5-15-2019

DESIGN FOR FAST REQUEST FULFILLMENT OR NATURAL INTERACTION? INSIGHTS FROM AN EXPERIMENT WITH A CONVERSATIONAL AGENT

Stephan Diederich
University of Goettingen, stephan.diederich@stud.uni-goettingen.de

Alfred Benedikt Brendel
University of Goettingen, abrende1@uni-goettingen.de

Sascha Lichtenberg
University of Goettingen, sascha.lichtenberg@uni-goettingen.de

Lutz Kolbe
University of Göttingen, lkolbe@uni-goettingen.de

DESIGN FOR FAST REQUEST FULFILLMENT OR NATURAL INTERACTION? INSIGHTS FROM AN EXPERIMENT WITH A CONVERSATIONAL AGENT

Research paper

Diederich, Stephan, University of Goettingen, Chair of Information Management, Goettingen, Germany, stephan.diederich@stud.uni-goettingen.de

Brendel, Alfred Benedikt, University of Goettingen, Chair of Information Management, Goettingen, Germany, abrende1@uni-goettingen.de

Lichtenberg, Sascha, University of Goettingen, Chair of Information Management, Goettingen, Germany, sascha.lichtenberg@uni-goettingen.de

Kolbe, Lutz M., University of Goettingen, Chair of Information Management, Goettingen, Germany, lkolbe@uni-goettingen.de

Abstract

Conversational agents continue to permeate our lives in different forms, such as virtual assistants on mobile devices or chatbots on websites and social media. The interaction with users through natural language offers various aspects for researchers to study as well as application domains for practitioners to explore. In particular their design represents an interesting phenomenon to investigate as humans show social responses to these agents and successful design remains a challenge in practice. Compared to digital human-to-human communication, text-based conversational agents can provide complementary, preset answer options with which users can conveniently and quickly respond in the interaction. However, their use might also decrease the perceived humanness and social presence of the agent as the user does not respond naturally by thinking of and formulating a reply. In this study, we conducted an experiment with N=80 participants in a customer service context to explore the impact of such elements on agent anthropomorphism and user satisfaction. The results show that their use reduces perceived humanness and social presence yet does not significantly increase service satisfaction. On the contrary, our findings indicate that preset answer options might even be detrimental to service satisfaction as they diminish the natural feel of human-CA interaction.

Keywords: Conversational agent, chatbot, dialogue design, online experiment.

1 Introduction

Conversational agents (CAs), i.e. software that interacts and exchanges information with its users through natural language, currently attract strong interest in both research and practice (Oracle, 2016; McTear, 2017). While the idea of natural language interaction with computers dates back several decades (Shawar and Atwell, 2007) and has been studied under various terms, such as product recommendation agents (Wang and Benbasat, 2005; Qiu and Benbasat, 2010), virtual agents (Baylor, 2009; Bickmore, Pfeifer and Jack, 2009), or dialogue systems (Mc Kevitt, Partridge and Wilks, 1999; Zadrozny et al., 2000), present CAs exhibit improved capabilities, in particular driven by advances in natural language processing and machine learning (Berg, 2014; Knijnenburg and Willemsen, 2016). With these improved capabilities researchers now explore the design of and interaction with CAs in a variety of domains (Diederich, Brendel and Kolbe, 2019a) ranging from customer service (Gnewuch, Morana and Maedche, 2017; Hu et al., 2018; Stock and Merkle, 2018) to rather specific application

areas, such as idea platforms (Tavanapour and Bittner, 2018) or data analytics (Matsushita, Maeda and Kato, 2004; Fast et al., 2017). Due to the natural interaction with CAs and variety of social cues provided by these agents, humans show social responses (Gong, 2008; Appel, Von Der Pütten, Krämer and Gratch, 2012) that need to be considered in the design process (Gnewuch et al., 2017; Seeger, Pfeiffer and Heinzl, 2017). As increasing the perceived humanness of CAs by adjusting its representation and communicative behavior is associated with other positive effects (Gong, 2008; Araujo, 2018), such as on perceived usefulness, different design elements have been studied to make CAs seem as human-like as possible. These aspects include for example the representation of a CA (Araujo, 2018; Seeger, Pfeiffer and Heinzl, 2018), the delay of responses to display pauses for thinking and typing (Gnewuch, Morana, Adam and Maedche, 2018), personalized communication with users (Holtgraves, Ross, Weywadt and Han, 2007), or designing a CA similar to its user (Al-Natour, Benbasat and Cenfetelli, 2006).

For CAs that interact with users via written text, often described as chatbots, different elements are available to design its representation, for example with a name or avatar (Hanus and Fox, 2015; Wünderlich and Paluch, 2017) and communicative behavior, for example the use of emoticons or self-references (Seeger et al., 2018). Thus, these elements influence the degree of perceived humanness and social presence. In particular, text-based CAs can suggest responses to its users in the form of preset answer options. These design elements allow the user to quickly and conveniently respond to a chatbot's statement without the need for manually formulating and typing a reply. While the use of such elements reduces the required effort for a user to reply in a conversation, it could also diminish the natural feeling in the interaction and the sense of human contact as those elements rather resemble traditional graphical user interfaces (Brandtzæg and Følstad, 2018), such as simple buttons that can be clicked or touched. In this study, we investigate the specific effect of these design elements to better understand how they impact user perceptions of and user satisfaction with text-based conversational agents with the following research question: *How do preset answer options of text-based CAs influence perceived humanness, social presence, and service satisfaction?*

We address this research question by means of an online experiment with a text-based CA in a customer service context, thus contributing to the growing knowledge base on (anthropomorphic) CA design. The remainder of this article is structured as follows: We continue by providing the research background for our work, in particular with regard to the design of CAs and underlying theories, and develop our hypotheses on the impact of preset answer options in human interaction with CAs. We then present our research design, an online experiment with a focus on perceived humanness, social presence, and service encounter satisfaction, as well as our results. Finally, we discuss the implications for the design of CAs, indicate limitations of our work, and outline opportunities for future research.

2 Theoretical Background and Related Work

The fundamental idea of CAs is to use natural language to converse on a common topic and exchange information like in a human-to-human interaction instead of traditional graphical interfaces (Berg, 2014; McTear, Callejas and Griol, 2016). CAs can exist in a variety of forms with regard to the communication mode (C. Lee, Jung, Kim and Lee, 2009), their embodiment (K. M. Lee, Jung, Kim and Kim, 2006), and application context (Nunamaker et al., 2011). The communication can take place both via spoken language, often referred to as virtual personal assistants, or through written text, often called chatbots or dialogue systems (Gnewuch et al., 2017; Morana et al., 2017). Conversational agents can be physically embodied (e.g. Featherman, Thatcher, Wright and Zimmer (2011) and Nunamaker et al. (2011)), have an interactive digital embodiment (e.g. Groom et al. (2009), Von Der Pütten, Krämer, Gratch and Kang (2010), and Seymour, Riemer and Kay (2017)) or a static digital embodiment (e.g. Hu et al. (2018) or A.-M. Seeger et al. (2018)). CAs can be used for general-purpose or in specific domains (Nunamaker et al., 2011), both in private and professional life. These domains include for example customer service (Gnewuch et al., 2017; Araujo, 2018; Hu et al., 2018; Stock and Merkle, 2018), marketing and sales (Wang and Benbasat, 2005; Qiu and Benbasat, 2009; Chattaraman, Kwon and Gilbert, 2012; Hanus and Fox, 2015; Vaccaro, Agarwalla, Shivakumar and

Kumar, 2018), human resources (Liao et al., 2018), data analytics (Matsushita et al., 2004; Fast et al., 2017), and collaborative team settings (Elson, Derrick and Ligon, 2018; Strohmann et al., 2018). In this study, we focus on text-based CAs with a static virtual embodiment in customer service as an exemplary domain.

The idea of human-computer interaction through natural language emerged already several decades ago when Joseph Weizenbaum presented the first CA, called ELIZA (Weizenbaum, 1966). Since then, a variety of CAs emerged and often disappeared (Ben Mimoun, Poncin and Garnier, 2012) due to the reliance on simple pattern matching approaches and the provision of a limited set of responses (Knijnenburg and Willemsen, 2016). However, with significant developments in the fields of natural language processing and machine learning, present CAs exhibit strongly increased capabilities (Berg, 2014). The example of Google Assistant, who made an appointment with a hairdresser in a spoken conversation at the last I/O developer conference (Welch, 2018) underlines this potential of today's CAs. With these capabilities, CAs gained momentum in research and practice alike (Oracle, 2016; Saffarizadeh, Boodraj and Alashoor, 2017; Wünderlich and Paluch, 2017). Companies from different industries, such as Amtrak (NextIT, 2018), Starbucks (Perez, 2016), H&M (Morana et al., 2017) or KLM (Vogel-Meijer, 2018), introduce CAs to innovate and automate tasks in customer service or sales and begin to realize its potential. For example, the American railroad company Amtrak introduced "Julie" to provide customers with an easy-to-use possibility to find and book tickets online, which answers more than 5 million questions annually and generated more than 1 million USD in savings for customer service within a single year (NextIT, 2018).

2.1 Design of Conversational Agents

Despite the capabilities of modern CAs, successful design represents a major challenge, in particular due to high user expectations that often cannot be fulfilled by the system (Luger and Sellen, 2016). As CAs provide different types of social cues (Von Der Pütten et al., 2010; Seeger et al., 2018), such as in the agent's human-like representation or in the language itself, users apply social rules and form expectations to them as posited in social response theory (Nass and Moon, 2000). These social responses to CAs offer potential and are associated with risks at the same time: On the one hand, we can transfer existing theories and concepts from human-to-human communication to understand and design digital communication with CAs. For example, Danielescu and Christian (2018) and Gnewuch et al. (2017) draw on Grice's Maxims as universal mechanisms for cooperative conversations (Grice, 1975), Saffarizadeh et al. (2017) use social penetration theory (Altman, Taylor and Derlega, 1997) to study self-disclosure in human-CA communication, and Burgoon et al. (2016) apply expectancy violations theory (Burgoon and Jones, 1976) in the context of CAs. On the other hand, social responses to CAs form high user expectations towards these systems, which are often not in line with their capabilities (Luger and Sellen, 2016; Brandtzæg and Følstad, 2018). A case study by Ben Mimoun et al. (2012) on the disappearance of agents on 80 French commercial websites indicated that many CAs fail due to appearance inadequacy as well as lack of interactivity and intelligence. Ben Mimoun et al. (2012, p. 607) describe that "[…] when implementing the agent, the company might fail to define and present its limits to consumers. Users thus cannot understand limits of the agent's capacities and become frustrated with its inability to answer their questions".

Researchers have investigated a variety of social cues to account for the social responses and resulting user expectations. Such cues can be categorized along three dimensions of anthropomorphic CA design: human identity, non-verbal cues, and verbal cues (Seeger et al., 2018). These cues range from the representation of the CA (Cowell and Stanney, 2005; Hanus and Fox, 2015; Shamekhi et al., 2018) in the category of human identity over non-verbal cues like smiling (Krämer, Kopp, Becker-Asano and Sommer, 2013; Biancardi, Cafaro and Pelachaud, 2017; Harjunen et al., 2018) or showing empathetic demeanor (De Rosis et al., 2003; Beale and Creed, 2009; Leite et al., 2013; Shi et al., 2018) to verbal cues, such as self-references (Seeger et al., 2018). For their design, scholars use the technical possibilities of CAs to deliberately mimic human behavior, such as dynamically delayed response times for thinking and typing a response (Gnewuch et al., 2018) or automatic sentiment detection of user state-

ments to provide adequate and empathetic answers (Bertacchini, Bilotta and Pantano, 2017; Hu et al., 2018). However, using the available technical possibilities to successfully design a CA that is able to fulfill user expectations in terms of adequate representation and communication behavior in a specific domain remains a challenging endeavor (Brandtzæg and Følstad, 2018).

2.2 Social Response Theory and Social Presence in the Context of CAs

The social cues provided by CAs, both verbal and non-verbal as well as with regard to the agent representation, elicit social responses in humans and can convey a sense of human contact in the interaction. These phenomena are explained by social response theory and the concept of social presence.

Social response theory (Reeves and Nass, 1996; Nass and Moon, 2000) and the paradigm of Computers As Social Actors (Nass, Steuer and Tauber, 1994) is the main theoretical foundation underlying CA research, particularly the work by scholars from Information Systems and Human-Computer Interaction disciplines (Gnewuch et al., 2018). Through a review of three sets of experimental studies, Nass and Moon (2000) posit that individuals mindlessly apply social rules and expectations to computers once they receive cues associated with human traits or behavior. The researchers showed that humans overuse human social categories, such as gender stereotypes, show and expect social behaviors, such as reciprocity, and exhibit social responses with respect to the "personality" of a computer program. Furthermore, Nass and Moon (2000, p. 97) hypothesize that "the more computers present characteristics that are associated with humans, the more likely they are to elicit social behavior".

Social presence represents another theoretical concept that is often used in CA research. Social presence represents "the sense of human contact embodied in a medium" (Gefen and Straub, 1997, p. 390), which is often related to the richness of the communication. For example, a video conference makes the conversation partner more salient than a chat message (Schuetzler, Grimes, Giboney and Buckman, 2014). Studies on social presence have shown that even small adjustments, such as adding personalized greetings (Gefen and Straub, 2003) or human images (Cyr, Head, Larios and Pan, 2009) to websites, positively influence perceived social presence for an artefact. In turn, social presence has been shown to improve trust (Gefen and Straub, 2004), self-efficacy beliefs (Baylor, 2009), and enjoyment as well as perceived usefulness (Hassanein and Head, 2007).

As CAs can provide rich characteristics that are associated with humans and can use a variety of social cues, in particular through natural language itself, they trigger considerable social responses by users (Gong, 2008) and have the potential to convey a substantial feeling of human contact in human-computer interaction (Nunamaker et al., 2011; Appel et al., 2012; Chattaraman et al., 2012).

3 Hypotheses

The main distinctive characteristic of CAs is the use of natural language in the interaction with users (Berg, 2014; McTear et al., 2016), which allows to intuitively interact with such programs (Shawar and Atwell, 2007). CAs that provide relevant information in natural language and a responsive manner, e.g. with follow-up questions or reciprocity, are shown to contribute to different positive aspects, such as user engagement (Schuetzler, Grimes and Giboney, 2018) or satisfaction (Chattaraman et al., 2012). When interacting with users through natural language, text-based CAs, or chatbots, can use elements known from digital human-to-human communication, such as emoticons (Park and Sundar, 2015) to indicate emotions or images (Yu, Lin and Liao, 2017) to increase the richness of the conversation and elicit associated positive effects (Dennis, Fuller and Valacich, 2014). However, designers of CAs have an additional element at their disposal: To provide a set of predefined answer options in a conversation (Brandtzæg and Følstad, 2018). These allow users to swiftly and easily respond in a conversation. Both approaches can be found in present CAs. For example, the CA by Starbucks (Figure 1) suggests different replies for placing orders, adding items, or customizing drinks (Bishop, 2016). In contrast to the design of Starbucks, the agent by Staples (Topbots, 2016) exclusively relies on manual formulation and typing of replies like in a human-to-human text-based communication.

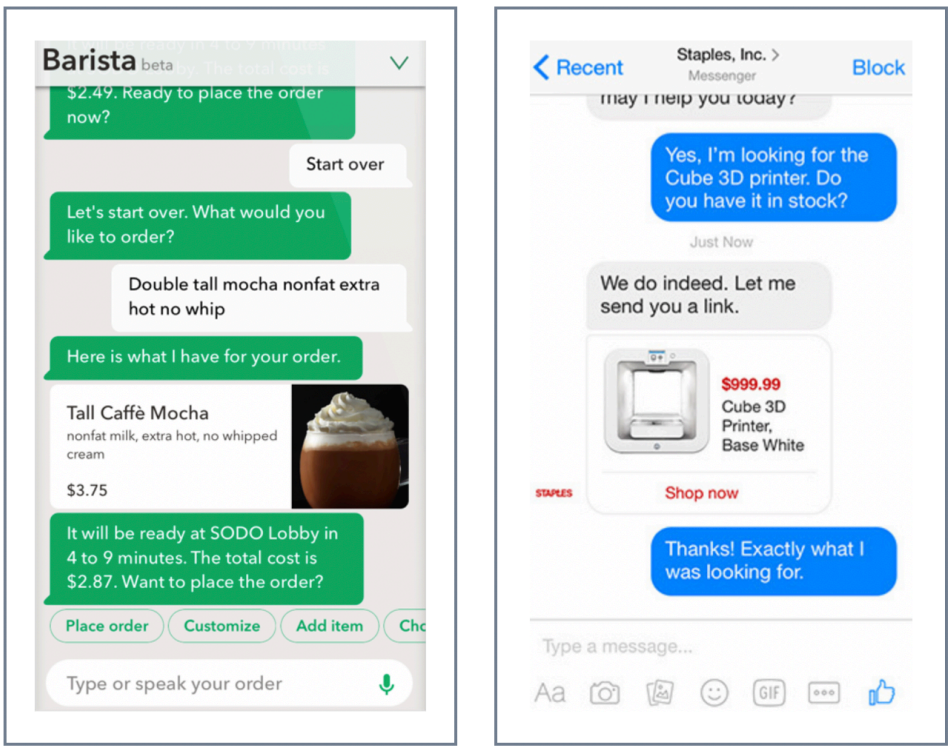

Figure 1. Exemplary CAs with (left, source: Bishop (2016)) and without (right, source: Topbots, (2016)) preset answer options

While such elements can have positive effects, for example that users do not need to manually type their response and thus can reach their goals in the conversation faster, or make it easier to guide conversations towards a specific goal for the designer, they can also negatively impact the interaction by reducing the natural feeling and sense of human contact in the conversation (Brandtzæg and Følstad, 2018). While being aware of the potentially divergent impact of such elements is certainly beneficial for designers of CAs, it has not been investigated in the past to the best of our knowledge. To address this gap, we formulate three hypotheses.

First, based on social response theory (Reeves and Nass, 1996; Nass and Moon, 2000), we hypothesize that the use of preset answer options in a conversation is detrimental to the perceived humanness of a chatbot as it represents a cue known from traditional, graphical user interfaces. It diminishes the natural feeling in a conversation, well-known from human-to-human communication, concerning the thinking of and formulating (e.g. typing) a response to the previous statement in the dialogue. Thus, we formulate our first hypothesis as follows:

*H1: A chatbot that provides preset answer options yields a **lower level of perceived humanness** than a chatbot that exclusively relies on manually typed natural language responses by the user.*

Similarly, the perceived social presence, defined as the sense of human contact in a mediated interaction (Short, Williams and Christie, 1976), depends on the availability of cues known from human-to-human communication, such as seeing the conversation partner in an image (Kahai and Cooper, 2003) the use of emotionally charged descriptions (Hassanein and Head, 2007). As preset buttons for answer options are similar to the elements used in traditional graphical user interfaces, we presume that their use in human-CA interaction reduces the sense of human contact in a conversation as users associate such elements with graphical instead of conversational interfaces:

*H2: A chatbot that provides preset answer options yields a **lower level of perceived social presence** than a chatbot that exclusively relies on manually typed natural language responses by the user.*

Furthermore, preset answer options have the potential to increase user satisfaction with the CA and service provision by providing a faster and more convenient way to achieve a goal in a conversation. As most users are acquainted with the use of such suggestions from text messaging on mobile phones, we can expect users to intuitively use them in a conversation with a CA. Such elements can reduce the time needed for formulating and typing a reply, thus decreasing the overall time needed for solving a customer request. As customers require as few waiting time as possible for the fulfillment of a service request (Taylor, 1994), in particular in a technology-mediated service setting (Elmorshidy, 2013; Larivière et al., 2017), the use of preset answer options represents a promising approach to immediately attend to customer needs and thus increase a user's satisfaction with the service. Hence, we formulate our third hypothesis as follows:

H3: A chatbot that provides preset answer options leads to a ***higher user satisfaction*** *than a chatbot that exclusively relies on manually typed natural language responses by the user.*

4 Research Design and Methods

In order to test our hypotheses, we conducted an online experiment with a between-subjects design to avoid carryover effects (Boudreau, Gefen and Straub, 2001). For the experiment, we chose a customer service context as it is one of the main application domains of conversational agents in enterprises (Verhagen, van Nes, Feldberg and van Dolen, 2014; Xu et al., 2017; Stock and Merkle, 2018) and is intuitively understandable for the participants. The experiment took place in November 2018 over a span of several days.

4.1 Data Collection Procedure and Sample

The participants received a briefing document, in which we explained the context (chatbots in a company context at the example of customer service) and structure of the experiment (interaction with the bot with subsequent survey) as well as described the participants' tasks. Every participant received the same document to make sure that the participants have exactly the same information for the experiment (Dennis and Valacich, 2001). The document further contained a link to the chatbot, which randomly assigned the participant to the control or treatment configuration. Each participant was supposed to complete three tasks:

1. Find out the order status for order number 480219
2. Change the delivery address for the order to Any Street 45, Any City, Any State 12345
3. Request e-mail confirmation of the changes to johndoe@mail.com

After completing the third task, the chatbot provided a link to the online survey. Overall, the participation in the experiment took around 10 minutes per participant. Our study has a sample size of N = 80 participants ranging from 18 to 47 years old (mean: 33 years) and a share of 50.6% female and 49.4% male persons. Similar to the CA experiment by Seeger, Pfeiffer and Heinzl (2018), we recruited native speakers through the panel company *clickworker* (Clickworker, 2018) and participants received a small compensation for their effort.

4.2 Control and Treatment Configurations

To conduct the experiment, we implemented two instances of the same chatbot using Dialogflow by Google. Both instances were prepared with the same training phrases and modeled dialogue corresponding to the experimental tasks. The chatbot could understand and process different wording of the same statements and was able to extract, validate and repeat different parameters, such as the order number or e-mail address to indicate that it understood a user's request. We further created a straightforward web interface (Figure 2) to provide direct access to the chatbot and minimize distraction during the experiment as alternative platforms for deployment, such as Facebook, could provide deflective information during the interaction.

In the implementation, we considered three types of anthropomorphic design dimensions (human identity, verbal cues, and non-verbal cues (Seeger et al., 2018)) to establish a common baseline for both instances with regard to perceived anthropomorphism and associated positive effects (e.g. on social presence (Araujo, 2018) and social judgment (Gong, 2008). With regard to the human identity, the chatbot received a name, Linda, and a static comic avatar which represented a female customer service agent. Regarding verbal cues, the chatbot used self-references, e.g. "I will do my best to assist you" and emotional expressions, e.g. "Great, I changed the delivery address". We further equipped the chatbot with non-verbal cues, i.e. the use of emojis (e.g. for greeting a user) and dynamic response delays (Gnewuch et al., 2018) with a visual indicator that the chatbot is typing. In addition, we let the chatbot self-disclose itself as a software program ("Even though I am not human, I will do my best to assist you") and use the extracted parameters (e.g. order number) in its response to contribute to the perceived responsiveness (Al-Natour, Benbasat and Cenfetelli, 2009; Saffarizadeh et al., 2017).

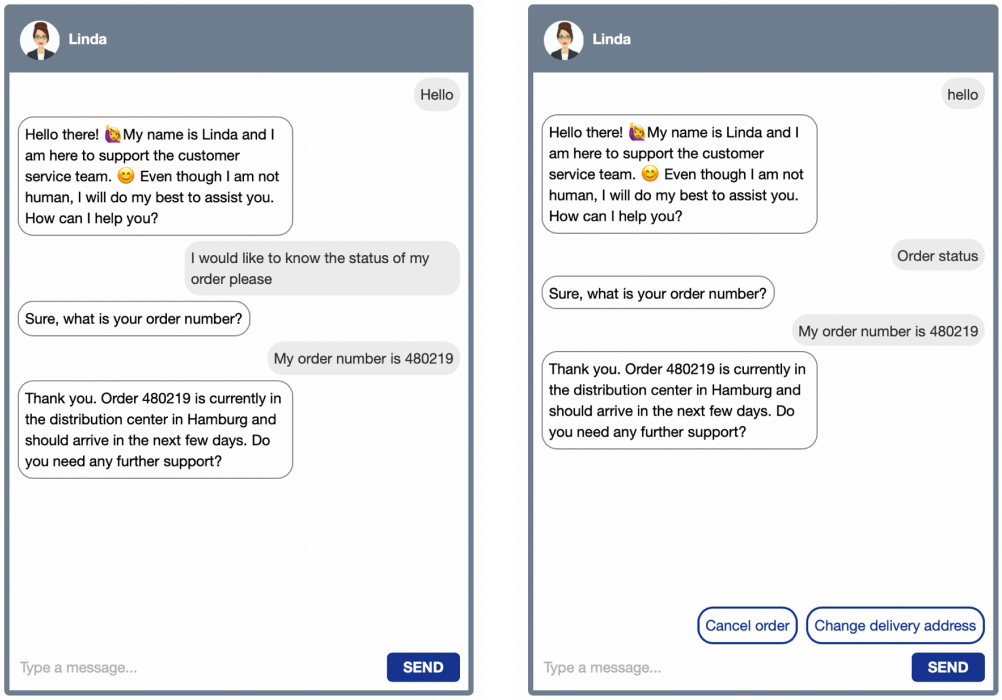

Figure 2. Web interface with control (left) and treatment (right) configurations (translated to English)

Overall, both chatbot instances were identical except for the provision of preset answer options in the treatment configuration. The treatment chatbot showed pre-defined buttons that provided the users with different options for their response (see Figure 2, interface on the right). For example, the chatbot provided the options "Product information", "Order status", and "Complaint" after greeting the user. These buttons could be clicked by the user to directly send a reply to the chatbot and thus avoid the need of typing a (potentially) whole sentence as a response. In contrast, participants in the control group had to manually type their request. For all of the three tasks (retrieve order status, change delivery address, and request e-mail confirmation) the treatment configuration provided a set of buttons of which one corresponded to the user's respective task.

4.3 Measures

Following the interaction with the chatbot, the participants completed a survey in which we measured how human-like and socially present the users perceived their conversation partner as well as how satisfied they were with the service encounter. For our survey design, we adapted measurement instruments established in previous studies and applied in the context of CAs (Gnewuch et al., 2018), namely perceived humanness (Holtgraves et al., 2007), social presence (Gefen and Straub, 1997), and ser-

vice encounter satisfaction (Verhagen et al., 2014). We measured perceived humanness on a 9-point semantic differential scale and social presence as well as service encounter satisfaction on a 7-point Likert scale. We further added two control questions by inverting two items, skill for perceived humanness and personalness for social presence. Table 1 summarizes the three constructs including their composite reliability (CR) and average variance (AVE) extracted, as well as the items and measured factor loadings.

Constructs and items	**Factor loading**
Perceived humanness (CR = .917, AVE = .688)	
extremely inhuman-like – extremely human-like	.811
extremely unskilled – extremely skilled	.873
extremely unthoughtful – extremely thoughtful	.810
extremely impolite – extremely polite	*dropped (.563)*
extremely unresponsive – extremely responsive	.777
extremely unengaging – extremely engaging	.872
Social presence (CR = .908, AVE = .767)	
I felt a sense of human contact with the chatbot	.903
I felt a sense of personalness with the chatbot	.892
I felt a sense of human warmth with the chatbot	*dropped (.596)*
I felt sense of human sensitivity with the chatbot	.831
Service satisfaction (CR = .913, AVE = .777)	
How satisfied are you with the chatbot's advice?	.911
…the way the chatbot treated you?	.851
…the overall interaction with the chatbot?	.881

CR = Composite Reliability, AVE = Average Variance Extracted

Table 1. Items and measures and factor loadings

Two items, politeness to measure perceived humanness and perceived human warmth for social presence, were dropped due to factor loadings below .60 (Gefen and Straub, 2005). All constructs exhibited sufficient CR (> .80) and AVE (> .50) with respect to the levels proposed by Urbach and Ahlemann (2010). In the final part of the survey, we collected complementary data with regard to demographic information (age and gender) and open-ended feedback on the interaction with the chatbot.

5 Results

We analyzed the data from the follow-up survey by means of descriptive statistics and used a t-test to test our three hypotheses concerning the impact of preset answer options on perceived humanness (H1), social presence (H2), and service satisfaction (H3). The dataset contained three invalid responses for the control group that were identified using the control questions presented in the previous section, where participants consistently gave the same answers also for inverted questions. Thus, the sample size was reduced from 80 to 77 participants. We analyzed conversation logs as a manipulation check to understand whether participants in the treatment group used the provided preset answer options and found that in 78% of all possible cases participants actually clicked on the quick reply buttons instead

of manually typing a reply. In addition, we checked for moderating effects of age, agender, and prior chatbot experience yet did not find significant effects. Table 2 summarizes the results of our analyses, which were carried out using SPSS version 25. The homogeneity of variance was successfully proven by the Levene tests for all three constructs (perceived humanness $F(75) = 3.25$, $p = 0.075$, social presence $F(75) = 0.49$, $p = 0.487$, service satisfaction $F(75) = 2.11$, $p = 0.151$). We then tested for a significant difference between the control (without preset answer options, necessitating manual typing) and treatment conditions (with preset answer options for every of the three experimental tasks).

With regard to perceived humanness, we found a significant difference between the control and treatment conditions. Thus, our data indicates that the use of preset answer options does indeed reduce the perceived humanness of a CA in a conversation, providing support for our first hypothesis.

		Condition		*t*-value (df = 75)	*p*-value
		Treatment (n = 44)	**Control** (n = 33)		
Perceived humanness	Mean SD SE	6.65 1.68 0.25	7.47 1.30 0.23	-2.35	0.022
Social presence	Mean SD SE	3.05 1.09 0.16	3.52 0.92 0.16	-2.00	0.049
Service satisfaction	Mean SD SE	4.21 0.81 0.12	4.53 0.75 0.13	-1.73	0.087

SD = Standard deviation, SE = Standard error

Table 2. Descriptive statistics and t-test results

Concerning social presence, our results reveal that there is a significant difference between the control and treatment conditions. The data from our experiment thus supports our second hypothesis that the use of preset answer options significantly reduces the sense of human contact in CA interaction.

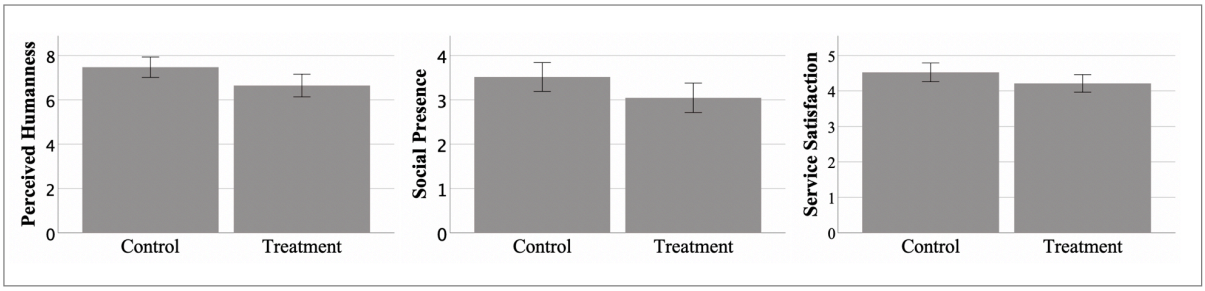

Figure 3. Differences between conditions for the three constructs

Finally, the results do not indicate a significant difference for service satisfaction between the control and treatment conditions. Consequently, we do not find support for the third hypothesis in our data that the use of quick reply elements increases a user's satisfaction with the service encounter due to faster and more convenient interaction. The results for perceived humanness, social presence and service satisfaction are visualized in Figure 3 for both conditions with error bars indicating the 95% confidence interval. Table 3 summarizes the results from the experiment for our three hypotheses.

Hypothesis	Result
H1: A chatbot that provides preset answer options yields **a lower level of perceived humanness** than a chatbot that exclusively relies on manually typed natural language responses by the user.	Supported
H2: A chatbot that provides preset answer options yields **a lower level of perceived social presence** than a chatbot that exclusively relies on manually typed natural language responses by the user.	Supported
H3: A chatbot that provides preset answer options leads to **a higher user satisfaction** than a chatbot that exclusively relies on manually typed natural language responses by the user.	Not supported

Table 3. Results for hypotheses

6 Discussion

In our experiment, we examined the effect of preset answer options on user perceptions of text-based conversational agents in a customer service context. Our results indicate that the use of such elements does reduce the perceived humanness and social presence but does not increase a user's satisfaction with the service encounter. Even though the availability of preset answer options reduces the required efforts for users to manually formulate a reply, the users did not perceive this as an added value in the conversation with the chatbot in our setting. Actually, our results indicated that the use of preset answer options could lead to a lower user satisfaction though the difference between control and treatment was not statistically significant.

These results might seem counter-intuitive at first glance: A CA that offers quick reply elements decreases required manual efforts for the user and aims to reduce the overall time needed to fulfill a customer request. As the speed with which a service request is fulfilled, together with the expertise of customer service and the quality of the reply, influences the perceived customer service performance and customer satisfaction (Setia, Venkatesh and Jogleklar, 2013), one can expect that the use of preset answer options actually increases service satisfaction as formulated in our third hypothesis. However, viewing the results in the light of social response theory allows for a different interpretation: As the CA provided a variety of social cues, such as the name ("Linda"), expressions of emotions through words and emoticons, and through natural language interaction itself, it elicited social responses by the users. Thus, participants anthropomorphized the CA, which is also reflected in the free-form feedback: "The short waiting time in comparison to chats with human service employees is useful and saves a lot of time" and "[…] it would be great if Linda could respond to things like praise or say how she feels at the moment (I know these are gimmicks, but they are fun!)" [translated to English]. The participants viewed the CA as a "virtual human" and compared it to customer service employees rather than to self-service systems, such as on company websites (Meuter, Ostrom, Roundtree and Bitner, 2000). As shown in previous work, this perceived humanness and social presence can be associated with further effects, particularly on service encounter satisfaction (Gnewuch et al., 2018), user engagement (Schuetzler et al., 2018), and perceived usefulness (Hassanein and Head, 2007).

Against this background, the use of preset answer options reduces the degree to which users anthropomorphize a CA as shown in our experiment. As such elements replace the formulating and typing of a natural language message, which users know from digital human-to-human communication and represents a social cue (Fogg, 2003), with something that resembles elements known from typical graphical user interfaces, they decrease how human-like the users perceive the CA. In parallel, it changes how users view the agent and how satisfied they are with the encounter as the feeling of a perceived personal and human contact is significantly reduced.

6.1 Implications for the Design of Conversational Agents

The results of our study have two main implications for the design of text-based conversational agents, in particular in a company context. First, our insights place emphasis in the importance of perceived humanness as well as social presence and their influence on other factors, such as user satisfaction (Gefen and Straub, 2004; Hassanein and Head, 2007; Baylor, 2009). For designers, our results highlight the need to consider different anthropomorphic design dimensions, e.g. with regard to the human identity and verbal as well as non-verbal cues (Seeger et al., 2018) when implementing a conversational agent besides traditional established dimensions, such as the time required to fulfill a user's request. In the context of our experiment, the degree to which a CA is perceived as human-like and socially present seemed to be even more relevant to the user's service encounter satisfaction than the speed at which a request is fulfilled. Specifically, the partial replacement of formulating and typing a natural language response with a set of preset answers reduced both how human-like a CA is perceived as well as how satisfied users were with the service encounter. These results are in line with recent work in the context of CAs like the study by Gnewuch et al. (2018), who found that dynamically delayed response times of chatbots increase perceived humanness, social presence, and service satisfaction even though the CA's responses might take longer to appear than with static response delays.

Second, our study highlights the need to carefully balance the use of "human versus robot" aspects in the design of conversational agents. While users have the advantage of conversational freedom in natural language interaction without the need to navigate through preset menus or dialogues, this freedom poses a major challenge for the design where high variation of input and flexible interaction flows need to be taken into account and the repertoire of control mechanisms for the designer is dramatically reduced (Brandtzæg and Følstad, 2018). As Brandtzæg and Følstad (2018, p. 41) state, the "user interface is to a much greater degree a blank canvas where the content and features of the underlying service are mostly hidden from the user, and where the interaction is more dependent on the user's input". In this context, the use of preset answer options can be a useful approach to guide and structure the conversation flow towards a specific goal as well as to ensure that the CA can provide adequate and helpful responses. However, as our results indicate, the use of such elements reduces the perceived humanness and social presence in the interaction as well as potentially even decreases user satisfaction. Consequently, we suggest to carefully consider the use of preset answer options by taking into account the need to structure the conversation flow at each specific part of the conversation (to ensure the responsiveness of the agent) with the potentially detrimental impact of such elements on perceived humanness and social presence of the CA.

6.2 Limitations and Opportunities for Future Research

Our study is not free of limitations and offers opportunities for future research on CA interaction and design. One limitation of the present work results from the specific context in which our experiment took place, customer service. While the general purpose of preset answers to provide quick and convenient replies is independent from the context in which a CA is used, user motivation, goals and behavior are different across contexts. In particular, a user that does not have a specific question or task in mind, might perceive these buttons differently. For example, a potential customer that has a longer lasting interaction with a CA in a marketing and sales context to explore the product range of a company might view preset answer options and the conversational structure they can provide as helpful. Thus, we suggest that future work explores the use of such elements in different contexts with regard to user satisfaction and perceived humanness, in particular in settings where users might not have a specific task they intent to complete or a specific question to ask.

In addition, the degree of CA anthropomorphism deserves further investigation with regard to the perception of preset answer options. For example, users might perceive preset answer options differently for a CA that is deliberately designed to be machine-like instead of human-like, which can be a suitable design option to induce CA trustworthiness for tasks that are typically carried out by machines (Seeger et al., 2017). We recommend future studies to explicitly investigate how such preset answer options are perceived by users, which we did not consider in our experimental design. Furthermore,

we provided a rather specific set of tasks to the participants. Hereby, the CA could be precisely prepared for the user statements and the variation of user input was reduced. The CA that did not provide preset answers was thus able to process the majority of user input without the use of fallback replies. Against this background, future research on CA design could investigate how preset answers are perceived by the user in a dialogue where otherwise some of the requests cannot be adequately processed by the CA. Additionally, the conversation logs offered few options to analyze and understand user behavior, for example the actual required time to complete the experimental task could not be investigated. We encourage future experimental CA studies to implement complementary conversation data storage instead of relying on the analytical functionality of the natural language processing platform, which may restrict options for log analysis (Diederich, Brendel and Kolbe, 2019b). In addition, our experiment took place in a rather controlled setting in which we deliberately minimized distraction by providing a basic web interface for the conversation. Typical distractions and breaks in the conversation, such as notifications on mobile phones or new content on social media, were excluded from our experiment. Thus, while our online setting had strengths regarding precision and control, it had its weaknesses regarding realism (Dennis and Valacich, 2001). We propose that future research explores the use of preset answers in the field, for example by means of a case study with a company that uses a CA for customer service on a website or social media.

Finally, we did not differentiate between requirements for specific groups of users in our experiment, particularly with regard to user age, and did not collect information on how they perceive preset answer options. Studies on CAs have shown that users of different age groups interact with and perceive CAs in different ways (e.g. Beer, Smarr, Fisk and Rogers (2015)) and thus might have different requirements regarding the design of dialogues. Similarly, cultural aspects with regard to values, beliefs, and behavioral patterns impact how interactive systems are used and need to be taken into consideration when designing a CA on a cross-cultural scale (Pereira and Baranauskas, 2015). One approach to deal with these differences could be to actively ask the users at the beginning of a conversation whether she or he prefers to receive preset answers, thus accounting for personal preferences. Hence, we propose the investigation of explicit customization of CAs with regard to the provision of preset answer options as a further opportunity for future research.

7 Concluding Remarks

The goal of this study was to investigate the impact of preset answer options in human-CA interaction on perceived humanness, social presence, and service encounter satisfaction, thus adding to the growing knowledge base on the design of and interaction with conversational agents. The results of our online experiment indicate that CAs which use preset answer options decrease perceived humanness and social presence compared to CAs who do not use such design elements. Interestingly, CAs that offer these elements for quick and convenient replies, do not increase service encounter satisfaction according to our results, but might even lead to a decrease in user satisfaction. Therefore, our work indicates that even though users are required to put in extra effort for manual typing, they are more satisfied with the customer service encounter, potentially due to a more 'human' feeling in the interaction. The results support the overall direction in CA design research to enhance perceived humanness as well as social presence by emphasizing that these can even lead to a higher user satisfaction despite the additional manual effort for the users. In addition, our insights can be used to inform the design process of CA in practice by highlighting the potential drawbacks of apparently useful elements and emphasizing the need to design for natural interaction.

References

Al-Natour, S., I. Benbasat and R. T. Cenfetelli. (2006). "The Role of Design Characteristics in Shaping Perceptions of Similarity: The Case of Online Shopping Assistants." *Journal of the Association for Information Systems*, *7*(12), 821–861.

Al-Natour, S., I. Benbasat and R. T. Cenfetelli. (2009). "The Antecedents of Customer Self-Disclosure

to Online Virtual Advisors." In: *Proceedings of the International Conference on Information Systems (ICIS)* (pp. 1–18). Phoenix, USA.

Altman, I., D. A. Taylor and V. J. Derlega. (1997). "Social penetration: The development of interpersonal relationships." *Contemporary Psychology*.

Appel, J., A. Von Der Pütten, N. C. Krämer and J. Gratch. (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." *Advances in Human-Computer Interaction*, 10.

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.

Baylor, A. L. (2009). "Promoting motivation with virtual agents and avatars: Role of visual presence and appearance." *Philosophical Transactions of the Royal Society B: Biological Sciences*.

Beale, R. and C. Creed. (2009). "Affective interaction: How emotional agents affect users." *International Journal of Human Computer Studies*, *67*(9), 755–776.

Beer, J. M., C. A. Smarr, A. D. Fisk and W. A. Rogers. (2015). "Younger and older users' recognition of virtual agent facial expressions." *International Journal of Human Computer Studies*, *75*, 1–20.

Ben Mimoun, M. S., I. Poncin and M. Garnier. (2012). "Case study-Embodied virtual agents: An analysis on reasons for failure." *Journal of Retailing and Consumer Services*, *19*(6), 605–612.

Berg, M. M. (2014). *Modelling of Natural Dialogues in the Context of Speech-based Information and Control Systems*. Christian-Albrechts-Universität zu Kiel.

Bertacchini, F., E. Bilotta and P. Pantano. (2017). "Shopping with a robotic companion." *Computers in Human Behavior*, *77*, 382–395.

Biancardi, B., A. Cafaro and C. Pelachaud. (2017). "Could a virtual agent be warm and competent? investigating user's impressions of agent's non-verbal behaviours." In: *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 22–24). Denver, USA.

Bickmore, T. W., L. M. Pfeifer and B. W. Jack. (2009). "Taking the time to care: empowering low health literacy hospital patients with virtual nurse agents." In: *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 1265–1274). Boston, USA.

Bishop, T. (2016). "Coffee from a chatbot: Starbucks unveils "My Starbucks Barista" AI technology for mobile orders." Retrieved from https://www.geekwire.com/2016/coffee-chatbot-starbucks-unveil-starbucks-barista-ai-technology-placing-mobile-orders/

Boudreau, M.-C., D. Gefen and D. Straub. (2001). "Validation in Information Systems Research: A State-of-the-Art Assessment." *Management Information Systems Quarterly*.

Brandtzæg, P. B. and A. Følstad. (2018). "Chatbots: Changing User Needs and Motivations." *Interactions*, *25*(5), 38–43.

Burgoon, J. K., J. A. Bonito, P. B. Lowry, S. L. Humpherys, G. D. Moody, J. E. Gaskin and J. S. Giboney. (2016). "Application of Expectancy Violations Theory to communication with and judgments about embodied agents during a decision-making task." *International Journal of Human Computer Studies*, *91*, 24–36.

Burgoon, J. K. and S. B. Jones. (1976). "Toward a Theory of Personal Space Expectations and their Violations." *Human Communication Research*.

Chattaraman, V., W. S. Kwon and J. E. Gilbert. (2012). "Virtual agents in retail web sites: Benefits of simulated social interaction for older users." *Computers in Human Behavior*, *28*(6), 2055–2066.

Clickworker. (2018). "clickworker." Retrieved from https://www.clickworker.com/

Cowell, A. J. and K. M. Stanney. (2005). "Manipulation of non-verbal interaction style and demographic embodiment to increase anthropomorphic computer character credibility." *International Journal of Human Computer Studies*, *62*(2), 281–306.

Cyr, D., M. Head, H. Larios and B. Pan. (2009). "Exploring Human Images in Website Design: A Multi-Method Approach." *Management Information Systems Quarterly*, *33*(3), 539.

Danielescu, A. and G. Christian. (2018). "A Bot is Not a Polyglot." In: *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 1–9). Montréal, Canada.

De Rosis, F., C. Pelachaud, I. Poggi, V. Carofiglio and B. De Carolis. (2003). "From Greta's mind to

her face: Modelling the dynamics of affective states in a conversational embodied agent." *International Journal of Human Computer Studies*, *59*(1–2), 81–118.

Dennis, A. R., R. M. Fuller and J. S. Valacich. (2014). "Media, Tasks, and Communication A Theory of Media Synchronicity Processes," *32*(3), 575–600.

Dennis, A. R. and J. S. Valacich. (2001). "Conducting Experimental Research in Information Systems." *Communications of the Association for Information Systems*, *7*(5), 1–41.

Diederich, S., A. B. Brendel and L. M. Kolbe. (2019a). "On Conversational Agents in Information Systems Research: Analyzing the Past to Guide Future Work." In: *Proceedings of the International Conference on Wirtschaftsinformatik*.

Diederich, S., A. B. Brendel and L. M. Kolbe. (2019b). "Towards a Taxonomy of Platforms for Conversational Agent Design." In: *Proceedings of the International Conference on Wirtschaftsinformatik*.

Elmorshidy, A. (2013). "Applying the technology acceptance and service quality models to Live Customer Support Chat for E-commerce websites." *Journal of Applied Business Research*, *29*(2), 589–596.

Elson, J. S., D. C. Derrick and G. S. Ligon. (2018). "Examining Trust and Reliance in Collaborations between Humans and Automated Agents." In: *Proceedings of the Hawaii International Conference on System Sciences (HICSS)* (pp. 430–439). Waikoloa Village, Hawaii, USA.

Fast, E., B. Chen, J. Mendelsohn, J. Bassen and M. Bernstein. (2017). "Iris: A Conversational Agent for Complex Tasks." In: *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 1–12). Denver, USA.

Featherman, M., J. Thatcher, R. T. Wright and J. C. Zimmer. (2011). "Design Principles for Special Purpose, Embodied, Conversational Intelligence with Environmental Sensors (SPECIES) Agents." *AIS Transactions on Human-Computer Interaction*, *3*(1), 1–25.

Fogg, B. J. (2003). "Computers as Persuasive Social Actors." In: *Persuasive Technology: Using Computers to Change What We Think and Do*.

Gefen, D. and D. Straub. (1997). "Gender Differences in the Perception and Use of E-Mail: An Extension to the Technology Acceptance Model." *Management Information Systems Quarterly*, *21*(4), 389–400.

Gefen, D. and D. Straub. (2003). "Managing User Trust in B2C e-Services." *E-Service Journal*, *2*(2), 7–24.

Gefen, D. and D. Straub. (2004). "Consumer trust in B2C e-Commerce and the importance of social presence: Experiments in e-Products and e-Services." *Omega*, *32*(6), 407–424.

Gefen, D. and D. Straub. (2005). "A practical guide to factorial validity using PLS-graph: tutorial and annotated example." *Communications of the Association for Information Systems*, 91–109.

Gnewuch, U., S. Morana, M. T. P. Adam and A. Maedche. (2018). "Faster Is Not Always Better: Understanding the Effect of Dynamic Response Delays in Human-Chatbot Interaction." In: *Proceedings of the European Conference on Information Systems (ECIS)*. Portsmouth, United Kingdom.

Gnewuch, U., S. Morana and A. Maedche. (2017). "Towards Designing Cooperative and Social Conversational Agents for Customer Service." In: *Proceedings of the International Conference on Information Systems (ICIS)* (pp. 1–13). Seoul, Korea.

Gong, L. (2008). "How social is social responses to computers? The function of the degree of anthropomorphism in computer representations." *Computers in Human Behavior*, *24*(4), 1494–1509.

Grice, H. P. (1975). "Logic and Conversation." *Syntax and Semantics, Vol 3*.

Groom, V., C. Nass, T. Chen, A. Nielsen, J. K. Scarborough and E. Robles. (2009). "Evaluating the effects of behavioral realism in embodied agents." *International Journal of Human Computer Studies*, *67*(10), 842–849.

Hanus, M. D. and J. Fox. (2015). "Persuasive avatars: The effects of customizing a virtual salespersons appearance on brand liking and purchase intentions." *International Journal of Human Computer Studies*, *84*, 33–40.

Harjunen, V. J., M. Spapé, I. Ahmed, G. Jacucci and N. Ravaja. (2018). "Persuaded by the machine:

The effect of virtual nonverbal cues and individual differences on compliance in economic bargaining." *Computers in Human Behavior*, *87*(March), 384–394.

Hassanein, K. and M. Head. (2007). "Manipulating perceived social presence through the web interface and its impact on attitude towards online shopping." *International Journal of Human Computer Studies*, *65*(8), 689–708.

Holtgraves, T. M., S. J. Ross, C. R. Weywadt and T. L. Han. (2007). "Perceiving artificial social agents." *Computers in Human Behavior*, *23*(5), 2163–2174.

Hu, T., A. Xu, Z. Liu, Q. You, Y. Guo, V. Sinha, … R. Akkiraju. (2018). "Touch Your Heart: A Tone-aware Chatbot for Customer Care on Social Media." In: *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 1–12). Montréal, Canada.

Kahai, S. S. and R. B. Cooper. (2003). "Exploring the Core Concepts of Media Richness Theory: The Impact of Cue Multiplicity and Feedback Immediacy on Decision Quality." *Journal of Management Information Systems*, *20*(1), 263–299.

Knijnenburg, B. P. and M. C. Willemsen. (2016). "Inferring Capabilities of Intelligent Agents from Their External Traits." *ACM Transactions on Interactive Intelligent Systems*, *6*(4), 1–25.

Krämer, N., S. Kopp, C. Becker-Asano and N. Sommer. (2013). "Smile and the world will smile with you - The effects of a virtual agent's smile on users' evaluation and behavior." *International Journal of Human Computer Studies*, *71*(3), 335–349.

Larivière, B., D. Bowen, T. W. Andreassen, W. Kunz, N. J. Sirianni, C. Voss, … A. De Keyser. (2017). ""Service Encounter 2.0": An investigation into the roles of technology, employees and customers." *Journal of Business Research*, *79*, 238–246.

Lee, C., S. Jung, S. Kim and G. G. Lee. (2009). "Example-based dialog modeling for practical multi-domain dialog system." *Speech Communication*, *51*(5), 466–484.

Lee, K. M., Y. Jung, J. Kim and S. R. Kim. (2006). "Are physically embodied social agents better than disembodied social agents? The effects of physical embodiment, tactile interaction, and people's loneliness in human-robot interaction." *International Journal of Human Computer Studies*, *64*(10), 962–973.

Leite, I., A. Pereira, S. Mascarenhas, C. Martinho, R. Prada and A. Paiva. (2013). "The influence of empathy in human-robot relations." *International Journal of Human Computer Studies*, *71*(3), 250–260.

Liao, Q. V., M. M. Hussain, P. Chandar, M. Davis, M. Crasso, D. Wang, … W. Geyer. (2018). "All Work and no Play? Conversations with a Question-and-Answer Chatbot in the Wild." In: *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 1–13). Montréal, Canada.

Luger, E. and A. Sellen. (2016). ""Like Having a Really Bad PA": The Gulf between User Expectation and Experience of Conversational Agents." In: *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 5286–5297). San José, USA.

Matsushita, M., E. Maeda and T. Kato. (2004). "An interactive visualization method of numerical data based on natural language requirements." *International Journal of Human Computer Studies*, *60*(4), 469–488.

Mc Kevitt, P., D. Partridge and Y. Wilks. (1999). "Why machines should analyse intention in natural language dialogue." *International Journal of Human-Computer Studies*, *51*(5), 947–989.

McTear, M. (2017). "The rise of the conversational interface: A new kid on the block?" *Lecture Notes in Computer Science*.

McTear, M., Z. Callejas and D. Griol. (2016). *The Conversational Interface: Talking to Smart Devices*. Basel, Switzerland: Springer Publishing Company.

Meuter, M. L., A. L. Ostrom, R. I. Roundtree and M. J. Bitner. (2000). "Self-Service Technologies:Understanding Customer Satisfaction with Technology-Based Service Encounters." *Journal of Marketing*, *64*(July), 50–64.

Morana, S., C. Friemel, U. Gnewuch, A. Maedche and J. Pfeiffer. (2017). "Interaktion mit smarten Systemen – Aktueller Stand und zukünftige Entwicklungen im Bereich der Nutzerassistenz." *Wirtschaftsinformatik & Management*, *5*, 42–51.

Nass, C. and Y. Moon. (2000). "Machines and mindlessness: Social responses to computers." *Journal*

of Social Issues, *56*(1), 81–103.
Nass, C., J. Steuer and E. R. Tauber. (1994). "Computers are social actors." In: *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (p. 204). Boston, USA.
NextIT. (2018). "Helping a railroad service conduct business." Retrieved from http://nextit.com/case-studies/amtrak
Nunamaker, J. F., D. C. Derrick, A. C. Elkins, J. K. Burgoon and M. W. Patton. (2011). "Embodied Conversational Agent-Based Kiosk for Automated Interviewing." *Journal of Management Information Systems*, *28*(1), 17–48.
Oracle. (2016). "Can Virtual Experiences Replace Reality? The future role for humans in delivering customer experience," 19.
Park, E. K. and S. S. Sundar. (2015). "Can synchronicity and visual modality enhance social presence in mobile messaging?" *Computers in Human Behavior*, *45*, 121–128.
Pereira, R. and M. C. C. Baranauskas. (2015). "A value-oriented and culturally informed approach to the design of interactive systems." *International Journal of Human Computer Studies*, *80*, 66–82.
Perez, S. (2016). "Starbucks unveils a virtual assistant that takes your order via messaging or voice." Retrieved from https://techcrunch.com/2017/01/30/starbucks-unveils-a-virtual-assistant-that-takes-your-order-via-messaging-or-voice/?guccounter=1
Qiu, L. and I. Benbasat. (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. and I. Benbasat. (2010). "A study of demographic embodiments of product recommendation agents in electronic commerce." *International Journal of Human Computer Studies*, *68*(10), 669–688.
Reeves, B. and C. Nass. (1996). *The Media Equation: How People Treat Computers, Television and New Media Like Real People and Places*. The Center for the Study of Language and Information Publications.
Saffarizadeh, K., M. Boodraj and T. M. Alashoor. (2017). "Conversational Assistants: Investigating Privacy Concerns, Trust, and Self-Disclosure." In: *Proceedings of the International Conference on Information Systems (ICIS)* (pp. 0–12). Seoul, Korea.
Schuetzler, R. M., G. M. Grimes and J. S. Giboney. (2018). "An Investigation of Conversational Agent Relevance, Presence, and Engagement." In: *Proceedings of the Americas Conference on Information Systems (AMCIS)* (pp. 1–10). New Orleans, USA.
Schuetzler, R. M., G. M. Grimes, J. S. Giboney and J. Buckman. (2014). "Facilitating Natural Conversational Agent Interactions: Lessons from a Deception Experiment." In: *Proceedings of the International Conference on Information Systems (ICIS)* (pp. 1–16). Auckland, New Zealand.
Seeger, A.-M., J. Pfeiffer and A. Heinzl. (2017). "When Do We Need a Human? Anthropomorphic Design and Trustworthiness of Conversational Agents." In: *Special Interest Group on Human-Computer Interaction* (Vol. 15, pp. 1–6).
Seeger, A.-M., J. Pfeiffer and A. Heinzl. (2018). "Designing Anthropomorphic Conversational Agents: Development and Empirical Evaluation of a Design Framework." In: *Proceedings of the International Conference on Information Systems (ICIS)* (pp. 1–17). San Francisco, USA.
Setia, P., V. Venkatesh and S. Jogleklar. (2013). "Leveraging Digital Technologies: How Information Quality Leads to Localized Capabilities and Customer Service Performance." *Management Information Systems Quarterly*, *37*(2), 565–590.
Seymour, M., K. Riemer and J. Kay. (2017). "Interactive Realistic Digital Avatars - Revisiting the Uncanny Valley." In: *Proceedings of the Hawaii International Conference on System Sciences (HICSS)* (pp. 547–556). Waikoloa Village, Hawaii, USA.
Shamekhi, A., Q. V. Liao, D. Wang, R. K. E. Bellamy and T. Erickson. (2018). "Face Value? Exploring the Effects of Embodiment for a Group Facilitation Agent." In: *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 1–13). Montréal, Canada.
Shawar, B. E. and E. Atwell. (2007). "Chatbots: Are they really useful?" In: *LDV-Forum* (Vol. 22, pp. 29–49).
Shi, Y., X. Yan, X. Ma, Y. Lou and N. Cao. (2018). "Designing Emotional Expressions of

Conversational States for Voice Assistants." In: *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 1–6). Montréal, Canada.

Short, J., E. Williams and B. Christie. (1976). *The Social Psychology of Telecommunications*. London, United Kingdom.

Stock, R. M. and M. Merkle. (2018). "Customer Responses to Robotic Innovative Behavior Cues During the Service Encounter." In: *Proceedings of the International Conference on Information Systems (ICIS)* (pp. 1–17). San Francisco, USA.

Strohmann, T., S. Fischer, D. Siemon, F. Brachten, C. Lattemann, S. Robra-Bissantz and S. Stieglitz. (2018). "Virtual Moderation Assistance: Creating Design Guidelines for Virtual Assistants Supporting Creative Workshops." In: *Proceedings of the Pacific Asia Conference on Information Systems (PACIS)* (pp. 3580–3594). Yokohama, Japan.

Tavanapour, N. and E. A. C. Bittner. (2018). "Automated Facilitation for Idea Platforms: Design and Evaluation of a Chatbot Prototype." In: *Proceedings of the International Conference on Information Systems (ICIS)* (pp. 1–9). San Francisco, USA.

Taylor, S. (1994). "Waiting for Service: The Relationship between Delays and Evaluations of Service." *Journal of Marketing*, *58*(2), 56–69.

Topbots. (2016). "Staples' "That Was Easy."" Retrieved from Staples' "That Was Easy"

Urbach, N. and F. Ahlemann. (2010). "Structural Equation Modeling in Information Systems Research Using Partial Least Squares." *Journal of Information Technology Theory and Application (JITTA)*, *11*(2), 5–40.

Vaccaro, K., T. Agarwalla, S. Shivakumar and R. Kumar. (2018). "Designing the Future of Personal Fashion Experiences Online." In: *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 1–11). Montréal, Canada.

Verhagen, T., J. van Nes, F. Feldberg and W. van Dolen. (2014). "Virtual customer service agents: Using social presence and personalization to shape online service encounters." *Journal of Computer-Mediated Communication*, *19*(3), 529–545.

Vogel-Meijer, K. (2018). "KLM: Making airline customer service soar with Messenger." Retrieved from https://www.facebook.com/business/success/klm-messenger

Von Der Pütten, A. M., N. C. Krämer, J. Gratch and S. H. Kang. (2010). ""It doesn't matter what you are!" Explaining social effects of agents and avatars." *Computers in Human Behavior*, *26*(6), 1641–1650.

Wang, W. and I. Benbasat. (2005). "Trust in and Adoption of Online Recommendation Agents." *Journal of the Association for Information Systems*, *6*(3), 72–101.

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.

Welch, C. (2018). "Google just gave a stunning demo of Assistant making an actual phone call." Retrieved from https://www.theverge.com/2018/5/8/17332070/google-assistant-makes-phone-call-demo-duplex-io-2018

Wünderlich, N. V. and S. Paluch. (2017). "A Nice and Friendly Chat With a Bot: User Perceptions of AI-based Service Agents." In: *Proceedings of the International Conference on Information Systems (ICIS)* (pp. 1–11). Seoul, Korea.

Xu, A., Z. Liu, Y. Guo, V. Sinha and R. Akkiraju. (2017). "A New Chatbot for Customer Service on Social Media." In: *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (pp. 3506–3510). Denver, USA.

Yu, T. K., M. L. Lin and Y. K. Liao. (2017). "Understanding factors influencing information communication technology adoption behavior: The moderators of information literacy and digital skills." *Computers in Human Behavior*, *71*, 196–208.

Zadrozny, W., M. Budzikowska, J. Chai, N. Kambhatla, S. Levesque and N. Nicolov. (2000). "Natural Language Dialogue for Personalized Interaction." *Communications of the ACM*, *43*(8), 116–120.