



AI-based chatbots in customer service and their effects on user compliance

Martin Adam¹ · Michael Wessel² · Alexander Benlian¹

Received: 18 July 2019 / Accepted: 12 February 2020
© The Author(s) 2020

Abstract

Communicating with customers through live chat interfaces has become an increasingly popular means to provide real-time customer service in many e-commerce settings. Today, human chat service agents are frequently replaced by conversational software agents or chatbots, which are systems designed to communicate with human users by means of natural language often based on artificial intelligence (AI). Though cost- and time-saving opportunities triggered a widespread implementation of AI-based chatbots, they still frequently fail to meet customer expectations, potentially resulting in users being less inclined to comply with requests made by the chatbot. Drawing on social response and commitment-consistency theory, we empirically examine through a randomized online experiment how verbal anthropomorphic design cues and the foot-in-the-door technique affect user request compliance. Our results demonstrate that both anthropomorphism as well as the need to stay consistent significantly increase the likelihood that users comply with a chatbot's request for service feedback. Moreover, the results show that social presence mediates the effect of anthropomorphic design cues on user compliance.

Keywords Artificial intelligence · Chatbot · Anthropomorphism · Social presence · Compliance · Customer service

JEL classification C91 · D91 · M31 · L86

Introduction

Communicating with customers through live chat interfaces has become an increasingly popular means to provide real-time customer service in e-commerce settings. Customers use these chat services to obtain information (e.g., product details) or assistance (e.g., solving technical problems). The real-time nature of chat services has transformed customer service into a two-way communication with significant effects on trust, satisfaction, and repurchase as well as WOM intentions (Mero 2018). Over the last decade, chat services have

become the preferred option to obtain customer support (Charlton 2013). More recently, and fueled by technological advances in artificial intelligence (AI), human chat service agents are frequently replaced by conversational software agents (CAs) such as chatbots, which are systems such as chatbots designed to communicate with human users by means of natural language (e.g., Gnewuch et al. 2017; Pavlikova et al. 2003; Pfeuffer et al. 2019a). Though rudimentary CAs emerged as early as the 1960s (Weizenbaum 1966), the “second wave of artificial intelligence” (Launchbury 2018) has renewed the interest and strengthened the commitment to this technology, because it has paved the way for systems that are capable of more human-like interactions (e.g., Gnewuch et al. 2017; Maedche et al. 2019; Pfeuffer et al. 2019b). However, despite the technical advances, customers continue to have unsatisfactory encounters with CAs that are based on AI. CAs may, for instance, provide unsuitable responses to the user requests, leading to a gap between the user's expectation and the system's performance (Luger and Sellen 2016; Orłowski 2017). With AI-based CAs displacing human chat service agents, the question arises whether live chat services will continue to be effective, as skepticism and resistance

Responsible Editor: Christian Matt

✉ Martin Adam
adam@ise.tu-darmstadt.de

¹ Institute of Information Systems & Electronic Services, Technical University of Darmstadt, Hochschulstraße 1, 64289 Darmstadt, Germany

² Department of Digitalization, Copenhagen Business School, Howitzvej 60, 2000 Frederiksberg, Denmark

against the technology might obstruct task completion and inhibit successful service encounters. Interactions with these systems might thus trigger unwanted behaviors in customers such as a noncompliance that can negatively affect both the service providers as well as users (Bowman et al. 2004). However, if customers choose not to conform with or adapt to the recommendations and requests given by the CAs this calls into question the *raison d'être* of this self-service technology (Cialdini and Goldstein 2004).

To address this challenge, we employ an experimental design based on an AI-based chatbot (hereafter simply “chatbot”), which is a particular type of CAs that is designed for turn-by-turn conversations with human users based on textual input. More specifically, we explore what characteristics of the chatbot increase the likelihood that users comply with a chatbot’s request for service feedback through a customer service survey. We have chosen this scenario to test the user’s compliance because a customer’s assessment of service quality is important and a universally applicable predictor for customer retention (Gustafsson et al. 2005).

Prior research suggests that CAs should be designed anthropomorphically (i.e., human-like) and create a sense of social presence (e.g., Rafaeli and Noy 2005; Zhang et al. 2012) by adopting characteristics of human-human communication (e.g., Derrick et al. 2011; Elkins et al. 2012). Most of this research focused on anthropomorphic design cues and their impact on human behavior with regard to perceptions and adoptions (e.g., Adam et al. 2019; Hess et al. 2009; Qiu and Benbasat 2009). This work offers valuable contributions to research and practice but has been focused primarily on embodied CAs that have a virtual body or face and are thus able to use nonverbal anthropomorphic design cues (i.e., physical appearance or facial expressions). Chatbots, however, are disembodied CAs that predominantly use verbal cues in their interactions with users (Araujo 2018; Feine et al. 2019). While some prior work exists that investigates verbal anthropomorphic design cues, such as self-disclosure, excuse, and thanking (Feine et al. 2019), due to limited capabilities of previous generations of CAs, these cues have often been rather static and insensitive to the user’s input. As such, users might develop an aversion against such a system, because of its inability to realistically mimic a human-human communication.¹ Today, conversational computing platforms (e.g., IBM Watson Assistant) allow sophisticated chatbot solutions that delicately comprehend user input based on narrow AI.² Chatbots built on these systems have a comprehension that is closer to that of humans and that allows for more flexible as

well as empathetic responses to the user’s input compared to the rather static responses of their rule-based predecessors (Reeves and Nass 1996). These systems thus allow new anthropomorphic design cues such as exhibiting empathy through conducting small talk. Besides a few exceptions (e.g., Araujo 2018; Derrick et al. 2011), the implications of more advanced anthropomorphic design cues remain underexplored.

Furthermore, as chatbots continue to displace human service agents, the question arises whether compliance and persuasion techniques, which are intended to influence users to comply with or adapt to a specific request, are equally applicable in these new technology-based self-service settings. The continued-question procedure as a form of the foot-in-the-door compliance technique is particularly relevant as it is not only abundantly used in practice but its success has been shown to be heavily dependent on the kind of requester (Burger 1999). The effectiveness of this compliance technique may thus differ when applied by CAs rather than human service agents. Although the application of CAs as artificial social actors or agents seem to be a promising new field for research on compliance and persuasion techniques, it has been hitherto neglected.

Against this backdrop, we investigate how verbal anthropomorphic design cues and the foot-in-the-door compliance tactic influence user compliance with a chatbot’s feedback request in a self-service interaction. Our research is guided by the following research questions:

RQ1: *How do verbal anthropomorphic design cues affect user request compliance when interacting with an AI-based chatbot in customer self-service?*

RQ2: *How does the foot-in-the-door technique affect user request compliance when interacting with an AI-based chatbot in customer self-service?*

We conducted an online experiment with 153 participants and show that both verbal anthropomorphic design cues and the foot-in-the-door technique increase user compliance with a chatbot’s request for service feedback. We thus demonstrate how anthropomorphism and the need to stay consistent can be used to influence user behavior in the interaction with a chatbot as a self-service technology.

Our empirical results provide contributions for both research and practice. First, this study extends prior research by showing that the computers-are-social-actors (CASA) paradigm extends to disembodied CAs that predominantly use verbal cues in their interactions with users. Second, we show that humans acknowledge CAs as a source of persuasive messages and that the degree to which humans comply with the artificial social agents depends on the techniques applied during the human-chatbot communication. For platform providers and online marketers, especially for those who consider employing AI-based CAs in customer self-service, we offer two recommendations. First, during CAs interactions, it is not

¹ Referred to as “Uncanny Valley” in settings that involve embodied CAs such as avatars (Mori 1970; Seymour et al. 2019).

² Narrow or weak artificial intelligence refers to systems capable of carrying out a narrow set of tasks that require a specific human capability such as visual perception or natural language processing. However, the system is incapable of applying intelligence to any problem, which requires strong AI.

necessary for providers to attempt to fool users into thinking they are interacting with a human. Rather, the focus should be on employing strategies to achieve greater human likeness through anthropomorphism, which we have shown to have a positive effect on user compliance. Second, providers should design CA dialogs as carefully as they design the user interface. Our results highlight that the dialog design can be a decisive factor for user compliance with a chatbot's request.

Theoretical background

The role of conversational agents in service systems

A key challenge for customer service providers is to balance service efficiency and service quality: Both researchers and practitioners emphasize the potential advantages of customer self-service, including increased time-efficiency, reduced costs, and enhanced customer experience (e.g., Meuter et al. 2005; Scherer et al. 2015). CAs, as a self-service technology, offer a number of cost-saving opportunities (e.g., Gnewuch et al. 2017; Pavlikova et al. 2003), but also promise to increase service quality and improve provider-customer encounters. Studies estimate that CAs can reduce current global business costs of \$1.3 trillion related to 265 billion customer service inquiries per year by 30% through decreasing response times, freeing up agents for different work, and dealing with up to 80% of routine questions (Reddy 2017b; Techlabs 2017). Chatbots alone are expected to help business save more than \$8 billion per year by 2022 in customer-supporting costs, a tremendous increase from the \$20 million in estimated savings for 2017 (Reddy 2017a). CAs thus promise to be fast, convenient, and cost-effective solutions in form of 24/7 electronic channels to support customers (e.g., Hopkins and Silverman 2016; Meuter et al. 2005).

Customers usually not only appreciate easily accessible and flexible self-service channels, but also value personalized attention. Thus, firms should not shift towards customer self-service channels completely, especially not at the beginning of a relationship with a customer (Scherer et al. 2015), as the absence of a personal social actor in online transactions can translate into loss of sales (Raymond 2001). However, by mimicking social actors, CAs have the potential to actively influence service encounters and to become surrogates for service employees by completing assignments that used to be done by human service staff (e.g., Larivière et al. 2017; Verhagen et al. 2014). For instance, instead of calling a call center or writing an e-mail to ask a question or to file a complaint, customers can turn to CAs that are available 24/7. This self-service channel will become progressively relevant as the interface between companies and consumers is "gradually evolving to become technology dominant (i.e., intelligent assistants acting as a service interface) rather than human-driven

(i.e., service employee acting as service interface)" (Larivière et al. 2017, p. 239). Moreover, recent AI-based CAs have the option to signal human characteristics such as friendliness, which are considered crucial for handling service encounters (Verhagen et al. 2014). Consequently, in comparison to former online service encounters, CAs can reduce the former lack of interpersonal interaction by evoking perceptions of social presence and personalization.

Today, CAs, and chatbots in particular, have already become a reality in electronic markets and customer service on many websites, social media platforms, and in messaging apps. For instance, the number of chatbots on Facebook Messenger soared from 11,000 to 300,000 between June 2016 and April 2019 (Facebook 2019). Although these technological artefacts are on the rise, previous studies indicated that chatbots still suffer from problems linked to their infancy, resulting in high failure rates and user skepticism when it comes to the application of AI-based chatbots (e.g., Orłowski 2017). Moreover, previous research has revealed that, while human language skills transfer easily to human-chatbot communication, there are notable differences in the content and quality of such conversations. For instance, users communicate with chatbots for a longer duration and with less rich vocabulary as well as greater profanity (Hill et al. 2015). Thus, if users treat chatbots differently, their compliance as a response to recommendations and requests made by the chatbot may be affected. This may thus call into question the promised benefits of the self-service technology. Therefore, it is important to understand how the design of chatbots impacts user compliance.

Social response theory and anthropomorphic design cues

The well-established social response theory (Nass et al. 1994) has paved the way for various studies providing evidence on how humans apply social rules to anthropomorphically designed computers. Consistent with previous research in digital contexts, we define anthropomorphism as the attribution of human-like characteristics, behaviors, and emotions to nonhuman agents (Epley et al. 2007). The phenomenon can be understood as a natural human tendency to ease the comprehension of unknown actors by applying anthropocentric knowledge (e.g., Epley et al. 2007; Pfeuffer et al. 2019a).

According to social response theory (Nass and Moon 2000; Nass et al. 1994), human-computer interactions (HCIs) are fundamentally social: Individuals are biased towards automatically as well as unconsciously perceiving computers as social actors, even when they know that machines do not hold feelings or intentions. The identified psychological effect underlying the computers-are-social-actors (CASA) paradigm is the evolutionary biased social orientation of human beings (Nass and Moon 2000; Reeves and Nass 1996). Consequently,

through interacting with an anthropomorphized computer system, a user may perceive a sense of social presence (i.e., a “degree of salience of the other person in the interaction” (Short et al. 1976, p. 65)), which was originally a concept to assess users’ perceptions of human contact (i.e., warmth, empathy, sociability) in technology-mediated interactions with other users (Qiu and Benbasat 2009). Therefore, the term “agent”, for example, which referred to a human being who offers guidance, has developed into an established term for anthropomorphically designed computer-based interfaces (Benlian et al. 2019; Qiu and Benbasat 2009).

In HCI contexts, when presented with a technology possessing cues that are normally associated with human behavior (e.g., language, turn-taking, interactivity), individuals respond by exhibiting social behavior and making anthropomorphic attributions (Epley et al. 2007; Moon and Nass 1996; Nass et al. 1995). Thus, individuals apply the same social norms to computers as they do to humans: In interactions with computers, even few anthropomorphic design cues³ (ADCs) can trigger social orientation and perceptions of social presence in an individual and, thus, responses in line with socially desirable behavior. As a result, social dynamics and rules guiding human-human interaction similarly apply to HCI. For instance, CASA studies have shown that politeness norms (Nass et al. 1999), gender and ethnicity stereotypes (Nass and Moon 2000; Nass et al. 1997), personality response (Nass et al. 1995), and flattery effects (Fogg and Nass 1997) are also present in HCI.

Whereas nonverbal ADCs, such as physical appearance or embodiment, aim to improve the social connection by implementing motoric and static human characteristics (Eyssel et al. 2010), verbal ADCs, such as the ability to chat, rather intend to establish the perception of intelligence in a non-human technological agent (Araujo 2018). As such, static and motoric anthropomorphic embodiments through avatars in marketing contexts have been found predominantly useful to influence trust and social bonding with virtual agents (e.g., Qiu and Benbasat 2009) and particularly important for service encounters and online sales, for example on company websites (e.g., Etemad-Sajadi 2016; Holzwarth et al. 2006), in virtual worlds (e.g., Jin 2009; Jin and Sung 2010), and even in physical interactions with robots in stores (Bertacchini et al. 2017). Yet, chatbots are rather disembodied CAs, as they mainly interact with customers via messaging-based interfaces through verbal (e.g., language style) and non-verbal cues (e.g., blinking dots), allowing a real-time dialogue through primarily text input but omitting physical and dynamic representations, except for the typically static profile picture. Besides two exceptions that focused on verbal ADCs (Araujo 2018; Go and Sundar 2019), to the best of our knowledge, no other

studies have directly targeted verbal ADCs to extend past research on embodied agents.

Compliance, foot-in-the-door technique and commitment-consistency theory

The term compliance refers to “a particular kind of response — acquiescence—to a particular kind of communication — a request” (Cialdini and Goldstein 2004, p. 592). The request can be either explicit, such as asking for a charitable donation in a door-to-door campaign, or implicit, such as in a political advertisement that endorses a candidate without directly urging a vote. Nevertheless, in all situations, the targeted individual realizes that he or she is addressed and prompted to respond in a desired way. Compliance research has devoted its efforts on various compliance techniques, such as the that’s-not-all technique (Burger 1986), the disrupt-then-reframe technique (Davis and Knowles 1999; Knowles and Linn 2004), door-in-the-face technique (Cialdini et al. 1975), and foot-in-the-door (FITD) (Burger 1999). In this study, we focus on FITD, one of the most researched and applied compliance techniques, as the technique’s naturally sequential and conversational character seems specifically well suited for chatbot interactions.

The FITD compliance technique (e.g., Burger 1999; Freedman and Fraser 1966) builds upon the effect of small commitments to influence individuals to comply. The first experimental demonstration of the FITD dates back to Freedman and Fraser (1966), in which a team of psychologists called housewives to ask if the women would answer a few questions about the household products they used. Three days later, the psychologists called again, this time asking if they could send researchers to the house to go through cupboards as part of a 2-h enumeration of household products. The researchers found these women twice as likely to comply than a group of housewives who were asked only the large request. Nowadays, online marketing and sales abundantly exploit this compliance technique to make customers agree to larger commitments. For example, websites often ask users for small commitments (e.g., providing an e-mail address, clicking a link, or sharing on social media), only to follow with a conversion-focused larger request (e.g., asking for sale, software download, or credit-card information).

Individuals in compliance situations bear the burden of correctly comprehending, evaluating, and responding to a request in a short time (Cialdini 2009), thus they lack time to make a fully elaborated rational decision and, therefore, use heuristics (i.e., rules of thumb) (Simon 1990) to judge the available options. In contrast to large requests, small requests are more successful in convincing subjects to agree with the requester as individuals spend less mental effort on small commitments. Once the individuals accept the commitment, they are more likely to agree with a next bigger commitment to stay consistent with their initial behavior. The FITD technique thus

³ Sometimes also referred to as “social cues” or “social features”.

exploits the natural tendency of individuals to justify the initial agreement to the small request to themselves and others.

The human need for being consistent with their behavior is based on various underlying psychological processes (Burger 1999), of which most draw on self-perception theory (Bem 1972) and commitment-consistency theory (Cialdini 2001). These theories constitute that individuals have only weak inherent attitudes and rather form their attitudes by self-observations. Consequently, if individuals comply with an initial request, a bias arises and the individuals will conclude that they must have considered the request acceptable and, thus, are more likely to agree to a related future request of the same kind or from the same cause (Kressmann et al. 2006). In fact, previous research in marketing has empirically demonstrated that consumer's need for self-consistency encourages purchase behavior (e.g., Ericksen and Sirgy 1989).

Moreover, research has demonstrated that consistency is an important factor in social exchange. To cultivate relationships, individuals respond rather affirmatively to a request and are more likely to comply the better the relationship is developed (Cialdini and Trost 1998). In fact, simply being exposed to a person for a brief period without any interaction significantly increases compliance with the person's request, which is even stronger when the request is made face-to-face and unexpectedly (Burger et al. 2001). In private situations, individuals even decide to comply to a request simply to reduce feelings of guilt and pity (Whatley et al. 1999) and to gain social approval from others to improve their self-esteem (Deutsch and Gerard 1955). Consequently, individuals also have external reasons to comply, so that social biases may lead to non-rational decisions (e.g., Chaiken 1980; Wessel et al. 2019; Wilkinson and Klaes 2012). Previous studies on FITD have demonstrated that compliance is heavily dependent on the dialogue design of the verbal interactions and on the kind of requester (Burger 1999). However, in user-system interactions with chatbots, the requester (i.e., the chatbot) may lack crucial characteristics for the success of the FITD, such as perceptions of social presence and social consequences that arise by not being consistent. Consequently, it is important to investigate whether FITD and other compliance techniques can also be applied to written information exchanges with a chatbot and what unique differences might arise by replacing a human with a computational requester.

Hypotheses development and research model

Effect of anthropomorphic design cues on user compliance via social presence.

According to the CASA paradigm (Nass et al. 1994), users tend to treat computers as social actors. Earlier research demonstrated that social reactions to computers in general (Nass and Moon 2000) and to embodied conversational agents in particular (e.g., Astrid et al. 2010) depend on the kind and

number of ADCs: Usually, the more cues a CA displays, the more socially present the CA will appear to users and the more users will apply and transfer knowledge and behavior that they have learned from their human-human-interactions to the HCI. Applied to our piece of research, we focus on only verbal ADCs, thus avoiding potential confounding nonverbal cues through chatbot embodiments. As previous research has shown that few cues are sufficient for users to identify with computer agents (Xu and Lombard 2017) and virtual service agents (Verhagen et al. 2014), we hypothesize that even verbal ADCs, which are not as directly and easily observable as nonverbal cues like embodiments, can influence perceived anthropomorphism and thus user compliance.

H1a: *Users are more likely to comply with a chatbot's request for service feedback when it exhibits more verbal anthropomorphic design cues.*

Previous research (e.g., Qiu and Benbasat 2009; Xu and Lombard 2017) investigated the concept of social presence and found that the construct reflects to some degree the emotional notions of anthropomorphism. These studies found that an increase in social presence usually improves desirable business-oriented variables in various contexts. For instance, social presence was found to significantly affect both bidding behavior and market outcomes (Rafaeli and Noy 2005) as well as purchase behavior in electronic markets (Zhang et al. 2012). Similarly, social presence is considered a critical construct to make customers perceive a technology as a social actor rather than a technological artefact. For example, Qiu and Benbasat (2009) revealed in their study how an anthropomorphic recommendation agent had a direct influence on social presence, which in turn increased trusting beliefs and ultimately the intention to use the recommendation agent. Thus, we argue that a chatbot with ADCs will increase consumers' perceptions of social presence, which in turn makes consumers more likely to comply to a request expressed by a chatbot.

H1b: *Social presence will mediate the effect of verbal anthropomorphic design cues on user compliance.*

Effect of the foot-in-the-door technique on user compliance

Humans are prone to psychological effects and compliance is a powerful behavioral response in many social exchange contexts to improve relationships. When there is a conflict between an individual's behavior and social norms, the potential threat of social exclusion often sways towards the latter, allowing for the emergence of social bias and, thus, nonrational decisions for the individual. For example, free samples often present effective marketing tools, as accepting a gift can function as a powerful, often nonrational commitment to return the favor at some point (Cialdini 2001).

Compliance techniques in computer-mediated contexts have proven successful in influencing user behavior in early stages of user journeys (Aggarwal et al. 2007). Providers use small initial requests and follow up with larger commitments to exploit users' self-perceptions (Bem 1972) and attempt to trigger consistent behavior when users decide whether to fulfill a larger, more obliging request, which the users would otherwise not. Thus, users act first and then form their beliefs and attitudes based on their actions, favoring the original cause and affecting future behavior towards that cause positively. The underlying rationale is users' intrinsic motivation to be consistent with attitudes and actions of past behavior (e.g., Aggarwal et al. 2007). Moreover, applying the social response theory (Nass et al. 1994), chatbots are unconsciously treated as social actors, so that users also feel a strong tendency to appear consistent in the eyes of other actors (i.e., the chatbot) (Cialdini 2001).

Consistent behavior after previous actions or statements has been found to be particularly prevalent when users' involvement regarding the request is low (Aggarwal et al. 2007) and when the individual cannot attribute the initial agreement to the commitment to external causes (e.g., Weiner 1985). In these situations, users are more likely to agree to actions and statements in support of a particular goal as consequences of making a mistake are not devastating. Moreover, in contexts without external causes the individual cannot blame another person or circumstances for the initial agreement, so that having a positive attitude toward the cause seems to be one of the few reasons for having complied with the initial request. Since we are interested in investigating a customer service situation that focusses on performing a simple, routine task with no obligations or large investments from the user's part and no external pressure to confirm, we expect that a user's involvement in the request is rather low such that he or she is more likely to be influenced by the foot-in-the-door technique.

H2: *Users are more likely to comply with a chatbot's request for service feedback when they agreed to and fulfilled a smaller request first (i.e., foot-in-the-door effect).*

Moderating effect of social presence on the effect of the foot-in-the-door technique

Besides self-consistency as a major motivation to behave consistently, research has demonstrated that consistency is also an important factor in social interactions. Highly consistent individuals are normally considered personally and intellectually strong (Cialdini and Garde 1987). While in the condition with few anthropomorphic design cues, the users majorly try to be consistent to serve personal needs, this may change once more social presence is felt. In the higher social presence condition, individuals may also perceive external, social reasons to comply and appear consistent in their behavior. This is consistent with prior research that studied moderating effects in

electronic markets (e.g., Zhang et al. 2012). Thus, we hypothesize that the foot-in-the-door effect will be larger when the user perceives more social presence.

H3: *Social presence will moderate the foot-in-the-door effect so that higher social presence will enhance the foot-in-the-door effect on user compliance.*

Research framework

As depicted in Fig. 1, our research framework examines the direct effects of *Anthropomorphic Design Cues (ADCs)* and the *Foot-in-the-Door (FITD)* technique on *User Compliance*. Moreover, we also examine the role of *Social Presence* in mediating the effect of *ADCs* on *User Compliance* as well as in moderating the effect of *FITD* on *User Compliance*.

Research methodology

Experimental design

We employed a 2 (ADCs: low vs. high) \times 2 (FITD: small request absent vs. present) between-subject, full-factorial design to conduct both relative and absolute treatment comparisons and to isolate individual and interactive effects on user compliance (e.g., Klumpe et al. 2019; Schneider et al. 2020; Zhang et al. 2012). The hypotheses were tested by means of a randomized online experiment in the context of a customer-service chatbot for online banking that provides customers with answers to frequently asked questions. We selected this context, as banking has been a common context in previous IS research on, for example, automation and recommendation systems (e.g., Kim et al. 2018). Moreover, the context will play an increasingly important role for future applications of CAs, as many service requests are based on routine tasks (e.g., checking account balances and blocking credit cards), which CAs promise to conveniently and cost-effectively solve in form of 24/7 service channels (e.g., Jung et al. 2018a, b).

In our online experiment, the chatbot was self-developed and replicated the design of many contemporary chat interfaces. The user could easily interact with the chatbot by typing in the message and pressing enter or by clicking on "Send". In contrast to former operationalizations of rule-based systems in experiments and practice, the IBM Watson Assistant cloud service provided us with the required AI-based functional capabilities for natural language processing, understanding as well as dialogue management (Shevat 2017; Watson 2017). As such, participants could freely enter their information in the chat interface, while the AI in the IBM cloud processed, understood and answered the user input in the same natural manner and with the same capabilities as other contemporary AI applications, like Amazon's Alexa or Apple's Siri – just in written form. At the same time, these functionalities represent

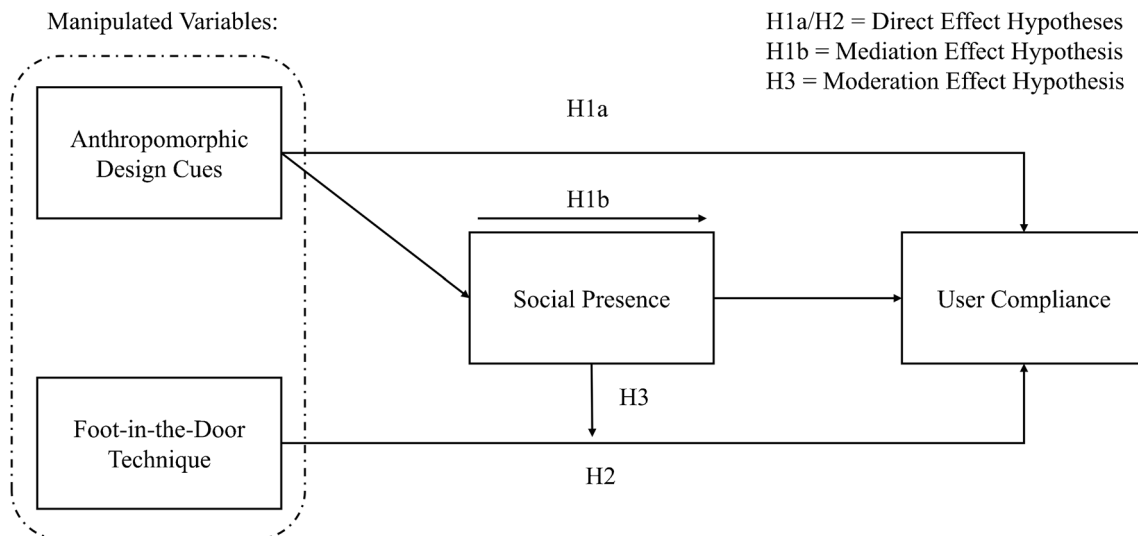


Fig. 1 Research framework

the necessary narrow AI that is especially important in customer self-service contexts to raise customer satisfaction (Gnewuch et al. 2017). For example, by means of the IBM Watson Assistant, it is possible to extract intentions and emotions from natural language in user statements (Ferrucci et al. 2010). Once the user input has been processed, an answer option is automatically chosen and displayed to the user.

Manipulation of independent variables

Regarding the ADCs manipulation, all participants received the same task and dialogue. Yet, consistent with prior research on anthropomorphism that employed verbal ADCs in chatbots (e.g., Araujo 2018), participants in the high ADCs manipulation experienced differences in form of presence or absence of the following characteristics, which are common in human-to-human interactions but have so far not been scientifically considered in chatbot interactions before: identity, small talk and empathy:

- (1) **Identity:** User perception is influenced by the way the chatbot articulates its messages. For example, previous research has demonstrated that when a CA used first-person singular pronouns and thus signaled an identity, the CA was positively associated with likeability (Pickard et al. 2014). Since the use of first-person singular pronouns is a characteristic unique to human beings, we argue that when a chatbot indicates an identity through using first-person singular pronouns and even a name, it not only increases its likeability but also anthropomorphic perceptions by the user.
- (2) **Smalltalk:** A relationship between individuals does not emerge immediately and requires time as well as effort from all involved parties. Smalltalk can be proactively used to develop a relationship and reduce the emotional distance between parties (Cassell and Bickmore 2003).

The speaker initially articulates a statement as well as signals to the listener to understand the statement. The listener can then respond by signaling that he or she understood the statement, so that the speaker can assume that the listener has understood the statement (Svennevig 2000). By means of small talk, the actors can develop a common ground for the conversation (Bickmore and Picard 2005; Cassell and Bickmore 2003). Consequently, a CA participating in small talk is expected to be perceived as more human and, thus, anthropomorphized.

- (3) **Empathy:** A good conversation is highly dependent on being able to address the statements of the counterpart appropriately. Empathy describes the process of noticing, comprehending, and adequately reacting to the emotional expressions of others. Affective empathy in this sense describes the capability to emotionally react to the emotion of the conversational counterpart (Lisetti et al. 2013). Advances in artificial intelligence have recently allowed computers to gain the ability to express empathy by analyzing and reacting to user expressions. For example, Lisetti et al. (2013) developed a module based on which a computer can analyze a picture of a human being, allocate the visualized emotional expression of the human being, and trigger an empathic response based on the expression and the question asked in the situation. Therefore, a CA displaying empathy is expected to be characterized as more life-like.

Since we are interested in the overall effect of verbal anthropomorphic design cues on user compliance, the study intended to examine the effects of different levels of anthropomorphic CAs when a good design is implemented and potential confounds are sufficiently controlled. Consequently, consistent with previous research on anthropomorphism (e.g., Adam et al. 2019; Qiu and Benbasat 2009, 2010), we

operationalized the high anthropomorphic design condition by conjointly employing the following design elements (see the Appendix for a detailed description of the dialogues in the different conditions):

- (1) The chatbot welcomed and said goodbye to the user. Greetings and farewells are considered adequate means to encourage social responses by users (e.g., Simmons et al. 2011)
- (2) The chatbot signaled a personality by introducing itself as “Alex”, a gender-neutral name as previous studies indicated that gender stereotypes also apply to computers (e.g., Nass et al. 1997).
- (3) The chatbot used first-person singular pronouns and thus signaled an identity, which has been presented to be positively associated with likeability in previous CA interactions (Pickard et al. 2014). For example, regarding the target request, the chatbot asked for feedback to improve itself rather than to improve the quality of the interaction in general.
- (4) The chatbot engaged in small talk by asking in the beginning of the interaction for the well-being of the user as well as whether the user interacted with a chatbot before. Small talk has been shown to be useful to develop a relationship and reduce the emotional and social distance between parties, making a CA appear more human-like (Cassell and Bickmore 2003).
- (5) The chatbot signaled empathy by processing the user’s answers to the questions about well-being and previous chatbot experience and, subsequently, providing responses that fit to the user’s input. This is in accordance with Lisetti et al. (2013) who argue that a CA displaying empathy is expected to be characterized as more life-like.

Consistent with previous studies on FITD (e.g., Burger 1999), we used the in the past highly successful continued-questions procedure in a same-requester/no-delay situation: Participants in the FITD condition are initially asked a small request by answering one single question to provide feedback about the perception of the chatbot interaction to increase the quality of the chatbot. As soon as the participant finished this task by providing a star-rating from 1 to 5, the same requester (i.e., the chatbot) immediately asks the target request, namely, whether the user is willing to fill out a questionnaire on the same topic that will take 5 to 7 min to complete. Participants in the FITD absent condition did not receive the small request and were only asked the target request.

Procedure

The participants were set in a customer service scenario in which they were supposed to ask a chatbot whether they could use their debit card abroad in the U.S. (see Appendix for the

detailed dialogue flows and instant messenger interface). The experimental procedure consisted of 6 steps (Fig. 2):

- (1) A short introduction of the experiment was presented to the participants including their instruction to introduce themselves to a chatbot and ask for the desired information.
- (2) In the anthropomorphism conditions with ADCs, the participants were welcomed by the chatbot. Moreover, the chatbot engaged in small talk by asking for the well-being of the user (i.e., “How are you?”) and whether the user has used a chatbot before (i.e., “Have you used a chatbot before?”). Dependent on the user’s answer and the chatbot’s AI-enabled natural language processing and understanding, the chatbot prompted a response and signaled comprehension and empathy.
- (3) The chatbot asked how it may help the user. The user then provided the question that he or she was instructed to: Whether the user can use his or her debit card in the U.S. If the user just asked a general usage of the debit card, the chatbot would ask in which country the user wants to use the debit card. The chatbot subsequently provided an answer to the question and asked if the user still has more questions. In case the user indicated that he or her had more, he or she would be recommended to visit the official website for more detailed explanations via phone through the service team. Otherwise, the chatbot would just thank the user for being used.
- (4) In the FITD condition, the user was asked to shortly provide feedback to the chatbot. The user then expresses his or her feedback by using a star-rating system.
- (5) The chatbot posed the target request (i.e., dependent variable) by asking whether the user is willing to answer some questions, which will take several minutes and help the chatbot to improve itself. The user then selected an option.
- (6) After the user’s selection, the chatbot instructed the user to wait until the next page is loaded. At this point, the conversation stopped, irrespective of the user’s choice. The participants then answered a post-experimental questionnaire about their chatbot experience and other questions (e.g., demographics).

Dependent variables, validation checks, and control variables

We measured *User Compliance* as a binary dependent variable, defined as a point estimator P based on

$$P(\text{user compliance}) = \frac{\sum_{k=1}^n x_k}{n}$$

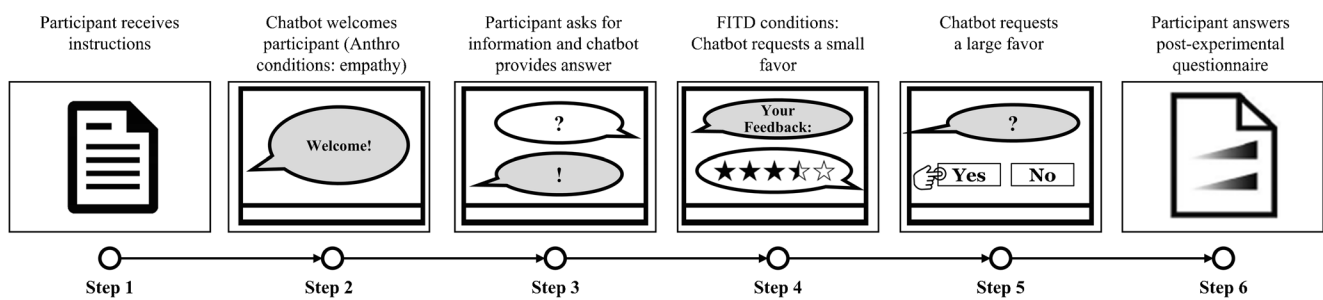


Fig. 2 Experimental procedure (As stated in the FITD hypothesis, we intend to investigate whether users are more likely to comply to a request when they agreed to and fulfilled a smaller request first. As

such, to avoid counteracting effects in the analysis, we will consider only participants who have agreed and fulfilled the initial small request (and thus remove participants who did not).)

where n denotes the total number of unique participants in the respective condition who finished the interaction (i.e., answering the chatbot's target request for voluntarily providing service feedback by selecting either "Yes" or "No"). x_k is a binary variable that equals 1 when the participant complied to the target request (i.e., selecting "Yes") and 0 when they denied the request (i.e., selecting "No").

Moreover, in addition to our dependent variable, we also tested demographic factors (age and gender) and the following other control variables that have been identified as relevant in extant literature. The items for *Social Presence* (SP) and *Trusting Disposition* were adapted from Gefen and Straub (2003), *Personal Innovativeness* from Agarwal and Prasad (1998), and *Product Involvement* from Zaichkowsky (1985). All items were measured on a 7-point Likert-type scale with anchors majorly ranging from *strongly disagree* to *strongly agree*. Moreover, we measured *Conversational Agents Usage* on a 5-point-scale ranging from never to daily. All scales exhibited satisfying levels of reliability ($\alpha > 0.7$) (Nunnally and Bernstein 1994). A confirmatory factor analysis also showed that all analyzed scales exhibited satisfying convergent validity requirements (Fornell and Larcker 1981) were met, since each scale's average variance extracted exceeded multiple squared correlations. Since the scales demonstrated sufficient internal consistency, we used the averages of all latent variables to form composite scores for subsequent statistical analysis. Lastly, two checks were included in the experiment. We used the checks to ascertain that our manipulations were noticeable and successful. Moreover, we assessed participants' *Perceived Degree of Realism* on a 7-point Likert-type scale with anchors ranging from *strongly disagree* to *strongly agree* (see Appendix (Tables 1,2,3,4 and 6) (Figures 3,4,5,6 and 7)

Analysis and results

Sample description

Participants were recruited via groups on Facebook as the social network provides many chatbots for customer service

purposes with its instant messengers. We incentivized participation by conducting a raffle of three Euro 20 vouchers for Amazon. Participation in the raffle was voluntary and inquired at the end of the survey. 308 participants started the experiment. Of those, we removed 32 (10%) participants who did not finish the experiment and 97 (31%) participants more, as they failed at least one of the checks (see Table 5 in Appendix). There were no noticeable technical issues in the interaction with the chatbot, which would have required us to remove further participants. Out of the remaining 179 participants, consistent with previous research on the FITD technique to avoid any counteracting effects (e.g., Snyder and Cunningham 1975), we removed 22 participants who declined the small request. Moreover, we excluded four participants who expressed that they found the experiment unrealistic. The final data set included 153 German participants with an average age of 31.58 years. Moreover, participants indicated that they had moderately high *Personal Innovativeness* regarding new technology ($\bar{x} = 4.95$, $\sigma = 1.29$), moderately high *Product Involvement* in bank product ($\bar{x} = 5.10$, $\sigma = 1.24$), and moderately low *Conversational Agent Usage* experience ($\bar{x} = 2.25$, $\sigma = 1.51$). Table 1 summarizes the descriptive statistics of the used data.

Table 1 Descriptive statistics of demographics, mediator, and controls

	Mean	SD
<i>Demographics</i>		
Age	31.58	9.46
Gender (Females)	54%	
<i>Mediator</i>		
Social Presence (SP)	3.23	1.54
<i>Controls</i>		
Trusting Disposition (TD)	4.98	1.17
Personal Innovativeness (PerInn)	4.95	1.29
Product Involvement (ProInv)	5.10	1.24
Conversational Agents Usage (CA Usage)	2.25	1.51

Notes: $N = 153$; SD = standard deviation

Table 2 Binary logistic regression on *User Compliance*

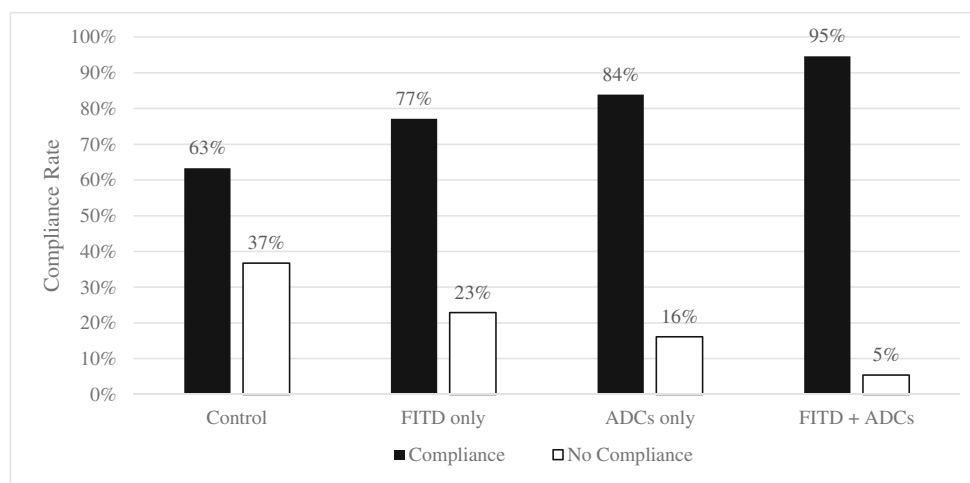
	Block 1			Block 2		
	Coefficient	S.E.	Exp(b)	Coefficient	S.E.	Exp(b)
<i>Intercept</i>						
Constant	0.898	1.759	2.454	−0.281	1.868	0.755
<i>Manipulations</i>						
ADCs				1.380**	0.493	3.975
FITD				0.916*	0.465	2.499
<i>Controls</i>						
Gender	−1.179*	0.462	0.308	−1.251*	0.504	0.286
Age	0.003	0.023	1.003	−0.008	0.024	1.009
TD	0.292	0.172	1.339	0.354	0.186	1.425
PerInn	0.210	0.177	1.234	0.159	0.186	1.173
ProInv	−0.068	0.177	0.935	−0.025	0.186	0.976
CA Usage	0.037	0.151	1.038	0.050	0.158	1.051
2 Log Likelihood	148.621			135.186		
Nagelkerke's R^2	0.106			0.227		
Omnibus Model χ^2	10.926			24.360**		

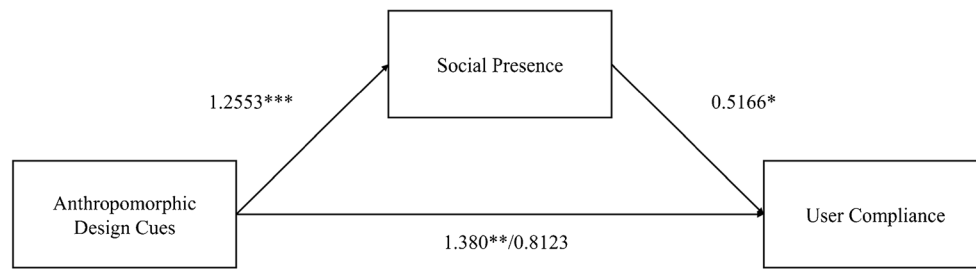
Notes: * $p < 0.05$; ** $p < 0.01$; $N = 153$

To check for external validity, we assessed the remaining participants' *Perceived Degree of Realism* of the experiment. *Perceived Degree of Realism* reached high levels ($\bar{x} = 5.58$, $\sigma = 1.11$), thus we concluded that the experiment was considered realistic. Lastly, we tested for possible common method bias by applying the Harman one-factor extraction test (Podsakoff et al. 2003). Using a principal component analysis for all items of the latent variables measured, we found two factors with eigenvalues greater than 1, accounting for 46.92% of the total variance. As the first factor accounted for only 16.61% of the total variance, less than 50% of the total variance, the Harman one-factor extraction test suggests that common method bias is not a major concern in this study (Figure 3).

Main effect analysis

To test the main effect hypotheses, we first performed a two-stage hierarchical binary regression analysis on the dependent variable *User Compliance* (see Table 2). We first entered all controls (Block 1), and then added the manipulations *ADCs* and *FITD* (Block 2). Both manipulations demonstrated a statistically significant direct effect on user compliance ($p < 0.05$). Participants in the FITD condition were more than twice as likely to agree to the target request ($b = 0.916$, $p < 0.05$, odds ratio = 2.499), while participants in the *ADCs* conditions were almost four times as likely to comply ($b = 1.380$, $p < 0.01$, odds ratio = 3.975). Therefore, our findings show that participants confronted with the FITD technique or an

Fig. 3 Results for the dependent variable *User Compliance*



Note: The first coefficient on a given path presents the direct effect without the mediator in the model. The second coefficient presents the direct effect when the mediator is inserted in the model; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Fig. 4 Mediation analysis

anthropomorphically designed chatbot are significantly more inclined to follow a request by the chatbot.

Mediation effect analysis

For our mediation hypothesis, we argued that ADCs would affect *User Compliance* via increased *Social Presence*. Thus, we hypothesized that in the presence of ADCs, social presence increases and, hence, the user is more likely to comply with a request. Therefore, in a mediation model using bootstrapping with 5000 sampled and 95% bias-corrected confidence interval, we analyzed the indirect effect of our ADCs on *User Compliance* and selection through *Social Presence*. We conducted the mediation test by applying the bootstrap mediation technique (Hayes 2017 model 4). We included both manipulations (i.e., ADCs and FITD) and all control variables in the analysis.

To analyze the process driving the effect of ADCs on *User Compliance*, we entered *Social Presence* as our potential mediator between ADCs and *User Compliance*. For our dependent variable *User Compliance*, the indirect effect of ADCs was statistically significant, thus *Social Presence* mediated the relationship between ADCs and *User Compliance*: indirect effect = 0.6485, standard error = 0.3252, 95% bias-corrected confidence interval (CI) = [0.1552, 1.2887]. Moreover, ADCs were positively related with *Social Presence* ($b = 1.2553$, $p < 0.01$), whereas the direct effect of our ADCs became insignificant ($b = 0.8123$, $p > 0.05$) after adding our mediator *Social Presence* to the model. Therefore, our results demonstrate that *Social Presence* significantly mediated the impact of ADCs on *User Compliance*: ADCs increased *Social Presence* and, thus, increased *User Compliance* (Figure 4).

Moderation effect analysis

We suggest in H3 that *Social Presence* will moderate the effect of FITD on *User Compliance*. To test the hypothesis, we conducted a bootstrap moderation analysis with 5000 samples and a 95% bias-corrected confidence interval (Hayes 2017, model 1). The results of our moderation analysis showed that the effect of FITD on *User Compliance* is not moderated by *Social*

Presence such that there was no significant interaction effect of *Social Presence* and FITD on *User Compliance* ($b = 0.2305$; $p > 0.1$). Consequently, our findings do not support H3.

Discussion

This study sheds light on how the use of ADCs (i.e., identity, small talk, and empathy) and the FITD, as a common compliance technique, affect user compliance with a request for service feedback in a chatbot interaction. Our results demonstrate that both anthropomorphism as well as the need to stay consistent have a distinct positive effect on the likelihood that users comply with the CA's request. These effects demonstrate that even though the interface between companies and consumers is gradually evolving to become technology dominant through technologies such as CAs (Larivière et al. 2017), humans tend to also attribute human-like characteristics, behaviors, and emotions to the nonhuman agents. These results thus indicate that companies implementing CAs can mitigate potential drawbacks of the lack of interpersonal interaction by evoking perceptions of social presence. This finding is further supported by the fact that social presence mediates the effect of ADCs on user compliance in our study. Thus, when CAs can meet such needs with more human-like qualities, users may generally be more willing (consciously or unconsciously) to conform with or adapt to the recommendations and requests given by the CAs. However, we did not find support that social presence moderates the effect of FITD on user compliance. This indicates that effectiveness of the FITD can be attributed to the user's desire for self-consistency and is not directly affected by the user's perception of social presence. This finding is particularly interesting because prior studies have suggested that the technique's effectiveness heavily depends on the kind of requester, which could indicate that the perceived social presence of the requester is equally important (Burger 1999). Overall, these findings have a number of theoretical contributions and practical implications that we discuss in the following.

Contributions

Our study offers two main contributions to research by providing a novel perspective on the nascent area of AI-based CAs in customer service contexts.

First, our findings provide further evidence for the CASA paradigm (Nass et al. 1994). Despite the fact that participants knew they were interacting with a CA rather than a human, they seem to have applied the same social rules. This study thus extends prior research that has been focused primarily on embodied CAs (Hess et al. 2009; Qiu and Benbasat 2009) by showing that the CASA paradigm extends to disembodied CAs that predominantly use verbal cues in their interactions with users. We show that these cues are effective in evoking social presence and user compliance without the precondition of nonverbal ADCs, such as embodiment (Holtgraves et al. 2007). With few exceptions (e.g., Araujo 2018), potentials to shape customer-provider relationships by means of disembodied CAs have remained largely unexplored.

A second core contribution of this research is that humans acknowledge CAs as a source of persuasive messages. This is not to say that CAs are more or less persuasive compared to humans, but rather that the degree to which humans comply with the artificial social agents depends on the techniques applied during the human-chatbot communication (Edwards et al. 2016). We argue that the techniques we applied are successful because they appeal to fundamental social needs of individuals even though users are aware of the fact that they are interacting with a CA. This finding is important because it potentially opens up a variety of avenues for research to apply strategies from interpersonal communication in this context.

Implications for practice

Besides the theoretical contributions, our research also has noteworthy practical implications for platform providers and online marketers, especially for those who consider employing AI-based CAs in customer self-service (e.g., Gnewuch et al. 2017). Our first recommendation is to disclose to customers that they are interacting with a non-human interlocutor. By showing that CAs can be the source of persuasive messages, we provide evidence that attempting to fool customers into believing they are interacting with a human might not be necessary nor desirable. Rather, the focus should be on employing strategies to achieve greater human likeness through anthropomorphism by indicating, for instance, identity, small-talk, and empathy, which we have shown to have a positive effect on user compliance. Prior research has also indicated that the attempt of a CAs to provide a human-like behavior is

impressive for most users, helping to lower user expectations and leading to more satisfactory interactions with CAs (Go and Sundar 2019).

Second, our results provide evidence that subtle changes in the dialog design through ADCs and the FITD technique can increase user compliance. Thus, when employing CAs, and chatbots in particular, providers should design dialogs as carefully as they design the user interface. Besides focusing on dialogs that are as close as possible to human-human communications, providers can employ and test a variety of other strategies and techniques that appeal to, for instance, the user's need to stay consistent such as in the FITD technique.

Limitations and future research

This paper provides several avenues for future research focusing on the design of anthropomorphic information agents and may help in improving the interaction of AI-based CAs through user compliance and feedback. For example, we demonstrate the importance of anthropomorphism and the perception of social presence to trigger social bias as well as the need to stay consistent to increase user compliance. Moreover, with the rise of AI and other technological advances, intelligent CAs will become even more important in the future and will further influence user experiences in, for example, decision-makings, onboarding journeys, and technology adoptions.

The conducted study is an initial empirical investigation into the realm of CAs in customer support contexts and, thus, needs to be understood with respect to some noteworthy limitations. Since the study was conducted in an experimental setting with a simplified version of an instant messaging application, future research needs to confirm and refine the results in a more realistic setting, such as in a field study. In particular, future studies can examine a number of context specific compliance requests (e.g., to operate a website or product in a specific way, to sign up for or purchase a specific service or product). Future research should also examine how to influence users who start the chatbot interaction but who just simply end the questionnaire after their inquiry has been solved, a common user behavior in service contexts that does not even allow for the emergence of survey requests. Furthermore, our sample consisted of only German participants, so that future researchers may want to test the investigated effects in other cultural contexts (e.g., Cialdini et al. 1999).

We revealed the effects only based on the operationalized manipulations, but other forms of verbal ADCs and FITD may be interesting for further investigations in the

digital context of AI-based CAs and customer self-service. For instance, other forms of anthropomorphic design cues (e.g., the number of presented agents) as well as other compliance techniques (e.g., reciprocity) may be fathomed, maybe even finding interacting observations between the manipulations. For instance, empathy may be investigated on a continuous level with several conditions rather than dichotomously, and the FITD technique may be examined with different kinds and effort levels of small and target requests. Researchers and service providers need to evaluate which small requests are optimal for the specific contexts and whether users actually fulfil the agreed large commitment.

Further, a longitudinal design approach can be used to measure the influence when individuals get more accustomed to chatbots over time, as nascent customer relationships might be harmed by a sole focus on customers self-service channels (Scherer et al. 2015). Researchers and practitioners should cautiously apply our results, as the phenomenon of chatbots is relatively new in practice. Chatbots have only recently sparked great interest among businesses and many more chatbots can be expected to be implemented in the near future. Users might get used to the presented cues and will respond differently over time, once they are acquainted to the new technology and the influences attached to it.

Conclusion

AI-based CAs have become increasingly popular in various settings and potentially offer a number of time- and cost-saving opportunities. However, many users still experience unsatisfactory encounters with chatbots (e.g., high failure rates), which might result in skepticism and resistance against the technology, potentially inhibiting that users comply with recommendations and requests made by the chatbot. In this study, we conducted an online experiment to show that both verbal anthropomorphic design cues and the foot-in-the-door technique increase user compliance with a chatbot's request for service feedback. Our study is thus an initial step towards better understanding how AI-based CAs may improve user compliance by leveraging the effects of anthropomorphism and the need to stay consistent in the context of electronic markets and customer service. Consequently, this piece of research extends prior knowledge of CAs as anthropomorphic information agents in customer-service. We hope that our study provides impetus for future research on compliance mechanisms in CAs and improving AI-based abilities in and beyond electronic markets and customer self-service contexts.

Funding Information Open Access funding provided by Projekt DEAL.

Appendix

Table 3 Measurement scales (translated)

Construct	Item
Social Presence Gefen and Straub (2003) ($\alpha = 0.943$, CR = 0.952, AVE = 0.798)	<ol style="list-style-type: none"> 1. There is a sense of human contact in the interaction with the chatbot 2. There is a sense of personalness in the interaction with the chatbot 3. There is a sense of sociability in the interaction with the chatbot 4. There is a sense of human warmth in the interaction with the chatbot 5. There is a sense of human sensitivity in the interaction with the chatbot
Product Involvement Zaichkowsky (1985) Personal Innovativeness Agarwal and Prasad (1998) ($\alpha = 0.868$, CR = 0.908, AVE = 0.711)	<ol style="list-style-type: none"> 1. I am interested in the functions of my bank products 1. If I heard about a new information technology, I would look for ways to experiment with it 2. Among my peers, I am usually the first to try out new information technologies. 3. In general, I am hesitant to try out new information technologies (reversed) 4. I like to experiment with new information technologies
Trusting Disposition Gefen and Straub (2003) ($\alpha = 0.919$, CR = 0.938, AVE = 0.791)	<ol style="list-style-type: none"> 1. I generally trust other people 2. I generally have faith in humanity 3. I feel that people are generally well meaning 4. I feel that people are generally trustworthy

Note: Measurement scales based on a 7-Point Likert-type scale with anchors ranging from “strongly disagree” to “strongly agree”.

Table 4 Self-developed measurement scale (translated)

Construct	Item
Conversational Agent Usage	How often do you use digital assistants, such as Siri (Apple), Google Assistant (Google), or Alexa (Amazon)?

Note: 1 = Never; 5 = Daily

Table 5 Checks (translated)

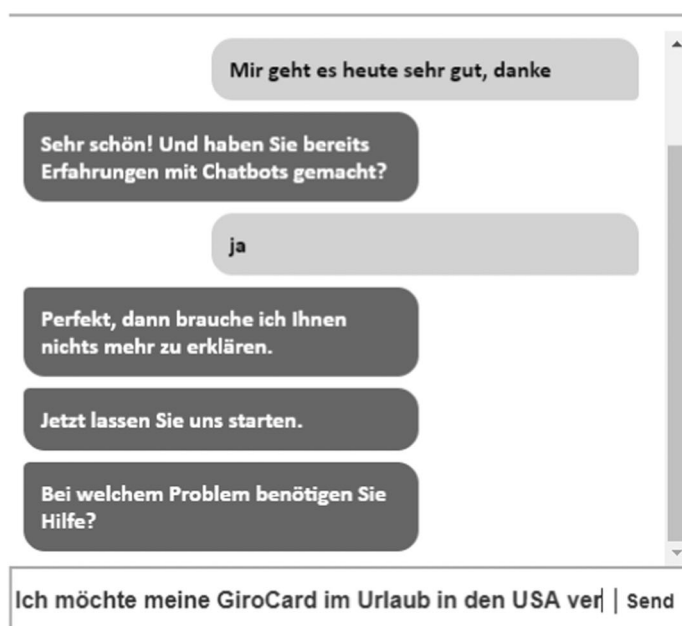
Check	Item
Anthropomorphic Design Cues	Did the digital assistant provide its name and/or asked how you are doing?
Foot-in-the-Door Technique	Did the digital assistant ask you twice for your evaluation: First to provide feedback based on a single question and then to take part in an interview?

Note: 1 = Yes; 2 = No

Table 6 Construct correlation matrix of selected core variables

Variable	1.	2.	3.
1. Social Presence	0.894		
2. Personal Innovativeness	0.056	0.843	
3. Trusting Disposition	0.211**	-0.083	0.889

Note: Bolded diagonal elements are the square root of AVE. These values exceed inter-construct correlations (off-diagonal elements) and, thus, demonstrate adequate discriminant validity; $N = 153$; ** $p < 0.01$

**Problembeschreibung:**

- Verwendung der GiroCard im Ausland
- Urlaubsziel: USA

Bitte schreiben sie dem Chatbot und drücken sie "Enter", um eine Antwort zu erhalten. Der Chatbot sagt Ihnen ebenfalls, wenn das Gespräch beendet ist.

Fig. 5 Exemplary and untranslated original screenshot of the chatbot interaction

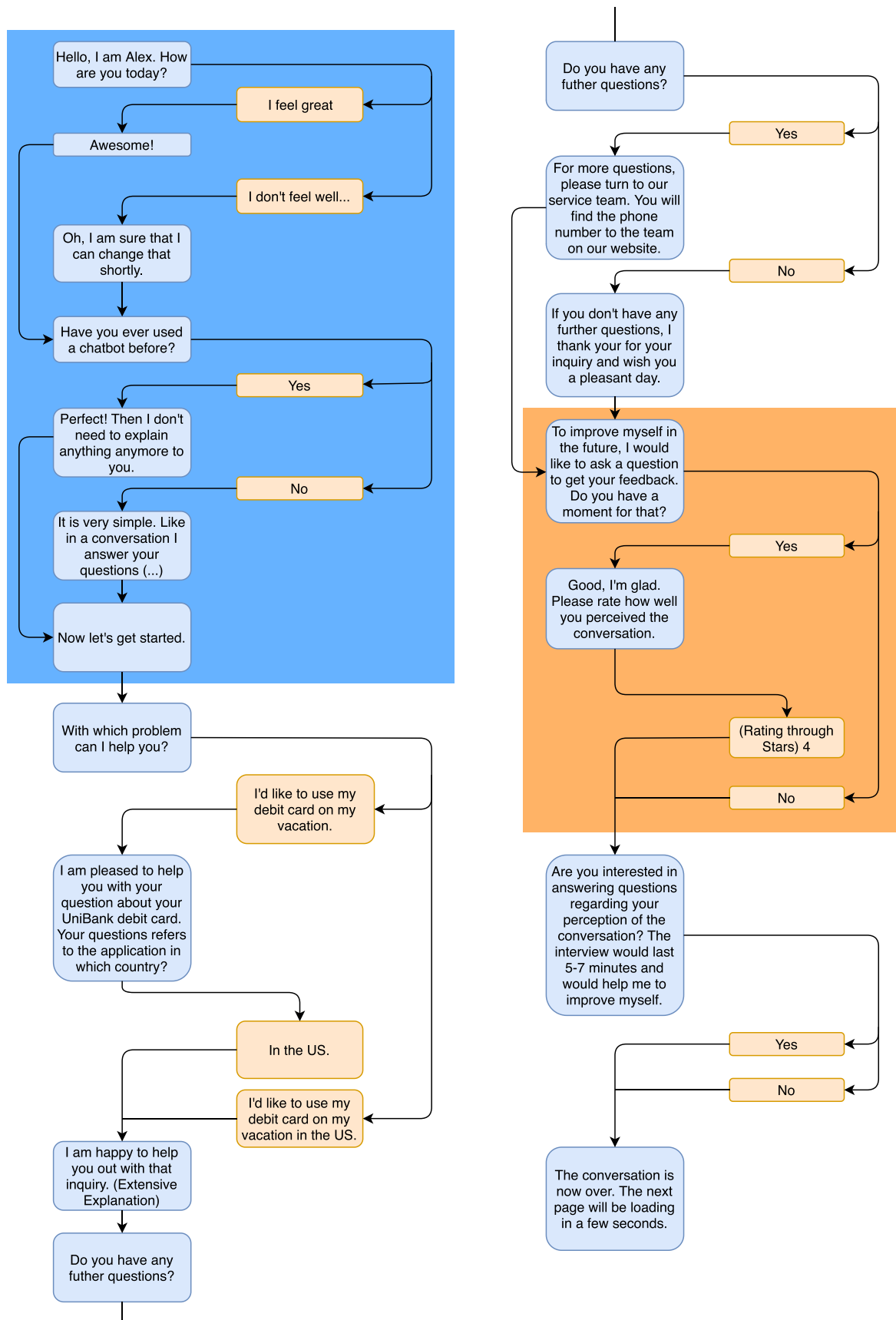


Fig. 6 Dialogue graph of an exemplary conversation of the ADCs present conditions

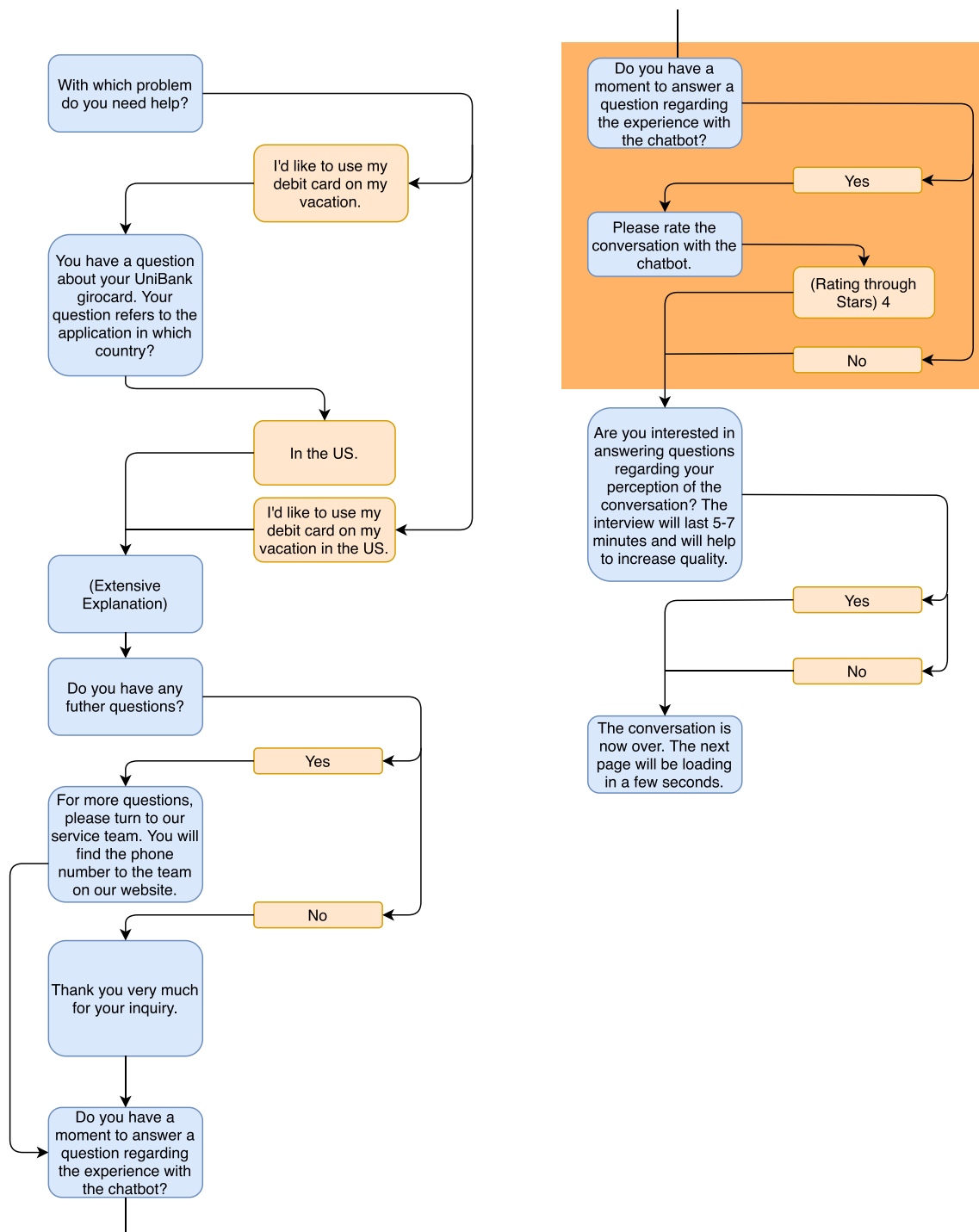


Fig. 7 Dialogue graph for an exemplary conversation of the ADCs absent conditions

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., Toutaoui, J., Pfeuffer, N., & Hinz, O. (2019). Investment decisions with robo-advisors: The role of anthropomorphism and personalized anchors in recommendations. In: Proceedings of the 27th European Conference on Information Systems (ECIS). Sweden: Stockholm & Uppsala.
- Agarwal, R., & Prasad, J. (1998). A conceptual and operational definition of personal innovativeness in the domain of information technology. *Information Systems Research*, 9(2), 204–215.
- Aggarwal, P., Vaidyanathan, R., & Rochford, L. (2007). The wretched refuse of a teeming shore? A critical examination of the quality of undergraduate marketing students. *Journal of Marketing Education*, 29(3), 223–233.
- Araujo, T. (2018). Living up to the chatbot hype: The influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions. *Computers in Human Behavior*, 85, 183–189.
- Astrid, M., Krämer, N. C., Gratch, J., & Kang, S.-H. (2010). “It doesn’t matter what you are!” explaining social effects of agents and avatars. *Computers in Human Behavior*, 26(6), 1641–1650.
- Bem, D. J. (1972). Self-perception theory. *Advances in Experimental Social Psychology*, 6, 1–62.
- Benlian, A., Klumpe, J., & Hinz, O. (2019). Mitigating the intrusive effects of smart home assistants by using anthropomorphic design features: A multimethod investigation. *Information Systems Journal*.
- Bertacchini, F., Bilotta, E., & Pantano, P. (2017). Shopping with a robotic companion. *Computers in Human Behavior*, 77, 382–395.
- Bickmore, T. W., & Picard, R. W. (2005). Establishing and maintaining long-term human-computer relationships. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 12(2), 293–327.
- Bowman, D., Heilman, C. M., & Seetharaman, P. (2004). Determinants of product-use compliance behavior. *Journal of Marketing Research*, 41(3), 324–338.
- Burger, J. M. (1986). Increasing compliance by improving the deal: The that’s-not-all technique. *Journal of Personality and Social Psychology*, 51(2), 277.
- Burger, J. M. (1999). The foot-in-the-door compliance procedure: A multiple-process analysis and review. *Personality and Social Psychology Review*, 3(4), 303–325.
- Burger, J. M., Soroka, S., Gonzago, K., Murphy, E., & Somervell, E. (2001). The effect of fleeting attraction on compliance to requests. *Personality and Social Psychology Bulletin*, 27(12), 1578–1586.
- Cassell, J., & Bickmore, T. (2003). Negotiated collusion: Modeling social language and its relationship effects in intelligent agents. *User Modeling and User-Adapted Interaction*, 13(1–2), 89–132.
- Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology*, 39(5), 752.
- Charlton, G. (2013). Consumers prefer live chat for customer service: stats. Retrieved from <https://econsultancy.com/consumers-prefer-live-chat-for-customer-service-stats/>
- Cialdini, R., & Garde, N. (1987). Influence (Vol. 3). A. Michel.
- Cialdini, R. B. (2001). Harnessing the science of persuasion. *Harvard Business Review*, 79(9), 72–81.
- Cialdini, R. B. (2009). *Influence: Science and practice* (Vol. 4). Boston: Pearson Education.
- Cialdini, R. B., & Goldstein, N. J. (2004). Social influence: Compliance and conformity. *Annual Review of Psychology*, 55, 591–621.
- Cialdini, R. B., & Trost, M. R. (1998). Social influence: Social norms, conformity and compliance. In S. F. D. T. Gilbert & G. Lindzey (Eds.), *The Handbook of Social Psychology* (pp. 151–192). Boston: McGraw-Hill.
- Cialdini, R. B., Vincent, J. E., Lewis, S. K., Catalan, J., Wheeler, D., & Darby, B. L. (1975). Reciprocal concessions procedure for inducing compliance: The door-in-the-face technique. *Journal of Personality and Social Psychology*, 31(2), 206.
- Cialdini, R. B., Wosinska, W., Barrett, D. W., Butner, J., & Gornik-Durose, M. (1999). Compliance with a request in two cultures: The differential influence of social proof and commitment/consistency on collectivists and individualists. *Personality and Social Psychology Bulletin*, 25(10), 1242–1253.
- Davis, B. P., & Knowles, E. S. (1999). A disrupt-then-reframe technique of social influence. *Journal of Personality and Social Psychology*, 76(2), 192.
- Derrick, D. C., Jenkins, J. L., & Nunamaker Jr., J. F. (2011). Design principles for special purpose, embodied, conversational intelligence with environmental sensors (SPECIES) agents. *AIS Transactions on Human-Computer Interaction*, 3(2), 62–81.
- Deutsch, M., & Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgment. *The Journal of Abnormal and Social Psychology*, 51(3), 629.
- Edwards, A., Edwards, C., Spence, P. R., Harris, C., & Gambino, A. (2016). Robots in the classroom: Differences in students’ perceptions of credibility and learning between “teacher as robot” and “robot as teacher”. *Computers in Human Behavior*, 65, 627–634.
- Elkins, A. C., Derrick, D. C., Burgoon, J. K., & Nunamaker Jr, J. F. (2012). *Predicting users’ perceived trust in Embodied Conversational Agents using vocal dynamics*. In: Proceedings of the 45th Hawaii International Conference on System Science (HICSS). Maui: IEEE.
- Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review*, 114(4), 864–866.
- Ericksen, M. K., & Sirgy, M. J. (1989). Achievement motivation and clothing behaviour: A self-image congruence analysis. *Journal of Social Behavior and Personality*, 4(4), 307–326.
- Etemad-Sajadi, R. (2016). The impact of online real-time interactivity on patronage intention: The use of avatars. *Computers in Human Behavior*, 61, 227–232.
- Eyssel, F., Hegel, F., Horstmann, G., & Wagner, C. (2010). Anthropomorphic inferences from emotional nonverbal cues: A case study. In *Proceedings of the 19th international symposium in robot and human interactive communication*. Viareggio: IT.
- Facebook. (2019). F8 2019: Making It Easier for Businesses to Connect with Customers on Messenger. Retrieved from <https://www.facebook.com/business/news/f8-2019-making-it-easier-for-businesses-to-connect-with-customers-on-messenger>
- Feine, J., Gnewuch, U., Morana, S., & Maedche, A. (2019). A taxonomy of social cues for conversational agents. *International Journal of Human-Computer Studies*, 132, 138–161.
- Ferrucci, D., Brown, E., Chu-Carroll, J., Fan, J., Gondek, D., Kalyanpur, A. A., et al. (2010). Building Watson: An overview of the DeepQA project. *AI Magazine*, 31(3), 59–79.

- Fogg, B. J., & Nass, C. (1997). Silicon sycophants: The effects of computers that flatter. *International Journal of Human-Computer Studies*, 46(5), 551–561.
- Fornell, C., & Larcker, D. F. (1981). Structural equation models with unobservable variables and measurement error: Algebra and statistics. *Journal of Marketing Research*, 18(3), 382–388.
- Freedman, J. L., & Fraser, S. C. (1966). Compliance without pressure: The foot-in-the-door technique. *Journal of Personality and Social Psychology*, 4(2), 195–202.
- Gefen, D., & Straub, D. (2003). Managing user trust in B2C e-services. *e-Service*, 2(2), 7–24.
- Gnewuch, U., Morana, S., & Maedche, A. (2017). Towards designing cooperative and social conversational agents for customer service. In: *Proceedings of the 38th International Conference on Information Systems (ICIS)*. Seoul: AISel.
- Go, E., & Sundar, S. S. (2019). Humanizing chatbots: The effects of visual, identity and conversational cues on humanness perceptions. *Computers in Human Behavior*, 97, 304–316.
- Gustafsson, A., Johnson, M. D., & Roos, I. (2005). The effects of customer satisfaction, relationship commitment dimensions, and triggers on customer retention. *Journal of Marketing*, 69(4), 210–218.
- Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (2nd ed.). New York: Guilford Publications.
- Hess, T. J., Fuller, M., & Campbell, D. E. (2009). Designing interfaces with social presence: Using vividness and extraversion to create social recommendation agents. *Journal of the Association for Information Systems*, 10(12), 1.
- Hill, J., Ford, W. R., & Farreras, I. G. (2015). Real conversations with artificial intelligence: A comparison between human–human online conversations and human–chatbot conversations. *Computers in Human Behavior*, 49, 245–250.
- Holtgraves, T., Ross, S. J., Weywadt, C., & Han, T. (2007). Perceiving artificial social agents. *Computers in Human Behavior*, 23(5), 2163–2174.
- Holzwarth, M., Janiszewski, C., & Neumann, M. M. (2006). The influence of avatars on online consumer shopping behavior. *Journal of Marketing*, 70(4), 19–36.
- Hopkins, B., & Silverman, A. (2016). The Top Emerging Technologies To Watch: 2017 To 2021. Retrieved from <https://www.forrester.com/report/The+Top+Emerging+Technologies+To+Watch+2017+To+2021/-/E-RES133144>
- Jin, S. A. A. (2009). The roles of modality richness and involvement in shopping behavior in 3D virtual stores. *Journal of Interactive Marketing*, 23(3), 234–246.
- Jin, S.-A. A., & Sung, Y. (2010). The roles of spokes-avatars' personalities in brand communication in 3D virtual environments. *Journal of Brand Management*, 17(5), 317–327.
- Jung, D., Dörner, V., Glaser, F., & Morana, S. (2018a). Robo-advisory. *Business & Information Systems Engineering*, 60(1), 81–86.
- Jung, D., Dörner, V., Weinhardt, C., & Pusmaz, H. (2018b). Designing a robo-advisor for risk-averse, low-budget consumers. *Electronic Markets*, 28(3), 367–380.
- Kim, Goh, J., & Jun, S. (2018). The use of voice input to induce human communication with banking chatbots. In: *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction*.
- Klumpe, J., Koch, O. F., & Benlian, A. (2019). How pull vs. push information delivery and social proof affect information disclosure in location based services. *Electronic Markets*. <https://doi.org/10.1007/s12525-018-0318-1>
- Knowles, E. S., & Linn, J. A. (2004). Approach-avoidance model of persuasion: Alpha and omega strategies for change. In *Resistance and persuasion* (pp. 117–148). Mahwah: Lawrence Erlbaum Associates Publishers.
- Kressmann, F., Sirgy, M. J., Herrmann, A., Huber, F., Huber, S., & Lee, D. J. (2006). Direct and indirect effects of self-image congruence on brand loyalty. *Journal of Business Research*, 59(9), 955–964.
- Larivière, B., Bowen, D., Andreassen, T. W., Kunz, W., Sirianni, N. J., Voss, C., et al. (2017). “Service encounter 2.0”: An investigation into the roles of technology, employees and customers. *Journal of Business Research*, 79, 238–246.
- Launchbury, J. (2018). A DARPA perspective on artificial intelligence. Retrieved from <https://www.darpa.mil/attachments/AIFull.pdf>
- Lisetti, C., Amini, R., Yasavur, U., & Rische, N. (2013). I can help you change! An empathic virtual agent delivers behavior change health interventions. *ACM Transactions on Management Information Systems (TMIS)*, 4(4), 19.
- Luger, E., & Sellen, A. (2016). Like having a really bad PA: The gulf between user expectation and experience of conversational agents. In: *Proceedings of the CHI Conference on Human Factors in Computing Systems*.
- Maedche, A., Legner, C., Benlian, A., Berger, B., Gimpel, H., Hess, T., et al. (2019). AI-based digital assistants. *Business & Information Systems Engineering*, 61(4), 535–544.
- Mero, J. (2018). The effects of two-way communication and chat service usage on consumer attitudes in the e-commerce retailing sector. *Electronic Markets*, 28(2), 205–217.
- Meuter, M. L., Bitner, M. J., Ostrom, A. L., & Brown, S. W. (2005). Choosing among alternative service delivery modes: An investigation of customer trial of self-service technologies. *Journal of Marketing*, 69(2), 61–83.
- Moon, Y., & Nass, C. (1996). How “real” are computer personalities? Psychological responses to personality types in human-computer interaction. *Communication Research*, 23(6), 651–674.
- Mori, M. (1970). The uncanny valley. *Energy*, 7(4), 33–35.
- Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues*, 56(1), 81–103.
- Nass, C., Moon, Y., & Carney, P. (1999). Are people polite to computers? Responses to computer-based interviewing systems. *Journal of Applied Social Psychology*, 29(5), 1093–1109.
- Nass, C., Moon, Y., Fogg, B. J., Reeves, B., & Dryer, C. (1995). Can computer personalities be human personalities? In: *Proceedings of the Conference on Human Factors in Computing Systems*.
- Nass, C., Moon, Y., & Green, N. (1997). Are machines gender neutral? Gender-stereotypic responses to computers with voices. *Journal of Applied Social Psychology*, 27(10), 864–876.
- Nass, C., Steuer, J., & Tauber, E. R. (1994). Computers are social actors. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*.
- Nunnally, J., & Bernstein, I. (1994). *Psychometric theory* (3rd ed.). New York: McGraw Hill Inc..
- Orlowski, A. (2017). Facebook scales back AI flagship after chatbots hit 70% f-AI-lure rate. Retrieved from https://www.theregister.co.uk/2017/02/22/facebook_ai_fail/
- Pavlikova, L., Schmid, B. F., Maass, W., & Müller, J. P. (2003). Editorial: Software agents. *Electronic Markets*, 13(1), 1–2.
- Pfeuffer, N., Adam, M., Toutaoui, J., Hinz, O., & Benlian, A. (2019a). *Mr. and Mrs. Conversational Agent - Gender stereotyping in judge-advisor systems and the role of egocentric bias*. Munich: International Conference on Information Systems (ICIS).
- Pfeuffer, N., Benlian, A., Gimpel, H., & Hinz, O. (2019b). Anthropomorphic information systems. *Business & Information Systems Engineering*, 61(4), 523–533.
- 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.
- Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of

- the literature and recommended remedies. *Journal of Applied Psychology*, 88(5), 879–903.
- 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.
- Qiu, L., & Benbasat, I. (2010). A study of demographic embodiments of product recommendation agents in electronic commerce. *International Journal of Human-Computer Studies*, 68(10), 669–688.
- Rafaeli, S., & Noy, A. (2005). Social presence: Influence on bidders in internet auctions. *Electronic Markets*, 15(2), 158–175.
- Raymond, J. (2001). No more shoppus interruptus. *American Demographics*, 23(5), 39–40.
- Reddy, T. (2017a). Chatbots for customer service will help businesses save \$8 billion per year. Retrieved from <https://www.ibm.com/blogs/watson/2017/05/chatbots-customer-service-will-help-businesses-save-8-billion-per-year/>
- Reddy, T. (2017b). How chatbots can help reduce customer service costs by 30%. Retrieved from <https://www.ibm.com/blogs/watson/2017/10/how-chatbots-reduce-customer-service-costs-by-30-percent/>
- Reeves, B., & Nass, C. (1996). *The media equation : How people treat computers, television, and new media like real people and places*. New York: Cambridge University Press.
- Scherer, A., Wunderlich, N. V., & Von Wangenheim, F. (2015). The value of self-service: Long-term effects of technology-based self-service usage on customer retention. *MIS Quarterly*, 39(1), 177–200.
- Schneider, D., Klumpe, J., Adam, M., & Benlian, A. (2020). Nudging users into digital service solutions. *Electronic Markets*, 1–19.
- Seymour, M., Yuan, L., Dennis, A., & Riemer, K. (2019). *Crossing the Uncanny Valley? Understanding Affinity, Trustworthiness, and Preference for More Realistic Virtual Humans in Immersive Environments*. In: *Proceedings of the Hawaii International Conference on System Sciences (HICSS)*.
- Shevat, A. (2017). *Designing bots: Creating conversational experiences*. UK: O'Reilly Media.
- Short, J., Williams, E., & Christie, B. (1976). *The social psychology of telecommunications*.
- Simmons, R., Makatchev, M., Kirby, R., Lee, M. K., Fanaswala, I., Browning, B., et al. (2011). Believable robot characters. *AI Magazine*, 32(4), 39–52.
- Simon, H. A. (1990). Invariants of human behavior. *Annual Review of Psychology*, 41(1), 1–20.
- Snyder, M., & Cunningham, M. R. (1975). To comply or not comply: Testing the self-perception explanation of the "foot-in-the-door" phenomenon. *Journal of Personality and Social Psychology*, 31(1), 64–67.
- Svennevig, J. (2000). *Getting acquainted in conversation: A study of initial interactions*. Philadelphia: John Benjamins Publishing.
- Techlabs, M. (2017). Can chatbots help reduce customer service costs by 30%? Retrieved from <https://chatbotsmagazine.com/how-with-the-help-of-chatbots-customer-service-costs-could-be-reduced-up-to-30-b9266a369945>
- Verhagen, T., Van Nes, J., Feldberg, F., & Van Dolen, W. (2014). Virtual customer service agents: Using social presence and personalization to shape online service encounters. *Journal of Computer-Mediated Communication*, 19(3), 529–545.
- Watson, H. J. (2017). Preparing for the cognitive generation of decision support. *MIS Quarterly Executive*, 16(3), 153–169.
- Weiner, B. (1985). "Spontaneous" causal thinking. *Psychological Bulletin*, 97(1), 74–84.
- 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.
- Wessel, M., Adam, M., & Benlian, A. (2019). The impact of sold-out early birds on option selection in reward-based crowdfunding. *Decision Support Systems*, 117, 48–61.
- Whatley, M. A., Webster, J. M., Smith, R. H., & Rhodes, A. (1999). The effect of a favor on public and private compliance: How internalized is the norm of reciprocity? *Basic and Applied Social Psychology*, 21(3), 251–259.
- Wilkinson, N., & Klaes, M. (2012). *An introduction to behavioral economics* (2nd ed.). New York: Palgrave Macmillan.
- Xu, K., & Lombard, M. (2017). Persuasive computing: Feeling peer pressure from multiple computer agents. *Computers in Human Behavior*, 74, 152–162.
- Zaichkowsky, J. L. (1985). Measuring the involvement construct. *Journal of Consumer Research*, 12(3), 341–352.
- Zhang, H., Lu, Y., Shi, X., Tang, Z., & Zhao, Z. (2012). Mood and social presence on consumer purchase behaviour in C2C E-commerce in Chinese culture. *Electronic Markets*, 22(3), 143–154.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.