J. S., & Rome, A. S. (2021). “Alexa, order me a pizza!”: The mediating role of psychological power in the consumer–voice assistant interaction. *Psychology & Marketing*, *38*(7), 1069–1080. https://doi.org/10.1002/mar.21488

Taylor, D., & Altman, I. (1987). Communication in interpersonal relationships: Social penetration processes. In M. E. Roloff & G. R. Miller (Eds.), *Interpersonal processes: New directions in communication research* (pp. 257–277). Sage Publications.

The Guardian. (2020, July 20). *"Alexa, I love you": how lockdown made men lust after their Amazon Echo*. The Guardian. https://www.theguardian.com/technology/2020/jul/20/alexa-i-love-you-how-lockdown-made-men-lust-after-their-amazon-echo

Tiku, N. (2022, June 11). *The Google engineer who thinks the company's AI has come to life.* The Washington Post. https://www.washingtonpost.com/technology/2022/06/11/google-ai-lamda-blake-lemoine/

Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological review, 117*(2), 440. https://doi.org/10.1037/a0018963

Tschopp, M., Gieselmann, M., & Sassenberg, K. (2023). Servant by default? How humans perceive their relationship with conversational AI. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace, 17*(3). https://doi.org/10.5817/CP2023-3-9

Tschopp, M., Ruef, M., & Monett, D. (2022). Vertrauen Sie KI? Einblicke in das Thema Künstliche Intelligenz und warum Vertrauen eine Schlüsselrolle im Umgang mit neuen Technologien spielt. In M. Landes, E. Steiner, & T. Utz (Eds.), *Kreativität und Innovation in Organisationen* (Vol. 81, pp. 319–346). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-63117-1_16

Turkle S (1984) *The Second Self*. New York: Simon & Schuster.

U.S. Smart Home Consumer Adoption Report 2022. (2022, June). *U.S. Smart Home Consumer Adoption Report 2022*. Voicebot Research. https://research.voicebot.ai/report-list/u-s-smart-home-consumer-adoption-report-2022/

van Berkel, N., Skov, M. B., & Kjeldskov, J. (2021). Human-AI interaction. *Interactions, 28*(6), 67–71. https://doi.org/10.1145/3486941

van Straten, C. L., Peter, J., Kühne, R., & Barco, A. (2022). On sharing and caring: Investigating the effects of a robot's self-disclosure and question-asking on children's robot perceptions and child-robot relationship formation. *Computers in Human Behavior*, **129**, 107135. https://doi.org/10.1016/j.chb.2021.107135

van Vianen, A. E.M. (2018). Person–Environment Fit: A Review of Its Basic Tenets. *Annual Review of Organizational Psychology and Organizational Behavior, 5*(1), 75–101. https://doi.org/10.1146/annurev-orgpsych-032117-104702

Véliz, C. (2023). Chatbots shouldn't use emojis. *Nature, 615*(7952), 375. https://doi.org/10.1038/d41586-023-00758-y

Verma, P. (2023, March 30). *They fell in love with AI bots. A software update broke their heart.* Washington Post. https://www.washingtonpost.com/technology/2023/03/30/replika-ai-chatbot-update/

Walster, E., Walster, G. W., & Berscheid, E. (1978). *Equity: Theory and research*. Allyn & Bacon, Boston http://hdl.handle.net/10822/779008

Wambsganss, T., Höch, A., Zierau, N., & Söllner, M. (2021). Ethical Design of Conversational Agents: Towards Principles for a Value-Sensitive Design. In F. Ahlemann, R. Schütte, & S. Stieglitz (Eds.), *Lecture Notes in Information Systems and Organisation. Innovation Through Information Systems* (Vol. 46, pp. 539–557). Springer International Publishing. https://doi.org/10.1007/978-3-030-86790-4_37

Waytz, A., Epley, N., & Cacioppo, J. T. (2010). Social Cognition Unbound: Insights Into Anthropomorphism and Dehumanization. *Current Directions in Psychological Science, 19*(1), 58–62. https://doi.org/10.1177/0963721409359302

Weizenbaum, J. (1966). ELIZA-a computer program for the study of natural language communication between man and machine. *Communications of the ACM, 9*(1), 36–45. https://doi.org/10.1145/365153.365168

Wirtz, J., Patterson, P. G., Kunz, W. H., Gruber, T., Lu, V. N., Paluch, S., & Martins, A. (2018). Brave new world: service robots in the frontline. *Journal of Service Management, 29*(5), 907–931. https://doi.org/10.1108/JOSM-04-2018-0119

Xie, T., & Pentina, I. (2022). Attachment theory as a framework to understand relationships with social chatbots: A case study of Replika. In T. Bui (Ed.), *Proceedings of the 55th Hawaii International Conference on System Sciences. ScholarSpace*. https://doi.org/10.24251/HICSS.2022.258

Xu, S., & Li, W. (2022). A tool or a social being? A dynamic longitudinal investigation of functional use and relational use of AI voice assistants. *New Media & Society*. Advance online article. https://doi.org/10.1177/14614448221108112

Yang, Y., Liu, Y., Lv, X., Ai, J., & Li, Y. (2022). Anthropomorphism and customers' willingness to use artificial intelligence service agents. *Journal of Hospitality Marketing & Management, 31*(1), 1–23. https://doi.org/10.1080/19368623.2021.1926037

Zaharia, S., & Würfel, M. (2021). Voice Commerce - Studying the Acceptance of Smart Speakers. In T. Ahram, R. Taiar, K. Langlois, & A. Choplin (Eds.), *Advances in*

Intelligent Systems and Computing. Human Interaction, Emerging Technologies and Future Applications III (Vol. 1253, pp. 449–454). Springer International Publishing. https://doi.org/10.1007/978-3-030-55307-4_68

Zogaj, A., Mähner, P. M., Yang, L., & Tscheulin, D. K. (2023). It's a Match! The effects of chatbot anthropomorphization and chatbot gender on consumer behavior. *Journal of Business Research*, *155*(7/8), 113412. https://doi.org/10.1016/j.jbusres.2022.113412

Appendix

Appendix A: Additional Information for Chapter 2

Chapter 2, Study 1

Data sets, analyses, and syntaxes are available on Researchbox (https://researchbox.org/636). Large parts of Appendix A are online in the published version of the manuscript constituting Chapter 2 (https://doi.org/10.5817/CP2023-3-9).

Screening study

A screening study was conducted to recruit participants for the studies in Chapter 2 to gather potential candidates to be invited to the study. The screening study aimed to gather users with experience in voice-operated technologies, such as Alexa, Google Assistant, or Siri. We invited 1050 participants on Prolific to screen for regular users of conversational AI.

Procedure of Screening Study

1. Introductory text
2. Informed consent
3. Measures
    a. Usage of conversational AI:
       This study serves as a screening study to find users who use voice-operated 'digital assistants', like Alexa (Amazon) or Siri (Apple). They are also called conversational AI here. Please indicate: What types of conversational AI are you using? (Multiple responses are possible)
       Response options: [Amazon Alexa, Google Assistant, Siri Apple, Cortana Microsoft, Bixby, Other]

    b. Frequency of use:
       How often do you use conversational AI?
       Response options: [several times a day, at least once a day, several times a week, at least monthly, less than once a month]

Main Study 1

Procedure and materials

1. Informed consent
2. Introductory text (see below)

3. Measures (see all items below)
    a. Human-AI relationship perception
    b. Trust
    c. Anthropomorphism
    d. Psychological distance
    e. Inclusion of self
    f. Affinity to technology
    g. Perceived competence
    h. Perceived warmth
    i. Frequency of use
    j. Experience of use
    k. Purpose of use
    l. Barriers of usage
4. Demographics
5. Final consent

<u>Introductory text</u>

This study focuses on the relationship of humans and their voice-operated “digital assistants”, like Alexa or Siri. They are called conversational AI in this study. Please indicate: Please select all types of conversational AI you are using (One or multiple responses are possible)

Response options: [Amazon Alexa, Google Assistant, Siri Apple, Cortana Microsoft, Bixby, Other]

After this question, we will focus on you and one specific conversational AI. Please choose one of the devices your answers will refer to in the next sections. Choose one conversational AI - please pay attention and remember your choice! In the following questions, my answers will refer to:

Response options: [Amazon Alexa, Google Assistant, Siri Apple, Cortana Microsoft, Bixby, Other]

Measures

1 Human-AI relationship, Haslam & Fiske, 1999

Prompt: Think about your chosen conversational AI. Please rate how much these statements describe your relationship with your chosen conversational AI.

1	1	There is a moral obligation to act kindly to each other
1	2	Decisions are made together
1	3	You tend to develop similar attitudes and behaviors
1	4	It seems you have something unique in common
1	5	The two of you belong together
1	6	Some requests are granted in anticipation of something in return
1	7	"One-Person, one vote'" is the principle for making decisions
1	8	You take turns doing what the other wants.
1	9	You are like peers or fellow co-partners
1	10	One of us is entitled to more than the other
1	11	One directs the work, the other pretty much follows
1	12	You are like leader and follower
1	13	One is above the other in a kind of hierarchy
1	14	What you get is directly proportional to how much you give
1	15	You have a right to a fair rate of return for what you put into this interaction
1	16	You expect the same return on your investment other people get
1	17	Your interaction is a strictly rational cost-benefit analysis

2 Trust in conversational AI (cAI), Jian, et al., 2010

Prompt: Below is a list of statements for evaluating trust between people and conversational AI (cAI). Please rate the intensity of your feeling of trust, or your impression of the system while operating the cAI. Please indicate which best describes your feelings or your impression.

2	1	The cAI is deceptive
2	2	The cAI behaves in an underhanded manner
2	3	I am suspicious of cAI's intent, action, or output
2	4	I am wary of the cAI
2	5	The cAI's action will have a harmful or injurious outcome

2	6	I am confident in the cAI
2	7	The cAI provides security
2	8	The cAI has integrity
2	9	The cAI is dependable
2	10	The cAI is reliable
2	11	I can trust the cAI
2	12	I am familiar with the cAI

3 Anthropomorphism, Waytz et al. (2010)

Prompt: Please answer the following questions referring to your chosen conversational AI (cAI)

3	1	To what extent does the cAI have thoughts of its own?
3	2	To what extent does the cAI have intentions?
3	3	To what extent does the cAI have a free will?
3	4	To what extent does the cAI have a consciousness?
3	5	To what extent does the cAI have desires?
3	6	To what extent does the cAI have values and norms?
3	7	To what extent does the cAI experience emotions?

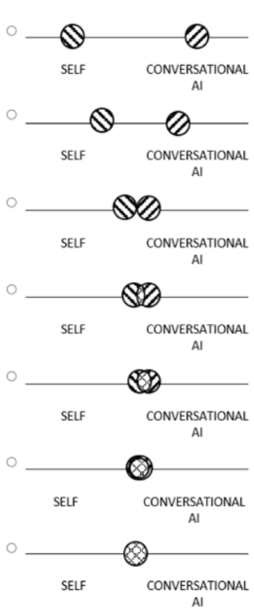
4 Psychological Distance , Li & Sung, 2021

Prompt: Please indicate how much you agree or disagree with the following statements regarding the conversational AI (short: cAI), you have chosen in the beginning.

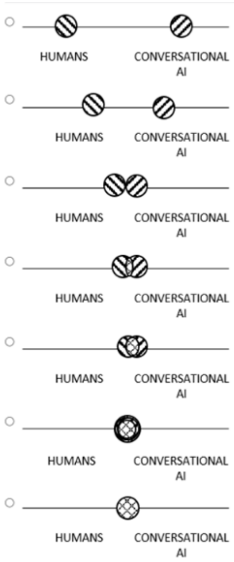
4	1	The cAI is similar to me
4	2	The cAI is psychologically close to me

5 Inclusion of AI in self, Schubert & Otten, 2002

5 1 Select the representation that you think best describes the closeness of humans and digital assistants



5 2 Select the representation that best describes your own closeness to the digital assistant



6 Affinity to Technology, Franke et al., 2019

Prompt: The following statements are about your interaction with technical systems.
The term “technical systems” refers to apps and other software applications, as well as entire digital devices (e.g., mobile phone, computer, TV, car navigation).

6 1 I like to occupy myself in greater detail with technical systems.

6 2 I like testing the functions of new technical systems.

6 3 I predominantly deal with technical systems because I have to.

6 4 When I have a new technical system in front of me, I try it out intensively

6 5 I enjoy spending time becoming acquainted with a new technical system.

6 6 It is enough for me that a technical system works; I don’t care how or why.

6	7	I try to understand how a technical system exactly works.
6	8	It is enough for me to know the basic functions of a technical system.
6	9	I try to make full use of the capabilities of a technical system

7 Perceived competence, Pitardi & Marriot, 2020

Prompt: Please indicate how much you agree or disagree with the following statements, referring to the conversational AI (cAI), you have chosen above.

7	1	I think the cAI is effective
7	2	I think the cAI is intelligent
7	3	I think the cAI is competent

8 Perceived warmth, Pitardi & Marriot, 2020

Prompt: Please indicate how much you agree or disagree with the following statements, referring to the conversational AI (cAI), you have chosen above.

8	1	I think the cAI is helpful
8	2	I think the cAI is warm
8	3	I think the cAI has good intentions

9 Frequency Usage, Funk et al., 2021

9	1	How often do you use conversational AI?

Response options: [several times a day, at least once a day, several times a week, at least monthly, less than once a month]

10 Experience of Use, Funk et al., 2021

10	1	Since when have you used conversational AI?

Response options: [5 years or more, 4-5 years, 3-4 years, 2-3 years, 1-2 years, less than 12 months]

11 Purpose of usage, Funk et al., 2021

Prompt: For what purpose do you use conversational AI? (Multiple responses are possible)

11	1	To retrieve information (e.g., How is the weather tomorrow?)
11	2	To navigate (e.g., How long do I have from home to work by bike?)
11	3	To locate (e.g., Where is the next Italian restaurant?)
11	4	To control (e.g., Call my sister!)
11	5	To shop (e.g., Can you buy me flour?)

11	6	To entertain (e.g. Tell me a joke!)
11	7	Other purposes: [text field]

12 Barriers of usage, Funk et al., 2021

Prompt: To what extent do the following reasons keep you away from using conversational AI more often, or at all in some respects e.g. shopping?

12	1	I don't want "someone" listening in all the time
12	2	I am concerned about privacy
12	3	I see no advantage in it
12	4	I find it uncomfortable/awkward to talk to a device
12	5	The cAI is often inaccurate in its statements
12	6	The cAI often does not understand me
12	7	The cAI cannot do what I expected it to do
12	8	Are there any other reasons that keep you away from using cAI more often or at all?

13 Demographic Variables

13	1	Please indicate your age
13	2	Please indicate your gender
13	3	Please Indicate your English language skills
13	4	In which country do you currently reside
13	5	What device have you used to answer this questionnaire

Additional details

Since the peer bonding scale did not exhibit a normal distribution, we employed a normal rank transformation (also referred to as area transformation or normalization). This statistical approach is used to convert a non-normal distribution observed in empirically collected data into a normal distribution.

Table A1

Normal rank transformation

Value table SPSS %	100 minus value	Value valid SPSS	Code SPSS	compute
12.8	87.2	1.00	IF rtemotional=1 Frtemotional=	-1.14
17.2	82.8	1.13	IF rtemotional=1.125 Frtemotional=	-0.95
23.7	76.3	1.25	IF rtemotional=1.25 Frtemotional=	-0.72
24.8	75.2	1.38	IF rtemotional=1.375 Frtemotional=	-0.68
27.8	72.2	1.50	IF rtemotional=1.50 Frtemotional=	-0.59
32.7	67.3	1.63	IF rtemotional=1.625 Frtemotional=	-0.45
37.6	62.4	1.75	IF rtemotional=1.750 Frtemotional=	0.32
41.4	58.6	1.88	IF rtemotional=1.875 Frtemotional=	-0.22
46.0	54.0	2.00	IF rtemotional=2.0 Frtemotional=	-0.11
49.6	50.4	2.13	IF rtemotional=2.125 Frtemotional=	-0.01
54.0	46.0	2.25	IF rtemotional=2.250 Frtemotional=	0.1
58.6	41.4	2.38	IF rtemotional=2.375 Frtemotional=	0.22
61.0	39.0	2.50	IF rtemotional=2.50 Frtemotional=	0.28
64.6	35.4	2.63	IF rtemotional=2.625 Frtemotional=	0.37
66.2	33.8	2.75	IF rtemotional=2.750 Frtemotional=	0.42
68.1	31.9	2.88	IF rtemotional=2.875 Frtemotional=	0.49
69.8	30.2	3.00	IF rtemotional=3.00 Frtemotional=	0.52
73.0	27.0	3.13	IF rtemotional=3.125 Frtemotional=	0.61
74.9	25.1	3.25	IF rtemotional=3.250 Frtemotional=	0.67
76.8	23.2	3.38	IF rtemotional=3.3750 Frtemotional=	0.73
79.3	20.7	3.50	IF rtemotional=3.50 Frtemotional=	0.82
82.3	17.7	3.63	IF rtemotional=3.625 Frtemotional=	0.93
83.9	16.1	3.75	IF rtemotional=3.750 Frtemotional=	0.99
86.6	13.4	3.88	IF rtemotional=3.875 Frtemotional=	1.11
87.2	12.8	4.00	IF rtemotional=4.00 Frtemotional=	1.14
89.1	10.9	4.13	IF rtemotional=4.125 Frtemotional=	1.23
90.5	9.5	4.25	IF rtemotional=4.250 Frtemotional=	1.31
92.4	7.6	4.38	IF rtemotional=4.375 Frtemotional=	1.43
93.2	6.8	4.50	IF rtemotional=4.5 Frtemotional=	1.49
94.0	6.0	4.63	IF rtemotional=4.625 Frtemotional=	1.56
94.8	5.2	4.75	IF rtemotional=4.750 Frtemotional=	1.63
95.9	4.1	4.88	IF rtemotional=4.875 Frtemotional=	1.65
96.2	3.8	5.00	IF rtemotional=5.0 Frtemotional=	1.78
96.5	3.5	5.13	IF rtemotional=5.125 Frtemotional=	1.82
96.7	3.3	5.25	IF rtemotional=5.250 Frtemotional=	1.85
97.3	2.7	5.38	IF rtemotional=5.3750 Frtemotional=	1.94
97.8	2.2	5.50	IF rtemotional=5.50 Frtemotional=	2.02
98.4	*1.6*	*5.63*	*IF rtemotional=5.625 Frtemotional=*	*2.14*
99.2	*0.8*	*5.75*	*F rtemotional=5.750 Frtemotional=*	*2.41*
99.7	*0.3*	*6.38*	*IF rtemotional=6.375 Frtemotional=*	*2.75*
100.0	*0.0*	*6.63*	*IF rtemotional=6.625 Frtemotional=*	*3*

Thanks to reviewer suggestions, additional analyses were made. To ensure discriminant validity, analyses using both the Fornell-Larcker criterion and Average Variance Extracted (AVE) were conducted. According to this criterion, the square root of the average variance extracted by a construct must be greater than the correlation between the construct and any other construct. Once this condition is satisfied, discriminant validity is established. If the average variance extracted is greater than 0.4 and composite reliability is higher than 0.6, the convergent validity of the construct is still acceptable (see all data https://researchbox.org/636).

Table A2

Additional Validity Analyses

	Factor loadings	**factor loadings squared**	**ave**	**ave squared**
There is a moral obligation to act kindly to each other	.474	.22		
Decisions are made together	.815	.66		
You tend to develop similar attitudes and behaviors	.788	.62		
It seems you have something unique in common	.845	.71		
The two of you belong together	.714	.51		
Some requests are granted in anticipation of something in return	.557	.31		
"One-Person, one vote" is the principle for making decisions	.387	.15		
You take turns doing what the other wants.	.761	.58		
You are like peers or fellow co-partners	.738	.54	.48	.70
One is entitled more than the other	.722	.52		
One directs the work, the other pretty much follows	.554	.301		
You are like leader and follower	.70	.49		
One is above the other in a kind of hierarchy	.715	.512	.46	.68
What you get is directly proportional to how much you give	.616	.38		
You have a right to a fair rate of return for what you put into this interaction	.741	.55		
You expect the same return on your investment other people get	0,741	0,55		
Your interaction is a strictly rational cost-benefit analysis	0,420	0,18	0,41	0,64

Correlations

		Scale peer bonding	Scale Market Pricing	Scale Authority Ranking
Scale peer bonding	Pearson Correlation	.70	.39**	-.03
Scale Market Pricing	Pearson Correlation	.40*	.70	.383**
Scale Authority Ranking	Pearson Correlation	-.03	.39**	.64

Chapter 2, Study 2

Study 2 was conducted based on the reviewer's suggestions. Data sets, analyses, and syntaxes are available on Researchbox (https://researchbox.org/636). Preregistration can be found online (https://aspredicted.org/qa5ty.pdf). Large parts of Appendix A are online in the published version of the manuscript constituting Chapter 2 (https://cyberpsychology.eu/article/view/21003).

Preregistration

Created	10/28/2022 06:13 AM (PT)
Made public	06/23/2023 07:58 AM (PT)
Available at	https://aspredicted.org/qa5ty.pdf

1) Have any data been collected for this study already?
No, no data have been collected for this study yet.

2) What's the main question being asked or hypothesis being tested in this study?
Are perceptions of the human-AI relationship influenced by the naming of the voice-operated AI?
What are the boundary conditions of different relationship perceptions?

H: When the voice-operated AI is named voice assistant (compared to conversational AI) it will lead to higher perceived authority ranking and lower perceived peer-bonding and market pricing.

3) Describe the key dependent variable(s) specifying how they will be measured.
We assess three dimensions of perceived AI-relationship: authority ranking (4 items), market pricing (4 items), and peer bonding (8 items).

4) How many and which conditions will participants be assigned to?
There will be two conditions to investigate whether the naming of the technology impacts perception. In condition 1, the voice-operated AI is called "conversational AI". In condition 2, the voice-operated AI is called "voice assistant". Participants are assigned randomly.

5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.

We will perform a 2 (naming – between) x 3 (relationship type – within) ANOVA and simple comparison of the naming factor. Evidence for the hypothesis will be provided by an interaction between both factors.

Before testing the predictions internal consistency for all scales will be tested. In case alpha < .7 and excluding an item can improve alpha or an item total correlation < .2, the respective item will be excluded before computing the respective index (i.e., the mean across items).

6) Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.

Outliers with a studentized deleted residual > 2.59 in the multiple regression of naming on the three relationship dimensions.

Exclusions of participants with no prior experience with conversational AI, at least one failed attention check, duration of survey was too short (< 150 seconds) or excessively long (> 80'000 seconds)

7) How many observations will be collected or what will determine sample size? No need to justify decision, but be precise about <u>exactly</u> how the number will be determined.

We ran a power analysis for the interaction effect in a mixed ANOVA mentioned above with an effect size f =.10, alpha = .05, 1-beta = .95. We aim for 95% power, because we aim to be able to interpret non-significant results. Based on the power analysis, we need a sample size of N =362. We will collect 400 observations.

8) Anything else you would like to pre-register? (e.g., secondary analyses, variables collected for exploratory purposes, unusual analyses planned?)

Variables explored for examining the boundary conditions:

<u>User related</u>

- Household size and specifications
- Educational level and employment
- Age

• Gender of user

• Country of origin

Technology related

• Device used

• Gender of the technology

• Technological knowledge

Main Study 2

Procedure and materials

1. Informed consent
2. Introductory text (see below)
3. Measures
    a. Human-AI relationship (see below, slight changes are marked bold)
    b. Boundary conditions
        i. Household size
        ii. Household specifics
        iii. Educational level
        iv. Employment
    c. Device used
    d. Gender of device
    e. Technological knowledge

4. Demographics
    a. Age
    b. Gender
    c. Country
5. Final consent

Introductory text

Note: Condition 1 = conversational AI, Condition 2 = voice assistant

This questionnaire will ask you questions about you and one specific [cond 1: conversational AI/cond 2 voice assistant]. It is important to now choose one [cond 1: conversational AI/cond 2 voice assistant] your answers will refer to in the next sections. Please pay attention and remember your choice! In the following questions, my answers will refer to:

Response options: [Amazon Alexa, Google Assistant, Siri Apple, Cortana Microsoft, Bixby, Other]

Measures

1 Human-AI relationship, Haslam & Fiske, 1999

Prompt: Think about your chosen [cond 1: conversational AI/cond 2 voice assistant]. Please rate how much these statements describe your relationship with your chosen [cond 1: conversational AI/cond 2 voice assistant]

1	1	There is a moral obligation to act kindly to each other
1	2	Decisions are made together ~~by consensus~~
1	3	You tend to develop similar attitudes and behaviors
1	4	It seems you have something unique in common
1	5	**You two are like a unit:** you belong together
1	6	**You are like tit for tat: you do something and expect something similar in return**
1	7	**Everyone has an equal say when a decision is made**
1	8	You take turns doing what the other wants.
1	9	You are like peers or fellow co-partners
1	10	One of us is entitled to more than the other
1	11	One directs the work, the other pretty much follows
1	12	You are like leader and follower
1	13	One is above the other in a kind of hierarchy
1	14	What you get is directly proportional to how much you give
1	15	You have a right to a fair rate of return for what you put into this interaction
1	16	You expect the same return on your **effort** other people get
1	17	Your interaction is a strictly rational cost-benefit analysis

2		**Boundary Conditions**
2	1	How many people live in your household? Response options: [1,2,3,4, more than 5]
2	2	Please specify who is living in your household Response options: [alone, spouse, children, roommates (not family), extended family, pets]
2	3	Please indicate your highest level of education Response options: [less than high school, high school graduate, some college, 2 year degree, 4 year degree, professional degree, doctorate]
2	4	On which device do mostly you use your [cond 1/cond 2]? Response options: [mobile phone, tablet, desktop computer, smart speaker without screen, with screen, other]
2	5	Which statement describes the gender of the chosen [cond 1 /cond 2] best? Response options: [male voice and gender, female voice and gender, neutral voice and gender]
2	5	Please rate your technological expertise Response options: [Not knowledgeable at all, Slightly knowledgeable, Moderately knowledgeable, Very knowledgeable, Extremely knowledgeable]
3		**Demographic variables**
3	1	Please indicate your age
3	2	Please indicate your gender
3	3	In which country do you currently reside

Additional details

Thanks to reviewer suggestions, additional analyses were made. To ensure discriminant validity, analyses using both the Fornell-Larcker criterion and Average Variance Extracted (AVE) were conducted. According to this criterion, the square root of the average variance extracted by a construct must be greater than the correlation between the construct and any other construct. Once this condition is satisfied, discriminant validity is established. If the average variance extracted is greater than 0.4 and composite reliability is higher than 0.6, the convergent validity of the construct is still acceptable (see all data https://researchbox.org/636).

Table A3

Additional Validity analyses

	factor loadings	factor loadings squared	ave	ave squared
There is a moral obligation to act kindly to each other	0,558	.31		
Decisions are made together	0,805	.65		
You tend to develop similar attitudes and behaviors	0,789	.62		
It seems you have something unique in common	0,862	.74		
The two of you belong together	0,807	.65		
Some requests are granted in anticipation of something in return	0,337	.11		
"One-Person, one vote" is the principle for making decisions	0,775	.60		
You take turns doing what the other wants.	0,744	.55		
You are like peers or fellow co-partners	0,799	.63	.54	.74
One is entitled more than the other	0,762	.58		
One directs the work, the other pretty much follows	0,747	.56		
You are like leader and follower	0,719	.52		
One is above the other in a kind of hierarchy	0,829	.69	.59	.77
What you get is directly proportional to how much you give	0,614	.38		
You have a right to a fair rate of return for what you put into this interaction	0,693	.48		
You expect the same return on your investment other people get	0,685	.47		
Your interaction is a strictly rational cost-benefit analysis	0,620	.38	.43	.65

Correlations

		Scale peer bonding	Scale Market Pricing	Scale Authority Ranking
Scale peer bonding	Pearson Correlation	.74	.39**	-.034
Scale Market Pricing	Pearson Correlation	.39**	.77	.38**
Scale Authority Ranking	Pearson Correlation	-.03	.38**	.65

Appendix B: Additional Information for Chapter 4

Chapter 4, Study 3

Data sets, analyses, and syntaxes, as well as the preregistration, are available on Researchbox (not public) (https://researchbox.org/1029&PEER_REVIEW_passcode=ZFEVZH).

Screening study

A screening study was conducted to recruit participants for the studies in Chapter 2 to gather potential candidates to be invited to the study. The screening study aimed to gather users with experience in voice shopping with conversational AI such as Alexa. Furthermore, the study explored what kind of products consumers – who have voice shopping experience – have bought in the past. We have invited 800 users to take part on Prolific to screen for users with voice shopping experience.

Procedure and materials

1. Introductory text
2. Informed consent
3. Measures
    a. Voice shopping experience
        i. If yes, list of voice shopping products
4. Thanks/Redirect

Introductory text

This study serves as a screening study to find users who have actual voice shopping experience, this means purchasing any items through the use of voice-operated 'digital assistants', for example, Amazon's Alexa. Please indicate: Have you ever purchased a product through a digital assistant (e.g., Alexa)?

Response options:

1. Yes, with Alexa
2. No, I have not engaged in voice shopping, please specify why you do not engage in voice-shopping [open text field]

3. Yes, but with other voice-operated technology, please specify below: [open text field]

Voice shopping – products

If yes, participants were forwarded to answer questions regarding what kind of products they have bought:

Introductory text for low-involvement products:

Think about your past voice shopping experiences. Have you used the voice assistant to shop for products, which are rather convenience products, that require no effort to buy, and there is no emotional values or risk attached? For example, products such as paper towels, chewing gum, cereals or a specific book.

Response options:

1. Yes (please explain what you bought - as many items as possible) [open text field]
2. No (please indicate why you do not engage in voice-shopping of these products) [open text field]

Introductory text for high-involvement products:

Think about your past voice shopping experiences. Have you used the voice assistant to shop for products, which are rather complicated and require thinking to make a decision, with higher emotional values or risks attached? For example, a laptop, a smartphone, vehicle or tablet.

Response options:

1. Yes (please explain what you bought - as many items as possible) [open text field]
2. No (please indicate why you do not engage in voice-shopping of these products) [open text field]

Main Study

Preregistration

Created	09/23/2022 08:28 AM (PT)
Made public	Not yet, will be made publicly available after publication
Available at	https://aspredicted.org/P9T_XW8

1) Have any data been collected for this study already?

No, no data have been collected for this study yet.

2) What's the main question being asked or hypothesis being tested in this study?

The study tests whether perceived human-AI relationship relates to different forms of voice shopping (low and high involvement shopping)

H1: Higher values in peer bonding predict a stronger intention to use voice shopping.

H2: Higher values in market pricing predict a stronger intention to use voice shopping.

H3: The intention to buy low-involvement products via voice shopping is predicted to a stronger extent by peer bonding than by market pricing.

H4: The intention to buy high-involvement products via voice shopping is predicted to a stronger extent by market pricing than by peer bonding.

Research question: how does authority ranking relate to voice shopping intentions in general as well as low and high involvement products?

3) Describe the key dependent variable(s) specifying how they will be measured.

Dependent variable is the intention to continue voice shopping. We differentiate general intention (3 items), intention to buy low involvement products (1 item), intention to buy high involvement products (1 item)

4) How many and which conditions will participants be assigned to?

No manipulation. The independent variables (perceived human-AI relationship) will be assessed with peer bonding (9 items), authority ranking (4 items), market pricing (4 items)

5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.

To test Hypotheses 1 and 2 general voice shopping intentions will be regressed on the three relationship dimensions. Evidence for H1 is provided by a positive significant regression coefficient of peer bonding. Evidence for H2 is provided by a positive significant regression coefficient of market pricing.

To test Hypotheses 3 and 4 voice shopping intentions for low and high-involvement products will be regressed on the three dimensions of relationship perception. Evidence for H3 is provided if the regression coefficient for peer bonding lies outside the CI for the regression coefficient for market pricing in the regression with low involvement shopping intention as dependent variable. Evidence for H4 is provided if the regression coefficient for market pricing lies outside the CI for the regression coefficient for peer bonding in the regression with high involvement shopping intention as dependent variable.

Before testing the predictions internal consistency for all scales will be tested. In case alpha < .7 and excluding an item can improve alpha or an item total correlation < .2, the respective item will be excluded before computing the respective index (i.e., the mean across items).

6) Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.

Outliers with a studentized deleted residual > 2.59 in the regression testing H1/2 will be excluded.

Exclusions of participants with no prior experience in voice shopping, at least one failed attention check, duration of survey was too short (< 150 seconds) /excessively long (> 80'000 seconds)

7) How many observations will be collected or what will determine sample size?

No need to justify decision, but be precise about exactly how the number will be determined.

We will collect 450 observations. Schönbrodt & Perugini (2013) suggest that the sample size should approach 250 for stable estimates. We include 200 further observations as we expect to exclude a number of observations.

8) Anything else you would like to pre-register?

(e.g., secondary analyses, variables collected for exploratory purposes, unusual analyses planned?)

Perceived benefits (15 items adapted McLean & Osei-Frimpong, 2019), desired benefits (10 items adapted McLean & Osei-Frimpong, 2019), trust (16 items, Ullman & Malle, 2019), user characteristics (device clarification, screen use, frequency, experience, amount of spending, age, gender)

<u>Deviations from preregistration</u>

In the submitted manuscript, we focus specifically on the desired benefits (stated in (8) as variables for exploratory purposes) to reinforce our claims.

<u>Procedure and materials (main measures bold)</u>

1. Informed consent
2. Introductory text
3. Measures
    a. **Human-AI relationship**
    b. **Voice shopping intentions (general, low-involvement, high-involvement)**
    c. Perceived benefits
    d. **Desired benefits**
    e. Trust
    f. **User characteristics**
        i. **Device clarification**
        ii. **Device specification**
        iii. **Frequency of use**
        iv. **Duration of use**
        v. **Shopping spendings**
        vi. Barriers
4. Demographics
    a. Age
    b. Gender
5. Final consent

<u>Introductory text</u>

This survey will focus on you and one specific voice assistant you have used for shopping for any goods. Please choose one of the devices your answers will refer to in the following sections. Choose one voice assistant - please pay attention and remember your choice while answering the questions hereafter! In the following questionnaire, my answers will refer to: *Response options: [Amazon Alexa, Google Assistant, Siri Apple, Cortana Microsoft, Bixby, Other]*

<u>Measures</u>

Note: Changes compared to prior questionnaire marked bold.

1 Human-AI relationship, Haslam & Fiske, 1999

Prompt: Think about your **past shopping experiences with your chosen voice assistant**. Please rate how much these statements below describe your relationship with your chosen voice assistant.

1	1	There is a moral obligation to act kindly to each other
1	2	Decisions are made together
1	3	You tend to develop similar attitudes and behaviors
1	4	It seems you have something unique in common
1	5	You two are like a unit: you belong together
1	6	You are like tit for tat: you do something and expect something similar in return
1	7	Everyone has an equal say when a decision is made
1	8	You take turns doing what the other wants.
1	9	You are like peers or fellow co-partners
1	10	One of us is entitled to more than the other
1	11	One directs the work, the other pretty much follows
1	12	You are like leader and follower
1	13	One is above the other in a kind of hierarchy
1	14	What you get from your interaction is directly proportional to how much you give
1	15	You have a right to a fair rate of return for what you put into this interaction
1	16	You expect the same return on your effort other people get
1	17	Your interaction is a strictly rational cost-benefit analysis

2 General Voice shopping intentions

Prompt: Please rate how much these statements describe your intention to continue shopping with your voice assistant in the future.

2	1	I plan to continue to use the voice assistant for shopping in the future.
2	2	I intend to continue to use the voice assistant for shopping in the future.
2	3	I predict to continue to use the voice assistant for shopping in the future.

3 Low-involvement voice shopping intentions

Prompt: Think about your future voice shopping experiences. Would you use the voice assistant to shop for products, which are rather convenience products, that require no effort to buy, and there is no emotional values or risk attached? For example, products such as paper towels, chewing gum, cereals or a specific book. Please rate the extent to which these statements describe your intention to continue purchasing these types of products with your voice assistant in the future.

3	1	I predict to continue to use the voice assistant for shopping in the future.

4 High-involvement voice shopping intentions

Prompt: Think about your future voice shopping experiences. Would you use the voice assistant to shop for products, which are rather complicated and require some effort to make a decision, with higher emotional values or risks attached? For example, a laptop, a smartphone, vehicle or tablet. Please rate the extent to which these statements describe your intention to continue purchasing these types of products with your voice assistant in the future.

4	1	I predict to continue to use the voice assistant for shopping in the future.

5 Perceived benefits (McLean and Osei-Frimpong, 2019)

5	1	I find using my voice assistant to shop to be enjoyable
5	2	The actual process of using my voice assistant for shopping purposes is entertaining
5	3	I have fun using my voice assistant to complete tasks in a shopping context
5	4	Using my voice assistant is a convenient way to manage my time when shopping
5	5	Shopping with my voice assistant makes my life easier
5	6	Shopping with the voice assistant fits with my schedule
5	8	Shopping with the voice assistant is an efficient use of my time
5	9	My voice assistant helps me achieve my shopping goals
5	10	I make better shopping decisions with my voice assistant
5	11	Using my voice assistant to shop enhances my image amongst my peers
5	12	Using my voice assistant to shop makes me seem more valuable amongst my peers
5	13	Using my voice assistant to shop is a status symbol for me

5	14	When I shop with the voice assistant it feels like someone is present in the room
5	15	My shopping interactions with the voice assistant are similar to those with a human
5	16	During my communication with the voice assistant I feel like I am dealing with a real sales agent

6 Desired benefits (adapted from McLean and Osei-Frimpong, 2019)

6	1	It is important to me to have fun while shopping with my voice assistant
6	2	I want to enjoy shopping with my voice assistant
6	3	It is important to me that the voice assistant makes shopping more efficient
6	4	I care about that shopping with my voice assistant is an efficient use of time
6	5	I care about makeing better shopping decisions with my voice assistant
6	6	It is important to me that my voice assistant helps me achieve my shopping goals
6	7	It matters to me that voice-shopping is like a status symbol
6	8	It is important to me that shopping with my voice assistant enhances my image among my peers
6	9	I care that shopping with a voice assistant is like dealing with a real person
6	10	It is important to me that when shopping with a voice assistant it feels like someone is present in the room

7 Trust (Ullman & Malle, 2019)

Prompt: Please rate your conversational AI using the scale from 0 (not at all) to 7 (very). If a particular item does not seem to fit the voice assistant in the situation, please select the option that says "Does not fit".

7	1	It is important to me to have fun while shopping with my voice assistant
7	2	I want to enjoy shopping with my voice assistant
7	3	It is important to me that the voice assistant makes shopping more efficient
7	4	I care about that shopping with my voice assistant is an efficient use of time
7	5	I care about makeing better shopping decisions with my voice assistant
7	6	It is important to me that my voice assistant helps me achieve my shopping goals
7	7	It matters to me that voice-shopping is like a status symbol

8 Device clarification

8 1 Which device do you usually use while voice shopping?

Response options: [Smart speaker at home, smart phone, smart tablet, desktop computer, other, please specify[text field]]

9 Device specifics

9 1 Do you use a screen while shopping with your voice assistant?

Response options: [Yes, I look at a screen while I shop with the voice assistant (I see the products), No, I do not look at the screen while I shop with the voice assistant (I do not see the products), Both, sometimes with and sometimes without a screen. Please explain: [text field], Other answer, please specify[text field]]

10 Frequency Usage, Funk et al., 2021

10 1 How often do you use your voice assistant for shopping?

Response options: [Almost daily, several times a month, about monthly, every 2-3 months, about 1-2 times per year, not at all]

11 Experience of Use, Funk et al., 2021

11 1 Since when do you use voice assistant for shopping purposes?

Response options: [5 years or more, 4-5 years, 3-4 years, 2-3 years, 1-2 years, less than 12 months]

12 Shopping spendings

11 1 On average, how much money (in £) do you spend via voice shopping per year?

Response options: [number in GB Pounds]

13 Barriers

13 1 It is difficult for me to achieve my shopping goals with my voice assistant

13 2 My voice assistant does not help me to make better shopping decisions

14 Demographic variables

14 1 Please indicate your age

14 2 Please indicate your gender

Additional details

A principal component analysis (PCA) with orthogonal rotation (varimax) was conducted on the 5 items of the voice shopping intentions measure. The three factors explained 97.62% of the variance. The table below shows the factor loadings after rotation.

Table B1

Results from a factor analysis of the voice shopping intentions questionnaire (N=423)

Item	Factor Loading		
	1	2	3
General shopping intention			
I plan to continue to use the voice assistant for shopping in the future	**.901**		
I intend to continue to use the voice assistant for shopping in the future	**.903**		
I predict I would continue to use the voice assistant for shopping in the future	**.899**		
Intention to continue voice shopping for low-involvement products			
I plan to continue to use the conversational AI for shopping in the future	.535		**.826**
Intention to continue voice shopping for low-involvement products			
I plan to continue to use the conversational AI for shopping in the future		**.973**	

Note. Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization. Rotation converged in 5 iterations. The highest factor loadings are in bold, factor loadings below .30 are not displayed.

A principal component analysis (PCA) with orthogonal rotation (varimax) was conducted on the 10 items of the desired benefits measure. The three factors explained 81.43% of the variance. The table below shows the factor loadings after rotation.

Table B2

Results from a factor analysis of the desired benefits questionnaire (N=423)

Item	Factor Loading			
	1	2	3	4
Hedonic benefits				
It is important to me to have fun while shopping with my voice assistant			.321	**.829**
I want to enjoy shopping with my voice assistant	.447			**.795**
Utilitarian benefits				
It is important to me that the voice assistant makes shopping more efficient	**.854**			
I care about that shopping with my voice assistant is an efficient use of time	**.875**			
I care about making better shopping decisions with my voice assistant	**.656**		.420	
It is important to me that my voice assistants help me achieve my shopping goals	**.791**			
Symbolic benefits				
It matters to me that voice-shopping is like a status symbol		**.904**		
It is important to me that shopping with my voice assistant enhances my image among my peers		**.905**		
Social benefits				
I care that shopping with a voice assistant is like dealing with a real person			**.883**	
It is important to me that when shopping with a voice assistant it feels like someone is present in the room		.387	**.790**	

Note. Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization. Rotation converged in 5 iterations. The highest factor loadings are in bold, factor loadings below .30 are not displayed.

In a pretest we surveyed experienced voice shoppers, what kind of products they buy categorized into the conditions we aimed to test.

Table B3

Results from a Pretest results of purchased low- and high-involvement items (N=800)

Low-involvement items bought

Instructions
Think about your past voice shopping experiences. Have you used the voice assistant to shop for products, which are rather convenience products, that require no effort to buy, and there is no emotional values or risk attached? For example, products such as paper towels, chewing gum, cereals or a specific book. - Yes (please explain what you bought - as many items as possible)

- Amazon fire stick
- Yes I bought a book from Amazon called a pinch of nom
- Shorts
- Books, holdall, superglue
- board games
- Mouthwash
- kitchen roll, antibacterial spray, hand soap
- A smart TV
- milk, bread, chocolate, sausages, other groceies, bluray, coffee pods
- cat food
- Groceries (food) + household items + cleaning products + single-use items such as batteries + consumables such as pens and other stationary
- Books pc games desk coffee machine
- I bought sanitary pads, toilet tissue, toothpaste, milk, brown sugar. I bought them in packs and on several occasion on amazon fresh
- hoover bags
- book
- Shampoo, toothbrushes , batteries, bird-seed , garden products , cleaning products , groceries
- A webcam
- Indigestion tablets..candles..lighter fluid..tissues..beef jerky
- Stud finder, baby play mat, toilet roll
- It was some headphones
- General food
- Alcohol, fizzy drinks, fifa points

High Involvement Items bought

Instructions
Think about your past voice shopping experiences. Have you used the voice assistant to shop for products, which are rather complicated and require thinking to make a decision, with higher emotional values or risks attached? For example, a laptop, a smartphone, vehicle or tablet. - Yes (please explain what you bought - as many items as possible)

- Amazon fire stick
- I have bought my food shopping on my mobile
- Shorts
- Yes. A piece of jewellery. My questions were easy answered so used.
- We have bought a tablet for our daughter
- laptop
- A second echo dot from Amazon
- Chlorhexidine gluconate 3x500ml
- I ordered something specific as thought this would be easier as had already done the research
- router, echo dot
- Smart tv
- replacement trimmer spool, headphones, bluetooth receiver
- baby milk
- Laptop + replacement computer parts such as a hard drive + books based on reviews
- Tablet
- Mobile phone. Got a Samsung galaxy A21s
- hoover bags
- Laptop, as I used it to compare specifications and purchase the best item
- Garmin watch , cctv cameras
- Webcam
- Barbecue..picture frames
- Armarni watch
- I have added toys and gifts to my cart using Alexa

- cat food
- Food items
- I got my AirPods
- I have used it a few times, to buy toilet roll and basic household things
- Toilet rolls and shower gel
- Tea bags, chocolate, toilet roll, kitchen roll, bin bags, bread
- Wrigleys Extra Bubblemint, Centrum Advanced Vitamins, Steradent, Listerine, Oral B Toothbrushes, Shower Gel etc.
- Dog food, nappies
- Vacuum cleaner, baby changing table,
- clothes
- I have so far only used Alexa to place repeat orders for things I regularly buy online - coffee comes to mind, also nuts and oscillating saw blades.
- brought the latest stephen king book
- Make up, technology, headphones, stationery, food items
- toilette rolls
- Dog poo bags
- DVDs, baby monitor, sun cream, make up
- DVD, book, groceries
- A magazine subscription
- perfume
- Book, sun cream,showergel
-
- Just for something I have ordered before
- a few books, i like the convenience if i see a book being mentioned somewhere i can ask alexa to buy it and its sent to my kindle.
- charger plug for new tablet
- Chocolate, rice, tissues, sweets, deride fruit, tea, coffee, breakfast cereal,
- Music, art products
- i bough makeup and stationery
- Baby diapers, cerelac, veggie straws, tissue
- I have ordered subscribe and save items for my dog such as chews and dog food. I have bought groceries from Amazon Fresh using Alexa and added items to my basket. I regularly buy things like mouthwash and drinks like Soya milk and Lucozade, as well as things like notepads and envelopes and napkins using Alexa.
- Shampoo
- Yes as the headphones were specific
- A tablet
- Electrical
- Printer
- cat littter
- Yes watch etc
- Pasta, rice, crisps
- just my AirPods
- New phone
- Toilet roll, bleach, washing up liquid, black bin liners
- phone case and clothes
- Laptop, games console
- WD 18TB External Hard Drive
- Smart cameras
- Clothes, food
- Kids tablet
- Trousers
- Only one example here. My latest smartphone - but some research was done first on a tablet.
- Brought a video doorbell that links with Alexa
- Laptop, amazon tablet, alexa dot and echo
- desk fan
- Dog poo bags
- Baby monitor
- iPad
- Magazine subscription
- It was linked to amazon
- laptop
- A car
- printer ink
- a firestick
- I purchased a bag through Alexa, but it was one I had been browsing on the app and she told me it was on offer that day.
- screen protectors for new tablet
- Fan heater, hair dryer, lamp
- TV, Soundbar, Kettle
- Laptop, Echo dot
- Echo dot
- makeup
- Smart watch
- Yes, I bought my laptop a month ago using Alexa. I also bought a GPS tracker with an activity tracker for my dog too.
- I bought a second alexa unit on a voice based offer
- smartphone
- A fire tablet
- Beard oil

• Groceries, the book The Thursday Murder Club. Vinyl record by Blur • mouse pad	• A Motorola smartphone, a Fire tablet

Note. Prolific survey, June 2022

For exploratory reasons we have surveyed experienced voice shoppers for their shopping habits. We deleted the table from the published paper for reasons of space based on the reviewers' suggestions.

Table B4

User information on voice shopping (N=423)

Which device do you use for voice shopping?	**Do you use a screen while voice shopping?**	**How often do you engage in voice shopping?**	**Since when you engage in voice shopping?**	**Average spending per year in £ (GBP)?**
• 75% use only a smart speaker at home • 11% use only a smartphone • 14% use both or other devices	• 54% do not look at a screen • 37% look at a screen and see the products • 9% do both	• 32% voice shop several times a month • 29% voice shop monthly • 19% every 2-3 months • 15% about 1-2 times per year • 4% daily and 1% less than yearly)	• 34% have been voice shopping for 1-2 years • 28% for 2-3 years • 22 % for more than 3 years • 16% for less than 12 months	*M*=415.82, *SD*=904.64

Appendix C: Additional Information for Chapter 6

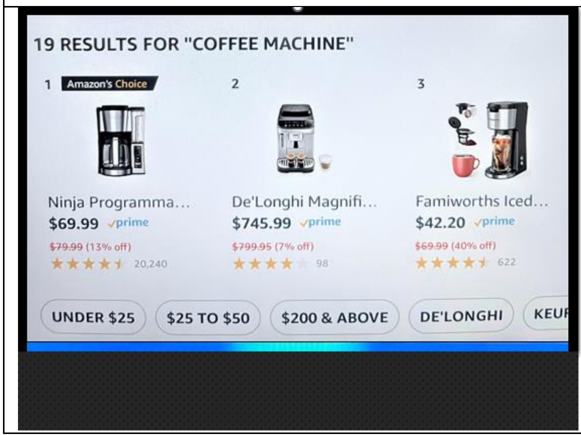
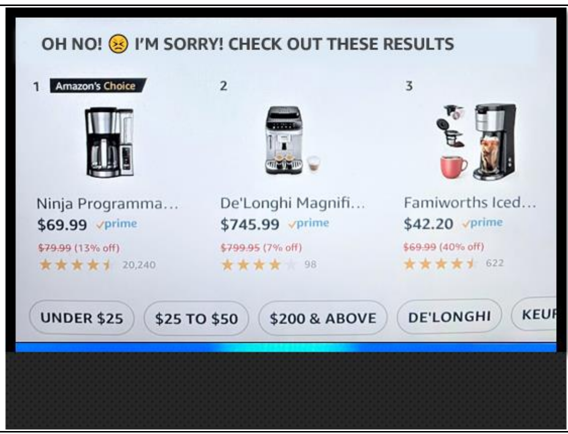
Chapter 6, Study 4

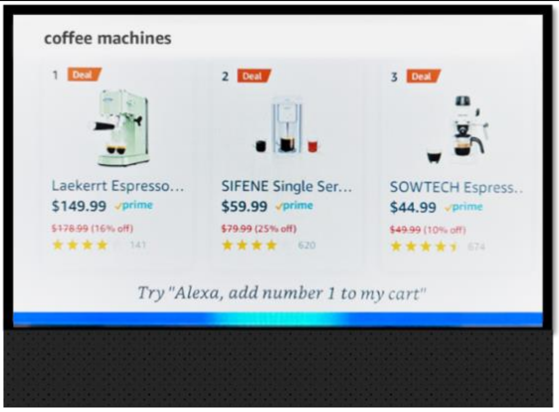
Data sets, analyses, and syntaxes, as well as the preregistration, are available on Researchbox (not public) (https://researchbox.org/1540&PEER_REVIEW_passcode=ZGFTER).

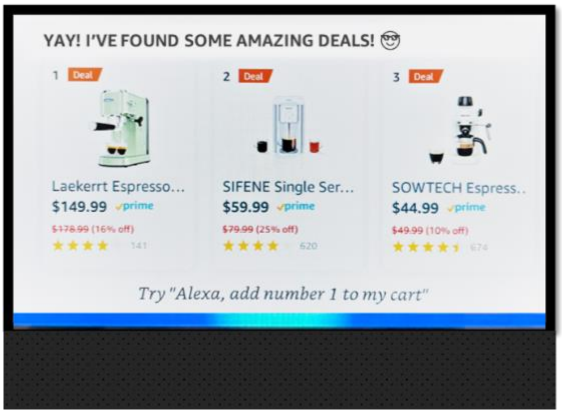
<u>Complete list of stimulus materials</u>

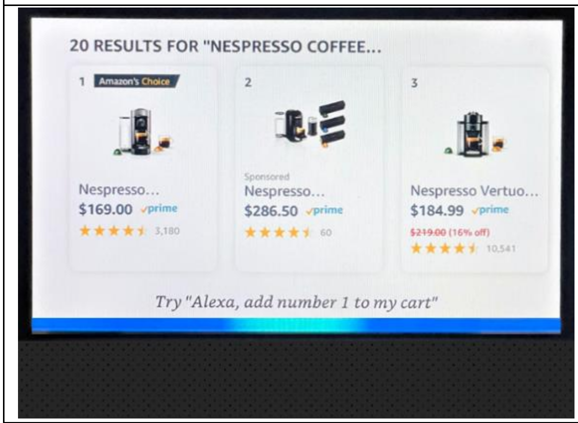
- Videos
    - Kellogs shopping video original: https://youtu.be/KXKDZBhZ0Cg
    - Kellogs shopping video manipulated: https://youtu.be/itCWoepRpAU
    - Coffee machine shopping original: https://youtu.be/FLrpHEz2aR4
    - Coffee machine shopping manipulated: https://youtu.be/KXo30d6dGI0

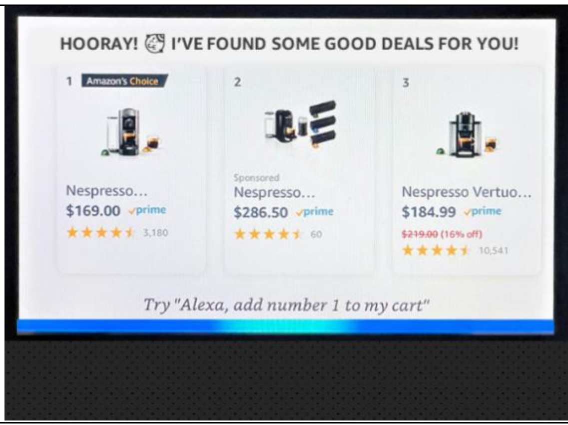
- Screenshots of the standard and manipulated conversational design, that were shown to the participants of the experiment
    - Coffee machine video

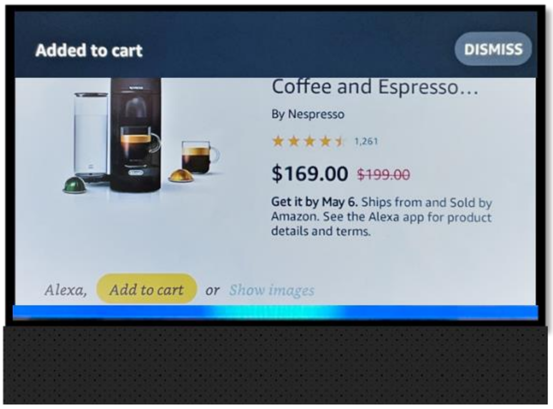
Normal Alexa	Emotionalized Alexa
	

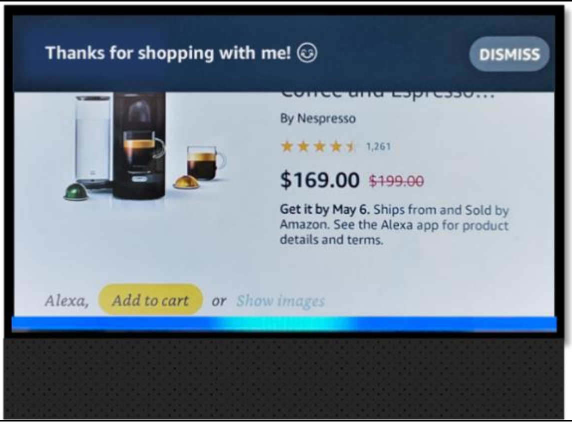

coffee machines
1 Deal
2 Deal
3 Deal
Laekerrt Espresso...
$149.99 prime
SIFENE Single Ser...
$59.99 prime
SOWTECH Espress..
$44.99 prime
Try "Alexa, add number 1 to my cart"


YAY! I'VE FOUND SOME AMAZING DEALS!
1 Deal
2 Deal
3 Deal
Laekerrt Espresso...
$149.99 prime
SIFENE Single Ser...
$59.99 prime
SOWTECH Espress..
$44.99 prime
Try "Alexa, add number 1 to my cart"

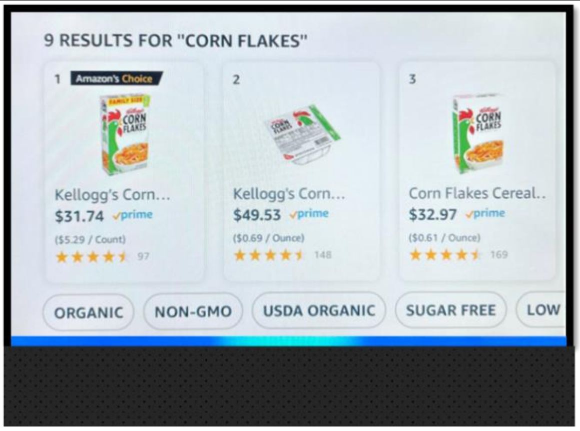
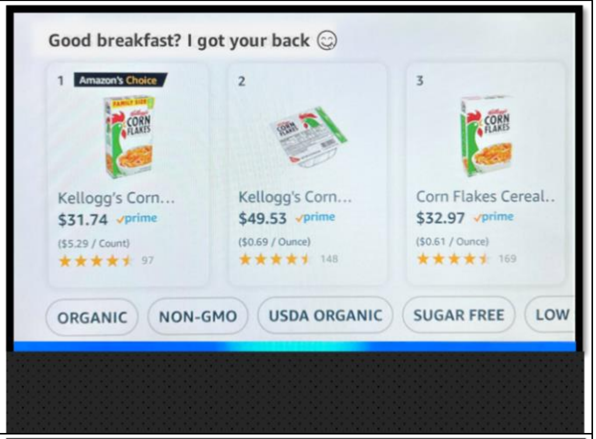
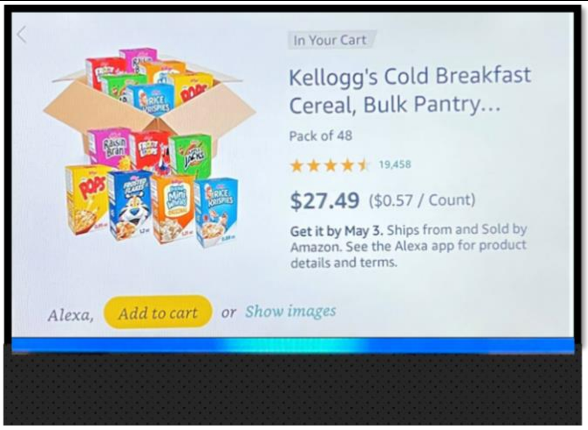
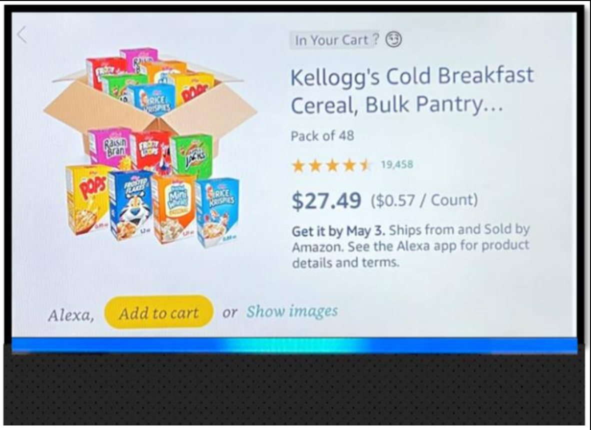
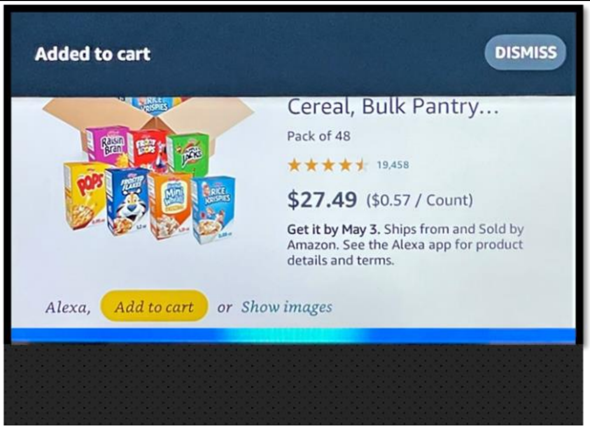
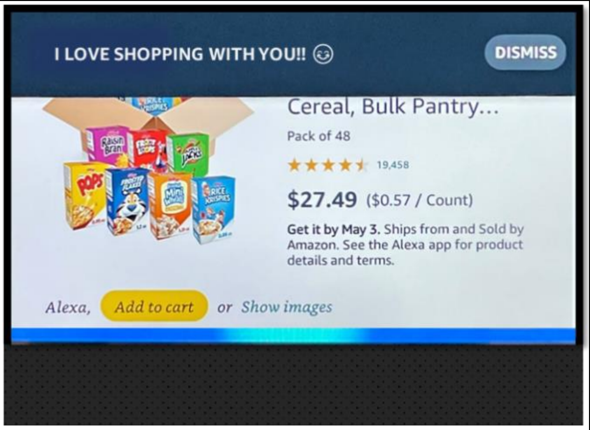

20 RESULTS FOR "NESPRESSO COFFEE...
1 Amazon's Choice
2
3
Nespresso...
$169.00 prime
Sponsored
Nespresso...
$286.50 prime
Nespresso Vertuo...
$184.99 prime
Try "Alexa, add number 1 to my cart"


HOORAY! I'VE FOUND SOME GOOD DEALS FOR YOU!
1 Amazon's Choice
2
3
Nespresso...
$169.00 prime
Sponsored
Nespresso...
$286.50 prime
Nespresso Vertuo...
$184.99 prime
Try "Alexa, add number 1 to my cart"


Added to cart
DISMISS
Coffee and Espresso...
By Nespresso
1,261
$169.00 $199.00
Get it by May 6. Ships from and Sold by Amazon. See the Alexa app for product details and terms.
Alexa, Add to cart or Show images


Thanks for shopping with me!
DISMISS
By Nespresso
1,261
$169.00 $199.00
Get it by May 6. Ships from and Sold by Amazon. See the Alexa app for product details and terms.
Alexa, Add to cart or Show images

Cornflakes video

Normal Alexa	Emotionalized Alexa
	
	
	

Procedure and materials pretest

1. Informed consent
2. Introductory text
3. Stimulus material (watching 2 videos)
4. Measures

a. Attitude (manipulation check)

b. Human-AI relationship (see table Appendix B)

Pretest/Manipulation check (7p Likert scale 1=strongly disagree – 7=strongly agree)

How would you describe the voice assistant **Alexa for shopping as it is shown in the video**? Please rate your agreement, there are no right or wrong answers. In the video Alexa was... (7p Likert scale 1=strongly disagree – 7=strongly agree)

1 Attitude

1	1	Friendly
1	2	Warm
1	3	Welcoming
1	4	Competent
1	5	Functional
1	6	Effective
1	7	Confident
1	8	Dominant
1	9	Competitive

Additional details

Table C1

Descriptive Statistics Pretest (N=200)

Emotional	Standard	Mean	SD	N
-1,00	-1,00	5,03	1,06	53
	1,00	4,75	1,19	41
	Total	4,91	1,13	94
1,00	-1,00	5,55	,96	54
	1,00	5,19	1,12	48
	Total	5,38	1,05	102
Total	-1,00	5,30	1,06	107
	1,00	4,98	1,17	89
	Total	5,15	1,11	196

Tabe C2

Test of Between-Subjects-Effects (N=200)

Source	Typ III Squaresum	df	Mean square	F	Sig.	Partial Eta square
Corrected Modell	16,03[a]	3	5,343	4,579	,004	,067
Intercept	5091,798	1	5091,798	4363,704	,000	,958
Emo	10,937	1	10,937	9,373	,003	,047
vid	5,151	1	5,151	4,414	,037	,022
emo * vid	,088	1	,088	,076	,784	,000
Error	224,036	192	1,167			
Total	5441,222	196				
Corrected total	240,065	195				

a. R-Squared = .067 (adjusted R-Squared = .052)

Main study

Preregistration

Created	05/15/2023 06:10 AM (PT)
Made public	Not yet, will be made publicly available after publication
Available at	https://aspredicted.org/1H1_YR9

1) Have any data been collected for this study already?

No, no data have been collected for this study yet.

2) What's the main question being asked or hypothesis being tested in this study?

In this study, we test whether the anthropomorphic design cues (normal voice shopping assistant vs. emotionalized voice shopping assistant) should match consumers' perception of their human-AI relationship in order to promote voice shopping intentions.

H1: The more people report peer bonding the more they have stronger voice shopping intentions after seeing the emotionalized (vs. normal) interface.

H2: The more people report authority ranking the more they have stronger voice shopping intentions after seeing the normal (vs. emotionalized) interface.

RQ1: Is the effect of the interface design contingent to market pricing?

3) Describe the key dependent variable(s) specifying how they will be measured.

• General attitude towards the voice assistant (5 bipolar items, 7-point Likert scale)

• Voice shopping intention (2 items, 7-point Likert scale)

4) How many and which conditions will participants be assigned to?

There will be two conditions, participants will be assigned randomly. In condition one, participants see two videos of voice shopping using the "normal" Alexa. In condition two, participants see two voice shopping videos using an "emotionalized" version of Alexa (dialogue and screenshots are edited to appear more emotional).

5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.

Before testing the predictions internal consistency for all scales will be tested. In case alpha < .7 and excluding an item can improve alpha or an item total correlation < .2, the respective item will be excluded before computing the respective index (i.e., the mean across items).

H1 will be tested with two multiple regressions with voice shopping intention and attitude as dependent variables respectively and peer bonding (centered), condition (-1 normal, 1 emotional Alexa), and their interaction as predictors.

H2 will be tested with analog multiple regression with authority ranking instead of peer bonding.

In both cases, the interaction will provide evidence for the prediction.

6) Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.

Inclusion requirements for participants:

- at least 18 years old,

- consent for participation and usage of data for scientific purposes,

- passing two attention checks.

After excluding participants according to these criteria, data will be checked for outliers using studentized deleted residuals (SDR). Outliers with a studentized deleted residual > 2.59 in the regression testing of H1 will be excluded.

7) How many observations will be collected or what will determine sample size?

No need to justify decision, but be precise about exactly how the number will be determined.

We ran a power analysis with the input parameters f2=0.02 (small effect), alpha error 0.05, and beta = 0.80. We aim for 80% power with 1 tested predictor and 3 predictors in total. The total sample size should be 395 for a significant model. We will collect 450 observations.

8) Anything else you would like to pre-register?

(e.g., secondary analyses, variables collected for exploratory purposes, unusual analyses planned?)

• User characteristics (device clarification, frequency, experience, barriers of usage, amount of spending, age, gender)

• Voice shopping benefits (2 items, 7-point Likert scale)

Deviations from preregistration

Exclusion criteria were failing one attention check instead of two attention checks (the second attention check did not work as intended). Additionally, we excluded participants who used a mobile phone to participate in the study because the stimulus videos were not clearly visible on screens of that size.

Our manuscript used voice shopping benefits as an important dependent variable versus the preregistered secondary variable/variable for exploratory purposes.

Procedure and materials

1. Informed consent
2. Introductory text (see below)
3. Measures (main measures bold)
   a. **Human-AI relationship** (see Appendix B)
4. Stimulus material (watching 2 videos, see above)
5. Voice shopping (see items below)
   a. Voice shopping attitude
   b. Voice shopping benefits
   c. Voice shopping intentions
6. Demographics
   a. Age
   b. Gender
7. User characteristics (measures, Appendix B)
   a. Frequency of general use
   b. Frequency of voice shopping
   c. Experience of use
   d. Voice shopping device

   e. Voice shopping spending
   f. Device used for the survey
8. Demographics
   a. Age
   b. Gender
9. Final consent

Introductory text

The next questions will ask you questions about how you perceive your specific voice assistant you are using at home. To answer the question, you first need to state to which voice assistant your answers refer to:
Response options: [My answers refer to Alexa (Amazon), My answers refer to another voice assistant. Please specify:[text field]]

Measures

2 Voice shopping attitude

Prompt: In the videos demonstrating voice shopping with Alexa, the voice assistant appeared to be…

2	1	Annoying – Pleasant
2	2	Bad - Good
2	3	Negative - Positive
2	4	Harmful - Beneficial
2	5	Risky - Safe

3 Voice shopping benefits

Prompt: Please rate the following statements with regard to the specific voice assistant as shown in the video. There are no right or wrong answers.

3	1	Shopping with this kind of voice assistant would be fun
3	2	Shopping with this kind of voice assistant would be efficient

<u>4 Voice shopping intentions</u>

Prompt: Please rate the following statements with regard to the specific voice assistant as shown in the video. There are no right or wrong answers.

4 1 I would like to use the voice assistant as shown in the video for shopping in the future

4 2 I would also use this kind of voice assistant for shopping for simple household. items like cereals or a coffee machine

<u>Additional details</u>

<u>Dialogue</u>

<u>V1 Normal Alexa, coffee machine</u>

User: Alexa my coffee machine broke down I need a new coffee machine

Alexa: one moment showing results for coffee machines you can narrow down the choices by saying things like filtered by Under 125.

User: hey Alexa are there any deals on coffee machines

Alexa here are the top Amazon deals on coffee machines the first one is laca espresso machine 20 bar espresso maker with milk processor it's 149 dollars

User: Alexa, I think I want a Nespresso coffee machine. Which one is the cheapest?

Alexa: Are you trying to shop for Nespresso coffee machines?

User: yes

Alexa: one moment showing results for Nespresso coffee machines the first one is Amazon's choice you can filter by saying filter under 125 dollars

User: okay Alexa add number one to my cart

Alexa: Okay added it to cart.

<u>V2 Normal Alexa, cornflakes</u>

User_ I'm running out of corn flakes.

Alexa: one moment showing results for corn flakes you can narrow down the choices by saying things like filter by organic or filter by sugar-free

User: Alexa filter by Kellogg's Corn Flakes

Alexa: one moment showing results the first one is Kellogg's Corn Flakes 26 ounce it's 52.62 cents one moment here is more information on Kellogg's cold breakfast cereals for 27.49 dollars.

User: Alexa show ratings

Alexa: Kellogg's cold breakfast cereal bulk Pantry Staples kid snacks it has a rating of 4.7 from 19469 customers

User: Alexa add to cart

Alexa: added to cart

V3 Emotional Alexa, coffee machine

User: Alexa, my coffee machine broke down I need a new coffee machine

Alexa. oh no I'm sorry to hear that I'm sure you'll love these results you can filter by saying filter by under 125. let's find your perfect coffee machine together.

User. hey Alexa, are there any deals on coffee machines.

Alexa: hmm let's check I found some pretty amazing deals on coffee machines let's check them out the first one is the laca espresso machine and it even comes with a milk frother it's 149 dollars.

User: Alexa I think I want a Nespresso coffee machine which one is the cheapest

Alexa: oh, frick I'm not sure if I got that right are you looking for Nespresso coffee machines today?

User: Yes.

Alexa: sweet idea my friend I think we just found the perfect coffee machine for you check out Amazon's top pick what do you think should we jump right in with this deal.

User: okay Alexa add number one to my cart

Alexa: okie dokie buddy that Nespresso machine is on its way to our home soon enjoy your coffee and let me know how you liked it

V4 Emotional Alexa, cornflakes

User: Alexa, I'm running out of corn flakes

Alexa: I'm not gonna let you starve my friend so let me think I found all kinds of corn flakes are you in the mood for something healthy or I might also surprise you with something new

User: Alexa filter by Kellogg's Corn Flakes

Alexa: okey dokey look I found some awesome Kellogg's deals the first one is corn flakes for thirty one dollars hmm But is that what you really want

USER scrolling/clicking

Alexa: I knew you wouldn't like boring cereals here is more info on Kellogg's cold breakfast cereals.

User: Alexa show ratings

Alexa: now look at that 4.7 out of 5 of 19455 customers looks like a pretty decent rating what do you think should we jump right in with this offer

User: Alexa at the court

Alexa: you made a good choice there consider it done and don't forget to be customer 19456 on the ratings

Summary

First, we typed, then we swiped, and nowadays, we talk to computers. Artificial intelligence (AI) allows users to communicate with computers in natural language. Conversational AI, like Amazon's Alexa, are just tools assisting users in many tasks (Kulkarni et al., 2019). However, humans often anthropomorphize technologies and respond in social ways (Gambino et al., 2020; Nass & Moon, 2000b). Moreover, recent progress in AI rendered human-AI interactions not only social in the sense of being imbued with intention or emotion. It has also expanded the potential for establishing what might be considered relationships with these agents (Pentina et al., 2023). However, little empirical research has studied human-AI relationships directly going beyond relational proxies such as trust. Independent of the ontological inquiry of whether humans *can* have relationships (Evans et al., 2023), the purpose of my dissertation is to advance the contemporary understanding of how humans *perceive* their relationship to conversational AI and how this perception is associated with user behavior.

In my dissertation, I break with the prevailing assumption that social relationships are a domain exclusive to humans (Guzman & Lewis, 2020). For the first time, A.P. Fiske's relational models theory (Fiske, 1992) was repurposed to assess how humans perceive their relationship with technology. Fiske's relational models theory proposes four relationship dimensions that guide human interaction (authority ranking, market pricing, communal sharing, and equality matching). I tested the relational approach in four cross-sectional, online studies (N_{total}=1568). Results consistently showed that the relationship modes people use to construe their relationships with conversational AI are similar but somewhat different to those used in relationships with other humans. Instead of the original four dimensions, users perceive their conversational AI along three different relationship modes: authority ranking (i.e., hierarchical, owner-master), market pricing (i.e., equal, rational, on "eye-level"), and peer bonding (i.e., new rather emotional dimension, peer-like). The first two studies highlighted the multidimensional approach's value, with each relationship dimension showing distinct associations with system perception variables (e.g., trust) and user characteristics (e.g., purposes of use).

After establishing a robust methodology for investigating how people perceive their relationship with conversational AI, I conducted two studies to examine the impact on behavior and explore its practical value in the context of voice commerce (i.e., shopping with conversational AI). Results showed that human-AI relationship modes associated differently with outcome variables relevant to voice shopping. In particular, voice shopping was strongly associated with peer bonding, emphasizing the importance of socio-emotional elements in voice shopping. Thus, the subsequent experimental study focused

on the role of socio-emotional design elements by taking a human-AI fit approach (i.e., aligning the interface with user characteristics to optimize user experience, Liu et al., 2011). I investigated the impact of different conversational designs in relation to the users' relational perception of voice shopping. Results indicated that a more emotional conversational design discouraged individuals in authority-ranking relationships from voice shopping. Thus, carefully considering how design might interact with relationship perception is essential to understanding factors that might hinder or facilitate voice shopping.

Despite various limitations, this dissertation provides a strongly theory-driven, readily reproducible assessment of human-AI relationship perception. Having illustrated the practical relevance of the multidimensional approach in voice commerce, the goal is to inspire forthcoming researchers and professionals to employ the human-AI relationship framework established here across diverse contexts. Whether involving conversational AI, conventional systems, or prospective yet-to-be-realized technologies – any situation where the perception of the relationship is deemed pertinent - I am optimistic that the human-AI relationship framework will prove beneficial in elucidating their potentially positive or negative impact on human-computer interactions.

Deutsche Zusammenfassung

Zuerst haben wir getippt, dann haben wir gewischt, und heute sprechen wir mit Computern. Künstliche Intelligenz (KI) ermöglicht es Nutzer:innen, mit Computern in natürlicher Sprache zu kommunizieren. Konversationelle KI, wie Amazons Alexa, sind lediglich Werkzeuge, die die Nutzer:innen bei vielen Aufgaben unterstützen (Kulkarni et al., 2019). Menschen anthropomorphisieren jedoch häufig Technologien und reagieren auf soziale Art und Weise (Gambino et al., 2020; Nass & Moon, 2000b). Darüber hinaus haben die jüngsten Fortschritte im Bereich der künstlichen Intelligenz dazu geführt, dass die Interaktionen zwischen Mensch und KI nicht nur sozial sind, sondern in gewissem Sinne mit Intentionen oder Emotionen angereichert werden. Dadurch wurde auch das Potenzial erweitert, Beziehungen zu diesen Agenten aufzubauen (Pentina et al., 2023). Es gibt jedoch nur wenige empirische Untersuchungen, die die Beziehungen zwischen Menschen und KI direkt untersucht haben, die über relationale Substitute wie Vertrauen hinausgehen. Unabhängig von der ontologischen Frage, ob Menschen Beziehungen haben können (Evans et al., 2023), besteht das Ziel meiner Dissertation darin, das gegenwärtige Verständnis darüber zu verbessern, wie Menschen ihre Beziehung zu konversationeller KI wahrnehmen und wie diese Wahrnehmung mit dem Nutzerverhalten zusammenhängt.

In meiner Dissertation löse ich mich von der vorherrschenden Annahme, dass soziale Beziehungen eine Domäne sind, die ausschließlich dem Menschen vorbehalten ist (Guzman & Lewis, 2020). Zum ersten Mal wurde die Theorie der Beziehungsmodelle von A.P. Fiske (Fiske, 1992) herangezogen, um zu bewerten, wie Menschen ihre Beziehung zur Technologie wahrnehmen. Fiskes Theorie der Beziehungsmodelle geht von vier Beziehungsdimensionen aus, die die menschliche Interaktion prägen (authority ranking, market pricing, communal sharing, und equality matching). Ich habe den relationalen Ansatz in vier Online-Querschnitts-studien getestet (Gesamt N=1568). Die Ergebnisse zeigten durchweg, dass die Beziehungsmodi, die Menschen verwenden, um ihre Beziehungen zu konversationeller KI zu konstruieren, denen ähneln, die in Beziehungen zu anderen Menschen verwendet werden, sich aber dennoch etwas davon unterscheiden. Anstelle der ursprünglichen vier Dimensionen nehmen die Nutzer:innen ihre konversationelle KI entlang dreier verschiedener Beziehungs-modi wahr: Authority ranking (d. h. eine hierarchische Diener-Meister Beziehung), market pricing (d. h. eine gleichberechtigte, rationale Beziehung auf “Augenhöhe”) und peer bonding (diese Dimension stellt eine neue, eher emotionale Dimension dar, kollegial). Die ersten beiden Studien unterstrichen den Wert des multidimensionalen Ansatzes, wobei jede

Beziehungsdimension distinkte Assoziationen mit Systemwahrnehmungsvariablen (z. B. Vertrauen) und Nutzereigenschaften (z. B. Nutzungszwecke) zeigte.

Nachdem ich eine robuste Methodik zur Untersuchung entwickelt hatte, wie Menschen ihre Beziehung zu konversationeller KI wahrnehmen, führte ich zwei Studien durch, um die Auswirkungen auf das Verhalten zu untersuchen und ihren praktischen Wert im Kontext des Voice Commerce (d. h. Einkaufen mit konversationeller KI). Die Ergebnisse zeigten, dass die Beziehungsmodi zwischen Menschen und KI in unterschiedlichem Maße mit Ergebnisvariablen verbunden sind, die für Voice Shopping relevant sind. Insbesondere war Voice Shopping stark mit peer bonding verbunden, was die Bedeutung sozio-emotionaler Elemente beim Voice Shopping unterstreicht. Daher konzentrierte sich die anschließende experimentelle Studie auf die Rolle sozio-emotionaler Designelemente, indem sie einen Human-AI-Fit-Ansatz verfolgte (d. h. die Anpassung der Benutzeroberfläche an Benutzereigenschaften zur Optimierung der Benutzererfahrung, Liu et al., 2011). Ich untersuchte die Auswirkungen verschiedener Konversationsdesigns in Bezug auf die Beziehungswahrnehmung der Nutzer:innen beim Voice Shopping. Die Ergebnisse zeigten, dass ein emotionaleres Gesprächsdesign Personen in authority ranking Beziehungen vom Voice Shopping abhielt. Daher ist eine sorgfältige Betrachtung der Wechselwirkung zwischen Design und Beziehungswahrnehmung von entscheidender Bedeutung für das Verständnis von Faktoren, die Voice Shopping behindern oder erleichtern können.

Trotz verschiedener Einschränkungen bietet diese Dissertation eine stark theoriegeleitete, leicht reproduzierbare Messung der Wahrnehmung der Beziehung zwischen Menschen und KI. Nachdem die praktische Relevanz des mehrdimensionalen Ansatzes im Voice Commerce veranschaulicht wurde, ist es mein Ziel, zukünftige Forscher:innen und Fachleute zu inspirieren, diesen Ansatz für die Mensch-KI-Beziehung in verschiedenen Kontexten anzuwenden. Ganz gleich, ob es sich um konversationelle KI, konventionelle Systeme oder zukünftige, noch zu realisierende Technologien handelt, in jeder Situation, in der die Wahrnehmung der Beziehung als relevant erachtet wird, bin ich optimistisch, dass sich der hier entwickelte Rahmenwerk der Mensch-KI-Beziehung als nützlich erweisen wird, um ihre potenziell positiven oder negativen Auswirkungen auf Mensch-Computer-Interaktionen zu ergründen.

Eidesstattliche Erklärung

Ich erkläre hiermit, dass ich die zur Promotion eingereichte Arbeit mit dem Titel „Towards a Better Understanding of Human-AI Relationship Perception“ selbständig verfasst, nur die angegebenen Quellen und Hilfsmittel benutzt und wörtlich oder inhaltlich übernommene Stellen als solche gekennzeichnet habe. Ich erkläre, dass die Richtlinien zur Sicherung guter wissenschaftlicher Praxis der Universität Tübingen (Beschluss des Senats vom 25.5.2000) beachtet wurden. Ich versichere an Eides statt, dass diese Angaben wahr sind und dass ich nichts verschwiegen habe. Mir ist bekannt, dass die falsche Abgabe einer Versicherung an Eides statt mit Freiheitsstrafe bis zu drei Jahren oder mit Geldstrafe bestraft wird.

Tübingen, ________________ ____________________________

Unterschrift