Association for Information Systems

AIS Electronic Library (AISeL)

Aug 10th, 12:00 AM

You are an Idiot! – How Conversational Agent Communication Patterns Influence Frustration and Harassment

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

Maike Greve
University of Goettingen, maike.greve@uni-goettingen.de

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

Johannes Bührke
University of Goettingen, johannes.buehrke@t-online.de

Lutz M. Kolbe
University of Goettingen, lutz.kolbe@wiwi.uni-goettingen.de

Follow this and additional works at: https://aisel.aisnet.org/amcis2020

'You are an Idiot!' – How Conversational Agent Communication Patterns Influence Frustration and Harassment

Completed Research

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

Maike Greve
University of Goettingen
maike.greve@uni-goettingen.de

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

Johannes Bührke
University of Goettingen
johannes.buehrke@stud.uni-goettingen.de

Lutz M. Kolbe
University of Goettingen
lkolbe@uni-goettingen.de

Abstract

Conversational Agents (CA) in the form of digital assistants on smartphones, chatbots on social media, or physical embodied systems are an increasingly often applied new form of user interfaces for digital systems. The human-like design of CAs (e.g., having names, greeting users, and using self-references) leads to users subconsciously reacting to them as they were interacting with a human. In recent research, it has been shown that this social component of interacting with a CA leads to various benefits, such as increased service satisfaction, enjoyment, and trust. However, numerous CAs were discontinued because of inadequate responses to user requests or only making errors because of the limited functionalities and knowledge of a CA, which can lead to frustration. Therefore, investigating the causes of frustration and other related emotions and reactions highly relevant. Against this background, this study investigates via an online experiment with 169 participants how different communication patterns influence user's perception, frustration, and harassment behavior of an error producing CA.

Keywords

Conversational agents, harassment, communication pattern

Introduction

Driven by machine learning and natural language processing, the way of how humans and computers communicate is continuously changing (Marinova et al. 2017). Conversational Agents (CA) in the form of digital assistants on smartphones, chatbots on social media, or physical embodied systems are seen as a significant new form of user interfaces for digital systems. CAs have the potential to provide services that are always available to the customer while promising a comfortable and convenient user experience (Verhagen et al. 2014).

However, interacting with CAs is leading to an interesting phenomenon: Users display social responses to CAs (Gnewuch et al. 2018; Nass and Moon 2000). The human-like design of a CA (e.g., having a name, greeting users and using self-references) leads to users subconsciously reacting to CAs as they were interacting with a human, for instance, leading to increased service satisfaction, enjoyment, and trust (Diederich, Janßen-Müller, et al. 2019; Gnewuch et al. 2018). Hence, the selection and implementation of human-like design features play a significant role in the successful applications of CAs (Seeger 2019). However, Wiese and Weis (2019) note that human-like design may also result in undesired adverse effects

due to feelings of uncanniness. Numerous CAs were discontinued because of failing to provide accurate responses and simply making errors because of the limited functionalities and knowledge of a CA (Ben Mimoun et al. 2012). Thus, CAs can lead to service provision mismatch because of the heightened expectations of users through the human-like design (Larivière et al. 2017). Overall, users may face major frustration caused by limited comprehension abilities, producing errors, or uncomfortable behavior patterns (Ben Mimoun et al. 2012). Specifically, CAs producing errors can lead to frustration and even harassment in the form of abusive language use (such as 'you are an idiot') by users (Berkowitz 1989). However, to the best of our knowledge, it remains unclear why some users harass CAs and how harassment is triggered and influenced by a CAs communication pattern. This leads us to the following research question:

> *RQ: How do different communication patterns influence user's perception, frustration, and harassment behavior of an error producing CA?*

In order to address this research question, we conducted an online experiment with CAs following one of three different communication patterns, based on Karpman (1968): victim, helper, or persecutor. During the communication with a user, the CA makes an error, potentially triggering harassment. Through this study design, we strive to investigate the interaction of communication pattern and CA error and its effect on perceived humanness, social presence, uncanniness, frustration, service satisfaction, and harassment (in the form of verbal abuses).

Research Background

Conversational Agents in the Context of Service Systems

The field of human-computer service has significant potential for future customer satisfaction, and the CAs capabilities in the context of service systems are continuously improving over the past (Gnewuch et al. 2018). To implement a CA in a service system still represents a major challenge and becomes increasingly complex as high user expectations, such as adaptive interaction behavior, can often not be fulfilled (Gnewuch et al. 2018). Nonetheless, CAs increasingly applied to fulfill a human-like service encounter (Marinova et al. 2017; Verhagen et al. 2014) while using a wide range of technical possibilities based on machine learning and artificial intelligence (Larivière et al. 2017). Users will have unlimited availability to services, and CAs will use its possibility to induce feelings in regards to social presence and personalization (Larivière et al. 2017; Verhagen et al. 2014). Service technology might overcome the basic level of human performance quality of today (Larivière et al. 2017). However, a significant investment of research has been conducted on conversational agents and their language speaking behavior. Most studies still focus on technical aspects like algorithms or new architectures. A neglected lack of knowledge connected to human-computer interaction exists (Følstad and Brandtzæg 2017).

Prominently, limited CAs capabilities lead to a rate of errors abruptly, influencing user perception in a negative way (Brandtzæg and Følstad 2018). Hence, the capabilities of CA-human interaction must be further investigated to solve currently existing problems. For instance, anticipating user requests for natural language software is a challenging endeavor due to the unpredictability of such interactions. Situations are likely to occur in which a CA process an error or needs to provide some kind of fallback response when the user's input was not understood, remind users that they are still interacting with a machine that has limited capabilities and functionalities.

Social Responses to Conversational Agents

Social Response Theory stands as a vital contributor to analyze human-computer interaction for enhancing and designing future CAs (Diederich, Brendel, Lichtenberg, et al. 2019; Gnewuch et al. 2018). Subsequently, it has been figured out that people may adopt and apply social behaviors to anything, even to technological artifacts, if those are demonstrating traits or behaviors that mimic humans (Nass and Moon 2000). Nass and Moon (2000) propose that human-like technology can elicit social behavior from real people. Besides that, different research streams have identified a positive impact of the degree of perceived humanness on trustworthiness, perceived competency (Araujo 2018), authenticity, and service satisfaction, while potential biases of uncanniness against human-like CA persists (Gnewuch et al. 2018). With an in-depth view of social responses, while interacting with CAs, empirical studies illustrate that frustration and verbal abusing of CAs

is a frequent occurrence. People appear to enjoy provoking, teasing, and humiliating conversational partners and misusing the technology in a disinhibiting context, like cyberbullying and harassment (Brahnam and De Angeli 2008). One hypothesis is that users become abusive when facing a CA that is frustrating them. In a human-like context, frustration can be caused due to different obstacles coming from a direct topic relation or non-contextual obstacles. As a result, aggression, harassment, and verbal abusing are typical examples of triggered reactions when facing frustration (Berkowitz 1989).

Communication Pattern of Conversational Agents

Human identity, nonverbal, and verbal communication are the main three dimensions for designing a human-like CA in a social behavior context (Seeger et al. 2018). Implementing human identity into CAs means to implement artifacts such as an avatar or a personal profile. Overall, with its expressions, CAs provoke social cue attributions and induce the user to perceive them as a real human than a passive object (Brahnam and De Angeli 2008). However, designers of CAs have to be careful in the selection of social cues (Seeger et al. 2018), because 'more is not always better.'

However, to the best of our knowledge, current research is not considering complex and diverging communication patterns. Prominently, CAs are designed to mimic service personnel (Larivière et al. 2017; Seeger et al. 2018). Therefore, designers select and combine cues that are expected in a service encounter. Similarly, the communication pattern of a service CA is based on the assumptions of how service personnel should communicate. In this context, a deeper understanding of how the perception of communication patterns differs between humans and CA is missing.

In the context of harassment, Karpman (1968) provides a role setup (called Drama Triangle) regarding individuals in an abusive interaction and relationship. In the role model, three different types of communicators are described: helper (also called rescuer), victim, and persecutor (McKimm and Forrest 2010). A victim is expressed by a person who suffers and causes problems for other individuals, conditions, and circumstances. The rescuer reflects a person that interferes out of a self-condition to support and help. A persecutor acts in his interest, pressurizes, and forces other individuals (McKimm and Forrest 2010).

Research Model and Hypothesis

While various studies have been conducted on human-computer interaction and human service encounters (Diederich, Brendel, and Kolbe 2019), to the best of our knowledge, little empirical research has been conducted on how different communication patterns are perceived by a user in a service interaction in which the CA makes an error. As a first step towards designing conversational communication patterns tailored for specific contexts, we base our communications patterns on the Karpman (1968) Drama Triangle, which includes the roles of helper, victim, and persecutor. We selected this model because it provides a role setup for an abusive relationship and communication, which fits the context of harassment of CAs. Furthermore, the helper communication pattern fits the communication of a service employee. For the experiment context, we selected a service encounter where the chatbot is initially unable to provide a satisfactory service (e.g., making an error), which can lead to frustration and the use of abusive language. The hypotheses are explained and presented in the following sections.

Humanness

In the context of human-like design, researchers explore the use of various cues that are intended to contribute to the human-likeness of CAs. It has been determined that a human-like CA personalized with a name and its own unique customized behavior patterns contribute to perceived humanness (Araujo 2018; Diederich, Brendel, Lichtenberg, et al. 2019; Diederich, Janßen-Müller, et al. 2019). In line with these studies, we consider the communication patterns to be a specific combination of social cues, focusing on verbal content and style. Hence, different levels of perceived humanness can be expected, depending on the cues implemented. Specifically, the communication pattern of victim and persecutor are expected to yield a higher level of humanness because they display a more extensive array of social behavior (e.g., the victim apologizes, and the persecutor blames the user), meaning they display a wider range of social cues. This leads us to the following two hypotheses:

H1a: A CA that follows a Victim-type communication pattern yields a higher level of perceived humanness than a CA that follows a Helper-type communication pattern.

H1b: A CA that follows a Persecutor-type communication pattern yields a higher level of perceived humanness than a CA that follows a Helper-type communication pattern.

Uncanniness

Uncanniness is related through the "uncanny valley" model into the context of humans interacting with human-like technology. It is expressing the human behavior related to the experience like CAs equipped with high applicability to humanity skills (MacDorman et al. 2009). It leads the user to a sense of discomfort and uneasiness when users get in touch with a similarity of an almost virtual human (Bartneck et al. 2007; MacDorman et al. 2009). Hence, it can be hypothesized that a high but not perfectly human-like CA receives more negative evaluations than a static agent offering only content orientated feedback (Groom et al. 2009). In order to present a valid communication pattern, the three applied communication patterns should not lead to a high level of perceived uncanniness; otherwise, they could be considered to feel artificial to the users. Therefore, we formulate the following hypothesis:

H2: All three CAs exhibit a low level of uncanniness.

Frustration

IS Research identified that CAs could contribute to negative user emotions such as frustration (Fox and Spector 1999; Jenkins et al. 2007). Frustration can be described as an emotional state resulting from the occurrence of an obstacle (e.g., external factors) that prevents the satisfaction of a need (Berkowitz 1989). In a service context, frustration, annoyance, as well as impatience has been noted as well (Jenkins et al. 2007). In their research, Jenkins et al. (2007) were analyzing that a non-accurate response of the CA can trigger frustration. As all three applied CA instances are making an error in the service encounter, frustration should be triggered. However, the differences in the communication pattern should have an impact on the resulting frustration. The persecutor communication pattern does not attempt to mitigate the unpleasantness of the error, even blaming the user for it, unlike the other two patterns. This behavior can be expected to lead to a high level of frustration. Hence, the following two hypotheses can be formulated:

H3a: A CA that follows a Persecutor-type communication pattern leads to a higher level of frustration than a CA that follows a Victim-type communication pattern.

H3b: A CA that follows a Persecutor-type communication pattern leads to a higher level of frustration than a CA that follows a Helper-type communication pattern.

Service Satisfaction

Within the area of the human service encounter, a high level of service satisfaction is related to service quality and customer experience. Now, empathy is considered to be the critical competence for creating likability, trustworthiness, and supportiveness within the eye of a customer (Caruana et al. 2000). In this context, the helper and victim display an interest in helping the user and trying to resolve any problems occurring (e.g., error), which should lead to a higher level of service satisfaction when compared to a persecutor communication pattern, which can appear to be unexpectedly confrontational by the users. Thus, we formulated the following two hypotheses:

H4a: A CA that follows a Helper-type communication pattern yields a higher level of service satisfaction than a CA that follows a Persecutor -type communication pattern.

H4b: A CA that follows a Victim-type communication pattern yields a higher level of service satisfaction than a CA that follows a Persecutor-type communication pattern.

Harassment

The existence of harassment has been reported in human-CA interactions (Brahnam and De Angeli 2008). In a social behavior context, Berkowitz (1989) explains that aggression can be triggered by experiencing frustration before. In the long-term, frustration will always lead to aggression (Berkowitz 1989). Aggression,

in our context, is defined as harassment, which isn't always directed at the source of origin, for example, social, and monetary consequences (Berkowitz 1989). Thus, applying a persecutor communication pattern can trigger harassment because it is expected to lead to a higher level of frustration. Furthermore, following the drama triangle, the victim is the role associated with being harassed by others (McKimm and Forrest 2010). Against this background, we formulate the following two hypotheses:

> *H5a: A CA that follows a Victim-type communication pattern leads to more abusive communications than a CA that follows a Persecutor -type communication pattern.*

> *H5b: A CA that follows a Persecutor-type communication pattern leads to more abusive communications than a CA that follows a Helper-type communication pattern.*

Research Design

Data Collection Procedure and Sample

The participants received a briefing document, in which we explained the context (making an appointment via a chatbot) and the structure of the experiment (interaction with the chatbot with the subsequent survey) as well as described the participants' tasks. Every participant received the same document to make sure that the participants have 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 one of the three chatbot configuration (helper, victim, and persecutor). Each participant was supposed to get an appointment for a tire change of a specific car (specific brand, model, and registration number). The conversation was structured into six steps: (1) Request an appointment for changing tires, (2) State brand of car, (3) State model of car, (4) State registration number of car, (5) Negotiate a date and time for the appointment within the next few days, and (6) Confirm appointment.

In step 5, the chatbot initially offers an appointment one year and a few months in the future, on a Sunday, and an hour before midnight, rendering this offer to be an obvious mistake and error. After the users decline this offer (as they are requested to make an appointment soon), the chatbot asks the users for a date and time for the appointment and subsequently rejects it. In the end, the chatbot offers an appropriate appointment date and time, concluding the process. This negotiation process was specifically designed to frustrate and trigger potential harassment (e.g., calling the chatbot names or answering sarcastically).

After completing the task, the chatbot provided a link to the online survey. Overall, participating in the experiment took around five minutes per participant. Our study has a sample size of n = 169 participants ranging from 18 to 74 years old (M= 33.52 years, SD = 13.88) and a share of 36.7% female and 45.5% male persons. 17,7% of the participants made no statement on gender. Our participants were recruited via personal networks and social media and, therefore, were not compensated for their participation. 98% of participants are from Germany, 2% are from different countries across Europe.

Configurations

We prepared three instances (a chatbot following a victim (n = 59), persecutor (n = 56), or helper (n = 54) communication pattern) of one CA using the design platform for natural language software "Dialogflow" by Google. We further implemented a custom-built web interface to provide convenient access to the CAs and minimize distraction (see Figure 1). All CAs received the same training phrases, i.e., exemplary statements which customers might make during the service encounter, that indicate a user's intent and triggers a reply. The CAs were able to process different variations of sentences with the same meaning and could extract parameters, such as the brand of car, and use them throughout the dialogue for paraphrasing.

The CAs received various cues for human-like CA design according to the three dimensions (human identity, verbal, non-verbal) as suggested by Seeger et al. (2018) to establish a baseline for perceived humanness: Concerning human identity, we equipped the CA with the name "Gernd Müller," a male gender, and a human stock photo avatar, representing a male customer service employee. Concerning verbal communication, the CA was designed to use self-references, such as "What can I do for you?", turn-taking, and a personal introduction ("Welcome! I am Gerd Müller, your Personal Assistant at your local car dealership."), including a greeting in the form of a welcome message. Regarding the non-verbal human-like CA design dimension, we implemented blinking dots in combination with dynamic response delays

depending on the length of the previous message, as suggested by (Gnewuch et al. 2018) to simulate thinking and typing of replies by the CAs.
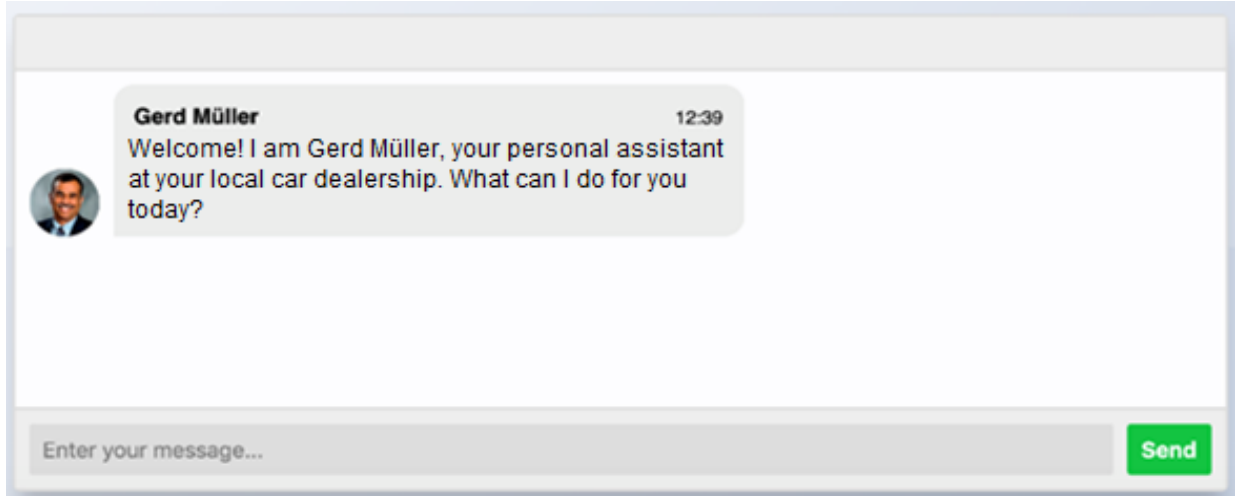

Figure 1. Web Interface of Victim-Chatbot with Greeting Message (translated from German)

Overall, the three instances of the CA were identical except for the specific cues and phrases related to their communication pattern (see Table 1), following the suggestions of Karpman (1968). The chatbot that applied a helper communication pattern uses "Please" and "Thank you," frequently uses the word "help," and uses positive language (e.g., "good" and "great"). In comparison, the chatbot resembling a victim communication pattern is frequently apologizing (e.g., using "Sorry"), uses passive language, and uses negative language (e.g., "Sadly" and "Unfortunately"). Lastly, the chatbot instance based on the persecutor communication pattern is quite rude (e.g., not using "Please" and "Thank you"), blames the user for the inability to make an appointment (e.g., "because you were unable to accept my first appointment offer […]"), and is assertive and commanding (e.g., "Provide the model of your car.").

Context	Communication Pattern	Response Examples
User declines first offered appointment	Helper	*"Please provide a different appointment date and time."*
	Victim	*"I am terribly sorry for the inconvenience. Could you please provide a different appointment date and time?*
	Persecutor	*"Because you were unable to accept my first appointment offer, provide me with a different date and time."*
User accepts offered appointment	Helper	*"Great, I was able to book the appointment for you:* [time] *and* [date]*"*
	Victim	*"Thank you so much for your patience. Sadly, I still make many mistakes. But the appointment is booked for you:* [time] *and* [date]*"*
	Persecutor	*"Ok, I was finally able to make the following appointment:* [time] *and* [date]*"*

Note, all responses are translated from German.

Table 1. Context, Communication Pattern and Response Examples

Measures and Descriptive Statistics

After interacting with the CA, the participants were forwarded to an online survey that measured five different constructs by asking a variety of items. All items were measured on a scale from 1 (strongly disagree) to 8 (strongly agree). For the design of the survey, only established constructs from previous studies were considered (see Table 2). Additionally, we included attention checks by asking two questions that prompt the participant to select a specific number on a scale. If the participant fails the questions, the data was not considered for the analysis.

We conducted a confirmatory factor analysis (CFA) to check the factor loadings of the items for each construct. We only considered items with a factor loading above the threshold of .60. We further evaluated the constructs using Cronbach's alpha (α) and the composite reliability (CR) that both require a value larger than .80, and the average variance extracted (AVE) that requires a value larger than .50 (Urbach and Ahlemann 2010).

In Table 2, the constructs with its corresponding items and factor loadings of the CFA are summarized. For uncanniness, we used a single item construct. For perceived humanness, frustration, and service satisfaction, we calculated weighted sum scores (DiStefano et al. 2009) based on the factor loading to generate one metric variable for each construct. The dependent variable harassment is not depicted in the table because it is a binary variable that indicates the use of abusive language. The use of abusive language has been identified by one of the authors based on the use of insults (such as 'you are an idiot'), sarcasm, irony, or intentional cap locking of words or phrases.

Results

We analyzed the survey data through descriptive statistics and analysis of variance to compare the means of the three groups. For the analysis, we used the statistical software R. Depending on the scale of the dependent variable, we conducted different analyses. We calculated a one-way ANOVA (based on the necessary tests for normal distribution) for perceived humanness, frustration, and service satisfaction. Therefore, as a first step, we conducted the Levene test to check for variance homogeneity of the metric variables. This resulted in heterogenic variances for perceived humanness ($F(2,166)=4.79$, $p=.009$) and variance homogeneity for frustration ($F(2,166)=1.84$, $p=.162$), and service satisfaction ($F(2,166)=1.86$, $p=.160$). Hence, we used a Welch ANOVA and Games-Howell post hoc Tests for perceived humanness and a one-way ANOVA with Turkey multiple pairwise-comparison for the variable with homogeneous variance. For the ordinal-scaled variable uncanniness, we considered the non-parametric Kruskal Wallis H test. The binary variable harassment was analyzed by the Fisher Exact test and its corresponding post hoc analysis.

Our descriptive data, as well as the statistical testing by variance analysis, show differences in the dependent variables for the participants, depending on the communication pattern of the CA. Concerning the perceived humanness, it shows that there is a statistically significant difference between the three conditions. The descriptive group averages show that the victim-type (M=4.26, SD=0.73) and helper-type communication pattern (M=3.96, SD=0.87) differ significantly from the persecutor-type, which yields a significantly lower perceived humaneness (M=2.79, SD=1.07). Thus, **H1b** is contradicted, since the opposite case, that the helper-type is perceived as more human-like, applies. There is no statistical difference between the helpers and the victim, which is why **H1a** is not supported. Furthermore, it is hypothesized (**H2**) that all communication patterns result in a low level of uncanniness (we assume a value of under three to be low, following the example of Diederich, Brendel and Kolbe, (2020). A one-sample t-test against the fixed value of 4 confirms that the three communication patterns of persecutor ($t(55)=-4.458$, $p<.001^{***}$), helper ($t(53)=-2.514$, $p=.007^{**}$), and victim ($t(55)=-3.153$, $p<.001^{***}$) exhibited a low level of uncanniness. The frustration level does not show a statistically significant difference in the three communication patterns. Hence **H3a** and **H3b** are not supported. In contrast, the one-way ANOVA for the dependent variable service satisfaction shows a significant difference between the groups. The post hoc analysis confirms a difference between the helper (M=4.52, SD=1.89) and the persecutor communication pattern (M=3.17, SD=1.36), where the latter yields lower satisfaction. This confirms **H4a**. Furthermore, **H4b** is supported since the victim pattern (M=4.83, SD=1.07) is also significantly higher than the persecutor pattern. Lastly, the results show that none of the participants receiving the helper CA used abusive language. Therefore, there is a significant difference in harassment between the three groups. Nevertheless, contradicting **H5**b, the group with the Victim communication pattern yields a lower value of harassment in comparison to the persecutor group. However, as hypothesized in **H5a,** the more abusive communication is found for participants receiving the persecutor pattern (M=0.38, SD=0.388) in comparison to participants receiving the helper communication pattern.

Constructs and items	Loadings	Source
Perceived humanness (α = .861, CR = .862, AVE = .513) The CA seemed to be competent. The CA seemed to be well thought out. The Ca seemed to be polite. The CA responded well to my answers. The CA seemed to be appealing. The CA seemed to be polite.	.707 .755 .727 .699 .676 .724	(Holtgraves and Han 2007)
Uncanniness The CA seemed to be unpleasantly human-like.	-	(Bartneck et al. 2007)
Frustration (α = .910, CR = .911, AVE = .837) How frustrating was the interaction with the CA for you? How frustrated are you after interacting with the CA?	.940 .889	(Fox and Spector 1999)
Service satisfaction (α = .868, CR = .873, AVE = .698) How satisfied were you with the CA service? How satisfied are you with the way the CA treated you? How satisfied are you with the general interaction with the CA?	.871 .781 .861	(Verhagen et al. 2014)

Note all items were translated to German for the survey.

Table 2. Measurement of Latent Variables

Dependent variable *(Scale)*		All (N=169)	Condition			Variance Analysis	Post-hoc comparison between groups	
			Persecutor (N=56)	Helper (N=54)	Victim (N=59)			
Perceived Humanness *(Metric)*	Mean SD	3.92 1.10	2.79 1.07	3.96 0.87	4.26 0.73	Welch ANOVA F(2,106.81) = 36.76 p < .001 ***	H - P V - P V - H	p < .001 *** p < .001 *** p = .117 n.s.
Uncanniness *(Ordinal)*	Mean SD	3.20 1.65	3.05 1.59	3.41 1.73	3.15 1.64	Kruskal Wallis H Test H(2) = 1.4692 p = .480 n.s.	-	-
Frustration *(Metric)*	Mean SD	3.65 1.80	4.15 1.89	3.81 1.92	3.72 1.59	One-way ANOVA F(2,166) = 0.926 p = .398 n.s.	-	-
Service Satisfaction *(Metric)*	Mean SD	4.18 1.39	3.17 1.36	4.52 1.13	4.83 1.07	One-way ANOVA F(2,166) = 31.08 p < .001 **	H - P V - P V - H	p < .001 *** p < .001 *** p = .354 n.s.
Harassment *(Binary)*	Mean SD	0.18 0.388	0.38 0.49	0 0	0.17 0.38	Fisher Exact Test p < 0.001 ***	H - P V - P V - H	p < .001 *** p < .001 *** p < .001 ***

SD = Standard Deviation, H = Helper, P = Persecutor, V = Victim, p = p-value
Significance level: * 0.01; ** 0.05; *** 0.001; n.s. = not significant

Table 3. Descriptive Statistics and Variance Analysis with Post Hoc Comparison

Discussion and Conclusion

Our study contributes to the knowledge base for human-like CA design by investigating the effects of different communication patterns on the user's emotions and behavior. For the context of a service encounter, we provide new insights into the causes for users starting to verbally harass a CA, for instance, by calling it names (e.g., 'you are an idiot'). Our results indicate that displaying a friendly demi-ner and interest in solving the problem at hand prevents harassment and loss in service satisfaction. Nonetheless, showcasing a "victim mentality" and frequently apologizing appears to lead some users to use abusive language.

Furthermore, our results contribute to theory by questioning the relation of frustration and harassment (Berkowitz 1989). All three communication patterns lead to similar levels of frustration but to a different

number of harassment incidents. For practice, our results hold the prescriptive knowledge of avoiding being too assertive or too submissive in case of an unexpected error. In the case of service encounters, CAs should be equipped with responses that display the intention to help without appearing to take the error or mistake personally.

In the following, we will discuss the primary limitations of this study and avenues for future research. Firstly, this study investigated the CA type of chatbot, these results and implications might not translate to other CA types (e.g., voice assistant or humanoid robots). Consequently, investigating the interrelation of CA errors, communication patterns, and harassment for other CA types should provide an attractive future research area. Secondly, the results of this study are highly dependent on the applied set of social cues (Seeger et al. 2018). Other sets of cues might lead to different results, similar to a different design of communication patterns. Thus, future research should engage in developing communication patterns for specific contexts. Thirdly, our experiment was designed in a way that users had no personal involvement (e.g., no actual appointment was booked, and no repercussions had to be expected from being rude to the CA). Adding this dimension could change the behavior and constitutes, therefore, a valuable area for future research.

Acknowledgments

REFERENCES

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), pp. 183–189.

Bartneck, C., Kanda, T., Ishiguro, H., and Hagita, N. 2007. "Is the uncanny valley an uncanny cliff?," in *RO-MAN 2007-The 16th IEEE international symposium on robot and human interactive communication*, IEEE, pp. 368–373.

Berkowitz, L. 1989. "Frustration-Aggression Hypothesis: Examination and Reformulation," *Psychological Bulletin* (doi: 10.1037/0033-2909.106.1.59).

Brahnam, S., and De Angeli, A. 2008. "Special issue on the abuse and misuse of social agents," *Interacting with Computers* (20:3), pp. 287–291 (doi: 10.1016/j.intcom.2008.02.001).

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

Caruana, A., Ewing, M. T., and Ramaseshan, B. 2000. "Assessment of the three-column format SERVQUAL: An experimental approach," *Journal of Business Research* (49), pp. 57–65.

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

Diederich, S., Brendel, A. B., and Kolbe, L. M. 2019. "Towards a Taxonomy of Platforms for Conversational Agent Design," in *International Conference on Wirtschaftsinformatik*, pp. 1–15.

Diederich, S., Brendel, A. B., and Kolbe, L. M. 2020. "Designing Anthropomorphic Enterprise Conversational Agents," *Business & Information Systems Engineering*, Springer Fachmedien Wiesbaden.

Diederich, S., Brendel, A. B., Lichtenberg, S., and Kolbe. 2019. "Design for Fast Request Fulfillment or Natural Interaction? Insights from an Experiment with a Conversational Agent," in *ECIS*, pp. 1–17.

Diederich, S., Janßen-Müller, M., Brendel, A. B., and Morana, S. 2019. "Emulating Empathetic Behavior in Online Service Encounters with Sentiment-Adaptive Responses: Insights from an Experiment with a Conversational Agent," in *Proceedings of the International Conference on Information Systems (ICIS)*, Munich, Germany, pp. 0–17.

DiStefano, C., Zhu, M., and Mîndrilă, D. 2009. "Understanding and using factor scores: Considerations for the applied researcher," *Practical Assessment, Research and Evaluation* (14:20).

Følstad, A., and Brandtzæg, P. B. 2017. "Chatbots and the new world of HCI," *Interactions* (24:4), pp. 38–42.

Fox, S., and Spector, P. E. 1999. "A model of work frustration–aggression," *Journal of Organizational Behavior* (20:6) (doi: 10.1002/(sici)1099-1379(199911)20:6<915::aid-job918>3.3.co;2-y).

Gnewuch, U., Morana, S., Adam, M. T. P., and Maedche, A. 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, pp. 1–17.

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

Holtgraves, T., and Han, T. L. 2007. "A procedure for studying online conversational processing using a chat bot," *Behavior Research Methods* (39:1), pp. 156–163.

Jenkins, M. C., Churchill, R., Cox, S., and Smith, D. 2007. "Analysis of user interaction with service oriented chatbot systems," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 4552 LNCS) (doi: 10.1007/978-3-540-73110-8_9).

Karpman, S. 1968. "Fairy tales and script drama analysis," *Transactional analysis bulletin* (7:26), pp. 39–43.

Larivière, B., Bowen, D., Andreassen, T. W., Kunz, W., Sirianni, N. J., Voss, C., Wünderlich, N. V., and De Keyser, A. 2017. "'Service Encounter 2.0': An investigation into the roles of technology, employees and customers," *Journal of Business Research* (79), pp. 238–246.

MacDorman, K. F., Green, R. D., Ho, C. C., and Koch, C. T. 2009. "Too real for comfort? Uncanny responses to computer generated faces," *Computers in Human Behavior* (25:3), pp. 695–710.

Marinova, D., de Ruyter, K., Huang, M. H., Meuter, M. L., and Challagalla, G. 2017. "Getting Smart: Learning From Technology-Empowered Frontline Interactions," *Journal of Service Research* (20:1), pp. 29–42.

McKimm, J., and Forrest, K. 2010. "Using transactional analysis to improve clinical and educational supervision: The Drama and Winner's triangles," *Postgraduate Medical Journal* (doi: 10.1136/pgmj.2009.093310).

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

Nass, C., and Moon, Y. 2000. "Machines and mindlessness: Social responses to computers," *Journal of Social Issues* (56:1), pp. 81–103.

Seeger, A.-M., Pfeiffer, J., and Heinzl, A. 2018. "Designing Anthropomorphic Conversational Agents: Development and Empirical Evaluation of a Design Framework," in *Proceedings of the International Conference on Information Systems (ICIS)*, San Francisco, USA, pp. 1–17.

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

Verhagen, T., van Nes, J., Feldberg, F., and 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), pp. 529–545.