

Studying the Anthropomorphic Visual Cues in Conversational Agents

Emilija Zlatkutė¹, Ujwal Gadiraju¹, Jie Yang¹, Sihang Qiu¹

¹Delft University of Technology

Abstract

With the increase of anthropomorphic conversational agents, the research on the effects of anthropomorphism in conversational agents is on the rise. However the studies cannot seem to agree on many of the raised questions and the research on the proper design of conversational agents is still lacking. This research paper attempted to contribute to filling this gap by analysing whether anthropomorphic visual cues used in conversational agents have an effect on the trust and satisfaction of the users. In particular, a between-subject experiment was performed, where the use of emojis and a profile image of four different levels of anthropomorphism were manipulated in a conversational agent based on Telegram. A total of 120 participants, split between eight experimental groups, had a conversation with the agent and filled in a survey about their experience. The results concluded that individual visual cues as well as combination of them did not have any significant effects on the trust and satisfaction of the users.

1 Introduction

The advancement in technological areas like natural language processing and machine learning has resulted in a rise of conversation agents (CAs) (Araujo, 2018; Diederich, 2020; Rapp et al., 2021). CAs, that can be defined as software that interacts with users in a natural human language (Diederich, 2020; Seeger et al., 2017), are now being employed to perform a wide variety of tasks, from completing simple transactions to complex advice giving (Araujo, 2018; Gnewuch et al., 2018; Seeger et al., 2017). CAs in the form of a chat-bot have become especially popular due to their easy integration into social media (Araujo, 2018; Rapp et al., 2021). Such a wide employment of these systems can be explained by their ability to provide a human-like interaction to users, which ensures a smoother and more intuitive experience (Pfeuffer et al., 2019; Seeger et al., 2021).

Because of their unique type of interaction with users, CAs go hand in hand with the theory of *anthropomorphism*. *Anthropomorphism* can be described as assigning human-like

attributes or traits to non-human agents or objects like computers or computer applications such as CAs (Araujo, 2018; Seeger et al., 2017, 2021). In CAs anthropomorphism can be expressed through different types of cues such as human identity (*e.g.* human-like visual representation and identification), verbal cues (*e.g.* human-likeness of the conversation), and non-verbal cues (Seeger et al., 2021).

With the increase in the use of CAs, the research on anthropomorphism and its role in CAs has also been on the rise (Diederich et al., 2019; Rapp et al., 2021). The research themes vary from examining how users assign humanness to the agents to what kind of effect that assignment can have and what factors might have influence on it (Diederich, 2020; Diederich et al., 2019; Go & Sundar, 2019; Gong, 2008; Rapp et al., 2021). However, a problem noticed by multiple studies is that many of the conducted researches have conflicting results (Catrambone et al., 2002; Diederich, 2020; Gong, 2008; Rapp et al., 2021; Seeger et al., 2021). Indeed, while many studies report positive effects of anthropomorphic CAs (Araujo, 2018; Gong, 2008), others have shown that increased level of anthropomorphism can lead to user frustration, confusion, and even feelings of eeriness (Diederich, 2020; Diederich et al., 2021; Pfeuffer et al., 2019).

Lastly, even though human-likeness of the conversation performed by the agent can be considered to be the most important factor in the success of the CA (Diederich, 2020), the importance of the design of the CA should not be overlooked. In fact, recent literature reviews concluded that there is a general lack of researches done on the design of CAs and suggested that future research should focus more on this area (Diederich et al., 2019; Rapp et al., 2021).

In summary, this research will contribute to filling the gap in the research of CA design. In particular, the main question that this paper will aim to answer is: *To what extent can a conversational agent with different levels of anthropomorphic visual cues improve the satisfaction and trust of the users?*

The rest of this research paper will be organised as follows. Section 2 will derive hypotheses about specific visual cues and their potential effects on user trust and satisfaction. Then, section 3 will describe the research method used to assess the influence of visual cues. In section 4 the results of the conducted research will be presented. Section 5 will consider the ethical implications of this study. Section 6 will discuss the effects of the gotten results as well as limitations of the study

and will propose ideas for future work. Finally, section 7 will conclude this research paper.

2 Related Work and Hypotheses

This section will first introduce the reasons behind measuring user trust and satisfaction. Then it will provide an overview of relevant research done on visual cues in CAs. In particular, due to the chat-bot nature of CAs, there is a limited number of visual cues that can be used for the design of the interface. Therefore, this paper will focus on the most common visual cues found in practice and research, which are profile images (Diederich, 2020; Go & Sundar, 2019; Gong, 2008; Seeger et al., 2021) and emojis (Beattie et al., 2020; Diederich, 2020; Seeger et al., 2021). Lastly, this section will present the hypotheses of this research.

2.1 The Importance of User Trust and Satisfaction

While there have been many different measures used to evaluate the effect anthropomorphic conversational agents have on users, user satisfaction has been one of the most popular choices in research (Rapp et al., 2021). User satisfaction has been defined as extent to which user needs are met by the system (Thong & Yap, 1996) and has been often considered to be indistinguishable from system success (Bano et al., 2017; McKeen et al., 1994). Indeed, it has been argued to be not only the most important measure for success (Thong & Yap, 1996; Zviran & Erlich, 2003) but also one of the easiest ways to do it (Rapp et al., 2021; Zviran & Erlich, 2003). Therefore, when it comes to evaluating design choices of CAs, user satisfaction can be a powerful indicator of whether the design has a positive effect on users.

Another prominent research focus when studying different CA interfaces has been user trust (Rapp et al., 2021). Trust is a complex concept studied by many disciplines and having a variety of different definitions (Marsh & Dibben, 2003; Söllner et al., 2016). One of the definitions of trust is as a belief that it is beneficial to be vulnerable based on the expectations of the actions and intentions of others (Söllner et al., 2016). With regards to CAs, user trust has been said to be determined by the beliefs about the competence and integrity of the agent (Seeger et al., 2017). Additionally, trust has been considered to be an important part of measuring system success (Srinivasan, 2004). High user trust has been said to lead to new purchases, increased intention to collaborate and share information (Söllner & Leimeister, 2013). It also has been shown to have an effect on the acceptance and use of technology (Söllner et al., 2016; Söllner & Leimeister, 2013) and has been considered to be a core element for technology usability (Seeger et al., 2017). Furthermore, trustworthy interfaces can lead to users wanting to interact with the system more, thus increasing the usability of the system (Marsh & Dibben, 2003). Indeed, without user trust, other system qualities like efficiency and productivity cannot be maximized and users might end up seeking other means to perform their tasks (Marsh & Dibben, 2003). Therefore, trust can be considered to be a valuable and important measure when assessing CAs and their effects on users.

2.2 Anthropomorphic Profile Images

Seeger et al. (2021) classifies profile images as a part of the human identity dimension, which "helps to identify a human being in a computer-mediated interaction context". The study also argues that CAs with human identity cues will lead to users experiencing a higher perception of anthropomorphism. According to Araujo (2018), such perception consequently should lead to a better emotional connection between the users and the agent and, therefore result in a positive effect on the user-agent relationship. This is further supported by Go and Sundar (2019), who state that the use of human identity indicating visual cues lead to users perceiving the chat-bot more human-like, therefore encouraging a more social response. Similarly, Gong (2008) argued that a more anthropomorphic computer representation will result in a more positive social judgement and a more trustworthy perception from the users. Gong's research is especially relevant for this paper, since the research manipulated computer representation through profile images. Indeed, the study used 85 computer generated images that all contained faces of different anthropomorphic level and concluded that higher anthropomorphic representation leads to more positive effects from the users in terms of trust, homophily attribution, and competency. The study was conducted using dilemma scenarios represented as multiple choice questions, thus the interaction between the user and computer was not human-like, as opposed to CAs. However, even though the results can only be partially applied to conversational agents, they still provide a valuable insight into the effects of anthropomorphic profile images on users. Therefore, the following hypotheses were formulated:

H1: The higher anthropomorphism level of a profile picture in a conversational agent, the higher user satisfaction will be.

H2: The higher anthropomorphism level of a profile picture in a conversational agent, the higher user trust will be.

2.3 The Use of Emojis

Emojis can be defined as small digital images that are used to express emotions or ideas (Oxford, 2021). Indeed, studies have shown that emojis can be a powerful indicators of emotions (Boutet et al., 2021) and can make the expressions of ideas less ambiguous (Riordan, 2017). A study done by Beattie et al. (2020) concluded that chat bots using emojis in text are perceived similarly to humans based on qualities such as social attraction, credibility, and competence. Therefore, the overall usage of emojis in text can be consider an anthropomorphic quality.

According to Seeger et al. (2021), emojis can be classified with the dimension of non-verbal anthropomorphic cues which includes all informative behavior that cannot be expressed in a purely linguistic form. Furthermore, Fadhill et al. (2018) performed a study on the role of emojis in health tracking CAs and concluded that emojis can increase enjoyment, attitude, and confidence towards the CA. Due to emojis unique ability to convey emotions in a text based environment and previous studies suggesting their positive effect on users of conversational agents, the following hypotheses were proposed:

H3: A conversational agent that uses emojis will have a higher user satisfaction than a conversational agent that does not use emojis.

H4: A conversational agent that uses emojis will have a higher user trust than a conversational agent that does not use emojis.

2.4 Combining Visual Cues

Assuming that both visual cues in a CA will increase the satisfaction and trust of the users, then, intuitively, it can be assumed that the combination of both should result in an even higher increase. Indeed, similar observations have been made by de Visser et al. (2016). The performed study manipulated both visual cues (agent appearance, visual indication of thinking) as well as social cues and concluded that adding such human-like features increases trust. Similarly, Diederich (2020) proposed a design for CAs, that, among other anthropomorphic cues, included both emojis and human avatar profile images. The study showed that such anthropomorphic CA design was perceived as more useful and enjoyable by the users. Due to the study using more than visual cues, it is not clear what was the exact effect of the combination of emojis and profile images on the users. However, it does give a strong suggestion of the possible positive effects, therefore this study proposes the following hypotheses:

H5: Use of both emojis and high anthropomorphic level profile pictures in conversational agents will lead to a higher user satisfaction, in comparison to conversational agents that use none or only one of the visual cues.

H6: Use of both emojis and high anthropomorphic level profile pictures in conversational agents will lead to a higher user trust, in comparison to conversational agents that use none or only one of the visual cues.

3 Method

This section will present the methodology of the research. In particular, it will discuss the used research design, the general procedure and tasks done during the experiment, the design of the conversational agent and its dialog, the way the participants were recruited for the experiment, and the measures used to evaluate user trust and satisfaction.

3.1 Study Design

In order to test the hypotheses and answer the research question, a 4 (none vs low vs medium vs high level anthropomorphism picture) x 2 (emojis vs no emojis) between-subjects experiment design was used. The group that did not have a profile picture and did not use emojis was considered to be the control group.

3.2 Procedure

Each participant received a set of instructions and a link to a conversational agent corresponding to their experimental group. All the participants were made aware of the bot nature of the agent. When connected to the CA, participants started the conversation, where the agent sent them questions related to their well-being. Most of the questions had a pre-defined set of possible answers, however a few required free

answers from the participants. The participants were able to skip any of the questions and were able to cancel the conversation at any point. After the participants went through all the questions given by the CA, they received a link to a post-experiment questionnaire about their trust and satisfaction with regards to the CA. The whole process of the experiment including the post-experiment survey took around seven minutes to complete.

3.3 Conversational Agent Design

The conversational agent used for the experiment was based on a *Telegram* bot called *Dandelion* (TU Delft, 2021). It contained a predetermined conversation structure about the well-being of the user.

Levels of anthropomorphism of profile images

The conversational agent used three images as profile images, each having a different level of anthropomorphism (see *Figure 1*). The levels of anthropomorphism and their representation were based on the study done by Gong (2008). For the agent that did not have a profile image, a letter 'D' was displayed by Telegram (see *Figure 1*). The high anthropomorphic level image was generated using Generated Photos (2021), which provides AI-generated images of people. This was done in order to not to use an image of a real, existing person. The medium anthropomorphic level image was made using Avatar Maker (2021) and was made to look similar to the high anthropomorphic level image, in order to control potential biases such as attractiveness. Lastly, the low anthropomorphism level image was designed by Freepik (2021) and taken from *flaticon.com*.

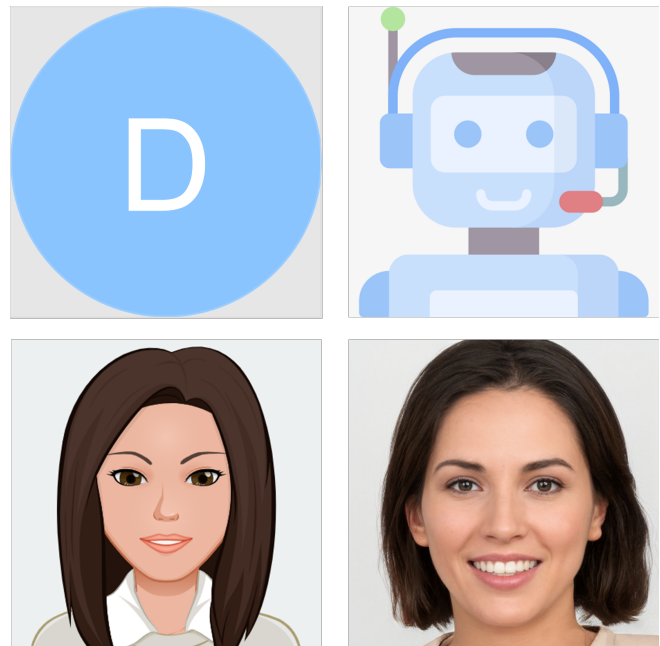


Figure 1: Levels of profile image anthropomorphism. *From left to right, top to bottom:* no image, low anthropomorphism, medium anthropomorphism, high anthropomorphism.

Use of Emojis and the Style of the Conversation

A previous study done by Seeger et al. (2021) showed that nonverbal cues result in a positive effect on perceived anthropomorphism only if they are combined with verbal cues. Furthermore, the interactivity of the conversation as well as meaningful responses have been said to be one of the most important factors for a positive user experience (Diederich, 2020; Diederich et al., 2021). Therefore, the human-likeness of the conversation was considered to be the most important confounding variable in this study.

In order to control the human-likeness variable, all the groups received messages containing the same text, except for the groups that included emojis as visual cues. In those cases, the text was still identical, however in some parts of the message emojis would be present. The type and place of emojis was also identical among all the groups that used them. Furthermore, a casual, non-formal conversation style was used, since it has been shown to result in a higher user engagement, as opposed to a formal one (Kim et al., 2019). Examples of the conversation style and the whole conversational agent can be seen in Figure 2. The full script of the conversation with and without emojis can be found at https://osf.io/hjtez/?view_only=257ed4774c984880a9d6fc3f1f6321c2.

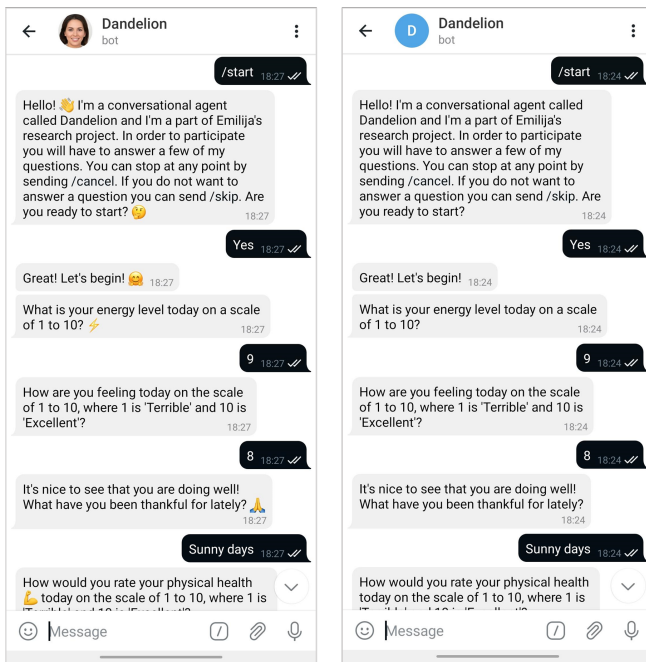


Figure 2: Example of conversational agent design used in the experiment. *On the left*: design with high anthropomorphic level profile picture and emojis. *On the right*: design with no profile picture and no emojis.

3.4 Participants

A total of 120 participants were recruited for this research, resulting in each group receiving 15 participants. The sample consisted of 37% females and 62% males. The participants were recruited via Prolific.co platform and were paid 1£ for their participation.

In order to take part in the experiment, the participants had to be fluent in English language. This was asked because the conversational agent and the post-experiment survey were both in English and it was important that the participants would fully understand the questions and tasks presented to them. Furthermore, the participants were screened on their usage of chat/messaging apps. In particular, only the people that had *Telegram* listed as one of the chat apps that they use regularly could have joined the study. This was done to ensure that participants were familiar with the way the application works and they did not need to install it on their device. This in turn would result in less confusion during the experiment and would increase the quality of the provided answers.

3.5 Measures

User satisfaction was measured by using items adapted from Barger and Grandey (2006) and Lee and Choi (2017) on a 7-point Likert scale (*i.e.* 1 - strongly disagree, 2 - disagree, 3 - somewhat disagree, 4 - neutral, 5 - somewhat agree, 6 - agree, 7 - strongly agree). The provided questions focused not only on the visual aspects of the CA but also on the overall interaction. Lee and Choi, 2017 used the questionnaire to study the influence of self-disclosure in conversational agents. Similarly, the questionnaire made by Barger and Grandey, 2006 has been previously adapted in various studies of conversational agents (Diederich et al., 2021; Verhagen et al., 2014). Therefore, the adaptation of the two resulted in a reliable way to measure user satisfaction. The full table of all items used to measure user satisfaction can be found in the *Appendix A, Table 2*.

Items for measuring user trust were adapted from Jian et al. (2000) and were measured on a 7-point Likert scale (*i.e.* 1 - strongly disagree, 2 - disagree, 3 - somewhat disagree, 4 - neutral, 5 - somewhat agree, 6 - agree, 7 - strongly agree). The items asked whether the user felt suspicious of the CA, if they could trust it with their information, if it seemed reliable, and similar. The full table of all items used to measure user trust can be found in the *Appendix A, Table 3*.

4 Results

The data collected from the experiment was analysed using a two-way ANOVA test, with significance level $\alpha = 0.05$. For both satisfaction and trust, Lavane's test showed no violation of the assumption of homogeneity: satisfaction p -value = 0.97, trust p -value = 0.12. *Table 1* presents the means and standard deviations of all experimental conditions for both satisfaction and trust. The full data set of the collected responses is made available at https://osf.io/hjtez/?view_only=257ed4774c984880a9d6fc3f1f6321c2. Overall, it can be said that no matter the condition, the participants were mostly neutral or on the more positive side towards the conversational agent with regards to satisfaction and trust.

With regards to satisfaction, the results revealed no significant difference between different levels of profile image anthropomorphism (p -value = 0.27). There was also no significant differences between using and not using emojis in the conversational agent (p -value = 0.71). Lastly, there was no

Table 1: Means and standard deviations (SD) of the dependent variables in all experimental conditions.

Condition*	Satisfaction		Trust	
	Mean	SD	Mean	SD
PN	6.12	0.74	5.30	0.84
PL	5.63	0.87	5.24	0.62
PM	5.64	0.74	5.05	0.72
PH	5.87	0.92	5.29	0.79
EN	5.97	0.64	5.04	0.53
EL	5.75	0.73	5.04	1.15
EM	6.15	0.74	5.27	0.81
EH	5.60	0.83	4.95	1.32

* Emoji state: P = Plain (no emojis), E = With emojis; Profile image anthropomorphism: N = No image, L = Low, M = Medium, H = High.

significant correlation between the two independent variables (p -value = 0.23).

Results on user trust were similar to the ones found on user satisfaction. In particular, no significant difference between different level of anthropomorphic profile images was found (p -value = 0.997). Furthermore, the usage of emojis also had no significant effect on the trust of the participants (p -value = 0.37). Finally, the correlation between the usage of emojis and anthropomorphic images was not significant (p -value = 0.62).

In summary, none of the proposed hypotheses (see Section 2) were confirmed.

5 Responsible Research

This research was performed following all the rules and regulations given by the TU Delft Human Research Committee. Each participant had to read and sign a consent form, which informed them about the purpose of this research as well as how and where the collected data will be used.

In order to ensure the well-being of the participants and their consensual participation in the experiment as well as their right to retract their data, the participants had the possibility to stop the experiment at any moment, for any reason. In particular, at the start of the conversation with the conversational agent, the participants were informed that they can send the command `/cancel` to the bot at any time, which would immediately stop their participation and any data that was collected up till that point would be discarded. Additionally, in the post-experiment survey, the participants were again asked if they consented to continue and had the ability to cancel their survey submission at any moment, in which case their data would not be recorded.

Special care was taken to ensure that the participants felt comfortable with the questions asked during the conversation

with the conversational agent. While the questions about the participants well-being can be considered sensitive, it was ensured that the submissions were fully anonymous. Due to the participants being recruited via third-party service *Prolif.co*, this study only received anonymous IDs of the participants, and, therefore had no data that could personally identify any of the participants. Furthermore, during the course of the experiment no additional data that could identify the participants was collected. Finally, if for any reason the participants found any of the questions asked by the conversational agent unacceptable, they were given the ability to skip them with the command `/skip` and could still continue to participate in the experiment if they wanted to.

Lastly, with regards to reproducibility of the present research, all the methods and data were disclosed in this paper. In particular, section 3 mentioned all the details about how the experiment was conducted and how participants were recruited and section 4 included the methods used to process the results as well as a link to the full data set of the experiment.

6 Discussion and Limitations

This research examined anthropomorphic visual cues in conversational agents and their effect on user trust and satisfaction. The conducted experiment results showed no significant increase in satisfaction (H1) or trust (H2) with increase in profile image anthropomorphism. Furthermore, the usage of emojis also did not have an effect on the two measures (H3, H4). Lastly, combining the two visual cues resulted in no significant effect on user satisfaction (H5) and trust (H6).

One of the reasons for insignificant results could be the sample size. A power analysis with medium effect size, $\alpha = 0.05$, and power = 0.80 has revealed that the sample size of the experiment should be at least 400 (Faul et al., 2007). However, due to limited funds, the sample size of this study was 120, leading to only 15 participants for each condition. Therefore, such a sample size can be considered to be one of the biggest limitations of this study, making the recorded results a poor representation of the general population.

6.1 Anthropomorphic Profile Image

The findings that higher anthropomorphism level of profile image in the CA does not increase trust and satisfaction of the users do not go in line with the findings of some of the other researches (de Visser et al., 2016; Go & Sundar, 2019; Gong, 2008). In particular, de Visser et al. (2016) and Gong (2008) in their experiments found that trust increases with the increase of anthropomorphism of agents. Both studies used profile images as one of the ways to express the anthropomorphism of the agent. The type of tasks and conversations that the participants had with the CA differed greatly between this study and the ones conducted by de Visser et al. (2016) and Gong (2008), which could be one of the reasons for different outcomes between the studies.

The tasks provided to the participants in both studies were of the advise-giving type. In particular, in the case of de Visser et al. (2016), the participants had to guess the next number in a sequence after getting a recommendation from the agent. Similarly, Gong (2008) in their study asked partic-

ipants to choose a response to a dilemma after the agent provided its recommendation. However, in the current study no advise was provided by the conversational agent. The participants only had to answer given questions about their current well-being, and, as opposed to the mentioned two studies, did not have to make any decisions. Therefore, it could be argued that the task of the current study did not require the user to form a trust-based relationship with the agent as in the de Visser et al. (2016) and Gong (2008) studies, thus trust was not significantly affected by different anthropomorphic properties.

Furthermore, a more significant difference between the mentioned studies and the current one is the amount of exposure to the profile picture, which might also be considered the main reason for contradicting result. de Visser et al. (2016) and Gong (2008) both used their own interfaces for the CA, allowing them more customisation on how the visual cues will be presented to the participants of the experiment. As a result, the image representing the CA took almost half of the whole CA interface, allowing the participants to clearly see the profile image during the whole conversation. This study, conversely, employed *Telegram* as the host for the CA, where the position and size of the profile image was fixed. In particular, the profile image was only visible in a small circle in the top left corner of the device, therefore making it not clearly visible when the participants were focusing on the conversation. Such a minimal exposure to the profile image might be the core reason for insignificant differences between different levels of anthropomorphism on trust and satisfaction of the users.

Similar conclusions about differences in results can be drawn when comparing this study with the study done by Go and Sundar (2019). While the researchers did not investigate trust or satisfaction with regards to anthropomorphic images in CAs, they did find that higher anthropomorphic visual cues (one of them being a profile image) result in a positive effect on the users of the CA. Similarly to de Visser et al. (2016) and Gong (2008), the researchers used a CA design where the profile image took a large portion of the whole interface. Furthermore, in their findings Go and Sundar (2019) mention that an ongoing exposure to a visual cue could improve the experience of the user. This further supports the argument that the less noticeable profile images used in this research could have been the reason why manipulation of anthropomorphic levels did not have a significant effect on users.

6.2 Use of Emojis

The recorded results that the use of emojis in CA has no significant effect on the user trust and satisfaction is not consistent with the research results presented by Fadhil et al. (2018). In their study, the researchers analysed a conversational agent that focused on well-being and found that the use of emojis in the CA has a positive effect on users. The theme of the conversation as well as the way the emojis were used and presented was very similar to the setup of the present study. However, Fadhil et al. (2018) measured user enjoyment, confidence, and attitude as opposed to trust and satisfaction measured in this study. Furthermore, the conversation length with the CA during Fadhil et al. (2018) experiment was longer than

the one in the present study, leading to the participants of the present study having less time to experience the CA and its anthropomorphic features. This in combination with the fact that the nature of some questions used in this study made it difficult to complement them with appropriate emojis, could be the main reasons for the variance in results between the studies.

6.3 Combining Visual Cues

Seeger et al. (2021) in their study found that using only one type of anthropomorphic cues in a CA does not increase the perceived anthropomorphism. In particular, the researchers found that using only emojis or only human identity cues, such as highly anthropomorphic profile image, would not increase anthropomorphism and in some cases might even harm it. The results of the present study could be considered in line with the findings of Seeger et al. (2021). In particular, following the conclusions presented by Seeger et al. (2021), the visual cues used in the CA of the present study could be considered not effective enough to increase the general anthropomorphism of the CA and, since anthropomorphism has been reported to have positive effects on users (Araujo, 2018; Gong, 2008), the effects of emojis and profile images on user trust and satisfaction were, therefore insignificant.

Similar reasoning can be applied to the effect of combined visual cues. Seeger et al. (2021) also showed that simultaneous employment of different cues might not necessarily result in higher anthropomorphism. This in combination with the already discussed potential issues of individual visual cues could explain why the usage of both emojis and different levels of anthropomorphic profile pictures did not produce significant effect on user trust and satisfaction.

7 Conclusions and Future Work

The research on the effects of anthropomorphism in conversational agents is still rising, however the studies cannot seem to agree on many of the raised questions and the research on the proper design of conversational agents is still lacking. This paper attempted to contribute to filling this gap by analysing whether anthropomorphic visual cues used in conversational agents have an effect on the trust and satisfaction of the users. In particular, the effects of anthropomorphic profile images and emojis were analysed in a text-based conversational agent.

The conducted experiments revealed that conversational agents that employ emojis or a profile image of any level of anthropomorphism do not provide a significantly higher satisfaction to the users nor they increase their trust. Furthermore, combining the two visual cues does not achieve such effects either. However, the limitations and observations of this research pave way for new studies that would help to better understand how and why particular interface choices might affect the users of the conversational agents. In particular, this research encourages to investigate the relationship between the amount of exposure to a visual cue and perceived anthropomorphism as well as its effects on the users. Further studies could also be done to test whether and how anthropomorphism and its effects depend on the type of task the conversational agent is performing. Lastly, emojis and their role

in anthropomorphic conversational agents could be explored even more. It would be interesting to analyse how different amounts of emojis as well as their placement and type (faces vs objects) can affect the anthropomorphism of a CA.

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, 183–189. <https://doi.org/10.1016/j.chb.2018.03.051>
- Avatar Maker. (2021). Avatar Maker. Retrieved May 27, 2021, from <https://avatarmaker.com>
- Bano, M., Zowghi, D., & da Rimini, F. (2017). User satisfaction and system success: An empirical exploration of user involvement in software development. *Empirical Software Engineering*, 22(5), 2339–2372. <https://doi.org/10.1007/s10664-016-9465-1>
- Barger, P. B., & Grandey, A. A. (2006). Service with a Smile and Encounter Satisfaction: Emotional Contagion and Appraisal Mechanisms [Publisher: Academy of Management]. *The Academy of Management Journal*, 49(6), 1229–1238. <https://doi.org/10.2307/20159829>
- Beattie, A., Edwards, A. P., & Edwards, C. (2020). A Bot and a Smile: Interpersonal Impressions of Chatbots and Humans Using Emoji in Computer-mediated Communication [Publisher: Routledge eprint: <https://doi.org/10.1080/10510974.2020.1725082>]. *Communication Studies*, 71(3), 409–427. <https://doi.org/10.1080/10510974.2020.1725082>
- Boutet, I., LeBlanc, M., Chamberland, J. A., & Collin, C. A. (2021). Emojis influence emotional communication, social attributions, and information processing. *Computers in Human Behavior*, 119, 106722. <https://doi.org/10.1016/j.chb.2021.106722>
- Catrambone, R., Stasko, J., & Xiao, J. (2002). Anthropomorphic Agents as a User Interface Paradigm: Experimental Findings and a Framework for Research, 7.
- de Visser, E. J., Monfort, S. S., McKendrick, R., Smith, M. A. B., McKnight, P. E., Krueger, F., & Parasuraman, R. (2016). Almost human: Anthropomorphism increases trust resilience in cognitive agents [Publisher: American Psychological Association]. *Journal of Experimental Psychology: Applied*, 22(3), 331–349. <https://doi.org/10.1037/xap0000092>
- Diederich, S. (2020). Designing Anthropomorphic Enterprise Conversational Agents. *Bus Inf Syst Eng*, 17.
- Diederich, S., Brendel, A., & Kolbe, L. (2019). On Conversational Agents in Information Systems Research: Analyzing the Past to Guide Future Work. *Wirtschaftsinformatik 2019 Proceedings*. <https://aisel.aisnet.org/wi2019/track13/papers/1>
- Diederich, S., Lembcke, T.-B., Brendel, A., & Kolbe, L. (2021). Understanding the Impact that Response Failure has on How Users Perceive Anthropomorphic Conversational Service Agents: Insights from an Online Experiment. *AIS Transactions on Human-Computer Interaction*, 82–103. <https://doi.org/10.17705/1thci.00143>
- Fadhil, A., Schiavo, G., Wang, Y., & Yilma, B. A. (2018). The Effect of Emojis when interacting with Conversational Interface Assisted Health Coaching System. *Proceedings of the 12th EAI International Conference on Pervasive Computing Technologies for Healthcare*, 378–383. <https://doi.org/10.1145/3240925.3240965>
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. <https://doi.org/10.3758/BF03193146>
- Freepik. (2021). Chatbot free vector icons designed by Freepik. Retrieved June 19, 2021, from https://www.flaticon.com/free-icon/chatbot_4233830?term=chat%20bot&page=1&position=13&page=1&position=13&related_id=4233830&origin=search
- Generated Photos. (2021). Generated Photos — Unique, worry-free model photos. Retrieved May 27, 2021, from <https://generated.photos/>
- Gnewuch, U., Morana, S., Adam, M., & Maedche, A. (2018). Faster is Not Always Better: Understanding the Effect of Dynamic Response Delays in Human-Chatbot Interaction. *Research Papers*. https://aisel.aisnet.org/ecis2018_rp/113
- Go, E., & Sundar, S. S. (2019). Humanizing chatbots: The effects of visual, identity and conversational cues on humanness perceptions. *Computers in Human Behavior*, 97, 304–316. <https://doi.org/10.1016/j.chb.2019.01.020>
- 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. <https://doi.org/10.1016/j.chb.2007.05.007>
- Jian, J.-Y., Bisantz, A. M., & Drury, C. G. (2000). Foundations for an Empirically Determined Scale of Trust in Automated Systems [Publisher: Routledge eprint: https://doi.org/10.1207/S15327566IJCE0401_04]. *International Journal of Cognitive Ergonomics*, 4(1), 53–71. https://doi.org/10.1207/S15327566IJCE0401_04
- Kim, S., Lee, J., & Gweon, G. (2019). Comparing Data from Chatbot and Web Surveys: Effects of Platform and Conversational Style on Survey Response Quality. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–12. <https://doi.org/10.1145/3290605.3300316>
- Lee, S., & Choi, J. (2017). Enhancing user experience with conversational agent for movie recommendation: Effects of self-disclosure and reciprocity. *International Journal of Human-Computer Studies*, 103, 95–105. <https://doi.org/10.1016/j.ijhcs.2017.02.005>
- Marsh, S., & Dibben, M. R. (2003). The role of trust in information science and technology. *Annual Review*

- of Information Science and Technology*, 37(1), 465–498. <https://doi.org/https://doi.org/10.1002/aris.1440370111>
- McKeen, J. D., Guimaraes, T., & Wetherbe, J. C. (1994). The Relationship between User Participation and User Satisfaction: An Investigation of Four Contingency Factors [Publisher: Management Information Systems Research Center, University of Minnesota]. *MIS Quarterly*, 18(4), 427–451. <https://doi.org/10.2307/249523>
- Oxford. (2021). Emoji. Retrieved April 30, 2021, from <https://www.lexico.com/definition/emoji>
- Pfeuffer, N., Benlian, A., Gimpel, H., & Hinz, O. (2019). Anthropomorphic Information Systems. *Business & Information Systems Engineering*, 61(4), 523–533. <https://doi.org/10.1007/s12599-019-00599-y>
- Rapp, A., Curti, L., & Boldi, A. (2021). The human side of human-chatbot interaction: A systematic literature review of ten years of research on text-based chatbots. *International Journal of Human-Computer Studies*, 151, 102630. <https://doi.org/10.1016/j.ijhcs.2021.102630>
- Riordan, M. A. (2017). Emojis as Tools for Emotion Work: Communicating Affect in Text Messages [Publisher: SAGE Publications Inc]. *Journal of Language and Social Psychology*, 36(5), 549–567. <https://doi.org/10.1177/0261927X17704238>
- Seeger, A.-M., Pfeiffer, J., & Heinzl, A. (2017). When Do We Need a Human? Anthropomorphic Design and Trustworthiness of Conversational Agents. Retrieved April 19, 2021, from <https://core.ac.uk/reader/301373590>
- Seeger, A.-M., Pfeiffer, J., & Heinzl, A. (2021). Texting with human-like conversational agents: Designing for anthropomorphism [Place: Atlanta, GA Publisher: AIS]. *Journal of the Association for Information Systems : JAIS*, tba, tba. Retrieved April 21, 2021, from https://aisel.aisnet.org/jais/all_issues.html
- Söllner, M., Hoffmann, A., & Leimeister, J. M. (2016). Why different trust relationships matter for information systems users. *European Journal of Information Systems*, 25(3), 274–287. <https://doi.org/10.1057/ejis.2015.17>
- Söllner, M., & Leimeister, J. M. (2013). *What We Really Know About Antecedents of Trust: A Critical Review of the Empirical Information Systems Literature on Trust* (SSRN Scholarly Paper No. ID 2475385). Social Science Research Network. Rochester, NY. Retrieved April 24, 2021, from <https://papers.ssrn.com/abstract=2475385>
- Srinivasan, S. (2004). Role of trust in e-business success [Publisher: Emerald Group Publishing Limited]. *Information Management & Computer Security*, 12(1), 66–72. <https://doi.org/10.1108/09685220410518838>
- Thong, J. Y. L., & Yap, C.-S. (1996). Information systems effectiveness: A user satisfaction approach. *Information Processing & Management*, 32(5), 601–610. [https://doi.org/10.1016/0306-4573\(96\)00004-0](https://doi.org/10.1016/0306-4573(96)00004-0)
- TU Delft. (2021). Dandelion. Retrieved April 24, 2021, from <https://dandelion1.gitlab.io/website/>
- Verhagen, T., van Nes, J., Feldberg, F., & van Dolen, W. (2014). Virtual Customer Service Agents: Using Social Presence and Personalization to Shape Online Service Encounters*. *Journal of Computer-Mediated Communication*, 19(3), 529–545. <https://doi.org/10.1111/jcc4.12066>
- Zviran, M., & Erlich, Z. (2003). Measuring IS User Satisfaction: Review and Implications. *Communications of the Association for Information Systems*, 12. <https://doi.org/10.17705/1CAIS.01205>

A Post-experiment Surveys

Table 2: Items of the user satisfaction survey with corresponding answer values.

Statement	Value						
	1	2	3	4	5	6	7
The responses of the conversational agent were appropriate	Strongly disagree	Disagree	Somewhat disagree	Neutral	Somewhat agree	Agree	Strongly agree
The appearance of the conversational agent was appropriate	Strongly disagree	Disagree	Somewhat disagree	Neutral	Somewhat agree	Agree	Strongly agree
How satisfied were you with the way the conversational agent treated you?	Very dissatisfied	Dissatisfied	Somewhat dissatisfied	Neutral	Somewhat satisfied	Satisfied	Very satisfied
How satisfied were with the way the conversational agent looked?	Very dissatisfied	Dissatisfied	Somewhat dissatisfied	Neutral	Somewhat satisfied	Satisfied	Very satisfied
How satisfied were you with the overall experience with the conversational agent?	Very dissatisfied	Dissatisfied	Somewhat dissatisfied	Neutral	Somewhat satisfied	Satisfied	Very satisfied

Table 3: Items of the user trust survey with corresponding answer values.

Statement	Value						
	1	2	3	4	5	6	7
The conversational agent is deceptive	Strongly agree	Agree	Somewhat agree	Neutral	Somewhat disagree	Disagree	Strongly disagree
I am suspicious of the conversational agent's intent, action and/or outputs	Strongly agree	Agree	Somewhat agree	Neutral	Somewhat disagree	Disagree	Strongly disagree
I am wary of the conversational agent	Strongly agree	Agree	Somewhat agree	Neutral	Somewhat disagree	Disagree	Strongly disagree
I can trust the conversational agent with my information	Strongly disagree	Disagree	Somewhat disagree	Neutral	Somewhat agree	Agree	Strongly agree
I can trust the information provided by the conversational agent	Strongly disagree	Disagree	Somewhat disagree	Neutral	Somewhat agree	Agree	Strongly agree
I am confident in the conversational agent	Strongly disagree	Disagree	Somewhat disagree	Neutral	Somewhat agree	Agree	Strongly agree
The conversational agent is reliable	Strongly disagree	Disagree	Somewhat disagree	Neutral	Somewhat agree	Agree	Strongly agree
The conversational agent is trustworthy	Strongly disagree	Disagree	Somewhat disagree	Neutral	Somewhat agree	Agree	Strongly agree