



Opposing Effects of Response Time in Human–Chatbot Interaction

The Moderating Role of Prior Experience

Ulrich Gnewuch · Stefan Morana · Marc T. P. Adam · Alexander Maedche

Received: 29 October 2020 / Accepted: 6 February 2022 / Published online: 30 May 2022
© The Author(s) 2022

Abstract Research has shown that employing social cues (e.g., name, human-like avatar) in chatbot design enhances users' social presence perceptions and their chatbot usage intentions. However, the picture is less clear for the social cue of chatbot response time. While some researchers argue that instant responses make chatbots appear unhuman-like, others suggest that delayed responses are perceived less positively. Drawing on social response theory and expectancy violations theory, this study investigates whether users' prior experience with chatbots clarifies the inconsistencies in the literature. In a lab experiment ($N = 202$), participants interacted with a chatbot that responded either instantly or with a delay. The results reveal that a delayed response time has opposing effects on social presence and usage intentions and shed light on the differences between novice users and experienced users – that is, those who have not interacted with a chatbot before vs. those who have. This study contributes to information systems literature by identifying prior experience as a key moderating factor that shapes users' social responses to

chatbots and by reconciling inconsistencies in the literature regarding the role of chatbot response time. For practitioners, this study points out a drawback of the widely adopted “one-design-fits-all” approach to chatbot design.

Keywords Chatbot · Response time · Response delay · Prior experience · Social response theory · Expectancy violations theory · Lab experiment

1 Introduction

Text-based conversational agents – commonly referred to as chatbots – are software systems designed to interact with humans using natural language (Dale 2016). Many organizations use chatbots to respond to customer service requests, provide personalized product information, or support customers in their purchase decisions (Sheehan et al. 2020). However, despite extensive interest in chatbots as a promising technology for customer interaction, their adoption and use by customers is growing much more slowly than expected (Grudin and Jacques 2019; Nordheim et al. 2019). A key reason cited for such slow adoption and limited use is that interacting with chatbots often does not feel natural and human-like (Schuetzler et al. 2014, 2020; Go and Sundar 2019).

A common way of rendering human–chatbot interactions more natural and human-like is to employ social cues in chatbot design. In the context of conversational agents, social cues are design features that trigger emotional, cognitive, or behavioral user reactions similar to those observed in the interaction between humans (Feine et al. 2019). Research has shown that users' natural tendency to respond to these cues promotes social presence perceptions giving users a sense of personal, sociable, and sensitive

Accepted after two revisions by Stefan Stieglitz.

U. Gnewuch (✉) · A. Maedche
Institute of Information Systems and Marketing, Karlsruhe
Institute of Technology, Kaiserstraße 89-93, 76133 Karlsruhe,
Germany
e-mail: ulrich.gnewuch@kit.edu

S. Morana
Junior Professorship of Digital Transformation and Information
Systems, Saarland University, Campus C3 1,
66123 Saarbrücken, Germany

M. T. P. Adam
School of Information and Physical Sciences, University of
Newcastle, University Drive, Callaghan, NSW 2308, Australia

human contact during the interaction, which in turn increases adoption and use (Hassanein and Head 2007; Hess et al. 2009; Qiu and Benbasat 2009). Research based on social response theory (SRT) (Nass et al. 1994; Nass and Moon 2000) has linked a broad range of social cues from chatbots, such as human-like avatars, small talk, or name tags, to positive user perceptions and behavior (e.g., Araujo 2018; Diederich et al. 2020; Benlian et al. 2020; Seeger et al. 2021).

However, the picture is less clear for response time – a vital social cue in technology-mediated interaction between humans (Walther and Tidwell 1995; Jacquet et al. 2019) and a key factor in website and mobile application usability (Galletta et al. 2006). According to the prevalent view in information systems (IS) literature, users do not readily tolerate slow response times of websites and mobile apps (Galletta et al. 2004; Yu et al. 2020). Nevertheless, some studies have found website delays to be beneficial, for example, in signaling the effort invested in generating product recommendations (Buell and Norton 2011; Tsekouras et al. 2022). In the context of chatbots, the role of response time is even less clear. Unlike human counterparts who need time to read a message and enter a response, chatbots can instantly process user input and generate a response (Schuetzler 2015; Følstad et al. 2018). Yet, some studies suggest that instant responses make a chatbot appear unhuman-like (Holtgraves and Han 2007), reduce the feeling of a natural conversation (Appel et al. 2012), and decrease user satisfaction (Gnewuch et al. 2018). In contrast, other studies show that chatbots with delayed responses elicit negative personality attributions (Holtgraves et al. 2007) and are perceived as less likeable (Schanke et al. 2021). Among practitioners there is also no consensus on this matter. While some designers intentionally delay chatbot responses to make them appear more human-like (e.g., Lufthansa's Mildred; see Crozier 2017), others stress the importance of instant responses (e.g., SysAid Technologies 2019). Hence, this study sets out to disentangle the opposing effects of chatbot response time in the extant literature. We particularly focus on users' prior chatbot experience as a factor that might explain how chatbot response time affects user perceptions.

Recent studies suggest that users transfer expectations from their prior experience with other chatbots to their current interaction with a chatbot (Moussawi et al. 2020; Grimes et al. 2021). Therefore, users with different prior chatbot experience are likely to have different expectations of chatbots in general and their response time in particular. For example, novice users who can only draw on their experience of chatting with humans might expect a longer response time than experienced users who have used other chatbots before. Consequently, a delayed response time may or may not violate users' expectations depending on

their prior chatbot experience. Since, according to expectancy violations theory (EVT) (Burgoon 1978, 1993), negative violations of expectations lead to negative outcomes, differences in users' prior chatbot experience could explain the inconsistent findings on the impact of chatbot response time in the literature. Against this backdrop, our study draws on SRT and EVT to investigate the questions of (1) how chatbot response time influences users' social presence perceptions and their chatbot usage intentions, and (2) how prior experience with chatbots moderates these relationships.

To address these questions, we conducted a lab experiment ($N = 202$) in which novice users (i.e., users who have not interacted with a chatbot before) and experienced users (i.e., users who have used chatbots before) interacted with a chatbot that responded either instantly or with a delay. In line with our reasoning, our results show that a delayed chatbot response time has opposing effects on social presence for novice and experienced users. While a delayed (as opposed to instant) response time positively influences novice users' social presence perceptions, the effect is negative for experienced users. Further, we find that social presence mediates the effect of chatbot response time on usage intentions, and that this mediation is moderated by prior chatbot experience such that the indirect effect of a delayed response time on usage intentions is positive for novice users and negative for experienced users. Finally, our results show that prior chatbot experience moderates the effect of social presence on usage intentions such that the effect is stronger for novice users than for experienced users.

Our study contributes to IS literature in three ways. First, we extend SRT by identifying prior experience as a key moderating factor that shapes users' social responses to chatbots. More specifically, by revealing opposing effects of a delayed response time for novice and experienced chatbot users, we shed a more differentiated light on SRT's core assumption that social cues more closely resembling human behavior (e.g., a delayed response time) trigger social responses in all users, regardless of their individual characteristics. Second, we offer an explanation for inconsistent findings regarding the role of response time in the context of chatbots, websites, and mobile apps by introducing users' prior experience as an important contingency factor. In particular, our study helps reconcile inconsistencies in the literature through clarifying the conditions under which a chatbot's instant or delayed response times result in positive outcomes, namely when its users are experienced users or novices. Finally, while previous research has mostly focused on verbal and visual cues (e.g., human-like avatars), we extend this literature stream by considering the impact of response time – a cue that falls in the category of chronemic cues (Feine et al. 2019).

Our study offers several practical implications for individuals and organizations who design, develop, or implement chatbots and other types of conversational agents (e.g., voice assistants). First, our findings suggest that a “one-design-fits-all” approach could be one reason for the ongoing struggle to meet user expectations. Since expectations differ among user groups and appear to evolve as users gain more experience with chatbots, practitioners might explore chatbots that can be adapted based on user characteristics and preferences. Second, our study highlights the need for sensitivity to seemingly minor chatbot design features and their impact on user perceptions. Since these features are easily overlooked, practitioners could, for example, extend their development teams to include language experts who can contribute knowledge of human conversation.

2 Theoretical Foundations and Related Work

2.1 Chatbots

Chatbots are conversational agents that rely on natural language in the form of text messages (Følstad and Brandtzæg 2017; Xu et al. 2017). Although the first chatbot ELIZA (Weizenbaum 1966) was already developed in the 1960s, it was not until the 2010s that chatbots gained broader organizational interest (Grudin and Jacques 2019; Seeger et al. 2021). Beginning in 2016, the growing excitement regarding artificial intelligence (AI) led to the development of more than 300,000 chatbots on Facebook Messenger alone (Araujo 2018; Facebook 2018). Common application areas are customer service, e-commerce, and health care (Følstad and Brandtzæg 2017; Adam et al. 2020), as well as workplace employee support (vom Brocke et al. 2018; Mirbabaie et al. 2021). However, despite their growing availability, customers are slow in adopting chatbots (Følstad et al. 2018; Nordheim et al. 2019). As a number of high-profile failures have demonstrated, many chatbots were unable to live up to their promises and have disappeared (Ashktorab et al. 2019; Grudin and Jacques 2019). Consequently, researchers and practitioners alike have realized that chatbot design and development is not only a technical challenge, but also needs to consider elements known to influence human–human interaction (Jenkins et al. 2007; Følstad and Brandtzæg 2017; Pfeuffer et al. 2019).

2.2 Social Response Theory (SRT)

In human–human interaction, a person perceives, interprets, and responds to a wide array of social cues (e.g., facial expressions, gestures) (Burgoon et al. 2010). Based

on the underlying Computers are Social Actors paradigm, SRT posits that users respond in a similar way to social cues from technology (e.g., natural language, human-like appearance) (Reeves and Nass 1996; Nass and Moon 2000). Even minimal social cues (e.g., a name tag) can trigger social responses in users (Nass et al. 1994; Nass and Moon 2000). These responses occur automatically and unconsciously without being “confined to a certain category of people” (Nass and Moon 2000, p. 98). When users respond socially to a technology, they unconsciously categorize it as a relevant social actor and ascribe human attributes to it, which enhances perceptions of social presence (Nass and Lee 2001).

Short et al. (1976) coined the term social presence to describe the extent to which a communication medium allows users to experience others as being psychologically present. It was originally defined as “the degree of salience of the other person in a mediated communication and the consequent salience of their interpersonal interactions” (Short et al. 1976, p. 65). However, its use has been extended since then to articulate how feelings of warmth, human contact, and sociability are created without actual human contact (Gefen and Straub 2004).

Building on SRT, various social cues from a host of different technologies have been linked to social responses in users (e.g., Hess et al. 2009; Qiu and Benbasat 2009). As for chatbots, researchers have primarily investigated verbal and visual cues (e.g., Araujo 2018; Moussawi and Benbunan-Fich 2020; Diederich et al. 2020; Seeger et al. 2021). For example, Moussawi and Benbunan-Fich (2020) showed that humorous comments can make chatbots appear more human-like. While these studies provide valuable knowledge on the impact of verbal and visual cues (e.g., human names, human-like avatars), less is known about other types of social cues such as response time (Feine et al. 2019).

2.3 Chatbot Response Time

Response time – also referred to as response latency or response delay – falls in the category of chronemic cues, which capture temporal aspects of communication (Walther and Tidwell 1995; Littlejohn and Foss 2009). It is an important social cue in human–human communication (Kalman et al. 2013; Schuetzler et al. 2019). In face-to-face interaction, response time is the time it takes a person to start speaking after another person has stopped. In technology-mediated interaction (e.g., instant messaging), it refers to the time it takes a person to respond to the other person’s message as well as the lag time between consecutive messages (Moon 1999). It includes the time needed to read, internalize, and make sense of another person’s

message, as well as the time needed to craft and edit a response (Derrick et al. 2013).

In contrast to humans, chatbots can respond almost instantly, as they need only fractions of a second to process user input and generate a response (Følstad et al. 2018; Schuetzler et al. 2021). However, some scholars suggest that instant responses make chatbots appear unhuman-like, reducing the feeling of a natural conversation (Holtgraves and Han 2007; Appel et al. 2012; Schanke et al. 2021). Schuetzler (2015) argues that “it introduces a non-negligible feeling of artificiality to interact with something that can respond instantly to anything you say” (p. 50). Consequently, some researchers and practitioners delay chatbot responses. For example, Holtgraves and Han (2007) employed dynamic delays based on the current message’s number of characters (i.e., 50 ms per character), while Appel et al. (2012) used static delays of 15–30 s. Further, Lufthansa delayed the responses of their chatbot Mildred after customers complained about its instant response time (Crozier 2017).

Unfortunately, existing research on the role of chatbot response time is scant and presents inconsistent findings that warrant further investigation. Moon (1999) showed that medium response times lead to higher persuasiveness (compared to instant and long response times). Gnewuch et al. (2018) found dynamically delayed responses to increase perceived humanness, social presence, and satisfaction. In contrast, Holtgraves et al. (2007) found that instant response times lead to more favorable personality perceptions. Schanke et al. (2021) showed that a chatbot with dynamically delayed (70 words per minute) rather than instant responses yielded lower likability. Against this backdrop, we investigate whether differences in users’ expectations based on prior chatbot experience can reconcile the inconsistent findings in the extant literature.

2.4 Expectancy Violations Theory (EVT)

Expectancy violations theory (EVT) (Burgoon 1978, 1993) relates to the impact of nonverbal behavior violations in human–human interaction. Although originally developed to understand proxemic violations (e.g., distancing), EVT has subsequently been expanded to cover other forms of nonverbal and verbal violations (e.g., addressing others by their first names without permission, putting a hand on another person’s shoulder) (Burgoon 2009). More recently, research has shown that EVT also extends to human–computer interaction (Burgoon 2015) and can explain how violating expectations associated with conversational agents and robots influences how users interact with and evaluate them (e.g., Spence et al. 2014; Burgoon et al. 2016; Grimes et al. 2021).

In essence, EVT posits that people have expectations regarding the nonverbal behavior of human and non-human counterparts. When these expectations are violated, people shift their attention toward the source of the violation and attempt to assign meaning to the violation. For example, an embodied conversational agent (ECA) could have a human-like visual appearance (e.g., an interactive 3D avatar) but use a mechanical-sounding voice (Burgoon et al. 2016). Then, the ECA’s voice might violate the expectations users formed based on the ECA’s visual appearance. In such a situation, users attempt to make sense of and interpret the violation, asking why the ECA appears human-like in one dimension but not in another. Based on this sense-making process, expectancy violations can be viewed as negative or positive, which influences attitudes and communication processes. For example, users can perceive the ECA design as unnatural and therefore react negatively to the violation of their expectations. Finally, EVT posits that positive violations produce more favorable outcomes than positive confirmations of expectations, whereas negative violations produce more negative outcomes than negative confirmations. For example, Burgoon et al. (2016) found that positive expectancy violations (e.g., adding a text transcript to an ECA) had more favorable effects on task attractiveness than positive confirmations.

Several studies have investigated chronemic expectancy violations in technology-mediated interaction between humans (e.g., using email or instant messaging). These violations relate to a person’s expectations about when to anticipate receiving a response to a message they sent (Sheldon et al. 2006; Kalman and Rafaeli 2011). For example, Kalman and Rafaeli (2011) showed that email response time affects a person’s evaluation of the email sender. Additionally, Sheldon et al. (2006) demonstrated that delayed response time in online collaboration tasks influenced a person’s evaluation of the collaborator. Thus, similar to expectations about other forms of nonverbal behavior, people hold expectations about response times. As EVT suggests, these expectations are formed through experiences a person has had (Burgoon 2009). Therefore, users’ expectations about a chatbot’s response time may depend on their prior experience with other chatbots.

3 Hypotheses Development

Drawing on SRT and EVT, this study investigates how chatbot response time and users’ prior chatbot experience influence social presence perceptions and chatbot usage intentions. Two types of response time are examined, namely, instant and delayed. For prior chatbot experience, we distinguish between novice users (i.e., users who have

not interacted with a chatbot before) and experienced users (i.e., users who have used chatbots before). Our research model shown in Fig. 1 captures our research hypotheses on the effect of response time on social presence (H1), the moderating role of prior chatbot experience (H2 & H4), and the mediating role of social presence in the relationship between response time and intention to use (H3).

3.1 Effect of Chatbot Response Time on Social Presence (H1)

According to SRT, humans tend to respond to technology similarly to how they do to other humans (Reeves and Nass 1996; Nass and Moon 2000). The fundamental assumption is that when technology is imbued with social cues (e.g., human-like appearance, natural language), users perceive the technology as a social actor and feel a sense of personal, sociable, and sensitive human contact during the interaction (Reeves and Nass 1996; Lee and Nass 2005). Based on this rationale, we hypothesize that during the interaction with a chatbot, a delayed response time more closely resembles how a human counterpart would respond and therefore serves as a social cue that triggers social responses in users. Consequently, a delayed response time might increase the feeling of a natural conversation and enhance users' social presence perceptions. In contrast, an instant response time could introduce a “non-negligible feeling of artificiality” (Schuetzler 2015, p. 50) because human counterparts would need some time to read a message, make sense of it, and enter their response. Consistent with this reasoning, reports from practice suggest that users were “irritated that [a chatbot] replied to their questions unnaturally fast” (Crozier 2017), indicating that an instant response time could reduce social presence perceptions. Therefore, based on SRT, we propose that:

H1 A chatbot with a delayed response time yields higher social presence than a chatbot with an instant response time.

3.2 Moderating Effect of Prior Chatbot Experience on Social Presence (H2)

SRT posits that users tend to generalize expectations from human–human interaction to human–computer interaction, regardless of individual user characteristics (Nass et al. 1994; Nass and Moon 2000). These expectations are especially pronounced when users interact with chatbots since they “use cues of humanness to create a sense of social presence that is not present in traditional information systems such as websites, applications, and databases” (Grimes et al. 2021, p. 2). However, recent studies suggest that users also transfer expectations from their previous interactions with other chatbots to their current interaction with a chatbot (Gambino et al. 2020; Moussawi et al. 2020; Cambre et al. 2021). Therefore, users who have prior experience with chatbots are likely to have different expectations of a chatbot than users who have never interacted with one before.

Against this backdrop, we draw on EVT (Burgoon 1978, 1993) to hypothesize on the moderating role that users' prior chatbot experience plays in how response time affects social presence perceptions. EVT not only offers a useful lens to theorize about expectancy violations in human–chatbot interaction (Grimes et al. 2021); it also provides deeper insight on violations of expectations that relate to response time – the so-called chronemic expectancy violations (Kalman and Rafaeli 2011). In fact, several studies have shown that users over time develop chronemic expectations about when to receive an anticipated response to a message they sent, and that they evaluate their counterpart based on these expectations (Sheldon et al. 2006; Kalman and Rafaeli 2011). Therefore, we propose that the effect of response time on social presence is contingent on users' prior chatbot experience because users with no previous chatbot interaction have different expectations about a chatbot's response time than users who have interacted with chatbots before.

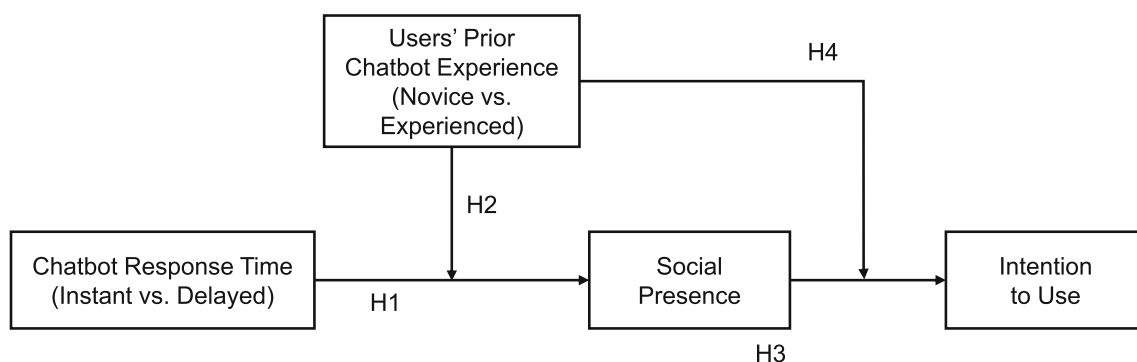


Fig. 1 Research model

SRT builds on the premise that users have pre-existing expectations regarding the interaction with another human in a given situation (e.g., chatting with friends via instant messaging applications such as WhatsApp, Telegram, or Facebook Messenger). These expectations have been developed over time as similar social situations humans encounter are stored in memory and activated when relevant situations arise (Gambino et al. 2020). Such situations include not only interactions with other humans but also interactions with technology (Nass and Moon 2000). Therefore, the same expectations – initially developed for human–human interaction – can extend to interactions with technology in general and chatbots in particular (Go and Sundar 2019). Here, importantly, recent studies suggest that this effect is especially pronounced in human–chatbot interaction because chatbots typically operate in the same environment – a simplistic chat window – that is frequently used to interact with other people (e.g., Araujo 2018; Beattie et al. 2020). In other words, given that human–chatbot interaction occurs in a chat window through natural language on a turn-by-turn basis (McTear 2017), users are more likely to draw on their experience from chatting with other people than from human–computer interactions with rich graphical user interfaces (e.g., websites, mobile apps) where the interaction occurs through clicking, scrolling, or swiping (Følstad and Brandtzæg 2017; Grimes et al. 2021).

Based on the above, we argue that in the absence of relevant experience, users who have not interacted with a chatbot before naturally draw on their experience of chatting with humans (who cannot respond instantly). Following this reasoning, a chatbot with an instant response time would violate novice users' chronemic expectations. According to EVT, this causes an attentional shift toward the chatbot as the source of the violation. In an attempt to interpret this violation, novice users shift from an automatic to a more deliberate way of thinking about the true nature of their counterpart. This shift, however, may reduce their tendency to respond socially because it interferes with the automatic and unconscious process that would trigger social responses (Nass et al. 1994). Conversely, a delayed response time might not violate novice users' chronemic expectations because it more closely resembles their previous interactions with humans. Therefore, an attentional shift is less likely and users' tendency to respond socially to the chatbot is not negatively affected. Consequently, we argue that the positive impact of a delayed (as opposed to instant) response time on social presence is particularly strong for novice users.

In contrast, users who have used chatbots before might have formed different response time expectations, since chatbots in practice usually respond instantly (Lester et al. 2004; Schuetzler et al. 2021). Although it is possible that not all chatbots have instant response times, reports from practice suggest that optimizing the response time to deliver fast

responses is an established chatbot design guideline (e.g., “Your chatbot needs to be fast; if it’s not it, won’t get used”; SysAid Technologies 2019). Therefore, experienced users' expectations would not be violated by an instant response time. Conversely, given that experienced users likely have experienced instant responses before, a delayed chatbot response time could violate their expectations. As Burgoon et al. (2016) noted, an “unexpected delay in response [...] may draw the human’s attention to the delay [who] will attempt to make sense of why the delay exists” (p. 7). In an attempt to make sense of the violation, experienced users might become more thoughtful and their tendency to (automatically and unconsciously) respond socially to the chatbot will be reduced. Consequently, experienced users' social presence perceptions might be reduced when they interact with a chatbot that has a delayed response time. Hence, based on EVT, we propose that:

H2 Users' prior chatbot experience moderates the effect of chatbot response time on social presence, such that the effect of a delayed (instant) response time is stronger for novice (experienced) users.

3.3 Mediating Effect of Social Presence on Intention to Use (H3)

As noted before, SRT posits that users treat technologies imbued with social cues as social actors rather than just tools to use. Therefore, a delayed response time, which more closely resembles human behavior than an instant response time, might serve as a social cue that enhances users' social presence perceptions when they interact with a chatbot. SRT further proposes that users develop social relationships with technologies that are treated as social actors (Nass et al. 1996; Moon 2000; Fogg 2002). These relationships create an emotional bond between users and technologies (Nass et al. 1996) and thereby increase users' intention to use them again in the future. Consequently, an increase in social presence might also lead to higher usage intentions toward a chatbot. In line with this reasoning, Xu et al. (2018) showed that users who have developed a social relationship with a recommendation agent exhibit higher usage intentions. Similarly, a number of studies have demonstrated that social presence is a key factor in driving usage intentions across various technology contexts (e.g., Cyr et al. 2007; Hassanein and Head 2007; Qiu and Benbasat 2009). Considering the relationships between chatbot response time, social presence, and intention to use together, we suggest that a chatbot with a delayed response time enhances users' social presence perceptions, which in turn increases their usage intentions. Therefore, based on SRT and in accordance with previous studies, we propose that:

H3 Social presence mediates the effect of chatbot response time on intention to use.

3.4 Moderating Effect of Prior Chatbot Experience on Intention to Use (H4)

Typically, users who have never interacted with a chatbot before “expect chatbots to have human-like communication skills” (Janssen et al. 2020, p. 222). Unsurprisingly, chatbots are unlikely to meet these high expectations, which results in user frustration and reduced usage intentions (Brandtzaeg and Følstad 2018; Schuetzler et al. 2021). In contrast, experienced users “have grown accustomed to chatbots’ limitations” and therefore do not expect a human-like level of social presence (Jain et al. 2018, p. 895). Such users typically focus on other criteria, such as functionality and features, when deciding to use the chatbot (Liao et al. 2016; Brandtzaeg and Følstad 2017). Against this backdrop, we again draw on EVT to hypothesize that the impact of social presence on chatbot usage intentions is stronger for novice users than for experienced users. According to EVT, negative violations of expectations exert a stronger influence than negative confirmations of expectations (Burgoon 2015; Burgoon et al. 2016). Therefore, novice users’ high expectations (i.e., of human-like social presence levels) might be more strongly violated than experienced users’ expectations. Consequently, experienced users’ usage intentions are less likely to be influenced by their social presence perceptions because they would not expect a chatbot to create a strong sense of social presence in the first place. Based on these considerations, we propose that:

H4 Users’ prior chatbot experience moderates the effect of social presence on intention to use the chatbot, such that the effect is stronger for novice users than for experienced users.

4 Method

4.1 Experimental Conditions

To test our hypotheses, we conducted a between-subjects lab experiment in which 202 participants interacted with a chatbot that responded either instantly (INST) or with a delay (DLY). Participants were randomly assigned to one of these conditions. In the instant response time condition (INST), the chatbot responded instantly without any delay (i.e., as fast as technically possible). Since sending a message involved a short network delay caused by physical limits of data transmission over the Internet, this corresponded to a response time of about 200 to 400 ms, which is comparable to chatbots with instant response times used in practice. A total of 67 participants (33.2%) interacted with the chatbot that had an instant response time. In the

delayed response time condition (DLY), the chatbot responded with a delay of 2.3 s on average. To ensure that the results do not rest on the specific type of delay, we considered both delay approaches described in the literature (i.e., static and dynamic delays). Therefore, this condition had twice as many participants as the INST condition (i.e., 135 participants, 66.8%). Our analysis confirmed that there was no difference in user perceptions between static and dynamic delays (see Appendix A, available online via <https://doi.org/10.1007/s12599-022-00755-x>). Hence, our analysis does not differentiate the two types of delay.

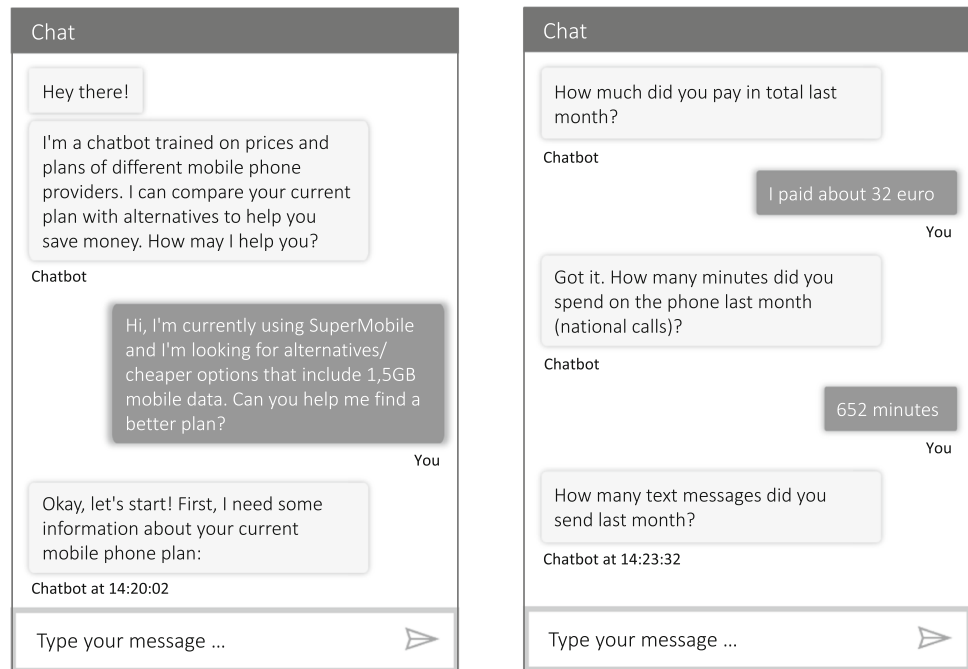
4.2 Experimental Procedure and Task

The experiment took place at the Karlsruhe Decision and Design Lab (KD2Lab), adhering to its procedural and ethical guidelines. First, upon arrival at the lab, participants read and signed an informed consent form. Subsequently, they sat down at a computer and received instructions on the experiment. These instructions introduced participants to the hypothetical scenario of using a chatbot to find out whether they could save money by changing their mobile phone plan. Additionally, we provided a fictitious copy of last month’s mobile phone bill, which indicated that the current mobile phone plan did not fit participants’ actual usage patterns (e.g., their data usage was much higher than the volume included in their plan, resulting in high additional costs) (see Figure A1 in Online Appendix A). Additionally, the instructions clarified that participants were about to interact with a chatbot and not with another human being. Subsequently, participants interacted with a chatbot in one of the experimental conditions. After identifying a more suitable mobile phone plan based on the chatbot’s recommendation and ending the conversation, participants completed a post-experiment questionnaire that captured the dependent variables social presence and intention to use, prior experience with chatbots, the manipulation check (i.e., perceived response time), and several controls. After completing the questionnaire, we debriefed participants and compensated them with €7 for their participation.

4.3 Experimental Chatbot

For the experiment, we developed different versions of a chatbot using the Microsoft Bot Framework (Microsoft 2021). The chatbots were able to advise participants on different mobile phone plans and help them save money by recommending plans that better fit their actual usage patterns. As illustrated in Fig. 2, the chatbots asked participants several questions about their current plan and usage patterns (e.g., data volume used last month), but they were

Fig. 2 Excerpts from an exemplary chatbot conversation during the experiment



also able to answer other questions related to mobile phone plans. Participants interacted with the chatbots by formulating and sending their own messages to increase the realism of the chatbot interaction. To process natural language user input (e.g., to recognize user intentions and extract entities in a user's message, such as the names of different mobile phone plans), the chatbots used Microsoft's Language Understanding Intelligent Services (LUIS). We trained all chatbots on the same language model and implemented identical dialogs. The only difference between the chatbots was their response time according to the experimental conditions.

4.4 Participants

We recruited participants from a European university student pool. Given the rather low chatbot adoption in the general population (SmartAction 2018; Inmar 2019), we considered students to be appropriate for this study because they are among the early chatbot adopters (Brandtzaeg and Følstad 2017; Tuzovic and Paluch 2018). An a priori power analysis using G*Power (Faul et al. 2007) with a significance level of 0.05 determined a minimum sample size of 207 participants to achieve a statistical power of 0.90 for detecting a medium effect size ($f = 0.25$). As we anticipated that some participants might encounter (technical) difficulties in their interaction with the chatbot, we aimed for a sample size of about 220 participants.

In total, 219 subjects participated in the study. We excluded 13 participants because they provided incorrect answers to the attention check questions in the

questionnaire (e.g., "If you are carefully filling out the survey, please select 'Strongly agree'"). Additionally, we screened the protocols of each conversation as prior chatbot studies did (e.g., Go and Sundar 2019) to filter out six participants who did not follow the instructions (e.g., ended the conversation before the chatbot had recommended a plan) and/or purposefully provided invalid inputs during the conversation with the chatbot (e.g., entered negative or unrealistically high values when asked about how much they would be willing to pay for a new mobile phone plan). Therefore, our final sample included 202 participants (77 female, 125 male). Participants were between 18 and 41 years old ($M = 23.21$, $SD = 3.45$). Online Appendix B shows the participants' demographic and personal characteristics.

4.5 Measures

We used previously validated measures for all constructs and slightly adapted them to the context of this study. We assessed social presence using the items from Gefen and Straub (1997) and intention to use using the items from Wang and Benbasat (2009) on seven-point Likert scales (1 = "strongly disagree"; 7 = "strongly agree"). Similar to previous studies (e.g., Ashktorab et al. 2019; Xu 2019; Moussawi et al. 2020), we measured prior chatbot experience and distinguished between novice and experienced users by asking participants whether and how often they use chatbots (five-point Likert scale; 1 = "never"; 5 = "daily"). This resulted in a roughly even split between novice users (i.e., users who have not interacted with a

chatbot before) and experienced users (i.e., users who have used chatbots before). Additionally, we collected demographic information from participants (i.e., age, gender) and assessed their level of sociability (Bruch et al. 1989) as potentially relevant control variables. Online Appendix B lists all measurement items.

Given that these items were measured with the same method, we tested for common method variance using Harman's single factor test (Podsakoff et al. 2003). This test showed that no single major factor emerges and that the first factor accounts for only 33.0% of the total variance. Moreover, our main independent variable – chatbot response time – represents an experimental manipulation and is therefore not measured via self-report. Thus, common method bias is not a serious concern in this study.

5 Results

5.1 Manipulation Check

To check our manipulation of the chatbots' response time, each participant rated the response time using a seven-point Likert scale (1 = "slow"; 7 = "fast") (Galletta et al. 2006). The results of a *t*-test showed that participants in the INST condition ($M = 6.67$, $SD = 0.70$) perceived the chatbot as significantly faster than participants in the DLY condition ($M = 5.81$, $SD = 1.15$; $t(200) = 5.61$, $p < 0.001$). Therefore, we conclude that our manipulation was successful in influencing participants' perception of the chatbots' response time as either instant or delayed.

To confirm the successful randomized assignment of participants to our experimental conditions, we conducted several *t*-tests and chi-square difference tests. There were no significant differences in prior chatbot experience ($\chi^2(1) = 0.087$, $p = 0.768$), age ($t(200) = -0.468$, $p = 0.640$), gender ($\chi^2(1) = 0.393$, $p = 0.530$), sociability ($t(200) = 0.778$, $p = 0.437$), education ($\chi^2(3) = 4.45$, $p = 0.216$), prior messenger experience ($t(200) = -1.80$, $p = 0.073$), and chat duration ($t(200) = 0.380$, $p = 0.704$) between the INST and DLY conditions.

5.2 Measurement Model Assessment

We assessed the measurement model by examining indicator reliability, internal consistency, convergent validity, and discriminant validity of all latent constructs (i.e., social presence, intention to use, and sociability). After dropping three items of the sociability control variable due to low factor loadings, all remaining items loaded significantly onto their intended factor with loadings ranging from 0.70 to 0.95, thus supporting indicator reliability (Gefen and Straub 2005) (see Online Appendix B). Next, to

demonstrate the constructs' internal consistency, we calculated the composite reliability (CR) and Cronbach's alpha (CA) for each construct. As shown in Table 1, all constructs exceeded the recommended threshold of 0.70 for both CR and CA (Nunnally and Bernstein 1994). Also, all average variance extracted (AVE) values exceeded the suggested threshold of 0.50 (Fornell and Larcker 1981). Finally, the square root of the AVE of each construct was higher than the correlations with other constructs, thus supporting discriminant validity (Fornell and Larcker 1981). Taken together, these results support the reliability and validity of the measurement model.

5.3 Descriptive Results

Table 2 provides descriptive statistics with means and standard deviation for all constructs in the instant and delayed response time conditions.

5.4 Hypotheses Testing

To test our hypotheses, we used Hayes' PROCESS macro for SPSS (Hayes 2018). Since our research model involved first and second stage moderated mediation (Edwards and Lambert 2007), we conducted a moderated mediation analysis with 5,000 bootstrap samples using Model 58 (Hayes 2018). In this analysis, chatbot response time served as the independent variable (0 = instant, 1 = delayed), prior chatbot experience (0 = novice, 1 = experienced) as the moderator, social presence as the mediator, and intention to use as the dependent variable (see Fig. 3). Finally, we included age, gender, and sociability as covariates in the analysis. Table 3 shows the analysis results in more detail.

5.4.1 Social Presence (H1 & H2)

The results show that while a delayed rather than an instant response time is associated with higher social presence (H1 supported: $b = 0.69$, $p < 0.05$), this relationship is moderated by prior chatbot experience (H2 supported: $b = -1.20$, $p < 0.01$). Specifically, the conditional effects reveal that delayed response time has a positive effect on social presence for novice users ($b = 0.69$, $p < 0.05$, 95% CI [0.14, 1.23]), whereas the effect is negative for experienced users ($b = -0.51$, $p < 0.05$, 95% CI [-1.02, -0.01]). Hence, it is important to note that although H1 is supported, this positive effect of a delayed (vs. instant) response time on social presence holds only for novice users, but not for experienced users. To clarify the nature of this interaction, we performed a spotlight analysis (Spiller et al. 2013). As depicted in Fig. 4, in the INST condition, social presence is significantly higher for experienced users than for novice users

Table 1 Internal consistency and discriminant validity of constructs

Construct		<i>M</i>	<i>SD</i>	<i>CR</i>	<i>CA</i>	<i>AVE</i>	1	2	3
1	Social presence	3.75	1.30	0.92	0.91	0.70	0.84	–	–
2	Intention to use	4.46	1.63	0.96	0.96	0.89	0.48***	0.94	–
3	Sociability ^a	5.58	0.90	0.72	0.72	0.56	0.09	0.06	0.75

M mean; *SD* standard deviation; *CR* composite reliability; *CA* Cronbach's alpha; *AVE* average variance extracted. Bolded diagonal elements are the square root of *AVE*; the values below represent the correlation between constructs

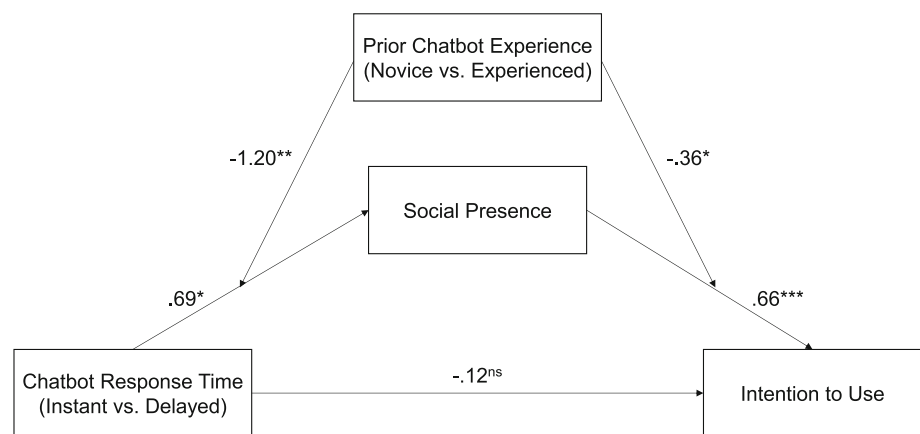
*** $p < .001$

^aControl variable

Table 2 Descriptive statistics

Experimental condition	Social presence			Intention to use		
	All users	Novice users	Experienced users	All users	Novice users	Experienced users
INST	3.71 (1.32)	3.12 (1.10)	4.25 (1.28)	4.44 (1.62)	4.20 (1.61)	4.66 (1.62)
DLY	3.76 (1.29)	3.80 (1.31)	3.74 (1.29)	4.47 (1.64)	4.34 (1.78)	4.58 (1.52)
Total sample	3.75 (1.30)	3.56 (1.27)	3.90 (1.30)	4.46 (1.63)	4.29 (1.71)	4.61 (1.54)

Means with standard deviations in parentheses. Measured on 7-point Likert scales

Fig. 3 Moderated mediation model with path coefficients

* $p < .05$; ** $p < .01$; *** $p < .001$; ns = not significant

($b = 1.13$, $t = 3.64$, $p < 0.001$), whereas in the DLY condition, the difference in social presence between experienced and novice users is not significant ($b = -0.06$, $t = -0.28$, $p = 0.78$). This shows that the opposing effects of a delayed response time on social presence are driven by the differences in how novice and experienced users perceive a chatbot that responds instantly.

5.4.2 Intention to Use (H3 & H4)

Regarding the right-hand side of the model, the analysis shows that social presence positively influences intention to use ($b = 0.66$, $p < 0.001$) subject to the user's prior

chatbot experience (H4 supported: $b = -0.36$, $p < 0.05$). Specifically, the conditional effects reveal a positive effect of social presence on intention to use for both novice users ($b = 0.66$, $p < 0.001$, 95% CI [0.41, 0.91]) and experienced users ($b = 0.30$, $p < 0.01$, 95% CI [0.08, 0.53]). Still, the effect of social presence on intention to use is twice as large for novice users than it is for experienced users. Providing support for the mediating role of social presence, the indirect effect of chatbot response time on intention to use is significant for both novice users ($b = 0.45$, $SE = 0.19$, 95% CI [0.12, 0.86]) and experienced users ($b = -0.15$, $SE = 0.10$, 95% CI [-0.37, -0.002]), while the direct effect is not

Table 3 Estimates of direct effects and interactions from the moderated mediation model

	Social presence				Intention to use					
	Coeff	(SE)	t	p	95% CI	Coeff	(SE)	t	p	95% CI
Chatbot Response Time ^a	0.686	(0.277)	2.482	.014*	[0.14, 1.23]	- 0.118	(0.231)	- 0.513	.609	[- 0.57, 0.34]
Prior Chatbot Experience ^b	2.332	(0.656)	3.553	<.001***	[1.04, 3.63]	1.466	(0.683)	2.145	.033*	[0.12, 2.81]
Chatbot Response Time ^a × Prior Chatbot Experience ^b	- 1.198	(0.378)	- 3.169	.002**	[- 1.94, - 0.45]	0.661	(0.127)	5.215	<.001***	[0.41, 0.91]
Social Presence						- 0.359	(0.172)	- 2.081	.039*	[- 0.70, - 0.02]
Social Presence × Prior Chatbot Experience ^b	2.278	(0.983)	2.318	.022*	[0.34, 4.22]	2.570	(1.141)	2.252	.025*	[0.32, 4.82]
Constant	- 0.029	(0.026)	- 1.109	.269	[- 0.08, 0.02]	- 0.008	(0.031)	- 0.249	.803	[- 0.07, 0.05]
Age	0.000	(0.184)	- 0.002	.998	[- 0.36, 0.36]	- 0.145	(0.220)	- 0.661	.509	[- 0.58, 0.29]
Gender ^c	0.147	(0.095)	1.540	.125	[- 0.04, 0.33]	- 0.003	(0.115)	- 0.029	.977	[- 0.23, 0.22]
Sociability	0.083					0.165				
R ²	2.939**			(df = 6; 195)		5.469***			(df = 7; 194)	
F										

N = 202. SE standard error; CI confidence interval

*p < .05; **p < .01; ***p < .001

^a1 = delayed. ^b1 = experienced user. ^c1 = male

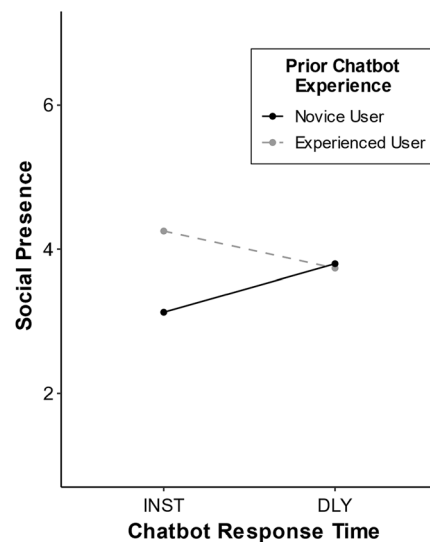


Fig. 4 Interaction effect between chatbot response time and prior chatbot experience on social presence (H2)

($b = -0.12$, $SE = 0.23$, 95% CI $[-0.57, 0.34]$) (see Table 4). This finding supports H3. In addition, the index of moderated mediation, which tests whether the indirect effect varies across levels of the moderator (Hayes 2015), is significant ($index = -0.61$, $SE = 0.21$, 95% CI $[-1.04, -0.22]$). This result indicates that the indirect effects of chatbot response time on intention to use via social presence significantly differ for novice vs. experienced users. More specifically, a delayed response time has a positive effect on novice users' intention to use (via social presence), whereas the mediated effect is negative for experienced users.

To understand the moderating role of chatbot experience (H4) in greater detail, we conducted a floodlight analysis using the Johnson–Neyman technique (Spiller et al. 2013; Finsaas and Goldstein 2021) to test at which social presence levels the impact on intention to use is different for novice compared to experienced users. We find that the difference between novice and experienced users is significant for social presence values between 1 and 2.74 (see Fig. 5). This result suggests that the usage intentions of novice users who perceive a low social presence are more strongly affected than the usage intentions of experienced users who perceive a similarly low social presence. In other

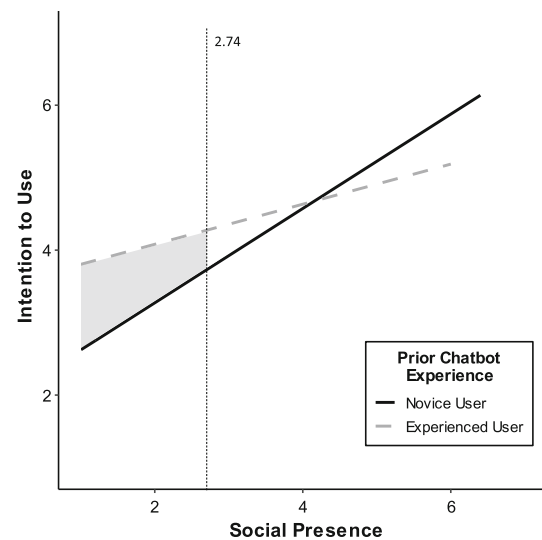


Fig. 5 Interaction effect between social presence and prior chatbot experience on intention to use (H4) *Note.* Shaded area represents levels of social presence, where the difference between novice and experienced users is significant at the .05 level

words, novice users are less likely to use a chatbot when social presence is low (≤ 2.74), while low social presence seems less critical for experienced users' intention to use a chatbot.

6 Discussion

In this study, we investigated how prior chatbot experience might explain the effect that chatbot response time has on users' social presence perceptions and chatbot usage intentions. Our results reveal opposing effects of a delayed response time and shed light on the differences between novice and experienced chatbot users. First, in line with our expectations, a delayed (as opposed to instant) response time enhances novice users' social presence perceptions. However, we find the opposite effect for experienced users: a delayed response time actually reduces their social presence perceptions. This finding not only highlights the important role of users' prior chatbot experience, but also suggests that under certain circumstances, incorporating artificial delays in a chatbot can backfire. Experienced users may find a delayed response time irritating or

Table 4 Direct effect and conditional indirect effects of chatbot response time on intention to use

	Direct effect		Indirect effects	
	Coefficient (SE)	95% CI	Coefficient (SE)	95% CI
Novice users	-0.12 (0.23)	$[-.57, .34]$	0.45 (0.19)	$ [.12, .86]$
Experienced users			-0.15 (0.10)	$ [-.37, -.002]$

SE standard error; *CI* confidence interval

annoying because they are aware that a chatbot can answer instantly, which could lead to social presence perceptions being lower than they would otherwise have been.

Second, our results show that social presence mediates the effect of chatbot response time on usage intentions, and that this mediation is moderated by prior chatbot experience. Corroborating our previous findings, the indirect effect of a delayed response time on chatbot usage intentions via social presence is positive for novice users but negative for experienced users. This finding suggests that a delayed response time not only has opposing effects on users' immediate social presence perceptions, but also leads to markedly different downstream consequences on usage intentions. Although a longitudinal study would be necessary to fully understand how the relationship between social cues, such as response time, and usage intentions evolves over time, there is reason to believe that the positive effect of social cues in chatbot design could disappear or even become negative as users gain experience through ongoing interactions with chatbots.

Third, our results reveal how the impact of social presence on usage intentions differs depending on users' prior chatbot experience. In general, the impact is stronger for novice than for experienced users, indicating that with increased experience, the importance of social presence as a determinant of chatbot usage intentions could decrease. Interestingly, we also find that when social presence perceptions are low, experienced users have higher usage intentions than novice users. This finding suggests that experienced users might be more tolerant of lower social presence perceptions than novice users when forming their intention to use a chatbot. Additionally, it could indicate that experienced users prefer less "humanized" chatbot designs because these users tend to focus on the utility of a chatbot rather than its ability to provide a human touch (Brandtzaeg and Følstad 2017). In the following section, we discuss the implications of our findings for theory and practice, highlight potential limitations, and outline directions for future research.

6.1 Theoretical Implications

Our findings contribute to theory in three ways. First, we extend SRT by identifying prior experience as a key moderating factor that shapes users' social responses to chatbots. This finding challenges a core assumption of SRT, namely that social responses occur automatically and unconsciously without being "confined to a certain category of people" (Nass and Moon 2000, p. 98). According to this view, chatbot response times that more closely resemble human behavior should trigger social responses in all users, regardless of their individual characteristics. However, our finding that for chatbot response time this

effect only holds for novice users and not for experienced users suggests that prior experience presents an important theoretical boundary condition that affects whether and how users respond socially to a chatbot, pointing to a more complex mechanism than previously thought. More broadly, it highlights the need for extending SRT to consider that individual differences between users in terms of prior experience with a particular technology can affect how they respond to specific design features (here, response time). Through integrating ideas from EVT (Burgoon 1978, 1993), our research can serve as a starting point to guide the extension of SRT by considering the expectations users bring to the interaction and how violating them can ultimately affect users' social responses to technology.

Second, we offer an explanation for earlier inconsistent findings regarding the role of response time in the context of chatbots (Holtgraves et al. 2007; Appel et al. 2012; Gnewuch et al. 2018; Schanke et al. 2021), as well as websites and mobile apps (Galletta et al. 2004; Buell and Norton 2011; Yu et al. 2020; Tsekouras et al. 2022) by introducing users' prior experience as an important contingency factor. Specifically, our findings help reconcile inconsistencies in the literature by clarifying the conditions under which instant or delayed chatbot responses result in positive outcomes (i.e., increased social presence, higher usage intentions), namely when its users are experienced users or novices. Tsekouras et al. (2022) made a similar observation in finding that recommendation agent users who were more familiar with a given product context (e.g., searching for cars on the Internet) reacted less positively to a delay in product recommendations being generated compared to users with low familiarity. Collectively, these findings suggest that investigations into the effects of response time should take users' prior experience with the technology itself or the context in which it is used into consideration.

Finally, existing research in IS and related disciplines has mostly focused on verbal and visual cues, such as human names or human-like avatars, and how they trigger social responses in users (e.g., Hess et al. 2009; Qiu and Benbasat 2009; Araujo 2018). Our study extends this research stream by examining response time – an important cue that falls into the category chronemic (i.e., time-related) cues which has received very little attention in the literature (Feine et al. 2019). Although chronemic cues are often overlooked due to their invisibility (Littlejohn and Foss 2009), our findings on the opposing effects of a delayed response time show that they can strongly impact users' perceptions of a chatbot. Further, related studies on voice assistants (Porcheron et al. 2018) and physical robots (Shiwa et al. 2009) suggest that the importance of response time as a social cue is not limited to the chatbot context.

Consequently, response time should be recognized as an important design feature that researchers and practitioners can deliberately design and control when implementing chatbots and other types of conversational agents.

6.2 Practical Implications

Our findings also have important practical implications. Currently, most organizations and individuals who design, develop, or implement chatbots and other types of conversational agents (e.g., voice assistants like Alexa) follow a “one-design-fits-all” approach by focusing on one design that meets the expectations of most potential users at that point in time. However, our findings suggest that a single design sooner or later will run into difficulties in meeting expectations because users have different expectations that can evolve as they gain experience. Therefore, a “one-design-fits-all” approach could be one reason why companies struggle to increase the adoption and use of their chatbot. For example, novice users might prefer a more “humanized” design (e.g., delayed responses), while experienced users could be irritated by a chatbot that aims to closely mimic human conversation; instead, they might prefer a more “machine-like” design (e.g., instant responses). Therefore, a key practical implication of our study is that chatbots need to be personalized based on the characteristics and preferences of their users. Although a discussion of personalization strategies for chatbots is beyond the scope of this paper, we recommend that practitioners explore how users could manually personalize a chatbot according to their preferences (e.g., allowing users to enable or disable certain design features) and how to develop chatbots that are capable of automatically adapting themselves to a user (e.g., based on available user data or previous interactions).

Second, our study highlights the need for sensitivity to seemingly minor design features and their impact on users’ perceptions of chatbots. Designers and organizations often put effort into giving their chatbot a “personality” with a human name (i.e., a verbal cue) and a human-like avatar (i.e., a visual cue) (Araujo 2018), yet chronemic cues that relate to temporal aspects of the interaction are easily overlooked. Our findings suggest that attention to such detail is important, particularly because chronemic cues, such as response time, are inherent to natural language interaction and therefore central to any technology equipped with a conversational user interface. Consequently, practitioners would be well advised to add language experts or psychologists to their development teams, thereby drawing on their knowledge of human conversation in designing chatbot conversations.

6.3 Limitations and Future Research

There are limitations to this study, which could open up avenues for future research. The first limitation relates to the operationalization and measurement of prior experience as a dichotomous variable (i.e., novice vs. experienced), which does not take different levels of experience with chatbots into account. However, we believe that the distinction between novice users and experienced users is appropriate for two main reasons. First, understanding the difference between novice and experienced users, rather than the difference between users with different levels of experience, is particularly important in the context of adoption and use (Thompson et al. 1994; Taylor and Todd 1995; Galletta and Dunn 2014). Moreover, recent chatbot studies specifically highlight the difficulties novice users face when making the “transition” to becoming experienced users (e.g., Jain et al. 2018; Muresan and Pohl 2019). Second, due to the low adoption of chatbots in the general population, a large part of our sample was novice users (46%) (see Online Appendix B). In situations with such an uneven distribution across categories, a reasonable strategy is to collapse a categorical variable by combining several of its categories (Babbie et al. 2018). Nevertheless, we acknowledge that future studies need to expand on our findings by conducting a finer examination along the continuum of chatbot experience in order to assess how different levels of experience influence users’ social responses to chatbots. Another promising avenue for future research could be to examine the impact of other individual user characteristics (e.g., age, personality, digital savviness) that may also lead to differences in their social responses.

A second limitation is the use of a student sample in our experiment, which might not be representative of the entire population (e.g., because students have higher levels of technical knowledge). Therefore, our findings might not be generalizable to the wider population. Nevertheless, we believe that using a student sample here was adequate for several reasons. First, students are among the early adopters of chatbots (Brandtzaeg and Følstad 2017), whereas a major part of the general population has not interacted with a chatbot before (Brandtzaeg and Følstad 2017; Inmar 2019). Therefore, a student population allowed us to recruit novice and experienced chatbot users to participate in the experiment. Second, organizations often implement chatbots to explicitly target the younger, tech-savvy generation (e.g., students) (Xiao and Kumar 2021). Therefore, insight into this group’s perceptions and intentions might be particularly relevant to IS research and practice. Third, a homogeneous group of participants generally helps to maximize internal validity in order to clearly identify the treatment effects without additionally controlling for many other factors (Lynch Jr. 1982; Price et al. 2015). However,

future research could further cross-validate our findings with user groups from different population segments.

A third limitation lies in our study design that did not allow us to capture participants' initial expectations about chatbots before the experiment. In experimental research, it is crucial that participants do not know the experiment's true purpose (e.g., treatment, group assignment) because this knowledge would influence their behavior and present a threat to internal validity (Orne 1962). Therefore, having asked participants about how quickly – based on their prior experience – they expected to receive an answer from a chatbot would have drawn their attention to this design feature during the experiment, which would have seriously biased participants' expectations and responses. Consequently, we were unable to capture participants' actual expectations in detail. For example, it is possible that experienced users had previous interactions with chatbots that did not respond instantly to their messages, even if reports from practice suggest that achieving a fast response time is an established chatbot design guideline (e.g., “Your chatbot needs to be fast; if it's not, it won't get used”; SysAid Technologies 2019). Although the approach of “assume[ing] that the expectations exist in the participant population and explore[ing] the consequences of violating the expectancies” is common in EVT studies (Kalman and Rafaeli 2011, p. 57), future research is needed to explicitly measure and analyze users' expectations about chatbots in more detail. This would be particularly interesting for experienced users whose previous interactions with chatbots might have been different, and therefore led to different expectations. Another promising research direction could be to analyze users' expectations about other types of social cues (e.g., verbal or visual cues) and the impact of violating expectations surrounding these cues.

Fourth, research suggests that there may be an inverted U-shaped relationship between response time and social presence, such that both very short and very long response times lead to negative user responses (Moon 1999; Tsekouras et al. 2022). However, since the delayed response time of 2.3 s on average in our experiment was still rather short, we could not examine how users evaluate an overly long response time of a chatbot. Therefore, future research should expand on our findings by investigating longer response delays and the potential inverted U-shaped effect of response time.

Fifth, a valuable extension of our research would be to explore users' assumptions about the delayed response time and how they make sense of the chatbot's delayed responses. Some users might attribute the response delay to the complexities of AI and natural language processing, which might need a certain amount of time to process a message sent to the chatbot. Other users might think that the delay is caused by a slow network or hardware

inefficiencies (e.g., poor internet connection, slow processing speed). Still others might believe that the delay was deliberately introduced to fool them into believing that they are interacting with a real person. Since our study did not investigate how users perceived and made sense of the delay, an in-depth analysis of the users' thinking process could be a promising avenue for future research. Similarly, in terms of system design, it would be interesting to investigate how these perceptions change if users are given the option to turn the delays on or off.

Finally, interactions between participants and chatbots in our experiment were rather short (i.e., five minutes on average) and did not allow us to track how perceptions and expectations evolve as users gain experience. Therefore, longitudinal studies are required to examine how perceptions and expectations evolve or change when people use a chatbot over a longer period of time.

7 Conclusion

Recent years have seen a rapid increase in the number of organizations that implement chatbots to automate customer service. Most, if not all, of the chatbots are intentionally designed to look and act like humans. Our findings, however, challenge the assumption that chatbots should always mimic human appearance and behavior. Taking users' prior experience into account, we find that a delayed response time of a chatbot – which more closely resembles human behavior than an instant response time – can have negative effects. While a delayed response time positively influences novice users' social presence perceptions and chatbot usage intentions, the effect is negative for experienced users. These findings not only highlight the important role of individual user characteristics in human–chatbot interaction, but could also help explain some of the high-profile failures of human-like chatbots (e.g., IKEA Anna) that could not sustain user engagement beyond an initial interest (Brandtzaeg and Følstad 2018). Therefore, a major implication of our study is that the current “one-design-fits-all” approach to chatbot design could be one reason for the ongoing struggle to meet users' expectations and increase adoption.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s12599-022-00755-x>.

Acknowledgements The authors thank the Department Editor, the Associate Editor, and the three anonymous reviewers for their valuable comments and helpful suggestions.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Adam M, Wessel M, Benlian A (2020) AI-based chatbots in customer service and their effects on user compliance. *Electron Mark* 9:1–19. <https://doi.org/10.1007/s12525-020-00414-7>
- Appel J, von der Pütten A, Krämer NC, Gratch J (2012) Does humanity matter? Analyzing the importance of social cues and perceived agency of a computer system for the emergence of social reactions during human-computer interaction. *Adv Hum-Comput Interact* 2012:1–10. <https://doi.org/10.1155/2012/324694>
- 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. *Comput Hum Behav* 85:183–189. <https://doi.org/10.1016/j.chb.2018.03.051>
- Ashktorab Z, Jain M, Liao QV, Weisz JD (2019) Resilient chatbots: repair strategy preferences for conversational breakdowns. In: *Proceedings of the 2019 CHI conference on human factors in computing systems*. Glasgow
- Babbie E, Wagner WE III, Zaino J (2018) *Adventures in social research: data analysis using IBM SPSS statistics*. Sage, Thousand Oaks
- Beattie A, Edwards AP, Edwards C (2020) A bot and a smile: interpersonal impressions of chatbots and humans using emoji in computer-mediated communication. *Commun Stud* 71:409–427. <https://doi.org/10.1080/10510974.2020.1725082>
- Benlian A, Klumpe J, Hinz O (2020) Mitigating the intrusive effects of smart home assistants by using anthropomorphic design features: a multimethod investigation. *Inf Syst J* 30:1010–1042. <https://doi.org/10.1111/isj.12243>
- Brandtzaeg PB, Følstad A (2017) Why people use chatbots. In: *Proceedings of the 4th international conference on internet science*. Thessaloniki, pp 377–392
- Brandtzaeg PB, Følstad A (2018) Chatbots: changing user needs and motivations. *Interactions* 25:38–43. <https://doi.org/10.1145/3236669>
- Bruch MA, Gorsky JM, Collins TM, Berger PA (1989) Shyness and sociability reexamined: a multicomponent analysis. *J Pers Soc Psychol* 57:904–915. <https://doi.org/10.1037/0022-3514.57.5.904>
- Buell RW, Norton MI (2011) The labor illusion: how operational transparency increases perceived value. *Manag Sci* 57:1564–1579. <https://doi.org/10.1287/mnsc.1110.1376>
- Burgoon JK (1978) A communication model of personal space violations: explication and an initial test. *Hum Commun Res* 4:129–142. <https://doi.org/10.1111/j.1468-2958.1978.tb00603.x>
- Burgoon JK (1993) Interpersonal expectations, expectancy violations, and emotional communication. *J Lang Soc Psychol* 12:30–48. <https://doi.org/10.1177/0261927X93121003>
- Burgoon JK (2009) Expectancy violations theory. In: Littlejohn SW, Foss KA (eds) *Encyclopedia of communication theory*. Sage, Thousand Oaks, pp 367–369
- Burgoon JK (2015) Expectancy violations theory. In: Berger CR, Roloff ME (eds) *The international encyclopedia of interpersonal communication*. Wiley Blackwell, New York, pp 1–9
- Burgoon JK, Guerrero L, Floyd K (2010) *Nonverbal communication*. Routledge, New York
- Burgoon JK, Bonito JA, Lowry PB et al (2016) Application of expectancy violations theory to communication with and judgments about embodied agents during a decision-making task. *Int J Hum Comput Stud* 91:24–36. <https://doi.org/10.1016/j.ijhcs.2016.02.002>
- Cambre J, Williams AC, Razi A, et al (2021) Firefox Voice: an open and extensible voice assistant built upon the web. In: *Proceedings of the 2021 CHI conference on human factors in computing systems (CHI '21)*. Yokohama
- Crozier R (2017) Lufthansa delays chatbot's responses to make it more "human." <https://www.itnews.com.au/news/lufthansa-delays-chatbots-responses-to-make-it-more-human-462643>. Accessed 10 May 2020
- Cyr D, Hassanein K, Head M, Ivanov A (2007) The role of social presence in establishing loyalty in e-service environments. *Interact Comput* 19:43–56. <https://doi.org/10.1016/j.intcom.2006.07.010>
- Dale R (2016) The return of the chatbots. *Nat Lang Eng* 22:811–817. <https://doi.org/10.1017/S1351324916000243>
- Derrick DC, Meservy TO, Jenkins JL et al (2013) Detecting deceptive chat-based communication using typing behavior and message cues. *ACM Trans Manag Inf Syst* 4:1–21. <https://doi.org/10.1145/2499962.2499967>
- Diederich S, Brendel AB, Kolbe LM (2020) Designing anthropomorphic enterprise conversational agents. *Bus Inf Syst Eng* 62:193–209. <https://doi.org/10.1007/s12599-020-00639-y>
- Edwards JR, Lambert LS (2007) Methods for integrating moderation and mediation: a general analytical framework using moderated path analysis. *Psychol Methods* 12:1–22. <https://doi.org/10.1037/1082-989X.12.1.1>
- Facebook (2018) F8 2018: David Marcus' Keynote. <https://www.facebook.com/business/news/david-marcus-f8-keynote-2018>. Accessed 22 Oct 2020
- Faul F, Erdfelder E, Lang A-G, Buchner A (2007) G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods* 39:175–191. <https://doi.org/10.3758/BF03193146>
- Feine J, Gnewuch U, Morana S, Maedche A (2019) A Taxonomy of social cues for conversational agents. *Int J Hum Comput Stud* 132:138–161. <https://doi.org/10.1016/j.ijhcs.2019.07.009>
- Finsaas MC, Goldstein BL (2021) Do simple slopes follow-up tests lead us astray? Advancements in the visualization and reporting of interactions. *Psychol Methods* 26:38–60. <https://doi.org/10.1037/met0000266>
- Fogg BJ (2002) *Computers as persuasive social actors. Persuasive technology: using computers to change what we think and do*. Morgan Kaufmann, San Francisco, pp 89–120
- Følstad A, Brandtzaeg PB (2017) Chatbots and the new world of HCI. *Interactions* 24:38–42. <https://doi.org/10.1145/3085558>
- Følstad A, Nordheim CB, Bjørkli CA (2018) What makes users trust a chatbot for customer service? An exploratory interview study. In: *Proceedings of the 5th international conference on internet science*. St. Petersburg, pp 194–208

- Fornell C, Larcker DF (1981) Evaluating structural equation models with unobservable variables and measurement error. *J Mark Res* 18:39. <https://doi.org/10.2307/3151312>
- Galletta DF, Henry RM, McCoy S, Polak P (2006) When the wait isn't so bad: the interacting effects of website delay, familiarity, and breadth. *Inf Syst Res* 17:20–37. <https://doi.org/10.1287/isre.1050.0073>
- Galletta DF, Dunn B (2014) Assessing smartphone ease of use and learning from the perspective of novice and expert users: development and illustration of mobile benchmark tasks. *AIS Trans Hum-Comput Interact* 6:74–91. <https://doi.org/10.17705/1thci.00062>
- Galletta DF, Henry R, McCoy S, Polak P (2004) Web site delays: How tolerant are users? *J Assoc Inf Syst* 5:1–28. <https://doi.org/10.17705/1jais.00044>
- Gambino A, Fox J, Ratan R (2020) Building a stronger CASA: extending the computers are social actors paradigm. *Hum-Mach Commun* 1:71–86. <https://doi.org/10.30658/hmc.1.5>
- Gefen D, Straub DW (2005) A practical guide to factorial validity using PLS-Graph: tutorial and annotated example. *Commun Assoc Inf Syst* 16:91–109. <https://doi.org/10.17705/1CAIS.01605>
- Gefen D, Straub DW (1997) Gender differences in the perception and use of e-mail: an extension to the technology acceptance model. *MIS Q* 21:389. <https://doi.org/10.2307/249720>
- Gefen D, Straub DW (2004) Consumer trust in b2c e-commerce and the importance of social presence: experiments in e-products and e-services. *Omega* 32:407–424. <https://doi.org/10.1016/j.omega.2004.01.006>
- Gnewuch U, Morana S, Adam MTP, Maedche A (2018) Faster is not always better: understanding the effect of dynamic response delays in human–chatbot interaction. In: Proceedings of the 26th European conference on information systems. Portsmouth, UK
- Go E, Sundar SS (2019) Humanizing chatbots: the effects of visual, identity and conversational cues on humanness perceptions. *Comput Hum Behav* 97:304–316. <https://doi.org/10.1016/j.chb.2019.01.020>
- Grimes GM, Schuetzler RM, Giboney JS (2021) Mental models and expectation violations in conversational AI interactions. *Decis Support Syst* 144:113515. <https://doi.org/10.1016/j.dss.2021.113515>
- Grudin J, Jacques R (2019) Chatbots, Humbots, and the quest for artificial general intelligence. In: Proceedings of the 2019 CHI conference on human factors in computing systems. Glasgow
- Hassanein K, Head M (2007) Manipulating perceived social presence through the web interface and its impact on attitude towards online shopping. *Int J Hum Comput Stud* 65:689–708. <https://doi.org/10.1016/j.ijhcs.2006.11.018>
- Hayes AF (2015) An index and test of linear moderated mediation. *Multivar Behav Res* 50:1–22. <https://doi.org/10.1080/00273171.2014.962683>
- Hayes AF (2018) Introduction to mediation, moderation, and conditional process analysis: a regression-based approach, 2nd edn. Guilford, New York
- Hess T, Fuller M, Campbell D (2009) Designing interfaces with social presence: using vividness and extraversion to create social recommendation agents. *J Assoc Inf Syst* 10:889–919. <https://doi.org/10.17705/1jais.00216>
- Holtgraves T, Han T-L (2007) A procedure for studying online conversational processing using a chat bot. *Behav Res Methods* 39:156–163. <https://doi.org/10.3758/BF03192855>
- Holtgraves T, Ross SJ, Weywadt CR, Han T-L (2007) Perceiving artificial social agents. *Comput Hum Behav* 23:2163–2174. <https://doi.org/10.1016/j.chb.2006.02.017>
- Inmar (2019) What marketers need to know about chatbots. <https://www.inmar.com/blog/thought-leadership/what-marketers-need-know-about-chatbots>. Accessed 1 Jul 2020
- Jacquet B, Baratgin J, Jamet F (2019) Cooperation in online conversations: the response times as a window into the cognition of language processing. *Front Psychol* 10:1–15. <https://doi.org/10.3389/fpsyg.2019.00727>
- Jain M, Kumar P, Kota R, Patel SN (2018) Evaluating and informing the design of chatbots. In: Proceedings of the 2018 designing interactive systems conference. Hong Kong, pp 895–906
- Janssen A, Passlick J, Rodríguez Cardona D, Breitner MH (2020) Virtual assistance in any context – a taxonomy of design elements for domain-specific chatbots. *Bus Inf Syst Eng* 62:211–225. <https://doi.org/10.1007/s12599-020-00644-1>
- Jenkins M-C, Churchill R, Cox S, Smith D (2007) Analysis of user interaction with service oriented chatbot systems. In: Jacko J (ed) Human-computer interaction. HCI Intelligent multimodal interaction environments. Springer, Heidelberg, pp 76–83
- Kalman YM, Rafaeli S (2011) Online pauses and silence: chronemic expectancy violations in written computer-mediated communication. *Commun Res* 38:54–69. <https://doi.org/10.1177/0093650210378229>
- Kalman YM, Scissors LE, Gill AJ, Gergle D (2013) Online chronemics convey social information. *Comput Hum Behav* 29:1260–1269. <https://doi.org/10.1016/j.chb.2012.12.036>
- Lee K, Nass C (2005) Social-psychological origins of feelings of presence: creating social presence with machine-generated voices. *Media Psychol* 7:31–45. https://doi.org/10.1207/S1532785XMEP0701_2
- Lester J, Branting K, Mott B (2004) Conversational agents. In: Singh MP (ed) The practical handbook of internet computing. Chapman & Hall/CRC, Boca Raton
- Liao QV, Davis M, Geyer W, et al (2016) What can you do? Studying social-agent orientation and agent proactive interactions with an agent for employees. In: Proceedings of the 2016 ACM Conference On Designing Interactive Systems. Brisbane, pp 264–275
- Littlejohn SW, Foss KA (2009) Encyclopedia of communication theory. Sage, Thousand Oaks
- Lynch JG Jr (1982) On the external validity of experiments in consumer research. *J Consum Res* 9:225–239. <https://doi.org/10.1086/208919>
- McTear MF (2017) The rise of the conversational interface: a new kid on the block? FETLT 2016. Springer, Cham, pp 38–49
- Microsoft (2021) Microsoft bot framework. <https://dev.botframework.com/>. Accessed 1 Aug 2021
- Mirbabaie M, Stieglitz S, Brünker F et al (2021) Understanding collaboration with virtual assistants – the role of social identity and the extended self. *Bus Inf Syst Eng* 63:21–37. <https://doi.org/10.1007/s12599-020-00672-x>
- Moon Y (1999) The effects of physical distance and response latency on persuasion in computer-mediated communication and human-computer communication. *J Exp Psychol Appl* 5:379–392. <https://doi.org/10.1037/1076-898X.5.4.379>
- Moon Y (2000) Intimate exchanges: using computers to elicit self-disclosure from consumers. *J Consum Res* 26:323–339. <https://doi.org/10.1086/209566>
- Moussawi S, Benbunan-Fich R (2020) The effect of voice and humour on users' perceptions of personal intelligent agents. *Behav Inf Technol*. <https://doi.org/10.1080/0144929X.2020.1772368>
- Moussawi S, Koufaris M, Benbunan-Fich R (2020) How perceptions of intelligence and anthropomorphism affect adoption of personal intelligent agents. *Electron Mark*. <https://doi.org/10.1007/s12525-020-00411-w>

- Muresan A, Pohl H (2019) Chats with bots: balancing imitation and engagement. In: Extended abstracts of the 2019 CHI conference on human factors in computing systems. Glasgow
- Nass C, Lee K (2001) Does computer-synthesized speech manifest personality? Experimental tests of recognition, similarity-attraction, and consistency-attraction. *J Exp Psychol Appl* 7:171–181. <https://doi.org/10.1037/1076-898X.7.3.171>
- Nass C, Moon Y (2000) Machines and mindlessness: social responses to computers. *J Soc Issues* 56:81–103. <https://doi.org/10.1111/0022-4537.00153>
- Nass C, Fogg BJ, Moon Y (1996) Can computers be teammates? *Int J Hum Comput Stud* 45:669–678. <https://doi.org/10.1006/jhch.1996.0073>
- Nass C, Steuer J, Tauber ER (1994) Computers are social actors. In: Proceedings of the SIGCHI conference on human factors in computing systems. Boston, pp 72–78
- Nordheim CB, Følstad A, Bjørkli CA (2019) An initial model of trust in chatbots for customer service – findings from a questionnaire study. *Interact Comput* 31:317–335. <https://doi.org/10.1093/iwc/iwz022>
- Nunnally JC, Bernstein IH (1994) Psychometric theory. McGraw-Hill, New York
- Orme MT (1962) On the social psychology of the psychological experiment: with particular reference to demand characteristics and their implications. *Am Psychol* 17:776–783. <https://doi.org/10.1037/h0043424>
- Pfeuffer N, Benlian A, Gimpel H, Hinz O (2019) Anthropomorphic information systems. *Bus Inf Syst Eng* 61:523–533. <https://doi.org/10.1007/s12599-019-00599-y>
- Podsakoff PM, MacKenzie SB, Lee J-Y, Podsakoff NP (2003) Common method biases in behavioral research: a critical review of the literature and recommended remedies. *J Appl Psychol* 88:879–903. <https://doi.org/10.1037/0021-9010.88.5.879>
- Porcheron M, Fischer JE, Reeves S, Sharples S (2018) Voice interfaces in everyday life. In: Proceedings of the 2018 CHI conference on human factors in computing systems. Montreal
- Price PC, Jhangiani RS, Chiang I-CA (2015) Research methods in psychology, 2nd edn. BCCampus
- Qiu L, Benbasat I (2009) Evaluating anthropomorphic product recommendation agents: a social relationship perspective to designing information systems. *J Manag Inf Syst* 25:145–182. <https://doi.org/10.2753/MIS0742-1222250405>
- Reeves B, Nass C (1996) The media equation: how people treat computers, television, and new media like real people and places. CSLI Publications, Cambridge
- Schanke S, Burtch G, Ray G (2021) Estimating the impact of “humanizing” customer service chatbots. *Inf Syst Res* 32:736–751. <https://doi.org/10.1287/isre.2021.1015>
- Schuetzler RM, Grimes GM, Giboney JS (2019) The effect of conversational agent skill on user behavior during deception. *Comput Hum Behav* 97:250–259. <https://doi.org/10.1016/j.chb.2019.03.033>
- Schuetzler RM, Grimes GM, Scott Giboney J (2020) The impact of chatbot conversational skill on engagement and perceived humanness. *J Manag Inf Syst* 37:875–900. <https://doi.org/10.1080/07421222.2020.1790204>
- Schuetzler RM, Grimes GM, Giboney JS, Buckman J (2014) Facilitating natural conversational agent interactions: lessons from a deception experiment. In: Proceedings of the 35th international conference on information systems. Auckland
- Schuetzler RM, Grimes GM, Rosser HK (2021) Deciding whether and how to deploy chatbots. *MIS Q Exec* 20:1–15. <https://doi.org/10.17705/2msqe.00039>
- Schuetzler RM (2015) Dynamic interviewing agents: effects on deception, nonverbal behavior, and social desirability. Dissertation, University of Arizona
- Seeger A-M, Pfeiffer J, Heinzl A (2021) Texting with human-like conversational agents: designing for anthropomorphism. *J Assoc Inf Syst* 22:931–967. <https://doi.org/10.17705/1jais.00685>
- Sheehan B, Jin HS, Gottlieb U (2020) Customer service chatbots: anthropomorphism and adoption. *J Bus Res* 115:14–24. <https://doi.org/10.1016/j.jbusres.2020.04.030>
- Sheldon OJ, Thomas-Hunt MC, Proell CA (2006) When timeliness matters: the effect of status on reactions to perceived time delay within distributed collaboration. *J Appl Psychol* 91:1385–1395. <https://doi.org/10.1037/0021-9010.91.6.1385>
- Shiwa T, Kanda T, Imai M et al (2009) How quickly should a communication robot respond? Delaying strategies and habituation effects. *Int J Soc Robot* 1:141–155. <https://doi.org/10.1007/s12369-009-0012-8>
- Short J, Williams E, Christie B (1976) The social psychology of telecommunications. John Wiley & Sons, London, UK
- SmartAction (2018) How demographics affect chatbot usage. <https://www.smartaction.ai/blog/demographics-affect-chatbot-adoption-use>. Accessed 1 Jul 2020
- Spence PR, Westerman D, Edwards C, Edwards A (2014) Welcoming our robot overlords: initial expectations about interaction with a robot. *Commun Res Rep* 31:272–280. <https://doi.org/10.1080/08824096.2014.924337>
- Spiller SA, Fitzsimons GJ, Lynch JG, McClelland GH (2013) Spotlights, floodlights, and the magic number zero: simple effects tests in moderated regression. *J Mark Res* 50:277–288. <https://doi.org/10.1509/jmr.12.0420>
- Taylor S, Todd P (1995) Assessing IT usage: the role of prior experience. *MIS Q* 19:561–570. <https://doi.org/10.2307/249633>
- SysAid Technologies (2019) Everything you need to know about IT support chatbot success. In: Joe IT Guy Blog. <https://www.joetheitguy.com/chatbot-success/>. Accessed 1 May 2021
- Thompson RL, Higgins CA, Howell JM (1994) Influence of experience on personal computer utilization: testing a conceptual model. *J Manag Inf Syst* 11:167–187. <https://doi.org/10.1080/07421222.1994.11518035>
- Tsekouras D, Li T, Benbasat I (2022) Scratch my back and I’ll scratch yours: the impact of user effort and recommendation agent effort on perceived recommendation agent quality. *Inf Manag* 59:103571. <https://doi.org/10.1016/j.im.2021.103571>
- Tuzovic S, Paluch S (2018) Conversational commerce – a new era for service business development? In: Bruhn M, Hadwich K (eds) Service business development. Springer, Wiesbaden, pp 81–100
- vom Brocke J, Maaß W, Buxmann P et al (2018) Future work and enterprise systems. *Bus Inf Syst Eng* 60:357–366. <https://doi.org/10.1007/s12599-018-0544-2>
- Walther JB, Tidwell LC (1995) Nonverbal cues in computer-mediated communication, and the effect of chronemics on relational communication. *J Organ Comput* 5:355–378. <https://doi.org/10.1080/10919399509540258>
- Wang W, Benbasat I (2009) Interactive decision aids for consumer decision making in e-commerce: the influence of perceived strategy restrictiveness. *MIS Q* 33:293–320. <https://doi.org/10.2307/20650293>
- Weizenbaum J (1966) ELIZA – a computer program for the study of natural language communication between man and machine. *Commun ACM* 9:36–45. <https://doi.org/10.1145/365153.365168>
- Xiao L, Kumar V (2021) Robotics for customer service: a useful complement or an ultimate substitute? *J Serv Res* 24:9–29. <https://doi.org/10.1177/1094670519878881>
- Xu K (2019) First encounter with robot Alpha: how individual differences interact with vocal and kinetic cues in users’ social responses. *New Media Soc* 21:2522–2547. <https://doi.org/10.1177/1461444819851479>

- Xu DJ, Benbasat I, Cenfetelli RT (2018) The outcomes and the mediating role of the functional triad: the users' perspective. *Inf Syst J* 28:956–988. <https://doi.org/10.1111/isj.12183>
- Xu A, Liu Z, Guo Y, et al (2017) A new chatbot for customer service on social media. In: *Proceedings of the 2017 CHI conference on human factors in computing systems*. Denver
- Yu M, Zhou R, Cai Z et al (2020) Unravelling the relationship between response time and user experience in mobile applications. *Internet Res* 30:1353–1382. <https://doi.org/10.1108/INTR-05-2019-0223>