

The influence of visual appearance and conversational style of text-based chatbots on UX and future interaction intention

Master Thesis

Msc. Communication Science



August 2020

Antonela Miruna Stan

First supervisor: Dr. J. Karreman

Second supervisor: Dr. A.D. Beldad

The influence of visual appearance and conversational style of text-based chatbots on UX and future interaction intention

UNIVERSITY OF TWENTE.

MASTER THESIS

Name:	Antonela Miruna Stan
Student number:	s1798618
E-mail:	a.m.stan@student.utwente.nl
Institution:	University of Twente
Faculty:	Behavioral Management and Social Sciences (BMS)
Master:	Communication Science
Specialization:	Technology and Communication
First Supervisor:	Dr. J. Karreman
Second supervisor:	Dr. A.D. Beldad
Date:	August 28, 2020

Abstract

Purpose

Everyday communication has drastically evolved over the years from phone calls to texts and now to messaging apps. Integrating chatbot services by companies is also evolving and is in an ascendant trend. Visual appearance and conversational tone of text-based chatbots are considered factors that have major influence in determining their success. This study aims to explore the influence of visual appearance and the conversational style of chatbots on user experience (UX) and future interaction intention. In addition, the possible mediating role of social presence and the moderating role of gender have been analyzed.

Method

The study was conducted with an online experiment where the users (N= 221) had to interact with one of the four conditions of chatbots, followed by a questionnaire. A part of the online experiment participants (N= 12) also took part in semi-structured interviews. A 2x2 experiment design was used where visual appearance (human vs. logo) and conversational style (human-like vs. machine-like) were manipulated. The effects on user experience, social presence and future interaction intention have been measured.

Findings

The results show that there is no significant influence of visual appearance and conversational style of chatbots on user experience and future interaction intention. There was also not enough evidence to support the hypothesis according to which social presence is a mediator between the independent and the dependent variables. Based on the results, when interacting with the chatbot with a human-like conversational style and a human visual appearance, users did not experience higher levels of social presence, a more positive UX or stronger future interaction intention. However, the interview results show that users did perceive the chatbots as humanlike by attributing human characteristics to them (e.g. empathy, logical thinking, calmness, a happy tone, being too talkative).

Conclusion

The quantitative results show that the visual appearance and the conversational style do not have a significant influence on social presence, UX and future interaction. On the one hand, the human-like conversational style was criticised most for aspects like the length of the messages. On the other hand, the machine-like conversational style was only criticised for the limited number of damage options. The human visual appearance also received critique for not looking real, while the logo appearance did not receive critique except for one interviewee. The results can help anyone interested in chatbots but more specifically chatbot developers, copywriters and dialog designers. The results show what is perceived as important and of value when it comes to visual appearance and conversational style. This can be used to develop chatbots more effectively and efficiently. More effective because it can be used to produce the wanted results, e.g. better satisfy the needs of the end-users. More efficient because resources can be allocated more specifically, e.g. time and money can be spend on what is considered important.

Keywords: Text-based chatbot, conversational style, visual appearance, UX, social presence, future interaction

Table of Contents

1. Introduction	6
2. Theoretical Framework	8
2.1 Anthropomorphism in the Chatbots era	8
2.2 User Experience (UX)	9
2.3 Future interaction intention	10
2.4 Visual appearance of chatbots	10
2.5 Conversational style of chatbots	11
2.6 Congruity between visual appearance and conversational style	12
2.7 The mediating role of Perceived Social Presence	12
2.8 The moderating role of gender	13
2.9 Research model	14
3. Methodology	15
3.1 Research Design	15
3.2 Stimulus material	15
3.3 Pre-test	17
3.4 Participants	18
3.5 Procedure	19
3.6 Measurements	20
3.6.1 Questionnaire	20
3.6.2 Interviews	21
3.7 Construct validity and reliability	22
3.7.1 Questionnaire	22
3.7.2 Interviews	24
4. Results	25
4.1 Results of the online experiment	25
4.1.1 Main effects	25
4.1.2 Interaction effects	27
4.1.3 Mediating role of social presence	28
4.1.4 Moderating effect of gender	28
4.2 Overview of the hypothesis	29
4.3 Results of the interviews	29
5. Discussion	33
5.1 Discussion of the results	33
5.2 Implications	35

5.3 Limitations and future research	37
6. Conclusion	39
7. References	40
8. Appendixes.....	49

1. Introduction

Chatbots, also called conversational agents (CAs) are software-based systems designed to interact with humans using natural language (Dale, 2016; McTear et.al.,2016). They can take different forms: text based, voice based, 3D, or even embodied forms. CAs are used in different fields such as: retail, healthcare, HR or education. They can help users find information or help to perform different routine tasks such as tracking inventory, making appointments or scheduling interviews (Feine, Gnewuch, Morana, Maedche, 2019).

Today, many companies are implementing chatbots as an extension of the services they provide to their customers. Having a good information flow in the customer service process is of extreme importance for the success of any business. Text based chatbots seem to take over because they provide benefits for companies as reduction of the response time and work overload, enhanced customer service, increased satisfaction and engagement (Radziwill & Benton, 2017; Rietz, Benke & Maedche, 2019). According to research, 57% of the companies already use or plan to implement a chatbot in the near future (Wang et. al., 2017).

Organizations opt for chatbots because their purpose is to provide efficient and fast service. They use platforms as Facebook Messenger, Slack and Skype among others that support the hosting of chatbots in the interaction with their users (Smestad, 2018).

A category of chatbots function on AI, which makes them capable of understanding natural language and they are also capable of getting smarter as they interact more due to their ability to maintain different states (Kar & Halder, 2016). However, at the moment most of the chatbots function based on rules, which are limited to be as smart as they are programmed to be and even like this, they are very helpful. For example, they are able of simplifying the way we search for information from multiple screens and physical materials (e.g. handbooks, catalogs) to simple conversational interfaces capable of delivering highly contextual and intelligible information within the flow of a chat app that has as a final result a good user experience (UX).

In order to determine the credibility of CAs, scholars have created taxonomies of social cues for chatbots. Social cues, are nonverbal characteristics associated with humans, for example they can make jokes, have a gender or even have facial expressions (Go & Sundar, 2019). Feine, Gnewuch, Morana and Maedche (2019) came with a classification of the social cues which are organized in four big categories: verbal, visual, auditory and invisible cues.

UX also has a subjective nature since it deals with the individual`s perception and thoughts (Nielsen & Norman, 2018). In this sense, chatbots with conversational interfaces and anthropomorphic cues, can easily engage users in performing tasks due to the fact that they pose humanlike features. Based on previous research, people tend to ascribe social attributes to computer interfaces (Reeves & Nass, 1996; Nass & Moon, 2000; Mimoun, Poncin, Garnier, 2017). In most academic literature the attention falls on usability and the ability of the chatbot to recognize the details in a user`s inquiry (Claessen, Schmidt, Heck, 2017). On the other hand, the user experience with chatbots based on the bot`s visual appearance and language aspects has not been debated enough and this is a reason why this research wants to explore this domain by actually engaging the user in an interaction with the chatbot and further analyzing the experience.

Designers implement elements as humanlike visual appearance and humanlike conversational style in order to compensate for the lack of social presence in the online environment. Research in this field has also found that social presence has a big positive impact on the user experience and overall user satisfaction (Park, Cho, Lee, 2019).

Visual appearance of a disembodied conversational agent is an important factor because it can influence the user to perceive it as being human-like or not (Araujo, 2018). Chatbots with a human visual appearance (real human or animated picture) are increasingly being used by big Dutch online operators such as ING, Ziggo, Blokker or the big retailers like Wehkamp (Beldad, Hegner & Hoppen, 2016). Furthermore, according to literature, the conversational style of chatbots is also a very important factor which determines the success of the interaction (Mehrabian & Ferries, 1967b). According to the literature, it is perceived as human-like when it shows empathy, affection, makes use of emoticons and is rather informal (Warwick & Shah, 2016; Liebrecht & van Hooijdonk, 2019).

The distant and computer-mediated nature of the internet, cause people to feel a diminished personalized approach. The role of anthropomorphic chatbots is to decrease the discrepancy created between human to human interaction and human to computer interaction (Go & Sundar, 2019). Designers are still facing the challenge of creating chatbots that can feel just human enough and are able to create a social presence feeling for the user which would further determine a good user experience and the intention to interact again with the software agent (Brandtzaeg & Følstad, 2017).

Due to the lack of empirical studies based on the actual interaction of participants with chatbot conditions, this study has the purpose to further research the implications of anthropomorphic cues of chatbots such as the influence of visual appearance and language style on user experience and future interaction intention in their communication with users. Another attribute of this research is that the participants had the opportunity to interact with unique chatbot conditions which were created and custom made for this research. The conditions incorporate humanlike and machinelike characteristics which were selected based on previous literature recommendations. An extra quality of this research is represented not only by its quantitative nature but also by its qualitative nature. The semi-structured interviews have the goal to collect more detailed information about the participants' user experience and the future interaction intention.

Based on the literature the following central research questions have been proposed:

RQ1: To what extent do visual appearance and conversational style of a text-based CA influence the UX and future interaction intention?

RQ2: To what extent are visual appearance and conversational style of a text-based CA on UX and future interaction intention mediated by social presence?

RQ3: To what extent are visual appearance and conversational style of a text-based CA on UX and future interaction intention moderated by gender?

RQ4: How do visual appearance and conversational style of a text-based chatbot interact with each other?

2. Theoretical Framework

The following sections describe the literature findings related to the CAs and the manipulations in visual appearance and conversational style with the purpose of understanding the way these elements can influence the social presence, the UX and the future interaction intention.

2.1 Anthropomorphism in the Chatbots era

Today, chatbots are so advanced that they are able to replace human agents in different fields as online-tutoring, customer-service and even cognitive therapy (Go & Sundar, 2019). Over the years, the technological developments from the first created virtual agent to the latest ones went through a long process. For example, ELIZA (Weizenbaum, 1966) and Mitsuku (Worswick, 2017) are both chatbots designed to mimic human behavior in text-based conversation, but the technological evolution can be seen while comparing the two chatbots. ELIZA was one of the first chatbots and considered to be the best back in the sixties; she was created to use natural language by using a template-based response mechanism in order to match a psychotherapist's conversational style (Weizenbaum, 1966). Recently, Mitsuku was deemed the most humanlike chatbot in the world by a judging panel. Mitsuku is based on machine learning and she helps people to deal with loneliness (Balch, 2020). Both chatbots are considered the best of their times, ELIZA back then and Mitsuku today. But the difference in the technology they use is huge and is based on decades of research and developments, which include plentiful open source code, different development platforms or implementation options via Software as a Service (Radziwill & Benton, 2017).

Today, chatbots are more advanced than back then and they become part of humans' everyday life. The way the customer service process looks has dramatically changed over the years. Chatbots have been introduced in fields as banking, commerce, government, education or technical support (Ferrara, Varol, Davis, Menczer & Flammini, 2016; Di Prospero, Norouzi, Fokaefs & Litoiu, 2017; Kowatsch, Nißen, Shih, Rüegger, Volland, Filler & Heldt, 2017; Simonite, 2017). Having a good information flow in the customer service process is of extreme importance for the success of any business. Text based chatbots seem to take over because they provide benefits for companies as reduction of the response time and work overload, enhanced customer service, increased satisfaction and engagement (Radziwill & Benton, 2017; Rietz, Benke & Maedche, 2019).

The Computers are Social Actors paradigm promotes the idea that during human- computer interaction, humans tend to respond to computers in a social way which is very similar to how they respond to other humans. This is even the case if they are aware of the fact that they are interacting with a computer (Nass, Steuer & Tauber, 1994). Humans tend to apply social rules in their conversations with chatbots when social cues are integrated in the conversational style (Liebrecht & Hooijdonk, 2019) as well as in the visual appearance (Qui & Benbasat, 2009). These type of social cues have a further positive influence on user experience and loyalty towards the chatbot (Hassaein & Head, 2007; Gefen & Straub, 2004).

When designing customer service chatbots, the anthropomorphism aspects are being addressed. In the HCI field, the anthropomorphism of chatbots is considered to be accounted

for and acknowledged when designing interfaces (Caporael, 1986). The idea of making virtual agents more human-like, stems from the Computers Are Social Actors paradigm (CASA). CASA states that social attributes are being ascribed to computer interfaces while interacting with humans (Reeves & Nass, 1996; Nass & Moon, 2000). This paradigm has been proven by several studies that were testing in detail the mindless responses but also the depth of the social responses, to the computer`s personality (Nass & Moon, 2000). The paradigm was further explored by Polzin and Waibel (2000). Based on studies, they concluded that computer interfaces should respond to this theory. More recent research also shows that anthropomorphism in chatbot design positively influences purchase intention, trust and required information input (Schanke, Burtch & Ray, 2020).

An important aspect that designers need to keep in mind is the uncanny valley effect. The human-computer interaction domain has had for decades as a major goal the design of natural and intuitive interaction modalities (Pirrone, Russo, Canella, Peri, 2008). In order to have a human-computer interaction that results in a good user experience, designers have to avoid the uncanny valley effect. If robots become too humanlike, they create the risk of inducing an uncanny feeling to the user which can be described as dislike, unease and unpleasantness (Mori, Macdorman, Kageki, 2012). Conversating with a software agent that pretends to be humanlike will be pleasant up to the point the user perceives the robot as too humanlike. At that point an abrupt shift in affinity will appear and the user will experience the interaction as unpleasant or strange (Skjuve, Haugstveit, Følstad & Brandtzaeg, 2019).

During the last decades, conversational systems have been created to simulate how a human would behave as a conversational partner (Schanke, Burtch & Ray, 2020). If a judge can not distinguish between the human and the machine participant they pass the Turing test. It would logically follow that when a machine can pass the Turing test the uncanny valley effect will not occur as long as the participant doesn`t know she or he is conversating with a machine. The Turing test was created to test the ability of online agents using natural language to converse with humans (Turing, 1950).

Because the virtual agents with incorporated anthropomorphic elements are so present in humans everyday life, this study especially analyses aspects as visual appearance and conversational style of text based chatbots and their impact on the user experience and the future interaction intention.

2.2 User Experience (UX)

According to the Nielson Norman Group, one of the oldest user experience consulting firm, UX is described as the feeling users have while interacting with a company, its services and its product (Norman, 2018). The ISO standard (2015), describes UX as being composed of all the aspects of usability and desirability of a product from the user`s perspective. Zarour and Alharbi (2017) also mention two terms that include the aspects that have influence on the usability and user experience: the pragmatic quality, which is directly related to the execution of a task, thus to the usability of a product and the hedonic quality, related to the intrinsic values of each user and their subjective perceptions, which refers to the user experience. It can be concluded that the user experience goes beyond the usability of a product, and is a broader

concept that is not only focused on the functionality of a product but also on the feelings of the user and the whole experience with the product.

A good user experience is crucial for the success of a chatbot, taking into consideration that a captivated user may repeat the interaction. According to Webcredible (2009) after interacting with a company, the user will leave with a positive or a negative emotion towards the brand. A satisfying experience is one that meets the particular needs of the user. The dynamism of the UX is related to the internal state of the user which can be modified by different aspects of use during and after the interaction with a product (Vermeeren et. al., 2010). The attention paid on various disciplines, including marketing, ethnography, interaction design, information design, technical writing and visual design will deliver in the end a good user experience (Sward, 2006).

2.3 Future interaction intention

Based on the Theory of Planned Behavior (Ajzen, 1991), the behavioral intention is the most influential predictor of the actual behavior. Indeed research also shows that the relationship between the behavioral intention to revisit an online website and the actual behavior is very strong (Jung, Kim & Kim, 2014). Further, a favorable attitude towards a virtual agent with humanlike elements, results in a greater behavioral intention for future interaction (Koda, 1996; Wexelblat, 1998; Sundar et al. 2016; Go & Sundar, 2019).

According to previous research, in online environments, revisit intention depends on the experience perceived by the user (Sivadas, & Baker-Prewitt, 2000; Brady & Robertson, 2001; Kabadayi & Gupta, 2011; Moriuchi, Landers, Colton & Hair, 2020).

2.4 Visual appearance of chatbots

Koh and Sundar (2010) have argued that the usability quality of a CA might not matter in evaluating the agent`s performance if the identity assigned to the agent is based on stereotypical judgements. Thus, if an agent is presumed to be a chatbot, the users are more likely to evaluate its performance based on their existing preconceptions about robots and machines. On the other hand, the use of human identity cues, will determine the users to evaluate the quality of the interaction with the chatbot based on their expectations of humans (Go & Sundar, 2019). This being said, designing and developing disembodied chatbots is about understanding what the user needs and the motivation behind his actions. The visual aspect of a chatbot has a big impact on the user`s behavior and designers have to take that into consideration when creating new conversational user interfaces (Appel, von der Pütten, Krämer & Gratch, 2012)

According to the social presence theory, a feeling of human contact can be created without the actual human contact (Gefen & Straub, 2004). In computer-mediated environments it is required to enhance and foster online interactions by creating a sense of connection between a chatbot and its users (Traphagan et. al., 2010). Human visual appearance of chatbots, contribute in creating the social presence effect mentioned above and this further creates a loyalty feeling that the users develop while interacting with the chatbot (Qui & Benbast, 2009; Følstad, Nordheim & Bjørkli, 2018; Rietz, Benke & Maedche, 2019). The creation of online social presence effect by making use of the human visual appearance of chatbots is considered

to be of extreme importance when building customer loyalty and a high level of customer experience (Gefen & Straub, 2003).

In order to compensate the lack of social presence in the online environment, designers implement human avatars for chatbots. An international study with 7000 participants, revealed that 46 percent of the customers want to interact with a chatbot that adopts a humanlike appearance while only 20 percent would like to see them as an avatar (Singh, 2017). Same study reveals that 36 percent of the participants prefer a female CA and only 14 percent would choose a male CA. The humanlike appearance seems to be preferred. Since based on the previous studies, the visual appearance has such an impact on the perception of the agent; relating this findings to the context of the present study, it is expected that the chatbots that adopt a human visual appearance will have a more positive effect on user experience, will determine a more positive effect on future interaction intention and a higher level of perceived social presence. Therefore, the next hypotheses are being proposed:

H1a: The chatbot with a human visual appearance will have a more positive effect on UX, than a chatbot that is not represented by a human visual appearance.

H1b: The chatbot with a human visual appearance will have a more positive effect on future interaction intention, than a chatbot that is not represented by a human visual appearance.

H1c: The chatbot with a human visual appearance will determine a higher level of perceived social presence, than a chatbot that is not represented by a human visual appearance.

2.5 Conversational style of chatbots

Today, there is an abundance of natural language interaction on the internet between humans and conversational agents, in contexts of customer service, marketing, e-commerce, e-learning, e-health etc. Because so much of this communication occurs using the digital technology rather than in person communication, this subject became an important area of exploration and research of this simulation of natural human language (Hill, Ford, Ferrares, 2015). Computer mediated communication (CMC) differs from spoken communication in its lack of body language, communicative pauses, and vocal tones (Hentschel, 1999). Despite this absence of specific social cues, CMC has been found to be able to communicate emotions as well as or sometimes even better than face-to-face communication (Derks, Fischer, & Bos, 2008).

The conversational style of a chatbot can influence the perception about of the experience with that chatbot. Humans desire to have a natural experience with computers that create a human-like feeling by means of conversational style (Araujo, 2017; Singh, 2017; Garcia, 2017). There are several factors that influence the human conversational style of chatbots, such as: typography styles, word frequency and responsiveness (Skjuve, Haugstveit, Følstad & Brandtzaeg, 2019). In the same note, Liebrecht and van Hooijdonk (2019) found through their research several key linguistic elements with increased anthropomorphism that can be incorporated in the human-like chatbot conversational style, these elements are: empathy, support, humor, informal attitude.

Other empirical research shows that the personality of a voice chatbot impacts user's perception and willingness to further interact with the virtual agent (Callejas, López-Cózar, Ábalos & Griol, 2011). In addition, it is considered that agents with a human-like conversational style are more likely to lead to a better user experience than those with a robotic conversational style. In previous research, users rated the human-like style better than the machine-like one (Hu et al., 2018). Furthermore, the perceived social presence level is increasing due to the back-and-forth nature of high message interactivity which is perceived as a dialog, the core element of human to human communication (Go & Sundar, 2019). Relating the findings mentioned above to this study, one could claim that chatbots that use a humanlike conversational style which include elements of human to human conversation have a positive influence on the user experience and on future interaction intention. Next, the perceived social presence level is also being positively affected by the presence of the human like elements in the conversational tone of text based chatbots. Therefore, the next hypotheses are being proposed:

H2a: The chatbot with a human-like conversational style will have a more positive effect on UX, than a chatbot that is using a machine-like conversational style.

H2b: The chatbot with a human-like conversational style will have a more positive effect on future interaction intention, than a chatbot that is using a machine-like conversational style.

H2c: The chatbot with a human-like conversational style will determine a higher level of perceived social presence, than a chatbot that is using a machine-like conversational style.

2.6 Congruity between visual appearance and conversational style

Previous research shows that the use of a human conversational style and a human visual appearance on chatbots result in triggering conscious evaluations of the chatbot as being humanlike (Laurel, 1997; Brahnam, 2009; Warwink & Shah, 2016; Araujo, 2017). Furthermore, a research by Baylor and Rosernberg-Kima (2006) showed that in an experiment, the presence of a human animated visual appearance and an apologetic or empathic message when an error message popped-up lead participants to attribute the cause of their frustration on the technology and not on themselves. Thus, this research will also investigate the relationship between the visual appearance and the conversational style of chatbots and analyze if there is an interaction effect between the two independent variables.

2.7 The mediating role of Perceived Social Presence

According to research based on the social presence theory, mediums that score high in social presence are more appropriate in carrying out interpersonal tasks (Steinfeld, 1986; Rice, 1993; Xu & Lombard, 2017). In their research Kear, Chetwynd and Jefferis (2014) also found that it is of value to create a social presence feeling in online mediated communication and that human profile pictures helped students feel more comfortable texting with other students. Janson, Degen and Schwede (2019) also found that chatbots that have anthropomorphic design elements influence the social presence feeling of users in a positive way.

Further, a high level of social presence helps visitors to establish loyalty towards e-Service websites and to feel more satisfied with the whole online experience (Cyr, Hassanein, Head and Ivanov, 2007; Etemad-Sajadi & Ghachem, 2015; Lu, Fan, Zhou, 2016; Meadows, 2017). In addition, one of the most recent studies concluded that chatbots should have a high level of social presence and that this will result in a more positive user experience (Hendriks, Ou, Amiri & Bockting, 2020). Based on the aforementioned literature findings, it is expected that the possible effects of visual appearance and conversational style on the user experience and future interaction intention are mediated by social presence. Therefore, the next hypotheses are being proposed:

H3a: The effects of visual appearance and conversational style on UX will be mediated by social presence.

H3b: The effects of visual appearance and conversational style on future interaction intention will be mediated by social presence.

2.8 The moderating role of gender

With regards to the gender of the user in the interaction with embodied chatbots, previous research has found that female users benefit from and are more sensitive to the nonverbal behavior while men are not affected that much by this type of behavior (Foster, 2007; Krämer, Hoffmann & Kopp, 2010).

Surprisingly, even though it is difficult to transmit nonverbal behavior through chat based conversation, further research shows that during the years, women became more frequent technology mediated users and are more interested in mediated communication than men are (Kimbrough, Guadagno, Muscanell & Dill, 2013). During research on text-based chatbots, female users used to rate the conversations they had more favorably than men did (Shah, Warwick, Vallverdú & Wu, 2016; Brandtzaeg & Følstad, 2017). Another research on user experience with chatbots, also shows that nearly 50% of the female online shoppers reported to enjoy using chatbots as a channel of communication, while only 36 % of the male users do the same. Further, almost 35% of men use chatbots if they can not find answers to simple questions while female users mostly use the virtual agents for online purchases (Jovic, 2020). For these reasons, the influence of gender as a moderator will be measured. It is assumed that gender will moderate the possible effect of visual appearance and conversational style on social presence, UX and future interaction intention. Therefore, the next hypothesis are being proposed:

H4a: The possible effects of human visual appearance and human conversational style on UX will be moderated by gender.

H4b: The possible effects of human visual appearance and human conversational style on future interaction intention will be moderated by gender.

2.9 Research model

Based on the literature findings and the hypothesis, the following research model is being proposed in Figure1:

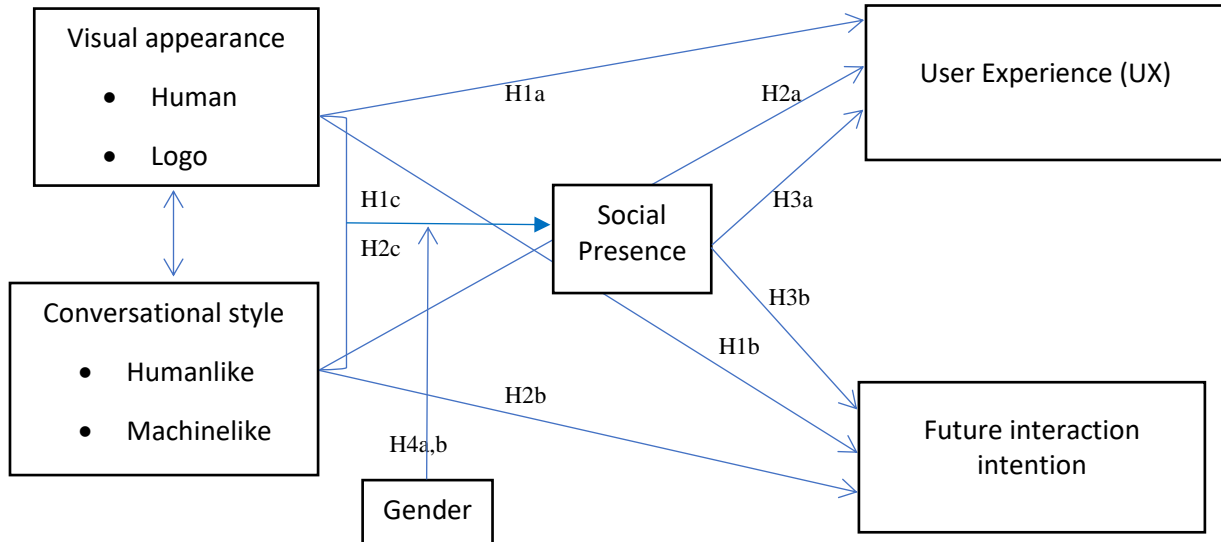


Figure 1. Research model

3. Methodology

In this section the methods and research instruments are elaborated and justified. The research design is being presented and tested by means of data collected through a combination of quantitative (online experiment) and qualitative (interviews) research.

3.1 Research Design

In this research an online experiment took place followed by semi-structured interviews for a small part of the participants. The online experiment had a 2 (visual appearance: human vs. logo) x 2 (conversational style: humanlike vs. machinelike) design. The participants were automatically assigned to one of the four conditions where the independent variables were manipulated. Table 1 presents the experimental conditions.

With the purpose of collecting more in-depth information, about the effects of the independent variables on the dependent variables interviews were conducted with some participants that have interacted with the chatbot conditions and have already completed the online experiment.

Table 1. Experimental conditions

Experimental conditions:

Conditions	Conversational style	Visual appearance
Condition 1	Humanlike (HC)	Logo
Condition 2	Machinelike (MC)	Logo
Condition 3	Humanlike (HC)	Human (HV)
Condition 4	Machinelike (MC)	Human (HV)

3.2 Stimulus material

For testing the four conditions, two different chatbot conversational styles were created and used. The humanlike conversational style included the key linguistic elements with increased anthropomorphism suggested by Liebrecht and Van Hooijdonk (2019): empathic, supportive, with humour, makes use of emoticons and rather informal. The machinelike conversational style did not include these anthropomorphic elements and is rather objective, straight to the point, with short answers, rather formal and basically the opposite of the model created by Liebrecht and van Hooijdonk (2019). For visual appearance a female human picture and a logo picture were used. The choice for a female picture is based on the pre-test where the majority of the users, 66.6% preferred a female looking avatar, while only 33.3% opted for a male looking avatar. After the gender choice was made, the chatbots that adopted a humanlike conversational style were named ‘Sarah’ while the agents that used machine-like conversational style displayed no name, as presented in Figure 2.

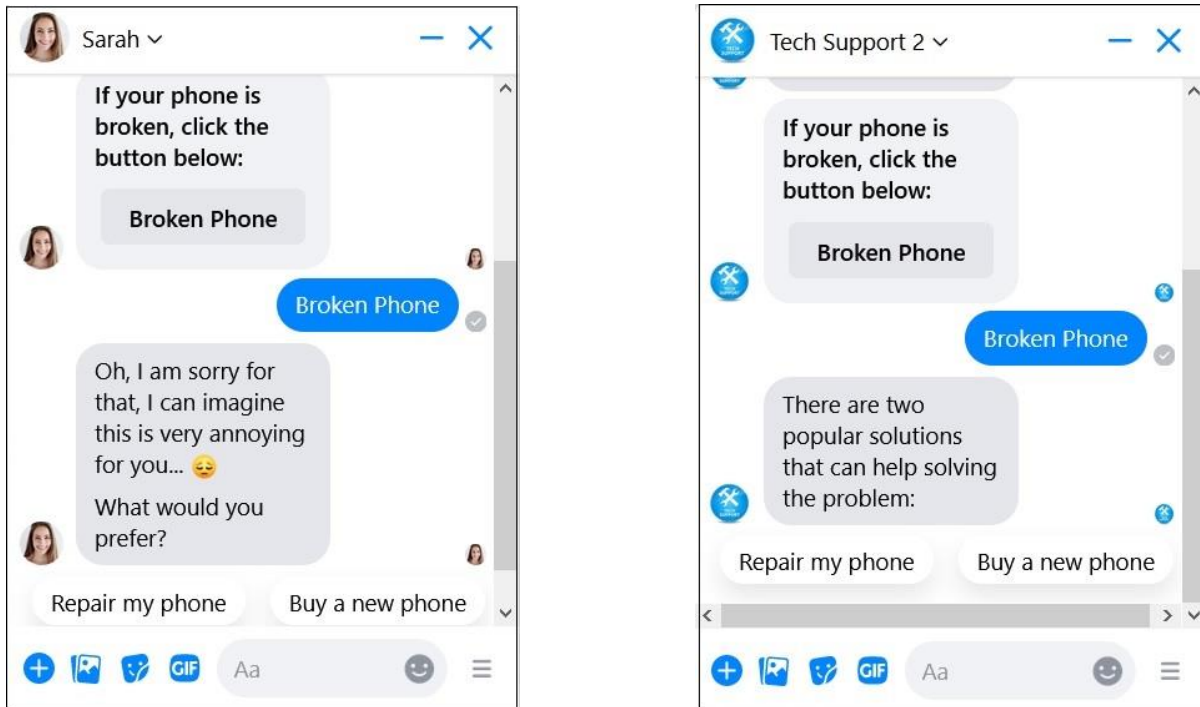


Figure 2. Showing a humanlike conversational style with a human visual appearance (left) and a machinelike conversational style with a logo appearance (right).

For measuring the effects of the independent variables, visual appearance and conversational style on the dependent variables, UX and future interaction intention, four different chatbot conditions were created. The conditions have been created with the help of the Snatchbot.me tool and for enabling the participants to take part in the experiment the chatbots have been connected to the Facebook platform where four additional pages were created. Through Messenger (an additional application which is connected to Facebook) the participants could interact with the chatbots. The participants have been randomly assigned to one of the four conditions. Following the advice of Fincher (2018) and Feng (2019) buttons have been assigned to each chatbot interface in order to avoid miscommunications during the human-computer interaction and to keep the interaction short and effective. One of the four Facebook pages and the dashboard with the four conditions can be seen in Appendix E and F.

The scenario presented a situation where the mobile phone of the participant was broken. Based on a friend's advice, the participant entered the Facebook page of a mobile phone service company and engaged into a conversation with the chatbot. The chatbot offered two options: "Repair my phone" or "Buy a new phone". The participant could choose the brand of the phone "Samsung", "Apple" or "Other" and the affected part of the gadget "Screen", "Battery" or "Charging Spot". The chatbot offered information about the price range of the reparation or the price of a new phone based on the brand choices made by the participant. When the choice was "Buy a new phone" the agent led the person to a website from a well-known Dutch electronics company "Media Markt" where the participant could choose a new phone. When the "Repair my phone" option was chosen the virtual agent led the participant to a local phone service shop in Enschede called "GSM Reparatie XL". In Figure 2 the difference in conversational style and visual appearance of the virtual agents is being presented. The conversation flow can be found in Appendix I.

3.3 Pre-test

In order to pinpoint eventual problem areas a pre-test was conducted. The respondents had to interact with one of the chatbot conditions. First condition had anthropomorphic elements included in the text and the second condition did not include these elements. Six of the participants interacted with the first condition and six with the second condition. After the interaction with the virtual agent the respondents indicated their preferences. The agent which had incorporated in the text the anthropomorphic elements of Liebrecht and van Hooijdonk (2019) received a higher score for anthropomorphism ($M = 4.30$, $SD = 0.20$). The chatbot that did not include the anthropomorphic elements in the messages received a lower score for anthropomorphism ($M = 2.63$, $SD = 0.34$).

In total 12 participants were part of the pre-test, 50% male and 50% female with the age ($M = 30$, $SD = 5.86$); level of education 58.3% Master, 16.7% Bachelor, 8.3% Highschool, 8.3% Pre-Master, 8.3% PHD. By making use of the Skype and WhatsApp call they had to verbalise the process they went through during the interaction (the think out loud protocol). The attendants gave their consent to be part of the pre-test and proceeded with the interaction. During the interaction information was gathered. Based on the received feedback, adjustments in the conversation flow of the chatbot were made when the answers were not clear enough or did not follow a logical path from the participant's point of view. The main key comments are mentioned on a list in the Appendix B.

Measuring the level of anthropomorphism in the conversational style:

For measuring the anthropomorphism in the conversational style the 5-point semantic differential scale of Bartneck, Croft and Kulic (2008) has been used. The three researchers define anthropomorphism as the attribution of a type of human form, characteristic or behaviour to nonhuman objects such as robots, computers or animals (Bartneck, Croft & Kulic, 2008). The scale is composed of five semantic differential items: Fake/Natural, Machinelike/Humanlike, Unconscious/Conscious, Artificial/ Lifelike and Moving Rigidly/ Moving Elegantly. It reported a Cronbach's Alpha value of .85; the alpha value is well above .7 thus it can be concluded that the scale has sufficient internal consistency reliability.

In order to find out if the difference between the two groups is significant, an independent sample T-test has been performed. The T-test result of $t(10) = -10.12$, $p < .001$ shows that in this sample, the two different groups do differ in the perceived anthropomorphism, and the H_0 which assumes equal variance can be rejected. Thus we can conclude that humanlike conversational style is perceived as being different from machinelike conversational style.

Choosing the gender of the agent:

After the interaction took place, the participants were asked to choose a gender for the chatbot they have just interacted with. In total 83.3% of the male participants preferred a female looking agent and 16.6% preferred a male photo. 66.6% of the female respondents chose a female and only 33.3% opted for the male. In total, 66.6% of the participants preferred a female looking avatar, while only 33.3% opted for a male looking avatar. Furthermore, this

results are in line with the findings of the creator of Sophia, Ben Goertzel from Hanson Robotics which stated that female robots are more popular and this is the reason why together with his team they created a young adult female CA which is eager to interact with humans (Goertzel, 2018).

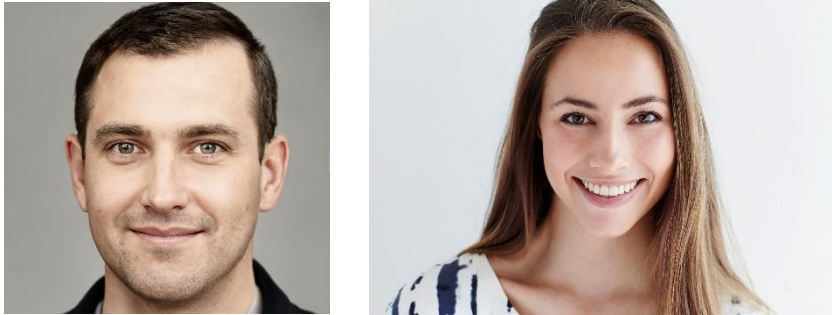


Figure 3. Male and female looking avatar

3.4 Participants

This study was focused on subjects that have a Facebook account and have the Messenger application installed. The Snatchbot platform offers 3 convenient options for the users to connect the chatbots to: Facebook, Slack and E-mail. Facebook was the most convenient choice for this research because it is a largely used platform where participants could easily get access the agents. For collecting participants, the snowball sampling method has been applied. The anonymous link has been shared through Facebook, LinkedIn, WhatsApp, Instagram. After the online experiment took place, a semi structured interview with 12 of the participants which were asked via Facebook Messenger if they wated to participate and were selected aleatory was conducted for collecting more in depth data. From the 385 participants, 140 did not finish the questionnaire. From the 245 of the remaining participants which did finish the survey, 24 were eliminated because they filled in the survey in less than three minutes. Three minutes were considered the minimum time needed to fill in the survey. The total final respondents which had their data used in the analysis were 221.

The participants are a total of (38.9%) men and (60.6%) women. Most of the participants have an age between 25 and 34 (46.2%) followed by the participants with an age range between 18 and 24 (24.4%). The highest level of education of the biggest group of the participants is master (47.1%), followed by bachelor (31.2%).

Table 2. Demographics across conditions:

Condition	N =	Age	Gender
Human conversational style + Human visual appearance	60	46.7% (25-34) 26.7% (18-24) 10% (35-44) 1.7% (45-54)	33.3% (m) / 66.7% (f)
Human conversational style + Logo	54	48.1% (25-34) 29.6% (18-24) 14.8% (35-44)	40.7% (m) / 57.4% (f)/ 1.7 (other)

Machinelike conversational style + Human visual appearance	51	41.2% (25-34) 17.6% (18-24) 5.9% (45-54)	39.2% (m) / 60.8% (f)
Machinelike conversational style + Logo	56	48.2% (25-34) 23.2% (18-24) 10.7% (35-44) 1.8% (45-54) 1.8% (65 or more)	42.9% (m) / 57.1% (f)

3.5 Procedure

The task that the participants had to perform was to interact with one of the chatbot conditions and to explore its possibilities. Based on a fictive scenario, the participants had to search for solutions after their phone broke and to inform the chatbot about their problem. The chatbot offered solutions by making use of different language styles and visual appearances.

After the participants accessed a link, agreed to join the experiment, read the scenario and the instructions they were transferred to the Facebook page where one of the four conditions of the chatbot was presented. An example of one of the pages can be found in Appendix E. Based on the experience they had with the chatbot all the participants had to fill in a questionnaire and some of them to be part of a short interview.

The subjects that were also part of the interview were informed and gave their approval to be audio recorded. They were directed to choose an emocard which represented their mood determined by the interaction they just had and based on the choice a set of questions were asked. The interviews took 15 minutes on average.

Next, after the raw quantitative data file was exported and prepared for the analysis, several statistical procedures in the SPSS software programme were performed. The more detailed analysis is further presented and explained in the paper.

The files which included audio recordings and memos have been prepared for analysis based on the steps advised by Boeije (2009). The audio recordings have been entirely transcribed in Microsoft Office WORD. Further, based on the transcripts the data was analysed in detail and segmented. Unclear, personal or irrelevant has been left out and the use of [...] marks this procedure. This process helped in making distinctions between the relevant fragments.

The resulted categories have been coded and a list of codes has been created. Finally, the data has been reassembled and the results were organised in categories (advantages, benefits, disadvantages, changes, future interaction intention). The list of codes can be found in Appendix H.

3.6 Measurements

In order to acquire a clear overview about the user experience of a product, both qualitative and quantitative data should be collected (Abel, 2010). UX evaluation represents more than just achieving the practical goals, it should also measure how participants feel about a product (Harpur, 2013). Based on these reasons, the research focused on analyzing also the nuances of UX by collecting qualitative and quantitative data from the participants.

3.6.1 Questionnaire

For building the questionnaire, the online platform Qualtrics was used. The survey was created in English and consisted of 43 items including statements, questions, semantic differential elements and demographics. The link with the questionnaire was shared through different social media channels (Facebook, LinkedIn, Instagram, Email, WhatsApp). The snowball sampling method was used for a more effective collection of participants. Because of the nature of the online experiment the participants had to respect two essential requirements: to be Facebook users and to have installed on their mobile phones the Facebook Messenger application. No other conditions for participation have been added.

For measuring the UX the Attrakdiff questionnaire developed by Hassenzahl, Burmester, Koller (2003) has been used. The questionnaire includes 7-point semantic differential scales which measure four dimensions that make up the UX. The scale is one of the most used user experience questionnaires, apart from the self-developed questionnaires (Bargas-Avila & Hornbaek, 2011).

In total 28 items are used to measure the user experience of a product. The dimensions are: pragmatic quality (*PQ*), hedonic quality - identity (*HQ - I*), hedonic quality - stimulation (*HQ - S*) and attractiveness (*ATT*). *The four different dimensions are further explained below:*

Pragmatic Quality:

Refers to how easy it is for the user to manipulate a product. Thinking pragmatically, requires the product to accomplish its meaning in helping the user to fulfill its goal (Bevan, 2008).

Hedonic identification:

Are the attributes that determine the users to identify themselves with the product in a social context. What do we transmit to other people by using this product? The products that help transmit what the user thinks is advantageous to others are preferred (Bevan, 2008).

Hedonic Stimulation:

Is related to the attributes of the product that allow the user to further develop its skills and knowledge. These attributes encourage the personal growth of the user. These are the features that are not used by the user but they represent a possibility of further development. (Bevan, 2008).

Attractiveness:

Users judge a product by summarizing the whole experience they had while interacting with the product. The global appeal of a product is measured through this dimension (Beven, 2008).

For measuring the social presence in this study a scale originally constructed by Gunawardena and Zittle (1997) has been used. In their research, social presence was used as a predictor of satisfaction within computer-mediated conferencing environments. The original scale has been adapted by adjusting the statements in order to address the human-chatbot interaction, and a 7-point Likert scale (1 = Strongly Disagree to 7 = Strongly Agree) has been used.

For measuring the future interaction intention, four statements were created and same 7-point Likert scale (1 = Strongly Disagree to 7 = Strongly Agree) scale has been used. The entire questionnaire can be found in the Appendix C.

To make sure that the order of the statements and questions do not influence the answer of the respondents and that they interpret them correctly a pre-test was conducted. In total a number of 11 respondents filled in the questionnaire. Based on their feedback phrasing of the statements and questions has been adjusted.

3.6.2 Interviews

For supporting and clarifying the results from the online experiment, 12 subjects (three per condition) which interacted with one of the agents and filled in the survey participated in a short semi-structured interview which took place via Skype immediately after the interaction with the agent took place.

Through the interviews the subjects that have already participated in the online experiment had the chance to express in their own words the opinion about the interaction with the chatbot and could explain their mood by making use of the emocards created by Desmet et. al. (2001). They picked a card that expressed how they feel about the interaction and gave argumentation for it by answering the interview questions mentioned at the end of this paragraph. The open-ended questions offered the participants the opportunity to express their opinion related to their experience. The questions were based on the dependent variables. The interviews took place in English, Romanian and one of them in Dutch. The interview answers have been translated entirely to English. No limitations were imposed in selecting the participants which are of different ages and have different backgrounds. Considering the gender, 50% of the participants are women and 50% are men with a mean age of 31.9.

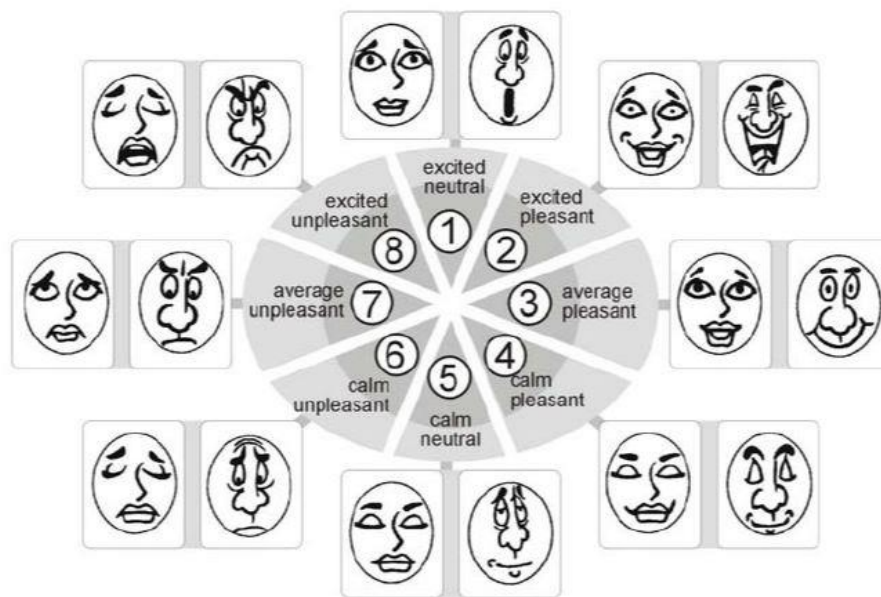


Figure 4. Eight emotional categories and Emocards (Desmet et. al., 2001, p. 6)

Interview questions:

1. Could you explain why you feel like this, which are the factors that determined this mood?
1. What did you enjoy about performing this task?
2. What did you dislike about performing this task?
3. Would you consider having a future interaction with this chatbot? If yes/no, why?
4. If you could, what would you change/improve at this chatbot?

3.7 Construct validity and reliability

3.7.1 Questionnaire

First, a factor analysis has been conducted for measuring the validity of the construct. Because the items come from different scales, a specific procedure had to be performed in the SPSS software. The orthogonal rotation option “varimax” was chosen. It assumes that the items are not related. The components with an eigenvalue over 1 explain the relationship between the items the best. Two factors have an eigenvalue under 1 but close enough to 1. Furthermore, the explained variance for the measurement items is above 50%. The components that were loading under the same construct have been deleted, in total 10 items. Six items from the UX construct and four from the social presence scale.

Second, the Cronbach`s alpha has been measured for each construct and for the UX for each dimension separately. The values are all above .700 which means that the constructs are reliable. In Table 3 the Cronbach`s alpha values are presented per construct and for UX also for each four dimensions, together with the factor analysis results.

The scale that measure the UX includes 28 items and reached a Cronbach`s alpha of .96. Measured separately, each dimension of the scale reached a Cronbach`s alpha over .80 as it can be seen in Table 3.

In order to raise the reliability score of the Social Presence construct which was under .70, the advice of the SPSS software has been followed and four indicated items were deleted and the new reliability for the seven left items of the scale reported a Cronbach`s alpha of .86. The deleted items are highlighted in red and can be found in Appendix C.

Finally, the items measuring the future interaction intention reported the highest Cronbach`s alpha value of .92.

Table 3. Factor analysis with 33 items and 3 constructs

Construct	Components	1	2	3	4	5	6
UX Cronbach`s α: .965	<i>Pragmatic Quality</i>						
	<i>Cronbach`s α: .886</i>						
	Complicated / Simple			.78			
	Impractical / Practical			.60			
	Cumbersome / Straightforward			.76			
	Unpredictable / Predictable			.70			
	Confusing / Understandable			.77			
	Unruly / Manageable			.73			
	<i>Hedonic Quality- Identity</i>						
	<i>Cronbach`s α: .874</i>						
	Unprofessional / Professional	.57					
	Tacky / Stylish	.52					
	Alienating / Integrating	.47					
	Separates me / Brings me closer	.40					
	<i>Hedonic Quality- Stimulation</i>						
	<i>Cronbach`s α: .893</i>						
	Conventional / Inventive				.74		
	Unimaginative / Creative				.69		
	Cautions / Bold				.70		
	Conservative / Innovative				.74		
	Dull / Captivating				.63		
	Undemanding / Challenging				.59		
	Ordinary / Novel				.68		
	<i>Attractiveness</i>						
	<i>Cronbach`s α: .967</i>						
	Unpleasant / Pleasant					.78	
	Ugly / Attractive					.77	
	Disagreeable / Likeable					.77	
Rejecting / Inviting					.76		
Bad / Good					.73		
Repelling / Appealing					.77		
Discouraging / Motivating					.68		
	Chatbot-mediated communication is an excellent		.71				

Social Presence	medium for this type of interaction						
Cronbach's α : .862	I felt comfortable conversing with this chatbot	.75					
	I felt comfortable interacting with the chatbot	.79					
	The chatbot created a feeling of an online community	.42					
	Overall the interaction with the chatbot met my expectations	.58					
Future interaction	I am likely to interact with this type of chatbot again						.84
Cronbach's α : .923	I am encouraged to interact with this type of chatbot in the near future						.83
	I look forward to interacting with this type of chatbot in the near future						.78
	I intend to interact with this type of chatbot in the next 3 months						.74
	Explained Variance	42.39 %	6.98%	6.97%	4.07%	3.14%	2.94%
	Eigenvalue	12.27	3.02	2.79	1.18	.91	.85

3.7.2 Interviews

For coding the transcripts, a codebook was created. The codebook was tested on interrater reliability. A second rater was assigned and both, the researcher and the second rater coded two interviews. The formula for the observed agreement was applied and resulted in a Cohen's kappa of .69. After some extra discussions the difference between the advantages and the benefits was more clear and the new result was a kappa of .89. A value between .81 and 1.00 is considered to be an almost perfect agreement thus the codebook was approved.

The codebook includes codes as "advantages" and "benefits" of the chatbots but also the negative sides of the interaction with the chatbots coded as "disadvantages". The "future interaction intention" and the eventual "requested changes" are also covered. The codebook can be found in Appendix H.

4. Results

In this section results from the online experiment and the interviews are further being presented and explained. For the online experiment, the main effects together with the interaction effects have been tested by means of a multivariate analysis of variance (MANOVA). A Wilk's Lambda test has been performed in order to explain the power of the independent variables on the dependent variables in the model. The possible mediation effect has been tested with the PROCESS v3.5 by Andrew F. Hayes, model number 4 and the possible moderation effect has been tested with the PROCESS v3.5 by Andrew F. Hayes, model number 1.

For the interviews, the responses from the 12 respondents have been analysed and they are further explored and explained based on the codebook.

4.1 Results of the online experiment

For studying the effects of the independent variables on the dependent variables, a factorial multivariate analysis of variance (MANOVA) has been performed. For investigating the effects of the visual appearance and conversational style on the dependent variables, a Wilk's Lambda (Λ) test has been performed. Based on the results of the test, it can be concluded that there is no significant main effect of the visual appearance and conversational style on the dependent variables nor between the two independent variables.

Based on a significance level of $p < .05$, the Wilk's Lambda in Table 4 shows no significant effects of the visual appearance style on the dependent variables $\Lambda = .98$, $F = 1.45$, $p = .227$. On the same note, the conversational style also shows no significant effects on the dependent variables $\Lambda = .97$, $F = 2.06$, $p = .107$. Looking at the interaction effect between the two independent variables, no significant results have been found $\Lambda = .99$, $F = .07$, $p = .972$.

Table 4. Multivariate results of independent variables

	Λ	F	p
Visual appearance	.980	1.457	.227
Conversational style	.972	2.060	.107
Visual appearance * Conversational style	.999	.077	.972

4.1.1 Main effects

Main effects of visual appearance

Table 5 shows there is no significant effect for the main effect of visual appearance on UX. It was hypothesized that the chatbots with a human visual appearance will have a larger effect on UX compared to the chatbots that do not use a human visual appearance. The difference in mean scores between human visual appearance ($M = 5.05$, $SD = 1.12$) and non-human visual appearance (logo) ($M = 4.97$, $SD = 1.05$) is not significant ($F = .364$, $p = .547$). Thus, based on the outcomes H1a was not supported.

There is no significant effect for the main effect of visual appearance on future interaction intention. It was expected that the chatbot with a human visual appearance will have a larger

effect on future interaction intention compared to the chatbots that do not use a human visual appearance. The difference in mean scores between human visual appearance ($M = 4.55$, $SD = 1.41$) and non-human visual appearance (logo) ($M = 4.64$, $SD = 1.35$) is not significant ($F = .252$, $p = .616$). Thus, H1b was also not supported.

On the same note as the other 2 hypothesis, there is no significant effect for the main effect of visual appearance on social presence. It was expected that the chatbots with a human visual appearance will have a larger effect on social presence compared to the chatbots that do not use a human visual appearance. The difference in mean scores between human visual appearance ($M = 4.90$, $SD = .806$) and non-human visual appearance (logo) ($M = 4.78$, $SD = .739$) is not significant ($F = 1.327$, $p = .251$). Thus finally, based on the outcomes H1c was also not supported.

Table 5. The mean scores of the independent variable visual appearance on the dependent variables social presence, UX and future interaction

Independent	Dependent	Manipulation of the independent variable	Mean	Std. Deviation
Visual appearance	Social Presence	Human	4.90	.80
		Logo	4.78	.74
	UX	Human	5.05	1.13
		Logo	4.97	1.05
	Future Interaction	Human	4.55	1.41
		Logo	4.64	1.35

Main effects of conversational style

According to Table 6 there is no significant effect for the main effect of conversational style on UX. It was hypothesized that the use of a humanlike conversational style will result in a larger effect on UX compared to the chatbots that do not use a human conversational style. The difference in mean scores between humanlike conversational style ($M = 5.11$, $SD = 1.08$) and machinelike conversational style ($M = 4.90$, $SD = 1.09$) is not significant ($F = 2.03$, $p = .155$). Based on the outcomes, H2a was not supported.

There is also no significant effect for the main effect of conversational style on future interaction intention. It was expected that the chatbots that use a humanlike conversational style will have a larger effect on future interaction intention compared to the chatbots that do not use a human conversational style. The difference in mean scores between human conversational style ($M = 4.65$, $SD = 1.39$) and machinelike conversational style ($M = 4.53$, $SD = 1.37$) is not significant ($F = .434$, $p = .511$). Thus, H2b was also not supported.

On the same note as the other 2 hypothesis, there is no significant effect for the main effect of conversational style on social presence. It was expected that the chatbots with a humanlike conversational style will have a larger effect on social presence compared to the chatbots that do not use a humanlike conversational style. The difference in mean scores between humanlike conversational style ($M = 4.81$, $SD = .698$) and machinelike conversational style ($M = 4.88$, $SD = .849$) is not significant ($F = .420$, $p = .518$). Finally, based on the outcomes H2c was also not supported.

Table 6. The mean scores of the independent variable conversational style on the dependent variables social presence, UX and future interaction

Independent	Dependent	Manipulation of the independent variable	Mean	Std. Deviation
Conversational style	Social Presence	Humanlike	4.81	.698
		Machinelike	4.88	.849
	UX	Humanlike	5.11	1.08
		Machinelike	4.90	1.09
	Future Interaction	Humanlike	4.65	1.39
		Machinelike	4.53	1.37

4.1.2 Interaction effects

An interaction effect between the two variables, visual appearance and conversational style did not take place. The MANOVA analysis in Table 4 showed that there is no significant interaction effect between the two independent variables ($F = .077$, $p = .972$). Based on this

results, it can be concluded that there was no interaction effect between the two independent variables.

4.1.3 Mediating role of social presence

A mediator variable is caused by the independent variable and is further a cause for the dependent variable. For measuring the moderation effect of social presence the PROCESS v3.5 by Andrew F. Hayes, Model number 4 has been performed. However, based on the conditions of Baron and Kenny (1986), when the independent variable (chatbot condition) does not affect the mediator (social presence), there is no ground for mediation. This means that social presence does not mediate the relationship between chatbot condition and UX. Based on the results, H3a is not supported.

On the same note, the possible mediation by social presence for the effect of chatbot conditions on future interaction intention has also been investigated. However, as in the case of UX, based on conditions of Baron and Kenny (1986) when the independent variable (chatbot condition) does not predict the mediator (social presence), there is no ground for mediation. This means that social presence does not mediate the relationship between chatbot condition and future interaction intention. Based on the results, H3b is not supported.

4.1.4 Moderating effect of gender

By introducing a moderating variable, it is expected that the direction or magnitude of the relationship between the independent and dependent variables will be changed. For measuring and testing the moderating effect of gender on the relationship between the independent and dependent variables, the PROCESS v3.5 by Andrew F. Hayes, Model number 1 has been performed. The results can be found in Table 7.

First, the results for chatbot condition * gender of the respondent on UX show that it is not a significant interaction $R^2 = .00$, $F(1.209) = .39$, $p = .47$. Thus H4a is not supported.

Further, the results for chatbot condition * gender of the respondent on future interaction with $R^2 = .01$, $F(1.916) = .75$, $p = .38$ and chatbot condition * gender of the respondent on social presence with $R^2 = .01$, $F(0.600) = .87$, $p = .85$ also show no significant interaction. Thus H4b is also not supported. According to the results the interaction term is not statistically significant which suggests that the predicted relationship between the independent variables and the dependent variables is not significantly moderated by gender.

Table 7. Gender as a moderator

Y	X	Moderator	R²	F	p
UX	Conditions	Gender	.0055	.39	.47
Future Interaction	Conditions	Gender	.0119	.75	.38
Social Presence	Conditions	Gender	.0124	.87	.85

4.2 Overview of the hypothesis

Table 8. Hypothesis overview

	Hypothesis	Supported
H1a	The chatbot with a human visual appearance will have a more positive effect on UX, than a chatbot that is not represented by a human visual appearance.	No
H1b	The chatbot with a human visual appearance will have a more positive effect on future interaction intention, than a chatbot that is not represented by a human visual appearance.	No
H1c	The chatbot with a human visual appearance will determine a higher level of perceived social presence, than a chatbot that is not represented by a human visual appearance.	No
H2a	The chatbot with a human-like conversational style will have a more positive effect on UX, than a chatbot that is using a machine-like conversational style.	No
H2b	The chatbot with a human-like conversational style will have a more positive effect on future interaction intention, than a chatbot that is using a machine-like conversational style.	No
H2c	The chatbot with a human-like conversational style will determine a higher level of perceived social presence, than a chatbot that is using a machine-like conversational style.	No
H3a	The effects of visual appearance and conversational style on UX will be mediated by social presence.	No
H3b	The effects of visual appearance and conversational style on future interaction intention will be mediated by social presence.	No
H4a	The possible effects of human visual appearance and human conversational style on UX will be moderated by gender.	No
H4b	The possible effects of human visual appearance and human conversational style on future interaction intention will be moderated by gender.	No

4.3 Results of the interviews

The emo-cards preferences are indicated, followed by the results from the interviews which are presented and explored per condition. A more detailed list of answers per condition, question and participant can be found in Appendix G.

There was no big variation among the emotions participants experienced during the interaction with the agents. The most chosen card was ‘calm pleasant’ (chosen by six interviewees), followed by ‘calm neutral’ (chosen by three interviewees). On the third place is ‘average pleasant’ (chosen by two users). All three are close to each other on the emo-cards disc spectrum which can be seen on Figure 4, page 19. Finally, one person chose ‘average unpleasant’ after interacting with the agent that had a machine-like conversational style and a human visual appearance.

A differentiation has been made between the advantages and benefits. Even though they sound as having the same meaning, in this study an advantage is recognised as being a position or condition that determines the user to feel better, while a benefit is a positive result or the effect which can be a consequence of an advantage.

When interacting with the CA that adopted a human conversational style and a human visual appearance, the interviewees appreciated the empathy and the logical way of thinking. Further, the easy buttons, the fact that the interaction was structured, straight forward and that the chatbot offered options that covered the problems also represented big advantages.

“The answers seemed to be human, empathic ‘oh, I am sorry to hear this...’ she seemed to understand what I wanted to say.”

The emoticons used by the agent in its answers were appreciated as humanlike and one respondent even compared them to real life facial expressions. Furthermore, the CA with human-like conversational style was perceived as an agent that understands the problem, gets fast to the point and needs no extra explanations.

“You feel that you are helped right away, not like when you call and you have to press 1, 2 or 3 and keep pressing numbers and keep waiting[...]With this system the interaction is faster and you get to the point faster.”

However, not everyone appreciated the human-like-empathic conversational style of the agent, in fact someone found her ‘too talkative’. In addition, the conversation was perceived as being too structured to believe that there is a human behind the computer. Also because the chatbot provides everyone with the same human-like answers the follow up reactions were less positive. One respondent also mentioned that the picture does not look real because it seems to be too professional. This created doubts and made people wonder if there is a real human behind the screen or not.

“She talks too much ‘your choice is good for the environment, bla, bla...’ I just want to repair my phone. If you are in a real crisis, you do not want to make conversation.”

Because this system gives the feeling of a faster service and the option to avoid direct human contact was appreciated and two of the respondents were positive about interacting again with it. However, one interviewee stated that she would come back to it just if there are no other options available.

Finally, one interviewee also mentioned to prefer shorter messages that were more straight to the point. Another stated that the virtual agent should not be connected to the Facebook

platform which is more seen as a social networking medium and not a customer support one. The privacy concerns were also mentioned.

For the chatbot condition with a human-like conversational style and a non-human visual appearance, the respondents mentioned the following advantages: calm, empathic and informal attitude enforced by a happy tone. The fact that she shows interest, seems to understand the situation and provides personalised attention by offering solutions was also highly appreciated.

“I liked the style of the text, it uses the spoken language, smiles and the encouragement related to the environment. [...] these elements show that you might be talking to a human.”

The benefits the user gets by making use of this system, from the participants` perspective, are the feeling of closeness and the sense that the problem will be solved and there is no need to worry. The effectiveness, the rapidness and the 24/7 accessibility were also mentioned.

“The chatbot comforted me and gave me a sense of security, I felt that my problem will be solved.”

Among the disadvantages, one respondent mentioned the lack of a human face. It was also mentioned that there were too many chat boxes and that the sentences were too long. Further, one user sensed the impossibility of the chatbot to respond spontaneous to written messages. Because it is effective and provides good information, respondents were positive about future interaction.

“Yes, I can not avoid it. It is happening already, it is cheap, they will use it. But if it can not help me, then I would like to get in contact with someone.”

Lastly, when asked about the improvements, the absence of a human face was noticed and its presence was required. In addition, the respondents also mentioned changes as: the use of shorter messages, the ability of the chatbot to respond when the user types back. Another possible improvement would be the personalized help where the user can introduce the type of phone and to immediately be directed to a webpage for her/his specific type of problem.

“To interact with ‘hi’ and ‘good bye’[...]it would also be nice if I could add the screen of my Samsung note 8 and to guide you immediately to the right page of the website.”

After the interaction with the CA that had a machine-like conversational style and a human visual appearance, the users mentioned among the advantages the clear tone, the speed and the correct answers. Surprisingly, even though the conversational style was machine-like, someone mentioned the empathy as an advantage. The presence of a human face was also appreciated because it gave a sense of comfort and trust.

“Because there was a photo, I got the idea that I am chatting with a person and it immediately transmitted a feeling of trust.”

Further, the chatbot was appreciated for offering a smooth conversation and leaving no place for doubts about the way it will solve the problem. However, two interviewees also mentioned

some disadvantages unrelated to the visual appearance and the conversational style. For example the presence of just three possible problems from where they could choose.

Related to the future interaction, two of the interviewees declared that they would like to interact in the future with the chatbot, because it proved to be helpful and because there was 'no extra noise'. However, one respondent mentioned that he would interact with a chatbot like this one just for basic questions, not for buying a product or a service. In that case, he would like to contact a human.

Finally, when asked the about changes, the participants mentioned the change of the visual representation because 'it looks fake'. The possibility to combine more options and the absence of information about the prices were also mentioned.

After interacting with the most machine-like condition, which had incorporated a machine-like conversational style and a logo as visual appearance, the users appreciated that this chatbot condition was very fast in solving the problem. It gave prompt and short answers, provided options and made use of easy buttons.

"[...] you just click on a button and automatically offers a new answer. The text is short, small sentences, you do not have to read a lot."

Among the benefits of making use of this CA the interviewees mentioned the absence of waiting time, that there was no room for doubt and also the possibility to leave the chat in the background and return when convenient which is impossible with a call centre. Two users also appreciated as a benefit the lack of need for human-to-human contact and the fact that there is no need to go to a store.

Among the disadvantages, one user stated that the fast answering ability of the chatbot gives a feeling of machinelike treatment which was not appreciated. In addition, when asked, the participants were positive about interacting again with the chatbot and stated that the reasons why they would interact with it in the future is because of its straight forwardness, speed and accessibility.

"You directly get your answer, usually you have to wait for your answer, they will tell you: 'in 24h we will react to your message...'"

Finally, the users did not mention changes related to the visual appearance or the conversational style. However, they did mention improvements related to the capabilities of the chatbot. For example, that they would prefer to get more technical information from the chatbot and also to have the possibility to respond in text. Based on the introduced information, the chatbot should be able to help them further. They would also appreciate it if the chatbot is able to give custom made advice, based on the type of phone they own.

5. Discussion

5.1 Discussion of the results

It was expected that the presence of human visual appearance and humanlike conversational style for text based chatbots would have a larger effect on the UX, increase the social presence level and would positively influence the future interaction intention. In this study, none of the formulated hypotheses were supported. Based on the experiment results, the presence of human conversational style and human visual appearance do not have a significant influence on UX, social presence and future interaction intention. However, several remarks that were made during the interviews are further explored and explained in this section.

Based on previous studies (Qiu & Benbasat, 2009; Gefen & Straub, 2004; Følstad, Nordheim & Bjørkli, 2018) it was expected that the use of human visual appearance will positively influence the future interaction intention, will determine a higher level of social presence, and according to Gefen & Straub (2004) will determine a larger effect on the UX. Based on the results from the multivariate analysis of variance there was no significant effect of visual appearance on the three dependent variable, thus all three hypothesis have been rejected.

However, during the interviews, participants that have interacted with the chatbots with a human visual appearance made remarks about the visual appearance of the chatbot. Two of them mentioned that the picture does not look real and raises questions of whether they are interacting with a human or a robot. In combination with a human conversational style, this can be a sign that the users did experience the uncanny valley effect where the presence of too much humanness creates a feeling of unpleasantness (Mori, Macdorman, Kageki, 2012). Because of its 'perfect' looks as one interviewee called it, the picture created a feeling of unease. However, one user that interacted with the machine-like conversational style but visualised a human picture mentioned that the presence of the picture immediately sent a feeling of trust and gave the feeling of chatting to a human being which is in line with some parts of the previous research.

In this study, based on the multivariate analysis of variance the results for the influence of the conversational style on the dependent variables show no significant effect. These results are in contradiction with previous research which states that a human conversational style will have a larger impact on UX (Michiels, 2017; Hu et. al., 2018), will determine higher levels of social presence (Go & Sundar, 2019) and will determine a more positive effect on the future interaction intention (Callejas et.al., 2011).

During the interviews, the participants which interacted with the chatbots that made use of a humanlike conversational style were satisfied with the happy tone, the feeling of closeness and empathy of the chatbots. At the same time they complained about it being too talkative in a situation where they tried to find a solution as fast as possible, offering only three options and it giving the same reaction to everyone. The ARCS model mentions the importance of attention, relevance, confidence and satisfaction in instructions material for enabling the users to perform a task (Keller, 1987). In this study the chatbot with a humanlike conversational style used encouragements which can be seen in the Appendix I during the task performance

while machinelike conversational style chatbot did not make use of these elements. There were complaints related to the length of the messages of the humanlike style and the relevance of the encouragements. These results can be explained by the findings of Loorbach, Karreman and Steehouder (2007) which tested the effect of motivational elements in task performance of students and seniors. At the end, the seniors appreciated the motivational elements but these had no effect on the students. The researchers concluded that the seniors, who are known for having a lower self-confidence with electronic devices benefited more from the motivational elements compared to younger people that have a higher level of self-confidence when operating this type of devices. In the present study, the big majority of the participants are young with an age between 25 and 34 years, are computer literate and experienced in making use of the internet and social media, which may explain the preference for shorter messages and more straight to the point answers.

During the interviews, four users mentioned the need for an option that allows the introduction of the type of phone they own and to have access to more technical details about the affected part. This type of requirements can be related to the need of users for declarative information which some users would appreciate to have access to. Even more, this is prove that users should have the possibility to decide for themselves whether they would like to read the declarative information or not (Karreman, 2004).

Comparing the two conversational styles, when asked what they disliked about the chatbot, the participants that interacted with the humanlike style mostly criticised the length of the messages. These results can be related to the minimalist approach (Carroll, 1990; van der Meij & Carroll, 1998) which is a framework about how designers should create instructions and training materials for computer users where the main ideas are: to have an action oriented approach, to build upon the learner's experience and to make instruction material simple but not simpler. As noticed during the interviews, users do ask for short and straight forward messages and this characteristic was appreciated in the chatbot with a machinelike conversational style where users were impressed by how fast and efficient the core of the problem is reached 'there is no extra noise'.

Additionally, the previous studies did show results in favour of social presence as an influencer for UX and future interaction intention (Cyr, Hassanein, Head and Ivanov, 2007; Etemad-Sajadi & Ghachem, 2015; Lu, Fan, Zhou, 2016; Meadows, 2017, Cyr et al., 2020), thus it was expected that social presence will mediate the relationship between the independent and the dependent variables. However, the results of this study are not statistically significant in order to assume this mediation. Based on the conditions of Baron and Kenny (1986), when the independent variables do not predict the mediator, there is no ground for mediation. These results do raise questions about the importance of perceived social presence for conversational user interfaces focused on customer service. In other words, further research on customer service chatbots should focus on further exploring if social presence really is the reason for the effect of visual appearance and conversational style on UX and future interaction intention.

Based on the previous studies conducted by several researchers (Araujo, 2017; Warwink & Shah, 2016; Hassaein & Head, 2007; Gefen & Straub, 2004) it was expected that the chatbot

with a human visual appearance and a human like conversational style would trigger evaluations as being humanlike and to further positively influence the social presence, determine a positive influence on the future interaction intention and have a larger effect on the user experience. However, the results are not significant. This results also help at answering the research question about the interaction effect between the two independent variables.

An explanation could be that in case of chatbots used for customer service, people`s attention is focused on solving the problem and not on aspects as visual appearance and conversational style. These might be more appreciated in chatbots used in previous studies which were mostly for marketing purposes. The manipulation check performed during the pre-test showed that the different manipulations have been recognised and in order to find out if the difference between the two groups were significant, an independent sample T-test has been performed. The T-test result also showed that the two different groups did differ in the perceived anthropomorphism.

In this study it was expected that gender will moderate the relationship between the independent variables and the dependent variables. Based on previous research (Jovic, 2020; Brandtzaeg & Føsland, 2017; Warwick, Vallverdú & Wung, 2016; Kimbrough et al., 2013) it was expected that the chatbot that uses a human conversational style and a human visual appearance will determine a higher level a social presence, a more positive UX and a stronger future interaction intention amongst women compared to men. Based on the results, in this study there is not a significant moderation effect of gender on this relationship. These results could also be explained by the role of the chatbot, which in this case has a more customer service purpose instead of a marketing purpose, and this means that users regardless of their gender are more purpose driven. These results are also in line with the results of Beldad, Hegner and Hoppen (2016) which showed that costumers` gender did not moderate the relationship between the CA gender-product gender congruence and the dependent variables.

5.2 Implications

Practical implications

The aim of this research was to find out if there are any significant effects of visual appearance and conversational style text based chatbots on the perceived social presence, UX and future interaction intention in a customer service context. The ultimate purpose is to help developers, copywriters, dialog designers and also the large public interested in chatbots to find out if aspects as visual appearance and conversational style do influence the experience and the future intention for interaction of the user.

First, the results of the online experiment show that the anthropomorphism in visual appearance and the use of human-like conversational style do not have a significant influence on the perceived social presence, user experience and future interaction intention. The interviews complement the statistical data by presenting the participants opinions. The information collected during the interviews show that the human-like elements were appreciated by some but they also created frustration and have been criticised. Based on these results, the field professionals should take into consideration that in a situation where the user

is looking for urgent solutions, the nuances related to appearance and conversational style tend to be ignored and the effectiveness and efficiency in solving the problem prevail. As mentioned before, these are aspects related to the minimalist approach (Carroll, 1990; van der Meij & Carroll, 1998) where the main focus is on action oriented instructions that are simple to follow but not simpler.

Second, it was also expected that female participants will be more sensitive towards the visual appearance and conversational style of the agent and would score higher levels of perceived social presence, larger effects on UX and a stronger future interaction intention when interacting with the agent with a human-like visual appearance and conversational style. However, the results were not statistically significant in order to accept this hypothesis. Experts in the field of chatbots can take this aspect into consideration when creating virtual agents because based on the online experiment and interviews, female and male participants made similar types of observations, which were mostly related to the effectiveness and the efficiency rather than to looks and conversational style preference.

Another aspect that should be taken into consideration is that in this study social presence was not the reason of the effect of human visual appearance and a human-like conversational on UX and future interaction intention. In addition, during the interview sessions participants perceived as ‘empathic’ and ‘understanding’ not only the agent with the most human-like elements but also the chatbot that used a logo as a visual appearance and a machinelike conversational style. This being said, it can be considered that in a technical user support context only having a logical back and forth text based interaction is enough for a chatbot to be perceived as human like and this discovery also represents an interesting starting point for further research.

Finally, other important additions were made by the users during their interaction with the chatbots. Firstly, the necessity of more technical information also known as declarative information which should be accessible in order to help the user to better understand the problem and to create a mental model which according to Kieras & Bovair (1984) has facilitative effects. Secondly, the capability of the chatbot to receive information about the type of mobile phone and to further direct the user to the correct website page was also required by the users. As mentioned by Brandztaeg and Følstad (2017) these type of abilities have a positive effect on perceived usefulness towards the agent.

Theoretical implications

Considering the theoretical implications of this research, even though the quantitative data was not statistically significant enough in order to support the hypothesis, the qualitative data plays an important role in offering important explanations about the interaction between the users and the virtual agents.

First, during the interviews the agents have been perceived as being humanlike and several human characteristics have been attributed to them (e.g. empathy, logical thinking, calmness, a happy tone, being too talkative etc.). These results are in line with the CASA paradigm

(Reeves & Nass, 1996; Nass & Moon, 2000) which pleads for the idea that humans apply social heuristics normally met in human to human interactions to computers even though they know that these machines do not have feelings or other human motivations.

Second, the chatbot with a human-like conversational style and the human-like visual appearance received appreciations but also the most critiques. Participants showed higher expectations from it, compared to the chatbot with a machine-like conversational style and a logo as visual appearance. Even though the chatbots have been built while following the literature recommendations (Chaves and Gerosa, 2019) concerning the humanness of conversational style and the anthropomorphism of the visual appearance, the most human-like agent might have been perceived as overly humanized and according to Gnewuch et. al (2017) overly humanized chatbots do determine higher expectations on users which can lead to more frustration if the chatbot fails. Even more, in their study Ciechanowski et al. (2019) show that simpler chatbots with no visual human identity result in lesser uncanny effect and a less negative effect, which is in line with the findings of this study where the less human like chatbot received less critique. The fact that the most human-like chatbot from them all received the most critique and caused the participants to experience frustration may be a sign that they have experienced the uncanny valley effect which according to Mori, Macdorman and Kageki (2012) can be described as dislike, unease and unpleasantness in the interaction with virtual agents.

Third, even though the chatbots did provide an introduction and an end message as was recommended by Chaves and Gerosa (2018), some of the users did try to interact with the chatbot by means of typed messages input. Even more, one of the interviewees did mention as an improvement of the chatbot, the possibility of interacting with 'hi' and 'good bye'. In their research, Jain et al. (2018) users mentioned that they would appreciate to be able to end a conversation which is considered a human-like conversational etiquette.

Finally, researchers and academics can use the results of this study when building and testing other chatbots about humanness of conversational style and anthropomorphism in visual appearance of chatbots. This research showed that in a customer support, a human like visual appearance and a human-like conversational style does not determine higher levels of social presence, larger effects on UX or a more positive effect on the future interaction intention. In contradiction with the expectations, the registered mean values for the conversational style are higher for a machinelike style in case of social presence and UX. Looking at the results of the interviews, the disadvantages of the machine like conversational style were related to the number of options offered by the agent but not to its conversational style which was appreciated as straight to the point, clear and even empathic. More research on chatbots taxonomies should follow.

5.3 Limitations and future research

The first limitation of this study is the total online nature of it which was determined by the pandemic restrictions. This also represented an impediment for the participants which sometimes had questions related to the interaction they had to perform and they were not always able to state them. Furthermore, performing the interviews 100% online did represent

a constrain because important body language signals and conversational aspects were not perceived by the interviewer and could have represented important data for the final results.

Another aspect that should be taken into consideration is related to the sample of participants. This study was initially conceived to be performed face to face. Due to the restrictions the final sample of participants is not homogenous and it is composed of participants with different background and circumstances from all over the world. A more homogeneous sample would decrease the possible implications of the variation in several sociodemographic factors on the final results.

A second limitation is represented by the measurement used for the social presence. This variable has been measured with the scale created by Gunawardena and Zittle (1997). However, due to the low reliability score six items of eleven had to be deleted in order to raise the Cronbach alpha value of the construct. It should be mentioned that future research should focus on the validation of the scales before data collection takes place. Even though frequently used, four out of the 28 items used to measure the UX also had to be deleted due to low Cronbach`s alpha values.

A third limitation is represented by the buttons as only mean of interaction between the user and the chatbot. Future research should offer participants the possibility to also interact with the chatbot by means of text input. Of course, text based interaction may ask a higher level of intelligence provided by the chatbot and may even include AI and machine learning.

An interesting topic for future research would be the influence of visual appearance and conversational style of chatbots on each dimension of UX from the Attrakdiff scale of Hassenzahl, Burmester and Koller (2003): pragmatic quality, hedonic quality - identity, hedonic quality - stimulation and attractiveness. In addition, expectations and attitudes toward chatbots, together with the uncanny valley effect also represent interesting variables that could be further explored, measured and tested together with the UX dimensions.

Another interesting topic for future research would be to replicate this research with two groups that differ in age. It would be of interest to find out if differences in visual appearance and conversational style do have an effect on social presence, UX and future interaction for digital natives that are considered to be early adaptors (Furini, 2014) and digital immigrants who are known for having different perceptions about the digital world and its functionality (Siriaraaya & Siang Ang, 2012). In addition, task complexity could also represent a subject of interest in this context.

Lastly, a comparison study between marketing vs. customer service focused chatbots and the implications of visual appearance and conversational style on the interactions with these agents would also represent an interesting topic of research.

6. Conclusion

This research aimed to investigate the influence of visual appearance and conversational style of text-based chatbots on social presence, user experience and future interaction intention.

Besides the main effects, this study also examined the interaction effect between visual appearance and conversational style which was not significant. Similarly, the effect of social presence as a mediator and gender as a moderator were also investigated but the tests showed no significant effects. Finally, the results show that there is not enough statistical evidence in order to support the hypothesis. Information related to the participants' preference on visual appearance and conversational style of chatbots in a customer service context has been collected from the interviews. Participants had higher expectations from the CA with a human visual appearance and a human-like conversational style which also received critiques related to looks and messages length. Furthermore, participants perceived as 'empathic' and 'understanding' not only the agent with the most human-like elements but also the chatbot that used a logo as a visual appearance and a machinelike conversational style. This being said, it can be concluded that in a customer support context only having a logical back and forth text based interaction might be enough for a chatbot to be perceived as human like. This discovery also represents an interesting starting point for further research. Based on these results, the field professionals should take into consideration that in a situation where the user is looking for an urgent solution to their problem, the nuances related to appearance and conversational style tend to be ignored and the effectiveness and efficiency in solving the problem prevail.

References

- Abel, T. (2010). Beyond usability: a methodology to evaluate the affective experience of interaction with e-commerce websites. *Proceedings of Measuring Behavior, 7th International Conference on Methods and Techniques*, 7(1), 257-260.
- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50(2), 179-211. doi.org/10.1016/0749-5978(91)90020-T
- 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(1), 183-189. doi: 10.1016/j.chb.2018.03.051
- Baier, D., Rese, A., & Röglinger, M. (2018, December). Conversational User Interfaces for Online Shops? A Categorization of Use Cases. In *ICIS*. Retrieved from <http://publica.fraunhofer.de/dokumente/N-534788.html>
- Balch, O. (2020). Meet the bots fighting loneliness. *Artificial Intelligence for Business*. Retrieved from: <https://www.raconteur.net/technology/ai-loneliness-mitsuku>
- Bargas-Avila, J. A., & Hornbæk, K. (2011, May). Old wine in new bottles or novel challenges: a critical analysis of empirical studies of user experience. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2689-2698). ACM. doi: 10.1145/1978942.1979336
- Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), 1173.
- Bartneck, C., Croft, E., & Kulic, D. (2008). Measuring the anthropomorphism, animacy, likeability, perceived intelligence and perceived safety of robots. *International Journal of Social Robotics*, 1(1), 71-81.
- Beldad, A., Hegner, S., & Hoppen, J. (2016). The effect of virtual sales agent (VSA) gender–product gender congruence on product advice credibility, trust in VSA and online vendor, and purchase intention. *Computers in Human Behaviour*, 60(1-3), 62-72.
- Bevan, N. (2008, June). Classifying and selecting UX and usability measures. *International Workshop on Meaningful Measures: Valid Useful User Experience Measurement 11(2)*, 13-18.
- Boeije, H. (2009). *Analysis in qualitative research*. Sage publications.

- Brady, M. K., & Robertson, C. J. (2001). Searching for a consensus on the antecedent role of service quality and satisfaction: an exploratory cross-national study. *Journal of Business Research*, 51(1), 53-60.
- Brahnam, S. (2009). Building character for artificial conversational agents: Ethos, ethics, believability, and credibility. *Psychology Journal*, 7(1), 9-47.
- Brandtzaeg, P. B., & Følstad, A. (2018). Chatbots: changing user needs and motivations. *Interactions*, 25(5), 38-43.
- Brandtzaeg, P. B., & Følstad, A. (2017, November). Why people use chatbots. *International Conference on Internet Science 1(1)*, 377-392.
- Callejas, Z., López-Cózar, R., Ábalos, N., & Griol, D. (2011). Affective conversational agents: the role of personality and emotion in spoken interactions. *Conversational agents and natural language interaction: Techniques and effective practices 1(1)*, 203-222. doi: 10.4018/978-1-60960-617-6.ch009
- Campbell, A. (2014). Identifying the reasons for consumers' non adoption of zero-emissions vehicles. *UTSG Newcastle*.
- Caporael, L. R. (1986). Anthropomorphism and mechanomorphism: Two faces of the human machine. *Computers in Human Behavior*, 2(3), 215-234.
- Carroll, J. M. (1990). The Nurnberg funnel: designing minimalist instruction for practical computer skill. *Educational Researcher* 20(9), 24-39. doi: 10.2307/1176249
- Chaves, A. P., & Gerosa, M. A. (2019). How should my chatbot interact? A survey on human-chatbot interaction design. *Human Computer Interaction Conference*. Retrieved from: <https://arxiv.org/abs/1904.02743v1>
- Ciechanowski, L., Przegalinska, A., Magnuski, M., & Gloor, P. (2019). In the shades of the uncanny valley: An experimental study of human–chatbot interaction. *Future Generation Computer Systems*, 92(-), 539-548.
- Claessen, V., Schmidt, A., & Heck, T. (2017). Virtual Assistants-A Study on the Usability and User Perception of Customer Service Systems for E-Commerce. *Proceedings of the 15th International Symposium of Information Science 1(4)*, 116-130.
- Cyr, D., Hassanein, K., Head, M., & Ivanov, A. (2007). The role of social presence in establishing loyalty in e-service environments. *Interacting with Computers*, 19(1), 43-56.
- Dale, R. (2016). The return of the chatbots. *Natural Language Engineering*, 22(5), 811-817.
- Danet, B., Ruedenberg-Wright, L., & Rosenbaum-Tamari, Y. (1997). Hmmm... where's that smoke coming from?. *Journal of Computer-Mediated Communication*, 2(4), JCMC246.

- Derks, D., Fischer, A. H., & Bos, A. E. (2008). The role of emotion in computer-mediated communication: A review. *Computers in Human Behaviour*, 24(3), 766-785.
- Desmet, P., Overbeeke, K., & Tax, S. (2001). Designing products with added emotional value: Development and application of an approach for research through design. *The Design Journal*, 4(1), 32-47.
- Di Prospero, A., Norouzi, N., Fokaefs, M., & Litoiu, M. (2017, November). Chatbots as assistants: an architectural framework. In *Proceedings of the 27th Annual International Conference on Computer Science and Software Engineering* (pp. 76-86). IBM Corp..
- Etemad-Sajadi, R., & Ghachem, L. (2015). The impact of hedonic and utilitarian value of online avatars on e-service quality. *Computers in Human Behavior*, 52, 81-86.
- Feine, J., Gnewuch, U., Morana, S., & Maedche, A. (2019). A Taxonomy of Social Cues for Conversational Agents. *International Journal of Human-Computer Studies*, 13(5), 138-161.
- Feng, D. (2019, February 1). The world of chatbots: customer service, business automation & scalability. Retrieved June 1, 2020, from <https://www.bigcommerce.com/blog/chatbots/>
- Ferrara, E., Varol, O., Davis, C., Menczer, F., & Flammini, A. (2016). The rise of social bots. *Communications of the ACM*, 59(7), 96-104.
- Fincher, M. (2018, August 20). 3 Ways to improve your chatbot traffic right now. Retrieved May 2, 2020, from <https://chatbotlife.com/3-ways-to-improve-your-chatbot-traffic-right-nowb07db4fa0d51?gi=c3eb224414ee>
- Følstad, A., Nordheim, C. B., & Bjørkli, C. A. (2018, October). What makes users trust a chatbot for customer service? An exploratory interview study. *International Conference on Internet Science* (pp. 194-208). Springer, Cham.
- Foster, M. E. (2007, July). Enhancing human-computer interaction with embodied conversational agents. In *International Conference on Universal Access in Human-Computer Interaction* (pp. 828-837). Springer, Berlin, Heidelberg.
- Furini, M. (2014). Users behavior in location-aware services: digital natives versus digital immigrants. *Advances in Human-Computer Interaction*, 2014.
- Garcia, M. (2017, July). Consumers want chatbots to feel human, not to look human. *Phocus Wire*. Retrieved from: <https://www.phocuswire.com/Consumers-want-chatbots-to-feel-human-not-look-human>
- Gefen, D., & Straub, D. W. (2004). Consumer trust in B2C e-Commerce and the importance of social presence: experiments in e-Products and e-Services. *Omega*, 32(6), 407-424.

- 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.
- Goertzel, B. (2018). Sophia. Retrieved from: <https://www.hansonrobotics.com/>
- Guest author (2009, June, 1st). User experience over usability. *Web Credible*. Retrieved from <https://www.webcredible.com>
- Gunawardena, C. N., & Zittle, F. J. (1997). Social presence as a predictor of satisfaction within a computer-mediated conferencing environment. *American Journal of Distance Education*, 11(3), 8-26.
- Hardman, S., Shiu, E., & Steinberger-Wilckens, R. (2016). Comparing high-end and low-end early adopters of battery electric vehicles. *Transportation Research Part A: Policy and Practice*, 88, 40-57.
- Harpur, P. A. (2013). *Evaluation of usability and user experience of an m-learning environment, custom-designed for a tertiary educational context* (Doctoral dissertation, University of South Africa).
- Hassenzahl, M., Burmester, M., & Koller, F. (2003). AttrakDiff hedonic and pragmatic evaluation scale for an interactive product. *Attrakdiff*. Retrieved from: <http://attrakdiff.de/science-en.html#publikationen>
- Hendriks, F., Ou, C. X., Amiri, A. K., & Bockting, S. (2020). The power of computer-mediated communication theories in explaining the effect of chatbot introduction on user experience. *Interaction*, 12, 15.
- Herring, S. C. (2002). Computer-mediated communication on the Internet. *Annual Review of Information Science and Technology*, 36(1), 109-168.
- Hill, J., Ford, W. R., & Farreras, I. G. (2015). Real conversations with artificial intelligence: A comparison between human–human online conversations and human–chatbot conversations. *Computers in Human Behavior*, 49, 245-250.
- Hu, T., Xu, A., Liu, Z., You, Q., Guo, Y., Sinha, V., ... & Akkiraju, R. (2018, April). Touch your heart: a tone-aware chatbot for customer care on social media. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (pp. 1-12).
- Jain, M., Kumar, P., Kota, R., & Patel, S. N. (2018, June). Evaluating and informing the design of chatbots. In *Proceedings of the 2018 Designing Interactive Systems Conference* (pp. 895-906).

- Janson, A., Degen, O., & Schwede, M. (2019, July). Designing for Social Presence and Leveraging the Outcomes of Customer Service Chatbots. *Academy of Management Proceedings* (Vol. 2019, No. 1, p. 11164). Briarcliff Manor, NY 10510: Academy of Management.
- Jung, N. Y., Kim, S., & Kim, S. (2014). Influence of consumer attitude toward online brand community on revisit intention and brand trust. *Journal of Retailing and Consumer Services*, 21(4), 581-589.
- Kabadayi, S., & Gupta, R. (2011). Managing motives and design to influence web site revisits. *Journal of Research in Interactive Marketing*.
- Kear, K., Chetwynd, F., & Jefferis, H. (2014). Social presence in online learning communities: The role of personal profiles. *Research in Learning Technology*, 22.
- Keller, J. M. (1987). Strategies for stimulating the motivation to learn. *Performance and Instruction*, 26(8), 1-7.
- Kieras, D. E., & Bovair, S. (1984). The role of a mental model in learning to operate a device. *Cognitive Science*, 8(3), 255-273.
- Kimbrough, A. M., Guadagno, R. E., Muscanell, N. L., & Dill, J. (2013). Gender differences in mediated communication: Women connect more than do men. *Computers in Human Behavior*, 29(3), 896-900.
- Koda, T., & Maes, P. (1996, November). Agents with faces: The effect of personification. In *Proceedings 5th IEEE International Workshop on Robot and Human Communication*, 189-194.
- Koh, Y. J., & Sundar, S. S. (2010). Effects of specialization in computers, web sites, and web agents on e-commerce trust. *International Journal of Human-Computer Studies*, 68(12), 899-912.
- Koh, Y. J., & Sundar, S. S. (2010). Heuristic versus systematic processing of specialist versus generalist sources in online media. *Human Communication Research*, 36(2), 103-124.
- Kowatsch, T., Nißen, M., Shih, C. H. I., Rügger, D., Volland, D., Filler, A., & Heldt, K. (2017). Text-based healthcare chatbots supporting patient and health professional teams: preliminary results of a randomized controlled trial on childhood obesity. *Persuasive Embodied Agents for Behavioral Change*, 1-10.
- Krämer, N. C., Hoffmann, L., & Kopp, S. (2010, September). Know Your Users! Empirical Results for Tailoring an Agent's Nonverbal Behavior to Different User Groups. *International Conference on Intelligent Virtual Agents*, 468-474.

- Kreijns, K., Kirschner, P. A., Jochems, W., & Van Buuren, H. (2011). Measuring perceived social presence in distributed learning groups. *Education and Information Technologies, 16*(4), 365-381.
- Laurel, B. (1997). Interface agents: Metaphors with character. *Human Values and the Design of Computer Technology, 1*(4), 207-219.
- Liebrecht, C., & van Hooijdonk, C. (2019, November). Creating humanlike chatbots: what chatbot developers could learn from webcare employees in adopting a conversational human voice. *International Workshop on Chatbot Research and Design 1*(1) 1-64.
- Loorbach, N., Karreman, J., & Steehouder, M. (2007, October). The effects of adding motivational elements to user instructions. *IEEE International Professional Communication Conference, 1-5*.
- Lu, B., Fan, W., & Zhou, M. (2016). Social presence, trust, and social commerce purchase intention: An empirical research. *Computers in Human Behavior, 56*(-), 225-237.
- McTear, M., Callejas, Z., & Griol, D. (2016). Conversational interfaces: Past and present. In *The Conversational Interface, 51-72*. doi:10.1007/978-3-319-32967-3_4
- Meadows, M. S. (2017). Artificial Intelligence & Artificial Trust – Chatbots Magazine. Retrieved from: <https://chatbotsmagazine.com/artificial-intelligence-artificial-trust-5e01d32ae39a>
- Mehrabian, A., & Ferris, S. R. (1967). Inference of attitudes from nonverbal communication in two channels. *Journal of Consulting Psychology, 31*(3), 248-252. doi.org/10.1037/h0024648
- Mimoun, M. S. B., Poncin, I., & Garnier, M. (2017). Animated conversational agents and e-consumer productivity: The roles of agents and individual characteristics. *Information & Management, 54*(5), 545-559.
- Mori, M., MacDorman, K. F., & Kageki, N. (2012). The uncanny valley. *IEEE Robotics & Automation Magazine, 19*(2), 98-100.
- Moriuchi, E., Landers, V. M., Colton, D., & Hair, N. (2020). Engagement with chatbots versus augmented reality interactive technology in e-commerce. *Journal of Strategic Marketing, 1-15*.
- Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues, 56*(1), 81-103.
- Nass, C., Steuer, J., & Tauber, E. R. (1994, April). Computers are social actors. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 72-78).

- Norman, D. (2018). The Definition of User Experience (UX) Nielsen Norman Group. *Nielsen Norman Group*. Retrieved from: <https://www.nngroup.com/articles/definition-user-experience>.
- Pirrone, R., Russo, G., Cannella, V., & Peri, D. (2008). GAIML: A new language for verbal and graphical interaction in chatbots. *Mobile Information Systems*, 4(3), 195-209.
- Polzin, T. S., & Waibel, A. (2000). Emotion-sensitive human-computer interfaces. *ISCA Tutorial and Research Workshop (ITRW) on Speech and Emotion*, 1-6.
- Porter, C. E., & Donthu, N. (2006). Using the technology acceptance model to explain how attitudes determine Internet usage: The role of perceived access barriers and demographics. *Journal of Business Research*, 59(9), 999-1007.
- Prensky, M. (2001). Digital natives, digital immigrants. *Cross Currents on the Horizon*, 9(5), 45-87.
- Preston, C. C., & Colman, A. M. (2000). Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica*, 104(1), 1-15.
- Qiu, L., & Benbasat, I. (2009). Evaluating anthropomorphic product recommendation agents: A social relationship perspective to designing information systems. *Journal of Management Information Systems*, 25(4), 145-182.
- Radziwill, N. M., & Benton, M. C. (2017). Evaluating quality of chatbots and intelligent conversational agents. *Computers and Society*. Retrieved from: *arXiv preprint arXiv:1704.04579*.
- Reeves, B., & Nass, C. I. (1996). *The media equation: How people treat computers, television, and new media like real people and places*. Centre for the Study of Language and Information; Cambridge university press.
- Rice, R. E. (1993). Media appropriateness: Using social presence theory to compare traditional and new organizational media. *Human Communication Research*, 19(4), 451-484.
- Rietz, T., Benke, I., & Maedche, A. (2019). The impact of anthropomorphic and functional chatbot design features in enterprise collaboration systems on user acceptance. *Proceedings of the 14th International Conference on Wirtschaftsinformatik*. Retrieved from: <https://aisel.aisnet.org/wi2019/track13/papers/7/>
- Schanke, S., Burtch, G., & Ray, G. (2020). Estimating the Impact of ‘Humanizing’ Customer Service Chatbots. *Information System Research*, 1(2), 2-50.
- Shah, H., Warwick, K., Vallverdú, J., & Wu, D. (2016). Can machines talk? Comparison of Eliza with modern dialogue systems. *Computers in Human Behavior*, 58, 278-295.

- Short, J., Williams, E., & Christie, B. (1976). *The social psychology of telecommunications*. John Wiley & Sons.
- Shum, H. Y., He, X. D., & Li, D. (2018). From Eliza to XiaoIce: challenges and opportunities with social chatbots. *Frontiers of Information Technology & Electronic Engineering*, 19(1), 10-26.
- Singh, N. (2017, September). Why consumers want your chatbot to be more human? *Entrepreneur*. Retrieved from: <https://www.entrepreneur.com/article/300412>
- Simonite, T. (2017). Customer Service Chatbots Are About to Become Frighteningly Realistic. *MIT Technology Review*. Retrieved from: <https://www.technologyreview.com/s/603895/customer-service-chatbots-are-about-to-become-frighteningly-realistic/>
- Siriaraya, P., & Siang Ang, C. (2012). Age differences in the perception of social presence in the use of 3D virtual world for social interaction. *Interacting with Computers*, 24(4), 280-291.
- Sivadas, E., & Baker-Prewitt, J. L. (2000). An examination of the relationship between service quality, customer satisfaction, and store loyalty. *International Journal of Retail & Distribution Management*, 28(2), 73-82. doi.org/10.1108/09590550010315223
- Skjuve, M., Haugstveit, I. M., Følstad, A., & Brandtzaeg, P. B. (2019). Help! Is my Chatbot Falling into the Uncanny Valley?: An Empirical Study of User Experience in Human-Chatbot Interaction. *Human Technology*, 15(1), 30-54. doi.org 10.17011/ht/urn.201902201607
- Smestad, T. L. (2018). *Personality Matters! Improving The User Experience of Chatbot Interfaces- Personality provides a stable pattern to guide the design and behavior of conversational agents* (Master's thesis, NTNU).
- Steinfeld, C. W. (1986). Computer-mediated communication in an organizational setting: Explaining task-related and socioemotional uses. *Annals of the International Communication Association*, 9(1), 777-804.
- Sundar, S. S., Bellur, S., Oh, J., Jia, H., & Kim, H. S. (2016). Theoretical importance of contingency in human-computer interaction: effects of message interactivity on user engagement. *Communication Research*, 43(5), 595-625.
- Sward, D. (2006). User experience design: A strategy for competitive design. *Americas Conference of Information Systems*, 2(1), 2-14.
- Traphagan, T. W., Chiang, Y. H. V., Chang, H. M., Wattanawaha, B., Lee, H., Mayrath, M. C., ... & Resta, P. E. (2010). Cognitive, social and teaching presence in a virtual world and a text chat. *Computers & Education*, 55(3), 923-936.
- Turing, A. M. (2004). *Computing machinery and intelligence. The Essential Turing: The Ideas that Gave Birth to the Computer*. Springer, Dordrecht.

- Van der Meij, H., & Carroll J. M. (1998). Principles and heuristics for designing minimalist instruction. *Minimalism beyond the Nurnberg Funnel*, J.M. Carroll, Ed. Cambridge, MA: MIT Press 42(2), 19-53.
- Vermeeren, A. P., Law, E. L. C., Roto, V., Obrist, M., Hoonhout, J., & Väänänen-Vainio-Mattila, K. (2010). User experience evaluation methods: current state and development needs. *Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries*, 6(1), 521-530
- Wang, X., Giron, F., Ask, J. A., Wise, J., Purcell, B., Stanhope, J., Nagel, B. (2017, May 24). Chatbots are transforming marketing. *Forrester*. Retrieved from <https://www.forrester.com>
- Warwinck, K., & Shah, H. (2016). The importance of a human viewpoint on computer natural language capabilities: A Turing test perspective. *AI & Society*, 31(2), 207e221.
- Weizenbaum, J. (1966). ELIZA—a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1), 36-45.
- Wexelblat, A. (1998). Don't make that face: A report on anthropomorphizing an interface. *Intelligent Environments*, 173(179), 98-02.
- Worswick, S. (2017). Mitsuku chatbot. Retrieved from <https://www.pandorabots.com/mitsuku/>
- Xu, K., & Lombard, M. (2017). Persuasive computing: Feeling peer pressure from multiple computer agents. *Computers in Human Behavior*, 74, 152-162.

Appendixes

Appendix A

Interview questions:

1. Could you explain why you feel like this, which are the factors that determined this mood?
1. What did you enjoy about performing this task?
2. What did you dislike about performing this task?
3. Would you consider having a future interaction with this chatbot? If yes/ no why?
4. If you could, what would you change/improve at this chatbot?

Appendix B

Pre-test remarks during the interaction with the chatbots:

Participant	Conversational style	
	Humanlike	Machinelike
1	“Trustworthy attitude. The speech feels personal”	“Straight to the point and short answers”
2	“Messages are too long, I lost my patience and focus”	“Interested in your problem and clear”
3	“Adding a phone number or an email address would help, in case I need to contact someone because my problem is complex”	“It is good, it is very fast”
4	“She talks too much, in a real situation I would get angry”	“It is very rapid but very strict, if my problem is different I would like to be able to contact a real agent, to explain my problem”
5	“It uses the everyday language, and empathy, I like it.”	“It does not allow you to give details”
6	“The encouragement I got, made me feel better “there are solutions, do not worry”	“Answers come too fast”
7	“It is clear and structured but the answers come too fast, this makes you think that it is not a human behind it”	“It offers options, I like that”
8	“I would add a phone number just in case before the goodbye at the end. Nice that understands what you feel.”	“Logical, fast and clear. I would like to type back”
9	“I like that there are options, but I would like to be able to contact a human too”	“I would like to be able to personalize my situation. Maybe my phone does not start, what do I do then?”
10	“I would prefer less information”	“The interaction was quite boring”

11 “Was nice to see emoticons,
humour, understanding from the
other side”

“Nothing special, just a commune
interaction with a chatbot”

12

“I prefer it without the typing option, it is
much easier and faster”

Appendix C

Survey:

Start of Block: Default Question Block

Dear participant,

I am a master student following the Technology and Communication program at the University of Twente and this survey is part of my final thesis.

Thank you for your interest in this study. When starting this survey, you will first have to read a brief scenario. Then, you will have a small text-based interaction with a chatbot via Facebook Messenger. Next, a survey related to the interaction you had will appear. The interaction with the chatbot and answering the survey will take between 5 and 10 minutes. Your participation in this survey is voluntarily and you have the right to withdraw at any time without further explanations.

The answers will remain anonymous and will be used just for this research and with the purpose to evaluate your interaction. For further questions and comments, please contact:

a.m.stan@student.utwente.nl Thank you, Miruna Stan

Do you agree to be part of this research?

Yes, I agree (1)

End of Block: Default Question Block

Start of Block: Scenario:

Scenario Imagine that a part of your smartphone broke (for example the screen, the battery or the charging spot). You hear about an online shop that repairs phones. You enter the shop's Facebook page and you engage yourself in a conversation on Facebook Messenger. The conversation is button based, **you just have to press buttons, no need for text and it takes approximately 30 seconds.**

Please follow the next steps:

1. Click the link from the next page and **start the interaction in the Facebook Messenger box.**
2. When finished, **please return to continue the survey.**

End of Block: Scenario:

Start of Block: 1. HC + MV

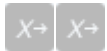


Q111 [For going to Facebook messenger, please click here](#)

- Please proceed just if you did interact with the agent on Facebook Messenger (1)

End of Block: 1. HC + MV

Start of Block: 2. MC+MV



Q222 [For going to Facebook Messenger, please click here](#)

- Please proceed just if you did interact with the agent on Facebook Messenger (2)

End of Block: 2. MC+MV

Start of Block: 3. HC+HV



Q333 [For going to Facebook Messenger, please click here](#)

- Please proceed just if you did interact with the agent on Facebook Messenger (3)

End of Block: 3. HC+HV

Start of Block: 4. MC+HV



Q444 [For going to Facebook Messenger, please click here](#)

- Please proceed just if you did interact with the agent on Facebook Messenger (4)

End of Block: 4. MC+HV

Start of Block: Attrakdiff

Q1 Where does the chatbot fall between the next categories?

	Neutral							
	1 (1)	2 (2)	3 (3)	4 (4)	5 (5)	6 (6)	7 (7)	
Technical	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Human
Complicated	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Simple
Impractical	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Practical
Cumbersome	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Straightforward
Unpredictable	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Predictable
Confusing	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Understandable
Unruly	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Manageable

Q2 What impression would you give by using this chatbot?

	Neutral	

	1 (1)	2 (2)	3 (3)	4 (4)	5 (5)	6 (6)	7 (7)	
Isolating	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Connective
Unprofessional	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Professional
Tacky	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Stylish
Cheap	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Premium
Alienating	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Integrating
Separates me	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Brings me closer
Unpresentable	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Presentable

End of Block: Attrakdiff

Start of Block: Attrakdiff

Q3

This chatbot is:

	Neutral	
--	---------	--

	1 (1)	2 (2)	3 (3)	4 (4)	5 (5)	6 (6)	7 (7)	
Conventional	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Inventive
Unimaginative	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Creative
Cautious	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Bold
Conservative	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Innovative
Dull	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Captivating
Undemanding	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Challenging
Ordinary	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Novel

Q4 The experience I just had with this chatbot was:

	Neutral	
--	---------	--

	1 (1)	2 (2)	3 (3)	4 (4)	5 (5)	6 (6)	7 (7)	
Unpleasant	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Pleasant
Ugly	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Attractive
Disagreeable	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Likeable
Rejecting	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Inviting
Bad	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Good
Repelling	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Appealing
Discouraging	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Motivating

End of Block: Attrakdiff

Start of Block: Social Presence + Future interaction intention

Q5 How would you rate the interaction?

	Strongly disagree (1)	Disagree (2)	Somewhat disagree (3)	Neither agree nor disagree (4)	Somewhat agree (5)	Agree (6)	Strongly agree (7)
The messages were impersonal (1)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Chatbot-mediated communication is an excellent medium for this type of interaction (2)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I felt comfortable conversing with this chatbot (3)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I felt comfortable interacting with the chatbot (4)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The chatbot created a feeling of an online community (5)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The chatbot facilitated the possibility of further discussion (6)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Discussions using a chatbot medium tend to be less personal than face-to-face discussions (7)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Text-based communication is less personal than audio based communication (8)

Text-based communication is less personal than video based communication (9)

I felt that my point of view was acknowledged by the chatbot (10)

Overall the interaction with the chatbot met my expectations (11)

Q6 Do you see yourself interacting with this chatbot in the future?

	Strongly disagree (1)	Disagree (2)	Somewhat disagree (3)	Neither agree nor disagree (4)	Somewhat agree (5)	Agree (6)	Strongly agree (7)
I am likely to interact with this type of chatbot again (1)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I am encouraged to interact with this type of chatbot in the near future (2)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I look forward to interacting with this type of chatbot in the near future (3)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I intend to interact with this type of chatbot in the next 3 months (4)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

End of Block: Social Presence + Future interaction intention

Start of Block: Gender + Educational Level

Q7 What is your gender?

Male (1)

Female (2)

Other (3)

Q8 What is your level of education?

Basic School (1)

Highschool (2)

Bachelor (3)

Pre-Master (4)

Master (5)

Doctorate (PHD) (6)

Q9 What is your age

- >18 (1)
- 18-24 (2)
- 25-34 (3)
- 35-44 (4)
- 45-54 (5)
- 55-64 (6)
- 65 or more (7)

End of Block: Gender + Educational Level

Appendix D:

List of codes:

Code number	Code name
1	Advantage about the system
2	Benefit of the system
3	Disadvantage about the system
4	Negative about the chatbot
5	Changes
6	Future interaction

Appendix F: Snatchbot- the chatbot development platform with 4 conditions

The screenshot shows the SnatchBot dashboard interface. On the left is a navigation sidebar with the following items: Dashboard, My Bots (highlighted), Plugins, Broadcast, Inbox, Reports, NLP, Activity logs, Advanced Stats, Upgrade to PRO, English (with a dropdown arrow), Balance (\$0), Notifications (5), SnatchBot Store, Documentation, Support Center, Miruna Stan, and Logout. The main content area is titled "My Bots" and contains four bot cards. Each card displays a profile picture or icon, the bot's name, its ID, and the last active time. The bots are: Sarah (id:93755, 27/02/2020 11:08), Sarah - Tech Support (id:96841, 19/03/2020 16:03), Tech Support (id:96355, 16/03/2020 14:27), and Tech Support 2 (id:96842, 19/03/2020 16:03). Each card also features a small UK flag icon and a set of action icons at the bottom. In the top right corner, there is a blue button labeled "+ Create Bot" with a help icon. At the bottom of the page, there is a footer with the text "© 2020 SnatchBot. All Rights Reserved" and a row of links: "Report an issue | Privacy Policy | Personal Data Request | Terms & Conditions".

Appendix G : Interview results per condition, question and participant

1. *Could you explain why you feel like this, which are the factors that determined this mood?*

HV + HC = Human visual appearance + Human conversational style

P1: HV + HC 5 calm neutral

- normal interaction that I normally have with a bot
- I prefer to speak to a real person because I can explain my problem
- a bot does not have all the answers and will not give me a satisfactory answer
- there is no robot yet that thinks like a human

P8: HV + HC 5 calm neutral

- I experienced bol.com system
- I have experienced this before
- You do not have to type this much

P6: HC + HV 4 calm pleasant

- I liked The interaction because it was structured
- I liked that I could choose an answer
- Options covered the problems
- The answers seemed to be human, empathic "oh, I am sorry to hear this..."
- She seemed to understand what I wanted to say
- too good structured to get the feeling that I am talking to a human
- Somehow I knew that this was the reaction that everyone gets
- This did not make me feel special (refers to the line above)

MV + HC = Logo + Human conversational style

P2: HC+ Logo 4 calm pleasant

- the chatbot gave me a sense of security
- I felt that my problem will be solved
- she talks very personal to me
- I feel her being close to me
- she offers me solutions

P5: HC + Logo 4 calm pleasant

- The style of the text, it uses the spoken language
- The use of similes, the encouragement related to the environment

P11: HC + Logo 4 calm pleasant

- The chatbot was Calm and had a happy tone
- The way it spoke to me, it was informal
- she understood the situation, she was empathic like "oh no, that must be so annoying for you"
- too many chat boxes after each other, if it goes very fast you need to scroll back

HV + MC = Human visual appearance + Machine-like conversational style

P3: MC + HV 4 calm pleasant

- I needed a service

P7: HV + MC 7 average unpleasant

- I prefer to talk to a human if I want to buy something or access a service
- I would use a chatbot for basic questions
- for a complex interaction like this one I would prefer a human to human interaction

P12: HV + MC 4 calm pleasant

- It was very clear
- Fast answers
- No place for doubts
- Correct answers

MV + MC = Logo + Machine-like conversational style

P4: MC + Logo 4 average pleasant

- solving the problem very fast
- the promptitude of the answers

P9: MC + Logo 5 calm neutral

- the answers were just good
- you did not have to rethink it

P10: MC + Logo 3 average pleasant

- because the options offers options
- you just click on a button and automatically offers a new answer
- the text is short, small sentences, you do not have to read a lot

2. *What did you enjoy about performing this task?*

HV + HC

P1: HV + HC 5 calm neutral

- enjoy is a big word
- it was a normal conversation with the bot
- she talked to much "oh, your choice is good for the environment" I do not care I just want to repair my phone
- if you are in a real crisis you do not want to make conversation
- maybe for another type of problem it would not bother me that much

P6: HC + HV 4 calm pleasant

- the answers with empathy
- the answers had a connection to my answer and then gave the solution
- the emoticons make you think of a real facial expression of a human

P8: HV + HC 5 calm neutral

- That you do not have to type every time
- That I did not have to explain
- straight forward
- you feel that you are helped right away, not like when you call and you have to press 1,2 or 3 and keep pressing numbers and keep waiting
- with this system the interaction is faster and you get to the point faster

HC+ Logo

P2: HC+ Logo 4 calm pleasant

- I liked that she is positive
- She uses words as "great", "thankful"
- I like that she asks what do I need
- She is interested in my problem
- She is empathic
- These elements show that you might be talking to a human

P5: HC + Logo 4 calm pleasant

- It was effective
- I got what I wanted

P11: HC + Logo 4 calm pleasant

- it was straight to the point first 2 options, than other 3 options and so on...

- getting to the core of the problem within one minute
- pretty efficient and happy tone

MC+ HV

P3: MC + HV 4 calm pleasant

- She was empathic
- She could feel me

P7: HV + MC 7 average unpleasant

- The rapidity
- It was fast

P12: HV + MC 4 calm pleasant

- that there was a photo, I got the idea that I am chatting with a person and it transmitted immediately a feeling of trust

MC+ Logo

P4: MC + Logo 4 average pleasant

- Did not take a lot of time
- Very rapid answers
- You do not need to be there personally or with another human
- Not even a phone call, very fast

P9: MC + Logo 5 calm neutral

- that it was really fast
- you do not have to wait, you get a fast answer back
- you do not have to wait like a couple of minutes as for example when it is a human at the other end and you have to wait

P10: MC + Logo 3 average pleasant

- the short text and the easy buttons
- but just 3 options, maybe there is another problem
- when chatting with a bot, you can just leave it aside and do something else, while with a human you can not do that

3. *What did you dislike about performing this task?*

HV + HC

P1: HV + HC 5 calm neutral

- That she gave only 3 options

P6: HC + HV 4 calm pleasant

- It gives this feeling that it is the same treatment for everyone

P8: HV + HC 5 calm neutral

- The face of the lady did not look real
- I would like to see a picture that is like less cosy. Like it can feel like more like a real person.
- The picture gives the feeling that it is a machine
- In this case it lets u wonder. It could be person, maybe it is a machine but you do not know

HC+ Logo

P2: HC+ Logo 4 calm pleasant

- Too long message about the cost
- It is a big boring to read the solutions which are too long. People loose patience when they see very long messages

P5: HC + Logo 4 calm pleasant

- Not being able to see a face
- When I typed good bye, it went in a loop. It did not know how to end a conversation which of core broke the illusion of personal interaction

P11: HC + Logo 4 calm pleasant

- Those text boxes (refers to those mentioned at Q1)

MC+ HV

P3: MC + HV 4 calm pleasant

- Nothing really

P7: HV + MC 7 average unpleasant

- That I could not combine more answers, it was set to let you choose just one option. For example I do not know what is broke but my phone does not work

P12: HV + MC 4 calm pleasant

- No, not really

MC+ Logo

P4: MC + Logo 4 average pleasant

- Because the answers come very fast, gives you a feeling of machinelike reaction, I can understand that because it is a software and I knew that.

P9: MC + Logo 5 calm neutral

- No , not really

P10: MC + Logo 3 average pleasant

- That you only have 3 options

4. *Would you consider having a future interaction with this chatbot? If yes/ no why?*

HV + HC

P1: HV + HC 5 calm neutral

- If it is the only option, yes
- I prefer to speak to a real person

P6: HC + HV 4 calm pleasant

- If they offer the option, sure I will use it. I prefer to avoid interactions with humans

P8: HV + HC 5 calm neutral

- Yes, because I get the feeling it is a faster service
- But (gives example with bol.com, machine keeps feeding tracking number when the problem was different) sometimes is kind of nice to have someone to talk right away to explain.

HC+ Logo

P2: HC+ Logo 4 calm pleasant

- Yes, I find it very useful and handy
- If you are at home and it is 3am you can enter the page and see what you have to do the second day and which service you have to contact
- I find it nice that you can find which are the costs

P5: HC + Logo 4 calm pleasant

- Yeah, it gives me the right information, it is effective, yes!

P11: HC + Logo 4 calm pleasant

- Yes, I can not avoid it, it is happening already. it is so cheap, they will use it
- If it can not help me then I would like to get in contact with someone because if my option is in between I would get frustrated

MC+ HV

P3: MC + HV 4 calm pleasant

- Yes, it is helpful

P7: HV + MC 7 average unpleasant

- Just if there are basic questions
- For example I ask Siri what is the weather like, how late it is.
- I would not choose it for accessing a service or to buy something
- .

P12: HV + MC 4 calm pleasant

- Yes, as long as the answers are correct
- Makes no difference if you talk to a human or a robot
- There is no extra noise. If you talk with a human sometimes you get extra information that is not needed and now it was extremely to the point

MC+ Logo

P4: MC + Logo 4 average pleasant

- Yes, for sure because of what I mentioned before, fast and prompt answers

P9: MC + Logo 5 calm neutral

- Yes because it is straight forward, fast
- You have directly your answer usually you have to wait for your answer they will tell you" in 24h we will react to your message" in this case it answers so fast

P10: MC + Logo 3 average pleasant

- Yes, it is a useful and easy way to get the right information and you can easily make an appointment
- It is fast, convenient and easy to contact

5. *If you could, what would you change/improve at this chatbot?*

HV + HC

P1: HV + HC 5 calm neutral

- Insert more options
- More straight to the point, because it is an urgent situation

P6: HC + HV 4 calm pleasant

- -

P8: HV + HC 5 calm neutral

- That it is connected via Facebook, I do not like the idea of this company having access to my information. I would prefer an online system but not via Facebook

HC+ Logo

P2: HC+ Logo 4 calm pleasant

- Shorter messages
- Direct the user to an online website where the problem will be solved

P5: HC + Logo 4 calm pleasant

- To interact with "hi" and "good bye" if I type back to her
- It would be nice if I could add the screen of my Samsung note 8 and to guide you immediately to the right page of the website

P11: HC + Logo 4 calm pleasant

- The boxes sizes are big for the amount of text, it needs scroll

MC+ HV

P3: MC + HV 4 calm pleasant

- The picture, because it looks fake

P7: HV + MC 7 average unpleasant

- More options
- The possibility to choose 2 options or to say "I do not know"
- Not give the price for the service

P12: HV + MC 4 calm pleasant

- No, for my problem it went smoothly
- - the tone was clear I mean more information is not needed here
- It did not give a cold feeling or robotic

MC+ Logo

P4: MC + Logo 4 average pleasant

- Some technical information should be given
- The ability of the user to insert information because there are just 3 options
- A 4th option where you can add a distinct problem

P9: MC + Logo 5 calm neutral

- -

P10: MC + Logo 3 average pleasant

- Integrate the option of choosing the type of the phone in the text box

6. *Would you like to add something else?*

HV + HC

P1: HV + HC 5 calm neutral

- If my phone is broken I would not write on Facebook to the company, I would go on the website and send an email before contacting via Facebook

P6: HC + HV 4 calm pleasant

P8: HV + HC 5 calm neutral

HC+ Logo

P5: HC + Logo 4 calm pleasant

- I would add a friendly face

P10: MC + Logo 3 average pleasant

- To extend the chatbot that can already ask for the type of phone

P11: HC + Logo 4 calm pleasant

MC+ HV

P3: MC + HV 4 calm pleasant

P7: HV + MC 7 average unpleasant

P12: HV + MC 4 calm pleasant

MC+ Logo

P2: MC+ Logo 4 calm pleasant

P4: MC + Logo 4 average pleasant

P9: MC + Logo 5 calm neutral

Appendix H: Codebook

Advantages - Benefits - Disadvantages

Human visual appearance + Human conversational style		
Advantages (qualities)	Benefits (effects)	Disadvantages
<ul style="list-style-type: none"> • not much typing • structured interaction • choices • problem covered by options • empathic • logic thinking • emoticons make you think of real facial expressions • straight forward • not much typing 	<ul style="list-style-type: none"> • she seemed to understand • getting fast to the point • extra explanations not needed 	<ul style="list-style-type: none"> • does not have all answers • only 3 options • talks too much about other things (e.g. environment) and creates frustration • too good structured to be a human • same reaction for everyone, thus you do not feel special • the face does not look real and gives the feeling that it is a machine, a less cosy picture would help

Logo + Human conversational style		
Advantages (qualities)	Benefits (effects)	Disadvantages
<ul style="list-style-type: none"> • personal speech • offers solutions • calm • happy tone • informal • empathic • understands the situation • interested • presents human elements • straight to the point 	<ul style="list-style-type: none"> • gives a feeling of closeness • gives a sense of security • effective • getting to the core of the problem very fast • accessible at any time of the day 	<ul style="list-style-type: none"> • too many chat boxes after each other • too long messages, boring to read • not seeing a face • goes in a loop, not able to react to "hi" and "goodbye"

Human visual appearance + Machine-like conversational style		
Advantages (qualities)	Benefits (effects)	Disadvantages
<ul style="list-style-type: none"> • clear tone • correct answers • empathic 	<ul style="list-style-type: none"> • fast answers • no place for doubts • transmits trust • smooth conversation 	<ul style="list-style-type: none"> • could not combine answers; set to let you choose just one option

Logo + Machine-like conversational style		
Advantages (qualities)	Benefits (effects)	Disadvantages
<ul style="list-style-type: none"> • solving problem very fast • prompt and short answers • options • easy buttons 	<ul style="list-style-type: none"> • leaves no doubts • no waiting time • the option to leave it aside and come to it later (impossible with a call center) • no need for human to human contact or to go personally to a shop 	<ul style="list-style-type: none"> • just 3 options • because the answers are fast it gives you a feeling of machinelike reaction

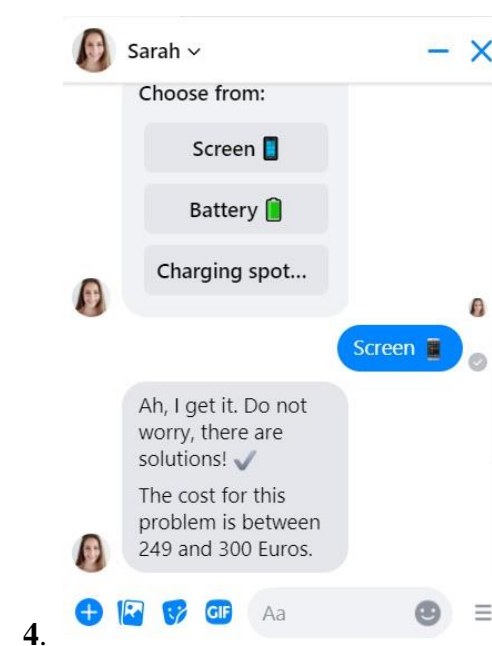
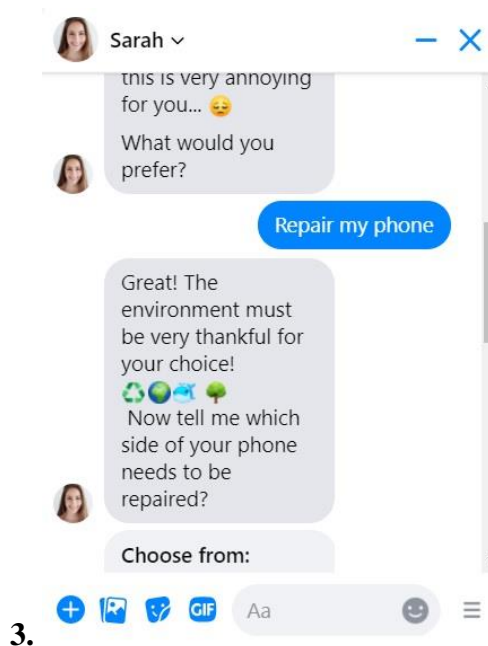
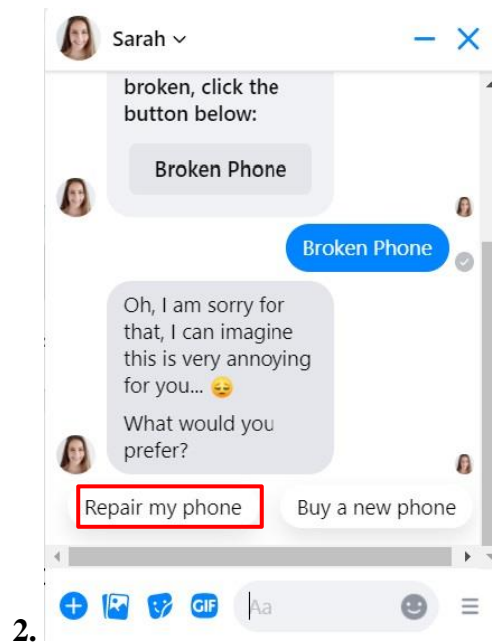
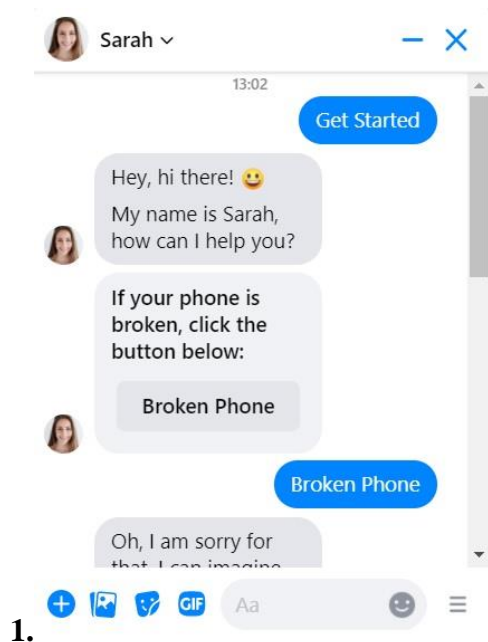
Requested changes – Future interaction

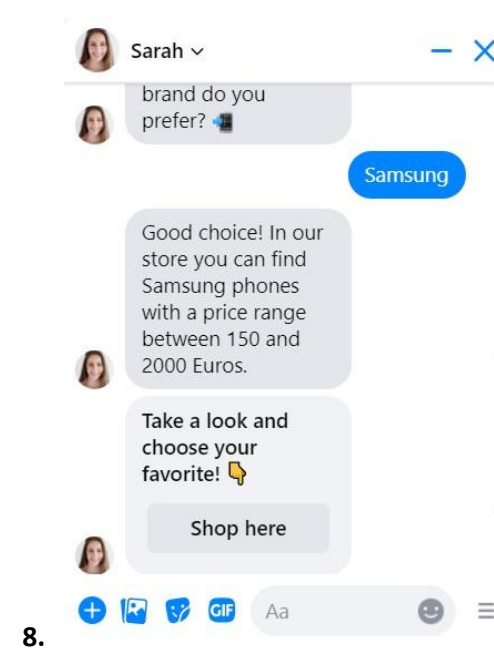
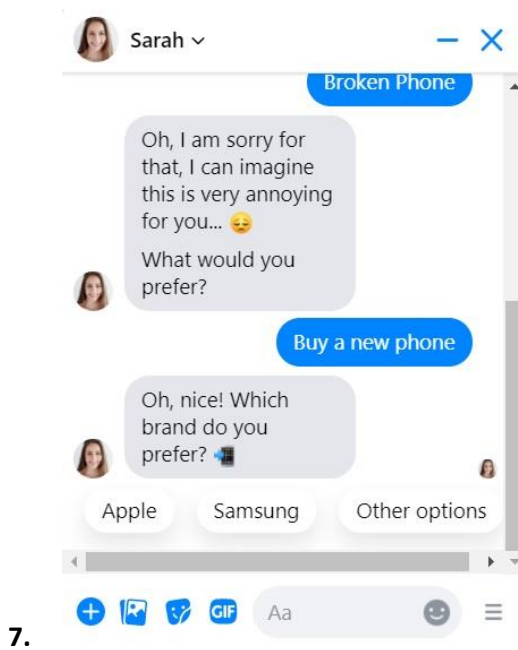
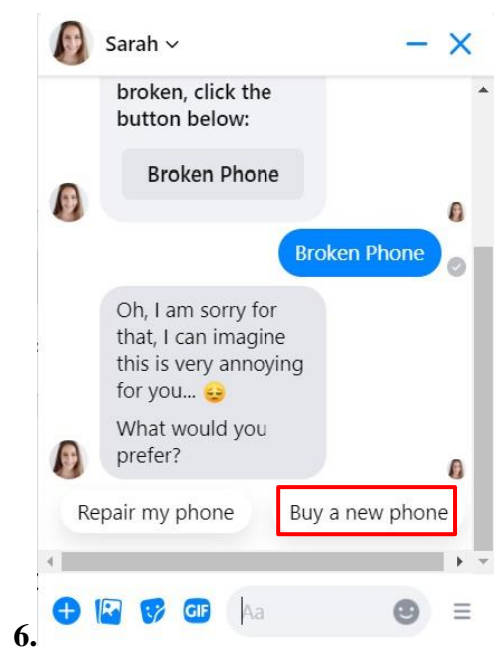
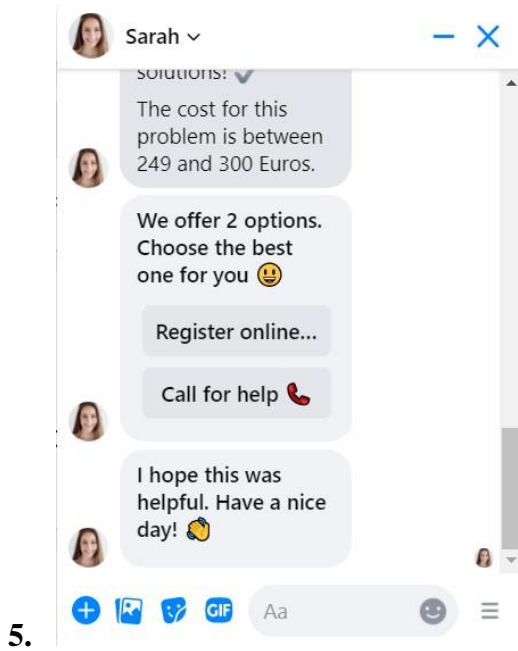
Required changes	Future interaction intention
HC + HV	
<ul style="list-style-type: none"> - more options - more straight to the point because the situation is urgent - connected to another online website, not Facebook 	P1: "Just if it is the only option." P6: "Yes, I prefer to avoid interactions with humans." P8: "Yes, it gives the feeling that it is a faster system."
HC + Logo	
<ul style="list-style-type: none"> - shorter messages - direct the user to a website where the problem will be solved - ability to interact with "hi" and "goodbye" if the user types back 	P2: "Yes, it is useful and handy. You can see the costs." P5: "Yes, it gives good information, it is effective." P11: "Yes, I can not avoid it. It is happening already, it is cheap, they will use it. But if it can

<ul style="list-style-type: none"> - personalized help (e.g. to introduce the type of phone and the problem and to directly direct the user to a website). - less text boxes - add a friendly face 	<p>not help me, then I would like to get in contact with someone.”</p>
MC + HV	
<ul style="list-style-type: none"> - picture, it looks fake - more options and possibility to combine 2 options - no prices 	<p>P3: “Yes, it is helpful.” P7: “Just for basic questions, not for buying a product or service.” P12: “Yes, no difference between this and a human; chatbot is better because of “no extra noise”.”</p>
MC + Logo	
<ul style="list-style-type: none"> - add some technical information - insert technical information option - add a 4th option - possibility to add the phone type you have 	<p>P4: “Yes because it is fast, prompt and straightforward.” P9: “Yes. You directly get your answers, usually you have to wait for your answer, they will tell you “in 24h we will react to your messages.” P10: “Yes, it is useful and an easy way to get information and you can easily make an appointment. It is fast, convenient and easy to contact.”</p>

Appendix I: Conversation flow

Human-like conversational style and human visual appearance





Machine-like conversational style and non-human visual appearance

