The power of computer-mediated communication theories in explaining the effect of chatbot introduction on user experience
Hendriks, Frank; Ou, Carol; K Amiri, Amin; Bockting, Sander

Published in:
Proceedings of the 53 Hawaii International Conference on System Sciences (HICSS 2020)

Document version:
Publisher's PDF, also known as Version of record

Publication date:
2020

Link to publication

Citation for published version (APA):
The Power of Computer-Mediated Communication Theories in Explaining the Effect of Chatbot Introduction on User Experience

Frank Hendriks  
KPMG The Netherlands  
hendriks.frank@kpmg.nl

Carol X.J. Ou  
Tilburg University  
carol.ou@uvt.nl

Amin Khodabandeh Amiri  
Tilburg University  
a.amiri@uvt.nl

Sander Bockting  
KPMG The Netherlands  
bockting.sander@kpmg.nl

Abstract

Chatbots have increasingly penetrated our lives as their behavior growingly imitates a human interlocutor. This paper examines the effect of different methods of self-presentation of a chatbot on the end-user experience. An interlocutor in a computer-mediated communication (CMC) environment can either introduce itself as a chatbot, a human being, or choose not to identify itself. We conducted an experiment to compare these three methods in terms of end-user experience that comprises of social presence, perceived humanness, and service encounter satisfaction. Our data demonstrate that a chatbot that discloses its virtual identity is scored significantly lower for social presence and perceived humanness than other two choices of self-presentation. Key findings and the associated implications are discussed.

1. Introduction

A chatbot is a piece of software that responds to natural language input and attempts to hold a conversation in a way that imitates a real person [1]. The consumer acceptance of chatbots has been growing vastly during the past years [2]. A substantial amount of internet users already has frequent contact with a chatbot. Consumers are also comfortable in using chat messaging channels for customer service interactions with companies [3]. Companies, also, start to see the advantages of chatbots, such as cost reduction, better service delivery, and improving the competitive position [4]. While nowadays, chatbots, mainly, play an advisory role [5] or help with relatively simple transactions with a lower level of intelligence [6], [7], they are expected to flourish in after-sales, customer care, and marketing environments in the future [7].

Research areas in chatbot development can be boiled down to two interesting streams: technical development and human-computer interaction. Whereas the first stream focuses on technological advances, such as natural language processing and artificial intelligence [8], [9], the interaction stream focuses on improving the experience and interaction with the end-user [10].

Research about chatbots has tended to focus on the technological advances, rather than on the interaction elements [11]. Such research development has led to the knowledge gaps in the second stream in literature, and thus heightened the need to research the factors influencing human-chatbot interaction.

Also, the first stream suggests that technological abilities are expanding vastly. In the near future, these technological expansions could lead to situations in which the end-user is unaware of the fact that s/he could be talking to a chatbot when expecting a real human being. Especially in chat communications, which involve written messages, the user cannot derive the identity of the interlocutor when a chatbot is able to perfectly imitate a human being. This described situation is imminent, as current chatbots are on the verge of passing the Turing Test, which is regarded as the last boundary for human-imitating computer-interaction [12]–[15].

These descriptions of future chatbot possibilities increase the urge for knowledge on how this interaction takes place, and which factors influence the interaction and user experience. This research focuses on the impact of a chatbot’s self-introduction in a computer mediated communication (CMC) environment. We present empirical results relevant for chatbot designers and contribute to relevant theories on chatbot interaction.
2. Literature Background

Users, unconsciously, apply the social norms and cues from social interaction to computer interaction [16]–[18]. The CMC theories, including the media richness theory [19] and reduced social cues approach [20], have successfully extended the implication of social norms and cues into the computer mediated environment. Chatbots work in an environment of computer and function based on algorithms. In this study, we attempt to examine whether chatbot studies can leverage current theories of social and communicational studies to explain how end-users perceive chatbots.

Particularly, CMC theories can be useful for our study context of online communication. CMC theories address communication between humans via a digital communication channel where the communication is mediated by computer software [21]. These theories suggest that digital communication channels have ‘social bandwidth’ that limits the amount of social cues and complexity that can be transferred in messages [22]. Therefore, non-verbal social cues, such as body language or vocal tone, are only conveyed to a limited extent.

This view drastically impacts chat conversations in comparison to normal conversations, as the reduction of social cues allows selective conveyance of specific social cues, such as a self-introduction.

3. Chatbot Interaction

Various aspects of chatbot interaction have been studied in the literature. An overview of research into these factors show that several aspects affect user experience, some in a counterintuitive manner. For instance, the level of friendliness, expertise [23] and tone-awareness of the chatbot [24] do contribute to the user-satisfaction in a way that was intuitively expected. However, a longer delay in message delivery impacts the user satisfaction in a positive way, which contrasts the main reasons for using a chatbot, that is faster and more efficient conversation [25].

Several studies have also addressed human likeness of chatbots. Neuroscience research shows that a more human-alike chatbot is perceived more competent [26]. Another study demonstrates that people perceive a dynamic chatbot to be more engaging and human-alike [27]. Human likeness is furthermore reflected in research towards trust in chatbots. Speech as expression gave a higher willingness to share personal information with machines [28]. However, trust in information that comes from a chatbot is treated in a different way, and end-users have high expectations from answers provided by chatbots [29]. Moreover, the visual depiction of typefaces influences the perceived humanness as well. A machine-like typeface makes the end-user perceive a chatbot as more machine-alike, but on the contrary, a handwriting mimicking typeface does not yield a higher perceived humanness [30].

Additionally, factors concerning the visual representation of chatbots have also been examined. The agency of a chatbot can have various appearances, represented in hierarchical levels. These different levels vary from chatbots with a profile picture, to chatbots with a virtual character [31]. Chatbots with a virtual character yield higher feelings of social presence with the end-user due to a higher number of social cues exchanged.

4. Theoretical Framework

4.1. Self-introduction in a Social Context

Self-introduction is a vital social norm in human communications and significantly influences the experience of the interlocutor. For example, a proper welcome to a social context positively influences engagement in social situations [32]. Nevertheless, no empirical research has been conducted that explains the effect of self-introduction of a virtual chatbot agent. It is however seen that if chatbots show similar anthropomorphistic traits a higher self-disclosure by the human interlocutor is provoked [33].

The social information processing theory [34] describes the process of building relationships via CMC. This theory claims that relationship building via CMC is arduous due to the limited ‘social bandwidth’ CMC communication offers. The limited bandwidth negatively affects the transfer of social cues and only permits a restricted expression of the complexity in messages [35].

Nevertheless, the social information processing theory argues that relationship building is possible if two restrictions are considered: the fact that extra time is needed, and that all social cues are transformed into written cues. In other words, what is not conveyed through body language and other non-verbal signs is conveyed in a written way [20].
The change to written cues instead of non-verbal cues opens the opportunity for selective self-presentation that allows making a controlled impression by managing the social cues shared with an interlocutor [19]. Selective self-presentation can occur in two ways: a proactive approach, where the interlocutor clearly states its identity, or a neutral approach, which keeps the identity enclosed.

Manipulation in selective self-presentation influences the course of the conversation, because it directly affects the amount and essence of the exchanged social cues. It can accelerate or slow down relationship building.

The effect various social cues have in CMC communication are reflected by the construct’s social presence, perceived humanness, and service encounter satisfaction. These constructs provide measures to assess the impact of selective self-presentation on end-user experience, as defined in Table 1 and further explained below.

### 4.2. Social presence

Social presence pertains to the degree to which a person is perceived as a ‘real person’ in CMC [36]. Social presence is mainly expressed in terms of human warmth, personalness, sociability and human sensitivity, as experienced by the interlocutor. These are highly influenced by the method of communication. For example, video communication has a higher degree of social presence than audio communication. Social presence is also influenced by the persons or machines involved in the communication [37]–[39].

If a chatbot introduces itself as a chatbot, as is the case with proactive self-presentation, the user will know the real identity of the interlocutor as the chatbot. As a result, a lower level of intimacy or warmth will be experienced by the user as knowing that him/herself is talking to a chatbot instead of a real person for the interaction. Such less amount of ‘real person’ perception in computer-mediated communication [22] will result to the feeling of less social presence. Therefore,

<table>
<thead>
<tr>
<th>Selective self-presentation [34]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Definition</strong>: The way the interlocutor presents itself in computer-mediated communication.</td>
</tr>
<tr>
<td><strong>Manipulation</strong>: either as “introduction as a chatbot”, “no introduction at all”, or “introduction as human being”</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Social presence [25], [36], [38], [40]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Definition</strong>: The degree to which an interlocutor is perceived as a ‘real person’ in computer-mediated communication. <strong>Scale</strong>: 7-point Likert scale ranging from “strongly disagree” to “strongly agree”. <strong>Items</strong>:</td>
</tr>
<tr>
<td>I felt a sense of human contact with the interlocutor.</td>
</tr>
<tr>
<td>I felt a sense of personalness with the interlocutor.</td>
</tr>
<tr>
<td>I felt a sense of sociability with the interlocutor.</td>
</tr>
<tr>
<td>I felt a sense of human warmth with the interlocutor.</td>
</tr>
<tr>
<td>I felt a sense of human sensitivity with the interlocutor.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Perceived humanness [22], [25], [27], [41], [42]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Definition</strong>: The degree to which somebody or something is perceived as a human being. <strong>Items</strong>: I found my interlocutor …. <strong>Scale</strong>: 7-point semantic differential scale from:</td>
</tr>
<tr>
<td>extremely inhuman-like - extremely human-like</td>
</tr>
<tr>
<td>extremely unskilled - extremely skilled</td>
</tr>
<tr>
<td>extremely unthoughtful - extremely thoughtful</td>
</tr>
<tr>
<td>extremely impolite - extremely polite</td>
</tr>
<tr>
<td>extremely unresponsive - extremely responsive</td>
</tr>
<tr>
<td>extremely unengaging - extremely engaging</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Service encounter satisfaction [23], [25], [43]–[46]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Definition</strong>: The degree to which the respondent is satisfied with the overall customer-care conversation. <strong>Scale</strong>: 7-point Likert scale from “extremely dissatisfied” to “extremely satisfied”. <strong>Items</strong>:</td>
</tr>
<tr>
<td>How satisfied are you with the interlocutor’s advice?</td>
</tr>
<tr>
<td>How satisfied are you with the way the interlocutor treated you?</td>
</tr>
<tr>
<td>How satisfied are you with the overall interaction with the interlocutor?</td>
</tr>
</tbody>
</table>

Table 1: Construct definitions and measures.
A chatbot with a self-presentation as a chatbot will yield a lower experienced social presence than a chatbot with a neutral-self presentation.

4.3. Perceived humanness

Perceived humanness is defined as the degree to which somebody or something is being experienced as a human being. It originates in the three-factor theory of anthropomorphism that defines humanness in terms of thoughtfulness, politeness and responsiveness of the interlocutor [41], [47]. Revealing the identity as a chatbot make interlocutors believe that they experience an artificial thoughtfulness, more automatic and less politeness and caring for the interlocutor. Hence, interlocutors experience lower level of humanity. Therefore,

H2: A chatbot with a self-presentation as a chatbot will yield a lower experienced perceived humanness than a chatbot with a neutral self-presentation.

4.4. Service Encounter Satisfaction

Service encounter satisfaction is related to measuring and understanding customer satisfaction of the service [44]. It is based on the comparison of expectations prior to the encounter and perceived evaluations after the encounter [48]. This satisfaction is influenced by the interaction that takes place during the service [43]. Knowing the service provider can allocate an employee instead of a chatbot to handle the service can arguably lead to a higher satisfaction with the service provided as that can accommodate for a better experienced and tailored interaction [49]. In this light, self-presentation of a chatbot in a CMC communication will lead to a lower level of satisfaction. Therefore,

H3: A chatbot with a self-presentation as a chatbot will yield a lower experienced service encounter satisfaction than a chatbot with a neutral self-presentation.

5. Methodology

This research has employed an experimental survey methodology approach making use of vignette research style [50], [51]. Participants were randomly assigned to three research groups and were exposed to web care chatbots with different introductions. The dependent variables, i.e., social presence, perceived humanness, and service encounter satisfaction were operationalized in line with their established literature (see Table 1).

The introduction message was manipulated based on the different levels of selective self-presentation ranging from neutral to identity revealing. Conditions involved a chatbot introducing itself as a chatbot (identity revealing self-presentation), a chatbot not introducing itself (neutral self-presentation), and a human introducing itself as a human (identity revealing self-presentation). The baseline in this experiment was an identity revealing chatbot.

According to these conditions, three different vignettes were produced. These vignettes were in the form of a video and showed excerpts of a chatbot conversation in an imaginary customer care setting. Participants were to act as if the conversations were of their own and that they were the ones chatting with the chatbot. Participants have bought a product online and wish to return it. During the chat, participants show such an intent and the interlocutor helps the participants with their wishes. The interlocutor does so by enumerating the criteria for a valid product return – the return terms and relevant instructions – and eventually assists the customers to take the right action.

A vignette style was chosen over a real chatbot interaction by the participants, in order to overcome several confounding factors and to ensure the research’s robustness. Using vignettes, the manipulation could be closely monitored and adjusted, while being able to keep it the same for all experiment groups [52]. Vignettes proved to be a useful way to implement a near-perfect human-imitating chatbot as participants could not derive information about the identity of the interlocutor from anything besides the introduction. Therefore, the vignettes design created a higher degree of control for our experiment as opposed to involving subjects in actually using the chatbot. In order to obtain the true understanding of self-introduction and ruling out other potential confounding effects of using the chatbot, in this study, a vignette design was chosen for our experiment.

In the pre-test, 19 participants were randomly assigned to three groups. Each group of participants received an introduction message from their interlocutor in a video according to the experiment group the participant was in (see table 2). After that introduction message, same set of interactions were displayed to all groups. In total, 6 messages were received by a participant, and 5 messages were sent on behalf of the
participant in the video (please see Figure 1). The video lasted for 45 seconds, and time between messages was held the same for each condition and was based on the length of the message.

After the video, we ask participants to indicate the interlocutor’s identity based on what they saw on the screen (1=chatbot, 2=no indication, and 3=human being). Their answers (i.e., 1, 2, and 3) were taken as the categorical outcome of our manipulation check. We checked the mean differences between the manipulated values - the interlocutors’ identity - using ANOVA. The ANOVA results indicate the significant difference of the three groups (p<0.001) in the way that participants can clearly indicate their interlocutors’ identity according to the group they were in. That means the self-presentation is effectively manipulated. As a result, the manipulation was successful. The self-presentation can be used as the IV for the formal experience and for further analyses later on.

| Condition 1 – Chatbot Revealing | **Hello, you are talking with a customer care chatbot. How can I help you?** |
| Condition 2 – Chatbot Neutral | **Hello, how can I help you?** |
| Condition 3 – Human Being | **Hello, you are talking with a customer care employee. How can I help you?** |

Table 2: Intro messages per condition.

6. **Data Collection and Analysis**

The formal experiment and the survey were operationalized in a Dutch company in which chatbots are designed. The current study was part of the chatbot projects of that company. In total, 159 useable survey responses were collected via business contacts to whom have a high probability of getting in touch with a chatbot. The sample size ensures a 0.99 power to detect an effect size of 0.5 on a Likert scale that is equal to the medium effect size [53]. The main experiment follows the procedures used in the pre-test and uses the same settings.

<table>
<thead>
<tr>
<th>Conditions</th>
<th>Social presence</th>
<th>Perceived humanness</th>
<th>SES¹</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cronbach’s Alpha = .906</td>
<td>Cronbach’s Alpha = .750</td>
<td>Cronbach’s Alpha = .886</td>
</tr>
<tr>
<td>Chatbot Revealing</td>
<td>Mean</td>
<td>SD²</td>
<td>Mean</td>
</tr>
<tr>
<td>Chatbot Neutral</td>
<td>54</td>
<td>4.3185</td>
<td>1.3199</td>
</tr>
<tr>
<td>Chatbot Neutral</td>
<td>52</td>
<td>5.1115</td>
<td>1.1282</td>
</tr>
<tr>
<td>Human Being</td>
<td>53</td>
<td>5.1434</td>
<td>1.3199</td>
</tr>
<tr>
<td>Test Statistics</td>
<td>p &lt; .000</td>
<td>p = .033</td>
<td>p = .319</td>
</tr>
<tr>
<td>Hypothesis</td>
<td>H₁ Supported</td>
<td>H₂ Supported</td>
<td>H₃ Not significant</td>
</tr>
</tbody>
</table>

¹ SES: Service encounter satisfaction. ² SD: Standard Deviation.

Table 3: Descriptive results and test statistics.
The entire experiment was then followed by the same manipulation check as in the pre-test. The manipulation checks (ANOVA, p=0.000) indicate the experiment results can be used for the main analyses.

Specifically, a one-way analysis of variance (ANOVA) was conducted on the influence of self-presentation on the dependent variables (see Table 3). Significant effects of self-presentation on social presence [F(2,156) = 8.383, p = 0.000] and perceived humanness [F(2,156) = 3.500, p = 0.033] are observed for the three types of self-introduction supporting the first two hypotheses. The effect of self-introduction on service encounter satisfaction was not significant [F(2,156) = 1.151, p = 0.319]. An independent-samples t-test was also conducted to compare the dependent variables between a chatbot with a self-presentation as a chatbot (chatbot revealing) and a chatbot with a neutral-self presentation (chatbot neutral). Similarly, the differences in social presence [t(104) = 3.70, p = 0.000, mean difference = 0.79] and perceived humanness [t(104) = 2.00, p = 0.048, mean difference = 0.26] were significant supporting the hypotheses. Again, the t-test indicates no significant effect of self-introduction on service encounter satisfaction [t(104) = 1.45, p = 0.90, mean difference = 0.34].

Considering the non-significant effect of self-introduction on service encounter satisfaction, we conduct a post-hoc robustness check. In a regression analysis, we treated Service Encounter Satisfaction as a dependent variable and Social presence and Perceived humanness as the independent variables. Results show a significant predictive power of social presence (b=0.424 p<0.001) and perceived humanness (0.557, p<0.001), with the total variance explained to service encounter satisfaction of 35.3%.

Therefore, while self-presentation has a significant effect on user experience in general, but our data supports the opposite of H1 and H2. That means the empirical data suggest the experiment subjects consider neutral-introducing chatbots as a better user experience than identity-revealing ones. We discuss the implications of such findings below.

7. Implications for Research and Practice

This paper contributes to the literature of CMC theories. More importantly, this study extends the findings of CMC theories in the context of chatbot. The paper demonstrates a clear difference between the effects of various social cues on the interlocutors’ experience of a conversation by a chatbot. In other words, our empirical results show that revealing the true identity of a chatbot does not equal to a human being’s self-presentation. Even if the conversations contain the same thoughtful, polite, and responsive answers, interlocutor’s self-presentation can make users’ perception and evaluation of the process completely different. If the current variables, i.e., social presence, humanness and satisfaction, are considered as a proxy of the overall user experience, our study indicates that participants still prefer talking to a real employee instead of a chatbot. Hence, the application of CMC theories for human-chatbot interactions need to be interpreted with care.

8. Future Research and Conclusion

This study has been primarily concerned with the impacts of a chatbot’s introduction on user experience. The study demonstrates that self-identification as a chatbot results in lower perceived social presence and perceived humanness. However, the study applied a vignette experiment design instead of using a real chatbot conversation to obtain a higher degree of control. Engaging users with a real chatbot and examine the entire user journey can be the next step of this study. With the availability of chatbot platforms such as Google DialogFlow and Microsoft Bot framework, future studies can expand the CMC theories to explain the interaction between chatbots and end-users and shed more light on the human-chatbot interaction.

9. Bibliography


