

How human should a chatbot be?

The influence of avatar presence and anthropomorphic characteristics in conversational tone regarding chatbots in the customer service field

Laury ten Donkelaar (S1877798)

University of Twente Faculty of Behavioural, Management and Social Sciences (BMS) Msc. Communication Studies – Marketing Communication First supervisor: Dr. J. Karreman Second Supervisor: MSc. K.R. Brunink Date: November 6, 2018 Enschede, 2018



Abstract

Over the past few years, different types of conversational agents and chatbots have been the target of increasing interest. With the rapid growth of new technologies and artificial intelligence, chatbot instruments are developing quickly. The prospect is that the use of chatbots will continue to grow, mostly in service focused industries.

It is expected that in the coming years, contact between people and organizations is mostly going to be handled by chatbots with humanoid characteristics. However, research about the appropriate use of humanoid characteristics regarding chatbots is still small in number. Therefore, this study has looked into conversational agents and their level of anthropomorphism in conversational tone in combination with the visual presence of an avatar. Moreover, this research has measured the effect of the independent variables on the dependent variables (likeability, perceived intelligence, user satisfaction, trust, perceived usefulness and intention to use) and the moderators (empathy and negative attitude towards robots). Furthermore, to test the hypotheses a survey has been conducted.

The findings showed that there is no significant effect of the independent variables on the dependent variables. Moreover, the results indicate that the negative attitude towards robots moderates the effect of the visual presence of an avatar and perceived usefulness. Additionally, a moderating effect of the negative attitude towards robots has been found between the conversational tone, trust and perceived intelligence. This information is important to investigate further to examine the reason why people with a negative attitude towards robots have less trust in the chatbot and perceive a chatbot as less intelligent. Additionally, more research is needed to give more insight into the effect of anthropomorphism in conversational tone and avatar appearance in different environments and purposes. The results of this study can serve as a guideline to a better chatbot design and is a stepping stone for further research.

Keywords

Anthropomorphism, chatbots, conversational agents, humanoid avatar, human-computer interaction

Table of Contents

Introduction	4
Theoretical framework	6
Anthropomorphism	6
Visual presence of a virtual agent	7
Anthropomorphism in conversational tone	8
Likeability	9
User satisfaction	10
Perceived intelligence	10
Trust	11
Perceived usefulness	11
Intention to use	12
Interaction between visual appearance of an avatar and anthropomorphism in conversational tone	12
Empathy towards robots	13
Negative attitude towards robots	14
Research model	15
Method	16
Research design	16
Stimulus material	16
Pre-tests	17
Main study	19
Participants	19
Procedure	20
Measurement instruments	21
Results	23
Multivariate analysis of variance (MANOVA)	23
Overview of the hypotheses	28
Discussion	30
References	35
Appendices	41

Introduction

In recent years, there has been an immense growth in the area of virtual conversational agents (Mindbowser, 2017). A conversational agent (chatbot) is a computer program that is designed to simulate a conversation with a user. Such technology is designed in order to assist people with everyday tasks or for entertainment purposes (Natason, 2017). Moreover, conversational agents can be used by companies, where the agent is used to assist with customer service questions or to advice in customer purchase decisions (Ferrara et al., 2016). The chatbot does not only answers questions of users, chatbots can be used to assist with learning, improve customer experiences or give support to elderly people (Radziwill & Benton, 2017; Ferrara et al., 2016). There are many different types of conversational agents for different purposes (Tsvetkova, 2017). Well known conversational agents are the digital assistants, Siri, Cortana, Google Assistant and Alexa. By just asking, the digital assistant answers questions, plays music, sends messages, tells the morning news and makes restaurant reservations (Natason, 2017).

The current study is focused on the customer service chatbot. Customer service chatbots are chatbots that are designed to increase customer experience, increase satisfaction, reduce time to respond to messages and increase engagement in the conversation (Radziwill & Benton, 2017).

An example of a customer service conversational agent is the chatbot O, from Dutch energy supplier Oxxio. This chatbot gives advice about a costumers' energy usage and gives answers or solutions to questions (Oxxio, 2017). In this example, the chatbot assists the users and human-to-human interaction is no longer necessary. This means that chatbots have the ability to disrupt the customer service department. Instead of a direct conversation with customers, employees may optimize the chatbot experience and only support human services when inevitable (Mindbowser, 2017).

Essentially, a conversational agent is a computer-based service assistant that interacts with an individual by means of a chat interface (Radziwill & Benton, 2017). Furthermore, a conversational agent simulates a conversation by responding to human input. Chatbot conversations are created by matching a stimulus (input) to a large collection of stored patterns and generating the most suitable response (output). Chatbots originate from the system ELIZA developed by Weizenbaum in 1966. Since then, many developers have tried to improve the system to create a more human imitating and intelligent conversational agent (McTear et al., 2016). A human imitating chatbot is necessary because, people expect that the chatbot will react in a reasonable manner, familiar to humanity. However, the development of chatbots is not ready to recognize context or subtle inferences that make a conversation human. Therefore, developers include human characteristics to their chatbot design (Foner, 1993). This process is an example of anthropomorphism, because the developers attribute human characteristics to a technical object (Hirsch et al., 2002). Foner (1993) states that it is natural and probably necessary to apply anthropomorphism to a chatbot design. If not, the chatbot cannot function because it is expected that the chatbots commands to reasonable actions. A test that is used to test if a robot or conversational agent can interact in an anthropomorphic and human imitating manner was to pass the Turing test (Turing, 1950) The Turing test is a challenge where the ability of a machine was tested to behave in a human-like manner. When a judge is not able to distinguish the human from the machine participant the test is considered passed.

Unfortunately, the chatbots of today are still vulnerable to technical complications and inconsistent responses (Newman, 2016). An example of a conversational agent that went off the rail is the Microsoft

chatbot "Tay". The chatbot turned racist by internet trolls within 24 hours. The bot is learning by human engagement and does not know what racism is. Trolls exploited this vulnerability (Price, 2016). To improve the chatbot, this research will investigate whether anthropomorphic characteristics have an effect on how the chatbot is perceived. This research will test anthropomorphism in avatar appearance and conversational tone. Another attribute that is investigated in this research is the feeling of empathy. The empathy a person feels can possibly affect the way a chatbot is perceived. Additionally, the negative attitude towards robots is investigated. When a person has a negative attitude towards robots, in general, the judgement towards a conversational agent may likewise negatively be affected. This research aims to investigate these matters in greater detail by looking into existing research on anthropomorphism and the use of avatars regarding conversation agents. If chatbots are going to be used more frequently in the customer environment there is also need to examine different conditions that can improve the conversational agent and customer experience. This research will look into the appearance of a chatbot and the conversational tone. These factors will be adjusted in different anthropomorphic levels. There is no academic study that tests these factors together, this makes this research an useful and interesting study. Additionally, the findings of this research can give new theoretical perspectives and can contribute to the design of a useful, efficient and user-friendly chatbot experience.

This study aims at providing more insights into the use of different types of avatar appearances combined with anthropomorphism in the conversational tone. Additionally, the empathy of the user and the negative attitude toward robots will be tested as a moderators. The study focuses on giving an answer to the following research question:

'What are the effects of avatar appearance and anthropomorphic characteristics in conversational tone regarding chatbot use in the customer service field?'

In the next section literature about chatbots, anthropomorphic characteristics in human-robot interactions and the importance of implementing an avatar will be discussed. Furthermore, literature about the moderator empathy and the moderator negative attitude towards robots will be examined. Five sets of hypotheses are formulated based on the literature presented. In the second chapter, the research method will be explained, followed by the results. Finally, the discussion section will cover the limitations and implementations of this study.

Theoretical framework

In this part the theoretical framework will be presented. First of all, an exploration on the characteristics of anthropomorphism will be given. Secondly, the visual presence of a virtual agent will be discussed. Additionally, anthropomorphism in conversational tone will be explained. Furthermore, the variables likeability, perceived intelligence, user satisfaction, trust, perceived usefulness and intention to use will be examined, followed by a description of the feeling of empathy and the negative attitude towards robots. At the end of this theoretical framework the research model will be introduced.

Anthropomorphism

The term anthropomorphism is defined by the new dictionary of cultural literacy as "The attributing of human characteristics to inanimate objects, animals, plants, or other natural phenomena, or to God. "To describe a rushing river as "angry" is to anthropomorphize it". (Hirsch et al., 2002, p. 86). A robot needs to have some type of human-like attributes for a meaningful social conversation (Duffy, 2003). The use of human-like characteristics can rationalize a robot's actions and can assist the progress of an individual's social understanding (Duffy et al., 2002).

Research by Laurel (1997) state that there are three important arguments that support the benefits of using anthropomorphic characteristic in human-robot interactions. First, using an agent that has a human persona invites the user to interact in a conversation. Second, a conversational agent that is personalized enhances a person's ability to make precise assumptions on how that agent is likely to act based on external cues. Thus, creating a better understanding. For example, when a conversational agent will act like a teenager based on external cues. These external cues can be, the use of certain modern slangs or using the avatar of a teenager as a virtual avatar. Third, the metaphor of an anthropomorphic character channels the attention of the user to the essential natural cues of an agent, which are, competence, responsiveness, the ability to perform actions and accessibility.

On the other hand, Erickson (1997) is skeptical about the use of humanoid interfaces. He thinks humanizing a robot leads to unnecessary complex systems. Erickson states that people want an efficient and simple interface without the hassle of a conversation with an overly emotive and 'fake' human being. Although, Erickson (1997) is not positive about anthropomorphic systems he acknowledged that "we may not have much of a choice" (Erickson, 1997, p.79). He states that people will react to computers in the same manner as to other people, this reaction is almost always unavoidable and automatic. This means that the interactions people will have with a computer will not much differ from interactions in everyday life. And thus, people may not have much of a choice than interact with computers as though they have some human-like entity. The theory that explains the tendency to respond to computers and other types of mechanical devices the same as to other people is called the Media Equation theory by Reeves and Nass (1996). These researchers concluded that people are polite to computers in both textual and verbal designs. People interact with computers just like interactions in everyday life. Individuals are being polite, attribute gender and/or personality characteristics (aggression, humor, intelligence) to the mechanical devices. To enhance this tendency developers add anthropomorphic characteristics to a chatbot design to create an even better understanding between machine and human (Brahnam, 2009). In addition, Polzin and Waibel (2000) argued that the design of an interface that reflects human-computer interaction should respond to this theory for effective communication. If people try to communicate with computers in the same manner as they interact with other people, then this nature should be reflected in the virtual design to facilitate a more 'natural' anthropomorphic interaction. The researchers argue that communication with a computer type interface is a two-way street. Stating that society needs interfaces that detects emotions in the user and not only express emotions. They state that effective communication is not efficient and can be confusing when the computer interface expresses emotions and the emotional state of the user is ignored.

The use of those human-like cues can be enhanced in human-robot interactions. According to Fong et al. (2003), robots that perform human-robot interactions have the following human-social characteristics:

- use of human-like natural characteristics
- communicate and/or grasp emotions
- may establish/develop social abilities
- establish/maintain social relationships
- observe/master models of other agents
- present unique personality and character
- · communicate with an effective and powerful dialogue

To conclude, building a valuable conversation chatbot requires some level of anthropomorphism to have a meaningful human-robot conversation.

Visual presence of a virtual agent

The level of anthropomorphism processed in the chatbot can be reflected in the visual representation of the virtual agent. The image of the conversational agent can be an important addition when designing a chatbot interface. An image can, for example, increase credibility, likeability, tele-presence (Nowak & Biocca, 2003) and user satisfaction (Holzwarth et al., 2006; Tinwell, 2009)

It is likely that using a virtual image in a virtual environment will influence how people feel about and react to the digital environment and the instrument itself. For example, an anthropomorphic and, thus, a more human-looking avatar can have a different effect on how people perceive a conversational agent than a digital avatar that is more abstract-looking (Nowak & Biocca, 2003). Research by Nowak and Biocca (2003) reported this phenomenon by investigating three different types of images; a virtual drawn head of a human, an image that displayed drawn eyes and a mouth or no image. Furthermore, the research manipulated the conditions by telling the participants that the images were controlled by a computer or controlled by a human being. The result shows that people respond equally social to the computer and human-controlled images. Furthermore, the presence of a virtual image enhanced the feeling of being in the virtual world (tele-presence). Additionally, the participants felt more presence from the company of others (co-presence) and felt more able to access to another mind (social presence) when presented with only an image displaying drawn eyes and a mouth either than no image or the image of a virtually drawn human head. This indicated that the virtually drawn human head sets higher expectations, which can lead to a lower feeling of presence when expectations were not met (Nowak & Biocca, 2003). This occurrence was later further investigated by Nowak (2004), where results showed that the less humanoid image compared to a high humanoid image and no image was seen as more likeable and credible. Mimoun et al. (2012) tried to explain in their case study why several embodied virtual agents were

disappearing from French websites. The research explained that an inadequate appearance, such as a visual virtual agent with too many anthropomorphic characteristics, can cause false expectation and disappointment. The researchers report two probable causes. First, the design of the website and the design and knowledge of the virtual agent may not be in balance and 'do not fit together'. Second, the appearance of the virtual agent is too anthropomorphic. When the human-like virtual agent does not respond as expected, costumers are disappointed. One of the participants stated: "the agent has a very sophisticated visual display but it is not able to answer all users questions'' (Mimoun et al., 2012, p. 608). In this example, the participant overestimated the virtual agent's capability based on appearance, which leads to unsatisfying experiences. This leads to the disappearance of embodied virtual agents on company websites.

An important theory to take into consideration when designing a virtual agent is the uncanny valley effect (Mori, 1970). This effect occurs when a robot appears so human-like that the effect is unsettling. People have a negative, creepy, eerie feeling when they encounter an entity that is almost human. When anthropomorphic characteristic increases, familiarity increases until this increase dips drastically. This results in a valley. The entity is human-like but misses some key attributes. examples of these entities are corpse, zombies or prosthetic hands.

Anthropomorphism in conversational tone

Using anthropomorphism in virtual agents invites the user to interact in a conversation (Laurel, 1997). Moreover, adding anthropomorphic cues to the conversation makes the chatbot more socially approachable (Brahnam, 2009) and meaningful (Duffy, 2003).

To satisfy the user and have a meaningful conversation with a conversational agent, two requirements are necessary. First, the agent must be effective and reliable. Second, the conversational agent must be able to process the text that the customer uses (Lester et al., 2004). The agent process the text and generates answers by connecting keywords of the user's conversational output to the database of the agents system, Additionally, the agent collects keywords and answers for that database by recording human-tohuman conversations and transferring these in a text format (Lester et al., 2004; McTear et al., 2016; Wang et al., 1997). It is essential to process the natural language of the user accurately and efficiently for an effective conversational agent. To respond adequately to a user's needs, the technology of the conversational agent must interpret the input of the user. Additionally, the agent generates the action that is suitable in response and perform that action. For example, if the users' input is: "I would like to order a peperoni pizza" The conversational agent analyses the meaning of the input. When the agent has determined the nature and meaning of the request or statement, the agent must decide how to act. The decision of the action depends on the goal of the agent, the dialog history and the information present in the database (Lester et al., 2004). When in the example the size of the pizza is not discussed in the dialog, the conversational agent may ask the user "What size of peperoni pizza would you like to order?". The agent generates answers by connecting keywords of the user's conversational output to the database of the agents' system. Some developers add a human-like personality to the text format to make the chatbot more anthropomorphic, which invites the user to interact in the conversation (Brahnam, 2009; Laurel, 1997).

Developers concede that it is important to design a virtual agent with a human-like personality (Brahnam, 2009). The personality of these agents is programmed to be shallow representations of a real human.

Many conversational agents have likes and dislikes, talk about their favorite movie, tv-show, celebrities and food. Some digital agents even have a boyfriend or girlfriend and some express different moods (Brahnam, 2009). The developers program key phrases with these aspects in their design, which can be activated depending on the conversation with the user (Brahnam, 2009). A human-like conversation can only have a negative effect when there is a high possibility of human judgement. When users are vulnerable they befit more from a logical and consistent conversation without emotion (Meyer et al., 2016).

To conclude, adding a human-like entity to a chatbot in a conversational context can mean many things. A personality can be created, specially designed for the conversational agent. Additionally, the digital agent can have its own emotional state or can express different moods. This can be expressed in the sentences the chatbot uses. The chatbot can, for example, provide a personalized greeting or mention a like or dislike for certain keywords (Brahnam, 2009; Morrissey & Kirakowski, 2013; Terzis et al., 2012). This research combines these elements to manipulate the conversational tone into a conversational agent with and without anthropomorphic conversational characteristics.

Dependent variables

The different applications of the independent variables can affect how the user perceived the conversational agent. The difference of anthropomorphism in conversational tone and visual appearance of an avatar can influence the dependent variables described below.

Likeability

People respond to and identify with virtual agents in different manners. One instrument to test if participants enjoyed the chatbot and the conversation with the chatbot is to measure the likeability. The attitude a person has towards something or someone can be evaluated by dimension such as likes-dislikes or enjoyable- unenjoyable (Ajzen & Fishbein, 2000). A positive impression, where a person is perceived as likeable, can lead to a positive evaluation of that person (Curley et al., 1986). Since, robots and computers are, to some extent, being treated as social actors, people will probably treat chatbots in the same manner (Wilson, 1999).

Brave et al. (2015) tested if emotional expressions used by chatbots could have an effect on likeability. They found that if an embodied conversational agent showed feelings of empathy in its conversational tone, participants gave the agent a more positive rating, including likeability, trustworthy and caring. Furthermore, adding an empathic text message to a design could be used to make a chatbot less annoying (Baylor & Rosenberg-Kima, 2006). This may be an indicator that an empathic text message could increase likeability. Additionally, the presence of an avatar can influence likeability. The results of Nowak (2004) showed the presence of an image was seen by participants as more likeable and more credible. Furthermore, the presence of an image reduces frustration and makes a chatbot more enjoyable to use (Baylor & Rosenberg-Kima, 2006). In addition, Nguyen and Masthoff (2009) tested the presence of a visual agent versus no visual agent present. Results showed that the participants had a more positive attitude towards the visual present agent. Furthermore, people perceived the visually present agent as more, enjoyable, likeable, empathic and caring.

Other possible factors that could increase likeability are the human-like cues used in the conversation or naming the conversational agent. Araujo (2018) has investigated that these factors were perceived as

more anthropomorphic. Moreover, using these factors resulted in an increased emotional connection in the service encounter.

User satisfaction

User satisfaction is typically considered in the information technology research as the attitude the user has toward a technical system (Wixom & Todd, 2005). In a more broadly used description of the word, satisfaction is described as "fulfilment of one's wishes, expectations, or needs, or the pleasure derived from this" (Oxford dictionary, 2018). In the case of user satisfaction, these factors are focused on the satisfaction that arises when an individual uses a device.

Researchers can analyze user satisfaction to get some perspective about the system and how the users see the technology. Furthermore, user satisfaction can be an influential factor in the intention to use the system (Lee & Choi, 2017). Additionally, developers can use this research to get a useful diagnostic for their system design. Thus, measuring the user satisfaction contributes to a useful foundation for examining and identifying the underlying structure of information and system characteristics (Wixom & Todd, 2005).

Radziwill and Benton (2017) named some research that reported enhanced user satisfaction. For example, giving conversational cues, providing a greeting and adding some type of personality to the chatbot can increase the user satisfaction (Morrissey & Kirakowski, 2013) Furthermore, it was reported that reading and responding to the moods of the participant (Meira & Canuto, 2015) and making tasks more fun and interesting (Eeuwen, 2017) can increase the satisfaction of the user. This can be an indicator that an anthropomorphic conversational tone where a personality and moods are added to the design increase the user satisfaction.

Moreover, the research by Holzwarth et al. (2006) examined the use of a visual representation of an avatar that delivered product information on a retailers webpage. They found that the use of an avatar leads to more satisfaction with the retailer. Furthermore, participants that saw the avatar were more satisfied than the participants that only saw the product information. In addition, Tinwell (2009) investigated that the human looking appearance of a character has a positive effect on the user satisfaction. Moreover, this research found that anthropomorphic looking characters were perceived as significantly more satisfactory than a photo-realistic human-like appearance. The research states that a creepy feeling can be attributed to a photo-realistic human-like agent, which can serve as a usability obstacle. This creepy feeling is called the uncanny valley theory (Tinwell, 2009).

Perceived intelligence

The virtual agent is struggling the most with formalizing human behavior. Virtual agents need human-like behavior to be perceived as intelligent (Turing, 1950). Researchers try to fake intelligent behavior, but how longer a person is chatting with a virtual agent, the more people get aware of their limitations. So far, the success of an intelligent virtual agent is limited to short conversations. When time extends, the user recognizes the patterns of the robot behavior and will eventually get bored (Bartneck et al., 2009).

Because of the fast technical development, virtual agents are getting smarter. Some agents are completely autonomous and function without the assistance of a human. In other words, the virtual agent is getting more intelligent. Intelligence is not just the ability to reason, but also the ability to be 'social smart', as the researchers Cassell et al. (2000) describe it. This means that the virtual agents need to be capable to

engage in human interest and use the appropriate speech and body behavior (Cassell et al., 2000). The more a virtual agent represents a real human the more intelligent the robot is perceived (Cassell, 2001).

Unfortunately, little is known about the appearance of an avatar and the influence of an avatar on the impression of the user, considering perceived intelligence. This study assumes that in line with the other independent variables an abstract image will be perceived as more intelligent than the anthropomorphic image or no image.

Trust

According to McKnight et al. (2002), trust is the belief that an entity has the ability, expertise, skills, benevolence and best interest in the person that evaluates that entity. Trust can encourage consumers to adopt a conversational agent. Additionally, trust in an agent can affect the perceived usefulness and the intention to use the virtual agent (Benbasat & Wang, 2005). Trust issues regarding online technological agents are complicated and is an understudied area, according to Benbasat and Wang (2005). Trust in technology can be a concern for users, because they think the technology of the virtual agents acts in the interest of the company rather than the interest of the user. Furthermore, trust issues can emerge when users are not familiar with the technology of the virtual agent (McKnight et al., 2002). A high level of trust in the technology can help overcome these concerns and encourages people to adopt the agent (Benbasat & Wang, 2005).

Trust and credibility are closely connected and credibility is a sub-phenomenon of trust. Trust is a personal judgement based on knowledge and experience and is used more often in everyday language. Credibility describes communication dimensions where it is described as a feature attributed to entities, individuals and communication products, for example, speeches or scientific literature. When the credibility is low there is no trust (Bentele & Seidenglanz, 2008). Credibility can be affected when the appearance of a virtual agent changes (Nowak, 2004). According to Nowak (2004), an agent that that has a less anthropomorphic appearance was perceived to be more credible than the agent with no image, which was more credible than an anthropomorphic image. However, research by Baylor and Ryu (2003) found that the presence of a visual agent was significantly more credible than when there was no visual agent present. Moreover, adding a visual representation of an agent next to a textual interface increases the credibility because characteristics such as gender and race are immediately recognizable. The recognizability of these factors gives a sense of trust to the user (Tseng & Fogg, 1999). Developers try to support virtual agents with personality cues. Many conversational agent more socially approachable. These aspects make a conversational agent more anthropomorphic (Brahnam, 2009).

Perceived usefulness

According to the Technology Acceptance Model (TAM), a virtual agent that is useful and easy to use will be used more frequently. Moreover, the usefulness of a conversational agent is an important predictor of the intention to use the technology (Davis, 1989). However, designing a useful system is complex due to the fast progress in technology (Hevner et al., 2004). Moreover, costumers are more willing to adapt to new technologies and accept innovations if the technology provides a unique benefit compared to existing technologies (Rogers, 2002).

A key question for the perceived usefulness is if people find a conversational agent more useful than other existing technologies or prefer an offline environment.

Intention to use

When designing a chatbot, it is important to investigate whether users have the intention to use the created conversational agents. The intention to use a conversational agent could predict the actual usage of the system (Davis, 1989). A study that tested the intention to use an embodied conversation agent is the study by Heerink et al. (2008). This study tested if elderly participants were intended to use the conversational agent. Results show that participants were more inclined to use the conversational agent when they enjoyed using the system. Furthermore, this research confirmed the assumption that the intention to use a system is a predictor of actual usage of the system. Additionally, the research of Terzis et al. (2012) tested the intention to use a conversational agent and the effect of emotional feedback with the use of a female embodied conversational agent. This study indicated that emotional feedback of the conversational agents has a positive influence on the intention to use the chatbot. Furthermore, the looks of an avatar could influence the intention to use a conversational agent. Research by Tinwell (2009) shows that a particular looking virtual character can be an obstacle to interact in a satisfying manner with the virtual character. This may be an indicator that people are not eager to use a technology based on the visual appearance of the character.

Hypotheses 1 & 2

Based on the findings, it is expected that the visual presence of an avatar will have a positive effect on the dependent variables. Moreover, this research presumes the visual representation of an abstract avatar will have the most positive influence on the dependent variables. Therefore, the following hypothesis is formulated:

• *H1 a/b/c/d/e/f: The visual representation of an abstract avatar will have the most positive influence on a) likeability, b) perceived intelligence, c) user satisfaction, d) trust, e) perceived usefulness and f) intention to use followed by the visual representation of a human face present, followed by no visual image present.*

Additionally, this study assumes that anthropomorphism in conversational tone regarding chatbots will have a positive influence the dependent variables. Therefore, the following hypothesis is formulated:

• *H2 a/b/c/d/e/f : Anthropomorphism in conversational tone will positively influence a) likeability, b) perceived intelligence, c) user satisfaction, d) trust e) perceived usefulness and f) intention to use.*

Interaction between visual appearance of an avatar and anthropomorphism in conversational tone

There is a possibility that the visual appearance of an avatar and anthropomorphism in conversational tone interact with each other. An example of a possible relation is the study of Mimoun et al. (2012). Their research explained that visual anthropomorphic characteristics can cause false expectations and disappointment. When an avatar is highly anthropomorphic and the conversation is 'robotic' an imbalance can cause disappointment, when expectations set by the anthropomorphic avatar are not met (Mimoun et al., 2012). Moreover, research by Baylor and Rosenberg-Kima (2006) has tested the presence of a visual agent combined with an empathic text message, which can be considered an anthropomorphic

characteristic (Nguyen & Masthoff, 2009). Results showed that adding an empathic text message could reduce frustration, particularly when a visual agent was present. Because the empathic text was present, the participants were able to attribute the cause of their frustration to the technology instead of to themselves.

Although, a visual present agent in a conversation leads to more socially approachable, enjoyable, valuable, credible, likeable, empathic and caring results (Baylor & Ryu, 2003; Brahnam, 2009; Nguyen & Masthoff, 2009; Tseng & Fogg, 1999). And people perceived a conversation with an anthropomorphic conversational tone as more likeable, credible, caring and empathic when anthropomorphic characteristics were used. (Araujo, 2018; Baylor & Rosenberg-Kima, 2006; Brave et al., 2015). There is no research that investigated both factors in the same research and concluded that both factors could enhance each other. Therefore, there cannot be concluded that interaction is taken place. This research will investigate the relationship between the appearance of an avatar and conversational tone in greater detail.

Therefore, the following hypotheses is formulated:

Hypotheses 3

• *H3*: There might be an interaction between the visual appearance of an avatar and anthropomorphism in conversational tone considering the dependent variables likeability, perceived intelligence, user satisfaction, trust, perceived usefulness and intention to use.

Empathy towards robots

Although specific academic research about conversational agents and emotional response is lacking, several research studies have examined the emotional response to different types of robots. Research by Carpenter (2013) investigated whether soldiers have an emotional attachment to their explosive disarm robots and the possibility that this could influence the decision-making process during a robot operation. The study found that many soldiers named the robot and were sad and angry when the robot lost his life. Some soldiers even had funerals for their robot 'friend'. However, the soldiers that got interviewed stated that attachment to the robot did not influence their decisions. Other research (Sung et al., 2007) discovered that emotional attachment to a robot vacuum cleaner increased the enjoyment for cleaning. People named their vacuum cleaner and had a guilty feeling when they felt their vacuum cleaner (iRoomba) had to work too hard.

One of the emotional expressions that can be used in a chatbot design is the feeling of empathy. Prendinger and Ishizuka (2005) studied the emotional state of empathy of a user and used this information in their chatbot interface design. The study examined the physiological data (skin conductance and electromyography) of people that used a conversational agent. The chatbot interface interprets the physiological data as emotions and gives the person the adjusted appropriate level of empathic feedback, e .g., "It seems you did not like this question so much," or "Maybe you felt a bit bad to be asked this kind of question." Users that received empathic feedback were less stressed during interaction with the conversational agent.

The feeling of empathy can be important for a chatbot design. Although, the empathy of the user towards the chatbot might likewise be an important attribute to take into consideration. Interestingly, there are essentially no studies investigating how the feeling of empathy from the user can affect the judgement towards the chatbot. The only exception is the research by Darling, Nandy and Breazeal (2015), who

investigated the role of empathy in human-robot interaction. They found that people with a high level of empathy hesitate longer to strike a robot. This hesitation was even stronger present when the robot had a story to tell. Furthermore, a study by Riek et al. (2009) concluded that a person felt more empathy towards a humanoid robot than to a mechanical looking robot.

Many studies (Brave et al., 2015; Klein et al., 2002; Prendinger & Ishizuka, 2005) is focused on investigating the reaction of participants to a chatbot that is capable of showing empathy. This research will reverse the factor of empathy and will test the empathy level of the user instead of manipulating the empathy factor of the virtual agent.

Hypotheses 4

Based on the findings, this study assumes that the feeling of empathy from the user toward a conversational agent will be a moderator and has an influence on: likeability, perceived intelligence, user satisfaction, trust, perceived usefulness and intention to use.

The following hypothesis is formulated:

• *H4 a/b/c/d/e/f: The feeling of empathy towards a chatbot might moderate the effects of visual appearance of an avatar and anthropomorphism in conversational tone and will positively influence a) likeability, b) perceived intelligence, c) user satisfaction, d) trust e) perceived usefulness and f) intention to use.*

Negative attitude towards robots

There is a concern that people have difficulties accepting robots and, therefore, create a negative attitude towards the technology (Nomura et al., 2005). When people have a negative attitude towards interactions with robots, they might also have a negative attitude towards conversational agents. The emotions and attitude towards robots affect the way people interact with them and, therefore, are important to investigate. Additionally, the negative attitude towards robots is considered a factor that prevents people from interacting with robots in everyday life (Nomura et al., 2005). Companies struggle to persuade customers to adapt to new technology and, therefore, it has become important to understand the attitude of the user.

Based on the findings, this study assumes that the negative attitude toward robots will be a moderator and moderates the relations between the dependent and independent variables.

Hypotheses 5

The following hypothesis is formulated:

• *H5 a/b/c/d/e/f: The negative attitude towards a chatbot might moderate the effects of visual appearance of an avatar and anthropomorphism in conversational tone and will positively influence a) likeability, b) perceived intelligence, c) user satisfaction, d) trust e) perceived usefulness and f) intention to use.*

Research model

The following figure shows the research model used in this research.



Method

The purpose of this study is to investigate whether the conversational tone and avatar appearance have an influence on likeability, user satisfaction, perceived intelligence, trust, perceived usefulness and intention to use. In this chapter, the research method will be discussed. First, the research design will be explained, followed by an explanation of the procedure. Thereafter, the pretests and the stimulus material will be described. Then, the characteristics of the respondents will be discussed and, lastly, the measurement instruments will be explained.

Research design

This study examined whether anthropomorphism and the visual appearance of an avatar have an influence on the dependent variables. The dependent variables in this research are, likeability, user satisfaction, perceived intelligence, trust, perceived usefulness and intention to use. Additionally, two moderators (empathy and the negative attitude towards robots) were taken into account. This study consist of a 3 x 2 design. In this research, anthropomorphism counters non-anthropomorphism in the conversational tone and the visual presence of an avatar is divided into three dimensions: no visual image, the visual presence of an abstract avatar, the visual presence of an human-looking avatar. An overview of the different conditions can be found in the table below.

Table 1Experimental conditions

Conditions	Conversational tone	Image appearance
1	Anthropomorphic	Abstract
2	Anthropomorphic	Anthropomorphic
3	Anthropomorphic	No image
4	Non-anthropomorphic	Abstract
5	Non-anthropomorphic	Anthropomorphic
6	Non-anthropomorphic	No image

Stimulus material

For the experiment, six different videos are created as stimulus material. The videos were made with the use of a website called botsociety.io and with the use of Adobe Premiere. The videos are each about 2 minutes long. In the videos a chatbot conversation on a telephone screen is shown, see figure 2. The conversation itself and the avatars were manipulated, see figure 1. The used avatars and conversation are determined by a pre-test.



Figure 2

Images of the video's where a non-conversational tone video is shown with image and without image



Pre-tests

Before the final questionnaire was distributed, three pre-tests were performed to decrease side effects as much as possible.

Pre-test 1: Determining the gender of the human-looking avatar

There are a few findings to take into consideration when designing an avatar for a human-looking conversational agent. The preferences of the user can vary in different environments. For example, in a learning environment, participants interacting with a pedagogical male agent had a higher interest in the learning task and the male agent was perceived as more intelligent than the female counterpart (Kim, Baylor, & Shen, 2007). Thus, it is necessary to determine the preference of gender, because this can vary in different conditions. Therefore, this study performed a pre-test to determine which gender the participants prefer in this particular circumstance.

In total, 18 respondents have filled in the pre-test questionnaire. In the pre-test, a short video of a chatbot conversation is displayed, where a gender-neutral avatar is present. After the video, the participants gave an answer to the following question: 'Which of the following images do you prefer?'. The participants made a choice which gender they preferred on the basis of two pictures (figure 3). The target was to get

responses from males and females, and from different ages. The results show the following composition: gender: 55.5% male and 44.5% female respondents; age (M=34.67, SD =14.1). According to the data, the male participants chose the female image (90%) more than the male image (10%). Interestingly, the decision of the female participants was equally divided, 50% chose the male image and 50% chose the female image.

Overall, the results show that 72.2% of the respondents prefer a female looking avatar for this chatbot conversation. Therefore, this research used the avatar of a female in the human-looking research condition.

Figure 3 Female and male human-looking avatar



Pre-test 2 & 3: Testing the suitable level of anthropomorphic and non-anthropomorphic characteristics in conversational tone and avatar appearance

In order to determine which conversational tone is perceived as anthropomorphic and which avatar is perceived as anthropomorphic or not, pre-test 2 and 3 were conducted. Testing the level of anthropomorphism in conversational tone is important to test to ensure that the respondents correctly perceive the conversation as anthropomorphic or non-anthropomorphic. Additionally, the human-looking avatar and abstract avatar were asked to evaluate with the use of the anthropomorphism scale. This test was conducted to examine if the human-looking avatar is correctly perceived as human and the abstract looking avatar is perceived correctly as less human. In the pre-test, the participant first saw the human-looking avatar, followed by the abstract looking avatar. For both avatars, question about the perceived anthropomorphism were asked. To test the perceived anthropomorphism the anthropomorphic scale from Bartneck et al. (2009) is used. This scale consists of five questions and uses a 5-point Likert scale, where '1=non-anthropomorphic and '5=anthropomorphic'. The scale is included in appendix.

Pre-test 2: anthropomorphism in avatar appearance

Pre-test 1 and 2 were conducted together and, therefore, both have 18 participants. The results show that, on average, participants perceived the human-looking avatar as more anthropomorphic (M = 3.48, SD = 0.76), than the abstract looking avatar (M = 2.22, SD = 0.83). The difference, 1.26, BCa 95% CI [0.86, 1.67], was significant t(17) = 6.54, p < .001. Therefore, there can be concluded that the human-looking avatar is significantly perceived as more anthropomorphic than the abstract looking avatar.

Pre-test 3: anthropomorphism in conversational tone

For the third pre-test, this study investigated whether the conversational tone is correctly perceived as anthropomorphic or non-anthropomorphic. In the pre-test participants watch a short video, the participants are randomly assigned to the anthropomorphic or non-anthropomorphic conversational tone condition. After watching the video, the participants evaluate the content by answering the following question: 'How did you think the conversation was going?'. To answer this question participants rated the conversation with the use of the anthropomorphism scale by Bartneck et al. (2009). This scale is a 5-point semantic differential scale with the following items: fake-natural, machinelike-humanlike, unconscious-conscious, artificial-lifelike.

In total, 20 respondents have completed the third pre-test questionnaire. The answers of the participants show the following composition: 50% male and 50% female participants; age (M = 35.0, SD = 14.3); education: 10% lower education, 35% intermediate education, 45% higher education and 10% university degree. On average participants gave the anthropomorphic video a higher anthropomorphism score (M = 3.42, SD = 0.87), than the non-anthropomorphic video (M = 2.64, SD = 0.64). This difference, -0.80, BCa 95% CI [-1.52, -0.81], is significant t(18) = -2.34, p = .031. Therefore, there can be stated that the conversational tone in the anthropomorphic video is correctly perceived as anthropomorphic and the non-anthropomorphic video is correctly perceived as not anthropomorphic.

Pre-test 4: Final questionnaire

All questions of the survey were pre-tested to ensure that participants do not perceive questions as incongruent and ambiguous. In total, five participants performed the pre-test and the feedback provided can be found in Appendix C. Amendments were made based on the feedback from the participants to ensure that all items are apparent and clear.

Main study

In the main study, a number of 155 respondents have completed the questionnaire. The participants were randomly selected into six different conditions. A questionnaire was used to measure the dependent variables. Moreover, questions about the moderators were asked. The main study tests if those variables are influenced by the different independent variable conditions.

Participants

For this experiment, a total of 155 participants have filled in the questionnaire. 32 questionnaires were deleted due to incomplete answers. The participants are Dutch consumers, male and female with a minimum age of 18 years. A Chi-square test showed that there were no significant differences X^2 (5) = 3.03, p = .70 between gender in the six conditions. Additionally, a one-way ANOVA shows that there are no significant differences F(5,149) = 0.26, p = .93 between the ages in the different conditions. In addition, a Chi-square showed X^2 (20) = 25.70, p = .18 that there was no significant differences between educational levels. So, The results show that there are no significant differences between participants in the different conditions. Therefore, the participants can be compared for further evaluation. Characteristics of the respondents can be found in table 2.

Table 2Distribution of sample within the experimental conditions

	Condition	Ν	Age	Gender	Level of education	
--	-----------	---	-----	--------	--------------------	--

		Μ	SD	Male	Female	1*	2*	3*	4*	5*	6*
Anthropomorphic conversation and abstract Image	26	43.42	13.31	38,5%	61,5%	0%	7.7%	30.8%	26.9%	23.1%	11.5%
Anthropomorphic conversation and anthropomorphic image	25	40.12	13.79	28%	72%	0%	4%	20%	20%	46%	0%
Anthropomorphic conversation and no image	25	43.44	15.35	40%	60%	0%	8%	28%	20%	24%	20%
Non- anthropomorphic conversation and anthropomorphic image	28	41.93	14.40	25%	75%	0%	17.9%	25%	3.6%	42.9%	10.7%
Non- anthropomorphic conversation and abstract image	26	43.35	15.10	42.3%	57.7%	0%	7.7%	26.9%	7.7%	38.5%	19.2%
Non- anthropomorphic conversation and no image	25	42.17	15.54	40%	60%	0%	4%	40%	8%	28%	20%
Total	155	42.17	14.42	35.5%	64,5%	0%	8.4%	28.4%	14.2%	35.5%	13.6%

Total15542.1714.4235.5%64,5%0%8.4%28.4%14.2% $3\overline{5.5\%}$ $1\overline{3.6\%}$ 1*= No or lower education, 2* = Intermediate vocational education 3*= higher education 4*= higher vocational education 5*=
Bachelor of Science degree 6*= University degree14.2% $3\overline{5.5\%}$ $1\overline{3.6\%}$

Procedure

The data for this research was collected by means of a questionnaire. The questionnaire was made with the online survey tool Qualtrics. The language of the survey is Dutch. The distribution of the questionnaire occurred with the use of online channels (email, social media, WhatsApp). A link was shared among the target group, which consist of Dutch people who use the internet.

The questionnaire starts with an introduction, followed by an explanation of the word chatbot. Thereafter, a video is shown, recording a staged chatbot conversation, which takes about 2 minutes to complete. There are six different videos for each of the various conditions. For example, a participant can be randomly assigned to the non-anthropomorphic conversational tone and abstract looking avatar. A script of the conversations can be found in the appendix and the visual designs can be found under the stimulus material heading.

After the video, the participants filled in the questionnaire. The questionnaire measured the effect of the independent variables on de dependent variables and manipulation checks were executed. Furthermore,

questions about the moderators were asked. Lastly, demographic questions were asked (gender, age, educational background).

Measurement instruments

The questionnaire in this study uses a semantic differentiation scale (5-point) and a 5-point Likert scale ranging from strongly disagree to strongly agree. The survey consists of 76 questions in total. Moreover, mostly existing scales for the measurement of variables were used. At the end of the survey, the demographic characteristics of the responders were asked (age, gender and level of education). The complete questionnaire can be found in the appendix.

Likeability ($\alpha = 0.91$)

To test the dependent variable likeability, the likeability construct of the Godspeed questionnaire from Bartneck et al. (2009) was used. The scale consists of five items. These items are: dislike-like, unfriendly-friendly, unkind-kind, unpleasant-pleasant, awful-nice.

Perceived intelligence ($\alpha = 0.88$)

The perceived intelligence was measured with the use of the perceived intelligence construct of the Godspeed questionnaire by Bartneck et al. (2009). The scale consists of five items. These items are: incompetent-competent, ignorant-knowledgeable, irresponsible-responsible, unintelligent-intelligent and foolish-sensible.

User satisfaction ($\alpha = 0.84$)

To test the dependent variable user satisfaction, the questionnaire for user interface satisfaction (QUIS) was used (Chin, Diehl, & Norman, 1998) This questionnaire consist of 27 items using the Likert scale (5-point). In this research, only the overall user satisfaction questions of this questionnaire were asked, consisting of 6 items. This study only tested the overall user satisfaction questions, because the conversation that took place could only be observed. The other items of the QUIS questionnaire could only be asked if the participant would be able to test the conversational agent for a long time period. Additionally, one item of the overall QUIS questionnaire was replaced, because this item could be ambiguous when translated in Dutch. Therefore the item about inadequate power – adequate power was replaced with the item: not useful –useful. An example item of the QUIS questionnaire scale was: 'Difficult – easy'.

Trust ($\alpha = 0.87$)

The level of trust can be measured using a trust scale (Cowell et al., 2003). The scale in this research is an adjusted scale of Mcknight et al. (2011) that tests trust in a specific technology. The scale consists of nine questions and uses a five-point Likert scale. An example item of this scale was: 'I have trust in the chatbot technology'.

Perceived usefulness ($\alpha = 0.76$)

The perceived usefulness is also measured with an adjusted SUS questionnaire (Brooke, 1986). The scale consists of ten statement and uses a five-point Likert scale. This research uses six question from the SUS questionnaire. The SUS questionnaire is used to test the usability of a system, because the system cannot be used and is only shown on a video the questions are adjusted to test if the system is perceived as

useful. The questions are: 'I found the chatbot unnecessary complex', 'I think the chatbot is easy to use', 'I think that I would need the learn a lot before I can use a chatbot', 'I would imagine that most people would learn to use a chatbot very quickly', 'I get why using a chatbot is useful', 'I think the chatbot is very troublesome to use', 'I think a chatbot is useful'.

Intention to use ($\alpha = 0.57$)

The intention to use is measured by means of four questions. The questions are loosely based on the System Usability Scale (SUS) by Brooke (1986). The item: 'I think I would like to use a chatbot frequently' is an adjusted item from the SUS scale. The other 2 items are asked to investigate whether participants would use the chatbot if it would be possible to use. These questions are : 'I don't think I would use a chatbot', 'I would like to use the chatbot myself'. Responses were recorded using a five-point Likert scale.

Negative attitude towards robots ($\alpha = 0.81$)

The Negative attitude towards robots scale is designed by Nomura et al. (2005). The negative attitude towards robots scale, also called the NARS, consist of 14 items. Each item is being rated with a 5-point Likert scale where 1=strongly disagree and 5= strongly agree. The NARS consists of three subscales: Attitude towards the interaction with robots, attitude towards the social influence of robots and attitude towards emotions in interaction with robots. An example item of this scale was: 'I feel that the future society will be dominated by robots'.

Empathy ($\alpha = 0.80$)

This questionnaire uses the interpersonal reactivity index from Davis (1983) to test the level of empathy of the respondents. This measurement instrument consists of 28-items and uses a 5-point Likert scale ranging from " does not describe me well' to "describes me very well". An example item of this scale was: 'I sometimes try to understand my friends better by imagining how things look from their perspective' The questionnaire uses 4 subscales with each 7 items. The subscales are: perspective taking, empathy, fantasy and personal distress.

Anthropomorphism ($\alpha = 0.89$)

To test the independent factor anthropomorphism, the human likeness item scale (also known as the Godspeed questionnaire) from Bartneck et al. (2009) was used. This questionnaire was used, because the questionnaire is designed to test different human-robot interactions (HRI). Bartneck et al. (2009) performed a literature review that measured five key concepts in de human-robot interaction (HRI) field: anthropomorphism, animacy, likeability, perceived intelligence and perceived safety. This study used the constructs, anthropomorphism, likeability and perceived intelligence of the Godspeed questionnaire. The anthropomorphism construct of Bartneck et al. (2009) is a semantic differentiation scale that consists of five items. These items are: fake-natural, machinelike-humanlike, unconscious-conscious, artificial-lifelike, moving rigidly- moving elegantly. The last item was deleted, because the avatars used in this research do not move. The anthropomorphism scale is the same scale that was used in the pre-tests to test the suitable level of anthropomorphic and non-anthropomorphic characteristics in the conversational tone and avatar appearance.

Results

In this study, six different conditions were created to examine the influence of the different conditions on the dependent variables: likeability, user satisfaction, perceived intelligence, trust, perceived usefulness and intention to use. The different conditions were created to examine the influence of anthropomorphism in conversational tone and avatar appearance. This section will present the results. First, an explanation of the manipulation check will be given, followed by a MANOVA analysis. Thereafter, the results for each dependent variable will be discussed. Additionally, the results of the moderators will be presented. Lastly, an overview of the hypotheses will be given.

Manipulation check

A manipulation check was conducted to test if the manipulations show a significant difference. An Independent sample T-test was conducted to determine the effect of anthropomorphism in the conversational tone manipulation. The respondents evaluated all items on a 5-point Likert scale where the results range from "totally disagree' (1) to 'totally agree (5). The analysis shows a significant difference between the anthropomorphic (M = 3.29, SD = 0.77) and non-anthropomorphic (M = 2.84, SD = 0.83) condition, with t(153)=-3.48, p < 0.001. These results suggest that the anthropomorphic video is perceived as significantly more anthropomorphic than the non-anthropomorphic video. Subsequently, an Independent sample T-test for the avatar appearance was performed. The analysis shows that there are no significant differences in anthropomorphism in the anthropomorphic (M = 3.00, SD = 0.74) and abstract looking avatar conditions (M = 3.09, SD = 0.67), with t(103) = 0.23, p = 0.817. Because the manipulation check reveals that there is no significant difference between the anthropomorphic image and the abstract image, they will be combined in an overall image variable. The results below will test a 2 x 2 design where an avatar or no avatar and an anthropomorphic conversational tone versus a non-anthropomorphic conversational tone will be presented.

Multivariate analysis of variance (MANOVA)

This study used a MANOVA analysis to examine the different effects of the independent variables (anthropomorphic conversational tone and avatar presence) on the dependent variables (likeability, user satisfaction, perceived intelligence, trust, perceived usefulness and intention to use). The results are presented separately for each dependent variable.

Before analyzing the results a Wilks' Lambda was conducted to test the general effect between the independent and dependent variables. The Wilks' Lambda reveals that there was a significant difference in conversational tone $\Lambda = 0.92$, F(6, 146) = 2.20, p = 0.046. Additionally, the Wilks' Lambda reveals no significant results in avatar presence $\Lambda = 0.98$, F(6, 146) = 0.51, p = 0.80. Furthermore, the Wilks' Lambda reveals no significantly interaction effect between the two independent variables $\Lambda = 0.96$, F(6, 146) = 0.97, p = 0.45.

Effects on likeability

The mean scores and standard deviation of likeability are displayed in table 5. The hypothesis of likeability was that likeability will be higher when an avatar was present. Additionally, it was hypothesized that an anthropomorphic conversational tone will positively influence likeability.

The results of the MANOVA show no significant effect between in the avatar presence conditions (*F* 1, 151) = 1.18, p = 0.63). A small difference was found in the conversational tone conditions, where the anthropomorphic conversational tone has a higher mean (M = 3.67) than the non-anthropomorphic conversational tone (M = 3.43). However, the difference is not significant (*F* 1, 151) = 3.10, p = 0.08). Furthermore, no interaction effect was found between the conversational tone and avatar presence on likeability (*F* 1, 151) = 0.74, p = 0.39).

Table 5

Mean scores and standard deviations of the scores on likeability

		Conversatio	onal tone	-		
Avatar presence	Anthropomorphic	N	Non- anthropomorphic	Ν	Total	Ν
Image	3.69 (0.71)	51	3.34 (0.88)	54	3.51 (0.82)	105
No image	3.64 (0.84)	25	3.52 (0.57)	25	3.58 (0.71)	50
Total	3.67 (0.75)	76	3.40 (0.80)	79		155

5-point scale from "totally disagree" (1) to "totally agree' (5)

Effects on user satisfaction

The MANOVA results show that the hypotheses regarding the positive effect of an avatar and anthropomorphic conversational tone are not supported. Results show that the highest user satisfaction is measured when a non-anthropomorphic conversational tone and no avatar is used (M = 3.61). However, the differences in results are too small to measure an effect in the conversational tone conditions (F 1, 151) = 0.00, p = 0.96) and the avatar presence conditions (F 1, 151) = 0.70, p = 0.40). Additionally, no interaction was found between the two independent variables (F 1, 151) = 0.34, p = 0.56).

Table 6

Mean scores and standard deviations of the scores user satisfaction

	C	conversational to				
Avatar presence	Anthropomorphic	Ν	Non- anthropomorphic	Ν	Total	N
Image	3.50 (0.67)	51	3.44 (0.70)	54	3.47 (0.68)	105
No image	3.53 (0.70)	25	3.61 (0.63)	25	3.57 (0.67)	50
Total	3.51 (0.68)	76	3.49 (0.68)	79		155

5-point scale from "totally disagree" (1) to "totally agree' (5)

Effects on trust

It was hypothesized that an anthropomorphic conversational tone and the presence of an avatar would have a positive effect on likeability. The results in table 7 show that there is nearly a difference between the means in the different conditions. No significant effect was found for the conversational tone (F 1, 151) = 0.16, p = 0.69). Additionally, no significant effect was found for avatar presence conditions (F 1, 151) = 0.11, p = 0.75). Together this resulted in no interaction effect between avatar appearance and conversational tone (F 1, 151) = 0.04, p = 0.84).

Table 7

Mean scores a	nd standard	deviations	of the	scores	trust
---------------	-------------	------------	--------	--------	-------

	C	onversation	nal tone	-		
Avatar presence	Anthropomorphic N		Non- anthropomorphic	Ν	Total	N
Image	2.22 (0.55)	51	2.20 (0.64)	54	2.21 (0.59)	105
No image	2.27 (0.59)	25	2.21 (0.45)	25	2.24 (0.52)	50
Total	2.24 (0.56)	76	2.20 (0.58)	79		155

5-point scale from "totally disagree" (1) to "totally agree' (5)

Effects on perceived intelligence

The formulated hypotheses state that an anthropomorphic conversational tone and the appearance of an image will increase the perceived intelligence. The results show that the condition where an anthropomorphic conversational tone and image were present scored higher for perceived intelligence than the condition with an image where a non-anthropomorphic conversational tone was used. However, no significant results were found in the conversational tone condition (*F* 1, 151) = 0.60, *p* = 0.44) and in the avatar appearance condition (*F* 1, 151) = 0.24, *p* = 0.63). Lastly, no interaction effect has been found between the independent variables (*F* 1, 151) = 1.10, *p* = 0.30).

Table 8

Mean scores and standard deviations of the scores perceived intelligence

		Conversatio	onal tone	-		
Avatar presence	Anthropomorphic	Ν	Non- anthropomorphic	Ν	Total	N
Image	3.71 (0.66)	51	3.50 (0.78)	54	3.60 (0.73)	105
No image	3.64 (0.67)	25	3.68 (0.51)	25	3.66 (0.59)	50
Total	3.69 (0.67)	76	3.56 (0.71)	79		155

5-point scale from "totally disagree" (1) to "totally agree' (5)

Effects on perceived usefulness

The hypothesis regarding the effect of perceived usefulness was that the use of an image will have a positive effect on the perceived usefulness. Additionally, this research hypothesized that an anthropomorphic conversational tone will have a positive effect on perceived usefulness. However, the results of the MANOVA analyses show that no significant results in the conversational tone condition (F 1, 151) = 1.80, p = 0.18) and the avatar appearance condition (F 1, 151) = 1.81, p = 0.18). Additionally, results reveal an interaction effect. However, the interaction effect is too small to support the hypothesis (F 1, 151) = 3.50, p = 0.06).

Table 9Mean scores and standard deviations of the scores on perceived usefulness

	C	onversation	al tone	_	
Avatar presence	Anthropomorphic	Ν	Non- anthropomorphic	Ν	Total N
Image	2.19 (0.42)	51	2.24 (0.56)	54	2.21 (0.49) 105
No image	2.23 (0.52)	25	1.96 (0.48)	25	2.10 (0.52) 50
Total	2.20 (0.45)	76	2.15 (0.55)	79	155

5-point scale from "totally disagree" (1) to "totally agree' (5)

Effects on intention to use

This research hypothesized that an anthropomorphic conversational tone and image appearance will positively influence intention to use. Furthermore, it is suspected that there might be an interaction effect regarding the conversational tone and avatar appearance.

The results reveal no significant differences in the conversational tone regarding likeability (F 1, 151) = 0.63, p = 0.43). Additionally, no significant results were found in the avatar appearance conditions (F 1, 151) = 2.21, p = 0.14). Lastly, no interaction effect has been found between the independent variables (F 1, 151) = 0.04, p = 0.84).

Table 10

Mean scores and standard deviations of the scores on intention to use

		Conversatio	onal tone	-		
Avatar presence	Anthropomorphic	Ν	Non- anthropomorphic	Ν	Total	N
Image	3.18 (0.80)	51	3.33 (1.02)	54	3.26 (0.92)	105
No image	2.97 (0.88)	25	3.07 (0.95)	25	3.02 (0.91)	50
Total	3.11 (0.83)	76	3.25 (1.00)	79		155

5-point scale from "totally disagree" (1) to "totally agree' (5)

Moderator: Negative attitude towards robots

Moderation is shown by a significant interaction effect. Results show that there is a significant interaction effect between the negative attitude towards robots and the dependent variables, trust, perceived intelligence and perceived usefulness. First, the results reveal that the explained variance of the regression model in trust is significant with $R^2 = 0.09$, F(3,151) = 5.20, p < .001. Additionally, the negative attitude towards robots moderates the relation between the conversational tone of the chatbot and trust with an interaction effect of b = 0.16, 95% CI [-0.70, -0.05], t = -2.30, p = 0.02.

Second, the results indicate that the explained variance of the perceived intelligence model is significant with $R^2 = 0.09$, F(3,151) = 5.26, p < .001. In addition, the relation between the conversational tone and the perceived intelligence is moderated by the negative attitude towards robots with an interaction effect of b= 0.20, 95% CI [-0.89, -0.12], t = -2.61 p = 0.01. The third significant moderating effect is that of avatar appearance and perceived usefulness interacting with the moderator, negative attitude towards robots (b = 0.16, 95% CI [0.03, 0.66], t = 2.19, p = 0.03), with an explained variance of the regression model for perceived usefulness, $R^2 = 0.17$, F(3,151) = 10.24, p < .001.

Moderator: Empathy

Results show that there is no significant interaction effect between the moderator empathy and the independent and dependent variables. The hypotheses where a possible moderation effect of empathy has been claimed is not supported.

Overview of the hypotheses

Table 11Overview of the hypotheses

hypotheses		supported
H1a	The visual representation of an abstract avatar will have the most positive influence on likeability , followed by the visual representation of a human face present, followed by no visual image present.	No
H1b	The visual representation of an abstract avatar will have the most positive influence on perceived intelligence , followed by the visual representation of a human face present, followed by no visual image present.	No
H1c	The visual representation of an abstract avatar will have the most positive influence on user satisfaction , followed by the visual representation of a human face present, followed by no visual image present.	No
H1d	The visual representation of an abstract avatar will have the most positive influence on trust , followed by the visual representation of a human face present, followed by no visual image present.	No
Hle	The visual representation of an abstract avatar will have the most positive influence perceived usefulness , followed by the visual representation of a human face present, followed by no visual image present.	No
H1f	The visual representation of an abstract avatar will have the most positive influence on intention to use , followed by the visual representation of a human face present, followed by no visual image present.	No
H2a	Anthropomorphism in conversational tone will positively influence likeability.	No
Н2Ь	Anthropomorphism in conversational tone will positively influence perceived intelligence.	No
H2c	Anthropomorphism in conversational tone will positively influence user satisfaction.	No
H2d	Anthropomorphism in conversational tone will positively trust.	No
H2e	Anthropomorphism in conversational tone will positively influence perceived usefulness.	No
H2f	Anthropomorphism in conversational tone will positively influence intention to use.	No
H3	There might be an interaction between the visual appearance of an avatar and anthropomorphism in conversational tone considering the dependent	No

	variables likeability perceived intelligence user satisfaction trust	
	perceived usefulness and intention to use	
H4a	The feeling of empathy towards a chatbot will moderate the effects of visual appearance (presence) of an avatar and anthropomorphism in conversational tone and will positively influence likeability .	No
H4b	The feeling of empathy towards a chatbot will moderate the effects of visual appearance (presence) of an avatar and anthropomorphism in conversational tone and will positively influence perceived intelligence .	No
H4c	The feeling of empathy towards a chatbot will moderate the effects of visual appearance (presence) of an avatar and anthropomorphism in conversational tone and will positively influence user satisfaction .	No
H4d	The feeling of empathy towards a chatbot will moderate the effects of visual appearance (presence) of an avatar and anthropomorphism in conversational tone and will positively influence trust .	No
H4e	The feeling of empathy towards a chatbot will moderate the effects of visual appearance (presence) of an avatar and anthropomorphism in conversational tone and will positively influence perceived usefulness.	No
H4f	The feeling of empathy towards a chatbot will moderate the effects of visual appearance of an avatar and anthropomorphism in conversational tone and will positively intention to use.	No
H5a	The negative attitude towards robots will moderate the effect of visual appearance (presence) of an avatar and anthropomorphism in conversational	No
H5b	tone and will negatively influence likeability. The negative attitude towards robots will moderate the effect of visual appearance (presence) of an avatar and anthropomorphism in conversational	Partially supported
Н5с	The negative attitude towards robots will moderate the effect of visual appearance (presence) of an avatar and anthropomorphism in conversational tone and will negatively influence user satisfaction	No
H5d	The negative attitude towards robots will moderate the effect of visual appearance (presence) of an avatar and anthropomorphism in conversational	Partially supported
Н5е	The negative attitude towards robots will moderate the effect of visual appearance (presence) of an avatar and anthropomorphism in conversational tone and will negatively influence paracived wasfulness.	Partially supported
H5f	The negative attitude towards robots will moderate the effect of visual appearance (presence) of an avatar and anthropomorphism in conversational tone and will negatively influence intention to use.	No

Discussion

The aim of this study was to gain more insight into the effects of anthropomorphism in conversational tone and avatar presence. Additionally, the effect of empathy and a negative attitude towards robots on the dependent variables in the different conditions were measured and explored. Six different conditions were created to test the effect of anthropomorphism in avatar appearance and conversational tone on likeability, perceived intelligence, user satisfaction, trust, perceived usefulness and intention to use. In this section, the results are explained and the implications and limitations of this study will be discussed.

Discussion of the results

The goal of this study was to answer the following research question: 'What are the effects of avatar appearance and anthropomorphic characteristics in conversational tone regarding chatbot use in the customer service field?'. This question was studied with the use of an online questionnaire, where six different conditions were created to test the effects of the independent variables on the dependent variables. In the conditions, anthropomorphism counters non-anthropomorphism in the conversational tone and the visual presence of an avatar is divided into three dimensions: no visual avatar, the visual presence of an abstract avatar, the visual presence of an human-looking avatar. Additionally, an interaction effect between the two independent variables was investigated. First, a manipulation check was conducted to investigate if the different manipulations in the conditions were significant. The manipulation check revealed that there was no significant difference between the abstract avatar appearance and anthropomorphic avatar appearance conditions. Therefore, this research has combined these conditions. This means that the results are analyzed using a 2 x 2 design, where the presence or absence of an avatar counters an anthropomorphic conversational tone or a non-anthropomorphic conversational tone.

The conversational tone of a chatbot

The results showed no significant relation between the anthropomorphic conversational tone and the non-anthropomorphic conversational tone considering the dependent variables. It was expected that an anthropomorphic conversational tone would be perceived as more positive than a non-anthropomorphic conversational tone. The results show that there is a positive trend for likeability (not significant) in the anthropomorphic conversational tone condition, yet no positive trend was detected for the other dependent variables. The findings might be explained by the way the participant is involved in the conversation. Using anthropomorphism in virtual agents invites the user to interact in a conversation (Laurel, 1997). Since this was a controlled situation, the participants could not really interact with the chatbot and not really order a pizza. There is a possibility that people might be less involved in the chatbot conversation and were less positive, because real interaction was not possible. Additionally, when there is less involvement in the conversation, participants might not care how the chatbot conversed with the user.

The results indicate a positive trend (not significant) in likeability when an anthropomorphic conversational tone was used. This indication can be researched further. There is a possibility that the result will be significant when interaction with the chatbot is possible. Another possible explanation for the fact no significant results were found is the product that was offered in this research. The participants were exposed to one type of product/company. There is a possibility that different products will give different results. For example, chatbots that are used as companions for the elderly might benefit from a more anthropomorphic conversational tone. Further research should, therefore, focus more on the effect of different types of products and/or companies.

The presence of a chatbot avatar

Based on the results of this study, it can be concluded that no significant effects were found in the different avatar presence conditions. This result is contradicting the study of Baylor and Rosenberg-Kima, (2006), who state that a chatbot with an avatar is more enjoyable, and the study of Nguyen and Masthoff (2009), who concluded that participants had a more positive attitude towards a visual present agent. There is a possibility that the uncanny valley effect could have caused some negative effect (Mori, 1970). It is possible that the participant will respond more negatively, when an avatar gives a participant a creepy feeling (uncanny valley). For example, it might be creepy for users that are new to the chatbot technology to see an avatar and realize that behind this avatar is not a real human. This research did not test a possible uncanny valley effect. Further research should, therefore, include the uncanny valley in their research model, for example as a moderator, to investigate this possibility. Additionally, different avatar appearances can be tested, for example, a picture of a real human. Different avatar images might lead to different results (Tinwell, 2009).

Interaction effect

This research examined the possible interaction effect between these two independent variables. There can be concluded that no interaction has taken place in the different conditions. However, more research is necessary to investigate this effect further. Especially since research about interaction effects between anthropomorphic characteristics in the conversational tone and avatar appearance is scarce. For further research, it might be of importance to measure the expectations someone has about the chatbot before and after presenting the virtual agent. It is possible that the avatar presented was not in line with the information the chatbot provided in the conversation, which caused false expectations and disappointment (Mimoun et al., 2012); This effect can be investigated further to gain more insight into a possible mismatch between the conversational tone and avatar appearance. When the limitations of this research are diminished, a possible interaction effect might still occur.

Moderator: Empathy

Results show that there is no significant interaction effect between the moderator empathy and the independent and dependent variables. This means that a user's level of empathy did not influence the way the chatbot was perceived. This information can give insight into the emotion, or lack of emotion, a user has towards a chatbot. Although, the results indicate that the empathy a user had towards the chatbot did not influence the way the chatbot was perceived, the study by Darling, Nandy and Breazeal (2015) found that people with a high level of empathy hesitate longer to strike a robot. It is possible that the results are different when a different type of robot is used. Moreover, this study used a physical present robot whereas the participant in this study was presented with a digital chatbot. However, it is possible that the variable 'negative treatment of a robot' has a stronger influence on the perception of the participant than the dependent variables used in this research. Since specific academic research about conversational

agents and the empathy of a user is lacking, further research is necessary to capture the full effect empathy has on the way the chatbot is perceived. Additionally, the empathy a person feels towards the chatbot can be further investigated. For example, research by Riek et al. (2009) investigated this and concluded that a person can feel more empathy towards a humanoid robot than to a mechanical looking robot.

Moderator: Negative attitude towards robots

The results of this study show that there is a significant interaction effect between the negative attitude towards robots and the dependent variables, trust, perceived intelligence and perceived usefulness. The results reveal that the negative attitude towards robots negatively moderates the relation between the conversational tone of the chatbot and the dependent variables trust and perceived intelligence. Moreover, a moderation effect has been found between the negative attitude towards robots, avatar presence and perceived usefulness.

The results confirm the concern that a high level of negative attitude towards robots gives a more negative result towards perceived intelligence and trust in the conversational tone and perceived usefulness in avatar presence. When people have difficulties accepting a robot, a negative attitude towards robots affect the way people interact with them and, therefore, are important to investigate. Additionally, the negative attitude towards robots is considered a factor that prevents people from interacting with robots in everyday life (Nomura et al., 2005). Companies struggle to persuade customers to adapt to new technologies and, therefore, it has become important to understand the attitude of the user. The results reveal that especially the perceived intelligence is affected in both conditions when there was a high negative attitude towards robots. This might mean that people will not use the robot. This information is important to investigate further to examine the reason why people with a negative attitude perceive a chatbot as less intelligent. There is a possibility that people are skeptical about the performance of the chatbot. It is possible that this skepticism is created by the vulnerability, technical complications and inconsistent responses the chatbots of today have (Newman, 2016).

Practical implementations

This study resulted in a better understanding of anthropomorphic characteristics in conversational agents. The results of this study can be used for marketing implementations. Marketers can use this study to adopt the appropriate level of anthropomorphic characteristics because, different levels of anthropomorphism can have an effect on how the conversational agent is perceived. The developers can use this insight to experiment with different levels of anthropomorphism in the conversational tone and avatar appearance to improve the user experience in the chatbot conversation. Developers can, for example, experiment with adding a personal greeting to a chatbot conversation. It is possible that adding a personal greeting results in a more positive attitude toward the conversational agent. Additionally, the use of conversational agents might increase when the chatbot offers a more satisfying experience. It is possible that the conversational agent improves and costumers choose the conversational agent over human contact, human to human interaction might no longer be required in customer service support.

Limitations and recommendations

Further research can improve this research model and reduce the limitations this research has. One of the limitations of this research is the questionnaire order in the pre-test. First, questions were asked about the anthropomorphism of the avatar. It is possible that confirmation bias occurred, because both avatars were shown. When the abstract avatar was shown followed by the anthropomorphic avatar, people could assume that it was expected that the second image represented a more human-like avatar. Confirmation bias can be prevented by splitting the participants into groups, where some will be presented with the abstract avatar questions and some will be presented with the anthropomorphic avatar questions.

Second, this research uses a video where a conversation between a customer and a chatbot is shown. Possibly, results will be more reliable when an actual conversation with the chatbot is possible. Moreover, there is a possibility that participants were not paying attention during the video. This can be prevented if participants have to interact with the robot. When interaction is possible, the participants have to pay attention to the chatbot and other variables can be tested, such as the ease of use. Another attribute that could be tested is the appearance of the chatbot itself. Different avatars could lead to different results. This research used an abstract avatar and human-looking avatar. For example, researchers can test a picture of a real human in their chatbot design. Additionally, if the chatbot avatar was able to move, research possibilities would extend. For example, the effect of animacy could be tested (Bartneck et al., 2009). The research of Darling, Nandy and Breazeal (2015) revealed that the movement of a robot increased the hesitation to strike the robot. Moreover, in this research, a conversation of a pizza order is shown. Different conversations could be tested with different products and outcomes.

Additionally, different environments could be tested. This research has tested business transactions for a pizza place. The preferences of the user for a certain type of chatbot can vary in different environments. For example, in a learning environment, participants interacting with a pedagogical male agent had a higher interest in the learning task and the male agent was perceived as more intelligent than the female counterpart (Kim, Baylor, & Shen, 2007). Moreover, when humans are vulnerable, for example in a medical environment, it is possible that a human-like conversation can have a negative effect. When users are vulnerable, they benefit more from a logical and consistent conversation without emotion (Meyer et al., 2016). It might, therefore, be important to test different environments.

Recommendations for future studies

This study is a step towards a better understanding of anthropomorphic characteristics in the conversational tone and avatar appearances for conversational agents. The results of this research reveal that people perceive a conversational agent as less intelligent and have less trust in the chatbot when a negative attitude towards robots was reported. It is important to investigate this further to examine the reason why people with a negative attitude perceive a chatbot as less intelligent and why people have less trust in the chatbot. Additionally, this research has tested some variables that can vary when anthropomorphic characteristics change. There are more variables that can be investigated and probably will be affected when avatar appearance and conversational tone change. Further research can extend this research and investigate other variables, conversations and avatar appearances.

This research addresses gaps in the literature on conversational agents, particularly considering the combination of anthropomorphic characteristics in chatbot appearance and conversational tone. Furthermore, it contributes to a better understanding of the effect of anthropomorphism in chatbots. To

conclude, more research is needed to give more insight into the effect of anthropomorphism in conversational tone and avatar appearance in different environments and purposes to provide the most satisfying and useful chatbot experience.

References

Ajzen, I., & Fishbein, M. (2000). Attitudes and the attitude-behavior relation: Reasoned and automatic processes. *European Review of Social Psychology*, *11*(1), 1-33. doi:https://doi.org/10.1080/14792779943000116

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

Bartneck, C., Kulić, D., Croft, E., & Zoghbi, S. (2009). Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International Journal of Social Robotics*, *1*(1), 71-81. doi:10.1007/s12369-008-0001-3

Baylor, A. L., & Rosenberg-Kima, R. B. (2006). Interface agents to alleviate online frustration. *Proceedings of the International Conference on Learning Sciences, 7,* 30–36. Retrieved from https://dl.acm.org/citation.cfm?id=1150039

Baylor, A. L., & Ryu J. (2003). The effects of image and animation in enhancing pedagogical agent persona. *Journal of Educational Computer Research, 28*(4), 373–395. doi: https://doi.org/10.2190/V0WQ-NWGN-JB54-FAT4

Bentele, G. & Seidenglanz, R. (2008). Trust and Credibility — Prerequisites for Communication Management. In Zerfass A., van Ruler B., Sriramesh K. (eds). *Public Relations Research - VS Verlag für Sozialwissenschaften* (pp. 49-62). Wiesbaden, Germany; Springer. doi:10.1007/978-3-531-90918-9_4

Benbasat, I., & Wang, W. (2005). Trust in and adoption of online recommendation agents. *Journal of the Association for Information Systems, 6*(3), 4. Retrieved from https://aisel.aisnet.org/cgi/viewcontent.cgi?referer=https://www.google.nl/&httpsredir=1&article=1454&c ontext=jais

Brahnam, S. (2009) Building Character for Artificial Conversational Agents: Ethos, Ethics, Believability, and Credibility. *PsychNology Journal 7*(1), 9-47. Retrieved from http://www.psychnology.org/File/PNJ7(1)/PSYCHNOLOGY_JOURNAL_7_1_BRAHNAM.pdf

Brave, S., Nass, C., & Hutchinson, K. (2005). Computers that care: investigating the effects of orientation of emotion exhibited by an embodied computer agent. *International Journal of Human-Computer Studies*, *62*(2), 161-178. doi:https://doi.org/10.1016/j.ijhcs.2004.11.002

Brooke, J. (1996). SUS-A quick and dirty usability scale. *Usability evaluation in industry*, *189*(194), 4-7. Retrieved from https://hell.meiert.org/core/pdf/sus.pdf

Carpenter, J. (2013). Just Doesn't Look Right: Exploring the impact of humanoid robot integration into Explosive Ordnance Disposal teams. In R. Luppicini (Ed.), *Handbook of Research on Technoself: Identity in*

a Technological Society (pp. 609-636). University of Washington, USA: IGI Global. doi:10.4018/978-1-4666-2211-1.

Cassell, J. (2001). Embodied conversational agents: representation and intelligence in user interfaces. *Al magazine, 22*(4), 67. doi:https://doi.org/10.1609/aimag.v22i4.1593

Cassell, J., Bickmore, T., Campbell, L., & Vilhjálmsson, H. (2000). Designing embodied conversational agents. In Cassel, J., Sullivan, J., Churchill, E. & Prevost, S. (eds.) *Embodied Conversational Agents* (pp. 29-63). Cambridge, USA: MIT Press. Retrieved from: https://books.google.nl/books?hl=nl&lr=&id=tHiKZGh9t7sC&oi=fnd&pg=PA29&ots=D-

XNsXiBzs&sig=Xnfpl_UfW85vRxVMJKRbhHOAFMc#v=onepage&q&f=false

Chin, J. P., Diehl, V. A., & Norman, K. L. (1988). Development of an instrument measuring user satisfaction of the human-computer interface. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 1*, 213-218. doi:10.1145/57167.57203

Cowell, A., & Stanney, K. (2005). Manipulation of non-verbal interaction style and demographic embodiment to increase anthropomorphic computer character credibility. *International Journal of Human-Computer Studies, 62*(2), 281-306. doi:10.1016/j.ijhcs.2004.11.008

Curley, S. P., Yates, J. S. & Abrams, R. A. (1986). Psychological sources of ambiguity avoidance. *Organizational Behavior and Human Decision Processes*, *38*(2), 230–256. doi:https://doi.org/10.1016/0749-5978(86)90018-X

Darling, K., Nandy, P., & Breazeal, C. (2015). Empathic concern and the effect of stories in human-robot interaction. *Robot and Human Interactive Communication 24*(1), 770-775. doi:10.1109/ROMAN.2015.7333675

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, *44*(1), 113–126. doi:http://dx.doi.org/10.1037/0022-3514.44.1.113

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319–340. doi:10.2307/249008

Duffy, B. R. (2003) Anthropomorphism and the social robot. *Robotics and Autonomous Systems, 42*(3-4), 177-190. doi:10.1016/S0921-8890(02)00374-3

Duffy, B. R., Joue, G., & Bourke, J. (2002). Issues in assessing performance of social robots. In *Proceedings of the WSEAS International Conference, 2*, 1-8. Retrieved from https://www.researchgate.net/profile/Brian_Duffy3/publication/228726145_Issues_in_assessing_performan ce_of_social_robots/links/0046352b0b04cc5193000000/Issues-in-assessing-performance-of-social-robots.pdf

Eeuwen, M. (2017). *Mobile conversational commerce: messenger chatbots as the next interface between businesses and consumers* (Master's thesis, University of Twente, Enschede, the Netherland). Retrieved from: http://essay.utwente.nl/71706/1/van%20Eeuwen_MA_BMS.pdf

Erickson, T. (1997) Designing agents as if people mattered. In J. Bradshaw, J (eds). *Software agents*. (pp. 79-96). Cambridge, USA: MIT Press. Retrieved from https://dl.acm.org/citation.cfm?id=267997

Ferrara, E., Varol, O., Davis, C., Menczer, F., & Flammini, A. (2016). The rise of social bots. *Communications of the ACM, 59*(7), *96-104.* doi:10.1145/2818717

Foner, L. (1993). What's an agent, anyway? a sociological case study. *Agents Memo, 93*(1). Boston, US: Media Lab, Massachusetts Institute of Technology. Retrieved from http://www.upv.es/sma/teoria/agentes/what%20is%20an%20agent-foner.pdf

Fong, T., Nourbakhsh, I., & Dautenhahn, K. (2003). A survey of socially interactive robots. *Robotics and Autonomous Systems*, *42*(3-4), 143-166. doi: https://doi.org/10.1016/S0921-8890(02)00372-X

Heerink, M., Kröse, B., Wielinga, B., & Evers, V. (2008). Enjoyment intention to use and actual use of a conversational robot by elderly people. *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction, 3*(1), 113-119. doi:10.1145/1349822.1349838

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *Management Information Systems Quarterly*, *28*(1), 75-105. Retrieved from https://www.researchgate.net/publication/201168946_Design_Science_in_Information_Systems_Research

Hirsch, E. D., Kett, J. F., & Trefil, J. S. (2002). *The new dictionary of cultural literacy*. Houghton Mifflin Harcourt

Holzwarth, M., Janiszewski, C. & Neumann, M. M. (2006). The Influence of Avatars on Online Consumer Shopping Behavior. *Journal of Marketing, 70*(4), 9-36. doi: https://doi.org/10.1509/jmkg.70.4.19

Klein, J., Moon, Y., & Picard, R. W. (2002). This computer responds to user frustration: Theory, design, and results. *Interacting with Computers*, *14*(2), 119-140. doi: https://doi.org/10.1016/S0953-5438(01)00053-4

Kim, Y., Baylor, A., & Shen, E. (2007). Pedagogical agents as learning companions: The impact of agent emotion and gender. *Journal of Computer Assisted Learning, 23*(3), 220-234. doi:10.1111/j.1365-2729.2006.00210.x

Laurel, B. (1997). Interface agents: metaphors with character. In *Software agents*, Jeffrey M. Bradshaw (Ed.). *MIT Press, Cambridge, MA, USA, 1*, 67-77.

Lee, S. & Choi, J. (2017). Enhancing user experience with conversational agent for movie recommendation: Effects of self-disclosure and reciprocity. *International Journal of Human-Computer Studies*, *103*(1), 95-105. doi:10.1016/j.ijhcs.2017.02.005

Lester, J., Branting, K., & Mott, B. (2004). Conversational agents. In. Munindar, P.S. (ed.) *The Practical Handbook of Internet Computing* (pp. 220-240). London, UK: Chapman & Hall

Meira, M. O., & Canuto, A. M. P. (2015). Evaluation of Emotional Agents' Architectures: an Approach Based on Quality Metrics and the Influence of Emotions on Users. In *Proceedings of the World Congress on Engineering*, *1*(1), 1-8. Retrieved from

https://pdfs.semanticscholar.org/3494/f710bbd35afeef2577234828f8545c42302a.pdf

Meyer, J., Miller, C., Hancock, P., de Visser, E. J. & Dorneich, M. (2016). Politeness in machine-human and human-human interaction. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 60(1), 279–283. doi:10.1177/1541931213601064

Mcknight, D. H., Carter, M., Thatcher, J. B., & Clay, P. F. (2011). Trust in a specific technology: An investigation of its components and measures. *ACM Transactions on Management Information Systems*, *2*(2). doi:10.1145/1985347.1985353

McKnight, D. H., Choudhury, V. & Kacmar, C. (2002). Developing and Validating Trust Measures for e-Commerce: An Integrative Typology. *Information Systems Research, 13*(3), 334-359. doi: https://doi.org/10.1287/isre.13.3.334.81

McTear, M., Callejas, Z., & Griol Barres, D. (2016). *The conversational interface : Talking to smart devices*. Switzerland: Springer, pp. 81-94.

Mimoun, M. S. B., Poncin, I., & Garnier, M. (2012). Case study—Embodied virtual agents: An analysis on reasons for failure. *Journal of Retailing and Consumer services*, *19*(6), 605-612. doi: https://doi.org/10.1016/j.jretconser.2012.07.006

Mindbowser (2017) Chatbot survey. Retrieved from http://mindbowser.com/chatbot-market-survey-2017/

Morgan, B. (2017) How chatbots improve customer experience in every industry an infographic. Retrieved from https://chatbotsjournal.com/how-chatbots-improve-customer-experience-in-every-industry-an-infograph-9a9b054ccc56

Mori, M. (1970). The uncanny valley. *Energy*, 7(4), 33-35. Retrieved from http://www.movingimages.info/digitalmedia/wp-content/uploads/2010/06/MorUnc.pdf

Morrissey, K., & Kirakowski, J. (2013). 'Realness' in Chatbots: Establishing Quantifiable Criteria. *Lecture Notes in Computer Science: vol 8007.* International Conference on Human-Computer Interaction (pp. 87-96). Berlin, Germany: Springer. doi: https://doi.org/10.1007/978-3-642-39330-3_10

Natanson, E. (2017) Artificial Intelligence smart assistants the next big thing in computing. Retrieved from https://www.forbes.com/sites/eladnatanson/2017/06/22/artificial-intelligence-smart-assistants-the-next-big-thing-in-computing/#247be79b4252

Newman (2016) Chatbots and the future of conversation based interface. Retrieved from https://www.forbes.com/sites/danielnewman/2016/05/24/chatbots-and-the-future-of-conversation-based-interfaces/#7aed3a8b1fc7

Nowak, K. L. (2004). The influence of anthropomorphism and agency on social judgment in virtual environments. *Journal of Computer-Mediated Communication*, *9*(2). doi:10.1111/j.1083-6101.2004.tb00284.x

Nowak, K. L., & Biocca, F. (2003). The effect of the agency and anthropomorphism on users' sense of telepresence, copresence, and social presence in virtual environments. *Presence: Teleoperators & Virtual Environments*, *12*(5), 481-494. doi: 10.1162/105474603322761289.

Nomura, T., Kanda, T., Suzuki, T., & Kato, K. (2005). People's assumptions about robots: Investigation of their relationships with attitudes and emotions toward robots. *In . IEEE International Workshop on Robot and Human Interactive Communication*, *1*, 125-130. doi: 10.1109/ROMAN.2005.1513768

Nguyen, H., & Masthoff, J. (2009). Designing empathic computers: the effect of multimodal empathic feedback using animated agent. *Proceedings of the International Conference on Persuasive Technology*, 4. doi: 10.1145/1541948.1541958

Oxxio (2017) Nooit meer bellen met je energieleverancier. Retrieved from https://www.oxxio.nl/over/nieuws/nooit-meer-bellen-met-je-energieleverancier/

Polzin, T. S., & Waibel, A. (2000). Emotion-sensitive human-computer interfaces. In *ISCA tutorial and research workshop (ITRW) on speech and emotion*. (pp. 1-6) Newcastle, Ireland: International Speech Communication Association. Retrieved from https://pdfs.semanticscholar.org/bb71/68e7b140a80f7df9dfff188dceac094869f9.pdf

Prendinger, H., & Ishizuka, M. (2005). The empathic companion: A character-based interface that addresses users' affective states. *Applied Artificial Intelligence 19*(3-4), 267–285. doi:10.1080/08839510590910174

Price, R. (2016) Microsoft deletes racist genocidal tweets from AI chatbot Tay. Retrieved from http://uk.businessinsider.com/microsoft-deletes-racist-genocidal-tweets-from-ai-chatbot-tay-2016-3?international=true&r=UK&IR=T

Radziwill, N. M, & Benton, M. C. (2017). Evaluating quality of chatbots and intelligent conversational agents. *ArXiv Software Quality Professional 19*(3), 25-36. Retrieved from https://arxiv.org/abs/1704.04579

Reeves, B., & Nass, C. (1996). *The media equation : How people treat computers, television, and new media like real people and places*. Stanford, California: CSLI Publications. Retrieved from https://www.researchgate.net/publication/37705092_The_Media_Equation_How_People_Treat_Computers_Television_and_New_Media_Like_Real_People_and_Pla

Riek, L. D., Rabinowitch, T. C., Chakrabarti, B., & Robinson, P. (2009). How anthropomorphism affects empathy toward robots. *International Conference on Human Robot Interaction, 4*(1), 245-246. doi:10.1145/1514095.1514158

Rogers, E. M. (2002). Diffusion of preventive innovations. *Addictive Behaviors, 27*(6), 989-993. doi: https://doi.org/10.1016/S0306-4603(02)00300-3

Spector, P. E. (1992). *Summated rating scale construction: an introduction*. Thousand Oaks, CA: SAGE.

Sung, J. Y., Guo, L., Grinter R.E. & Christensen, H.I. (2007). "My Roomba Is Rambo": Intimate Home Appliances. In J. Krumm J., Abowd G.D., Seneviratne A., Strang T. (eds). *Lecture Notes in Computer*

Science: Vol 4717. Ubiquitous Computing (pp. 145-162). Berlin, Germany: Springer. doi: 10.1007/978-3-540-74853-3_9

Terzis, V., Moridis, C., & Economides, A. (2012). The effect of emotional feedback on behavioral intention to use computer based assessment. *Computers & Education, 59*(2), 710-721. doi:10.1016/j.compedu.2012.03.003

Tinwell A. (2009) Uncanny as Usability Obstacle. In: Ozok A.A., Zaphiris P. (eds), *Lecture Notes in Computer Science: Vol 562. Online Communities and Social Computing* (pp. 622-631). Berlin, Germany: Springer. doi: https://doi.org/10.1007/978-3-642-02774-1_67

Tseng, H. & Fogg, B. J. (1999). The elements of computer credibility. *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, 1, 80-87. doi: 10.1145/302979.303001

Tsvetkova, M., García-Gavilanes, R., Floridi, L. & Yasseri, T. (2017). Even good bots fight: The case of Wikipedia. *PLOS ONE, 12*(2). doi: https://doi.org/10.1371/journal.pone.0171774.

Turing, A. M., (1950). Computing Machinery and Intelligence. *Mind, 49*(236), 433-460. Retrieved from http://phil415.pbworks.com/f/TuringComputing.pdf

Wang, H., Wang, F. J & Liu, N. Y. (1997). A conversational agent for food-ordering dialog based on VenusDictate. *Proceedings of the 10th Research on Computational Linguistics International Conference*, *10*, 325-334. doi:https://doi.org/10.1371/journal.pone.0171774

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. doi:10.1145/365153.365168

Wilson, G. S. (1999). Personality variables in levels of predicted and actual test anxiety among college students. *Educational Research Quarterly, 23*(3), 3–13. Retrieved from https://search.proquest.com/openview/694e37ad23b946168176b10965463622/1?pq-origsite=gscholar&cbl=48020

Wixom, B. & Todd, P. (2005). A theoretical integration of user satisfaction and technology acceptance. *Information Systems Research, 16*(1), 85-102. doi:10.1287/isre.1050.0042

Appendices

Appendix A: Survey questions

Anthropomorphism (5-point semantic differential scale)

- 1. Fake natural
- 2. Machinelike humanlike
- 3. Unconscious conscious
- 4. Artificial lifelike

Likeability (5-point semantic differential scale)

- 1. Dislike like
- 2. Unfriendly friendly
- 3. Unkind kind
- 4. Unpleasant pleasant
- 5. Awful nice

Perceived intelligence (5-point semantic differential scale)

- 1. Incompetent competent
- 2. Ignorant knowledgeable
- 3. Irresponsible responsible
- 4. Unintelligent intelligent
- 5. Foolish sensible

User satisfaction (5-point semantic differential scale)

- 1. Terrible wonderful
- 2. Difficult easy
- 3. Frustrating satisfying
- 4. Dull stimulating
- 5. Not useful useful
- 6. Rigid flexible

Intention to use (5-point Likert scale)

- 1. I think that I would like to use a chatbot frequently
- 2. I don't think I would use a chatbot*
- 3. I would like to use the chatbot myself

(*reversed items)

Perceived usefulness (5-point Likert scale)

1. I found the chatbot unnecessary complex*

- 2. I think the chatbot is easy to use
- 3. I think that I would need the learn a lot before I can use a chatbot*
- 4. I would imagine that most people would learn to use a chatbot very quickly
- 5. I get why using a chatbot is useful
- 6. I think the chatbot is very troublesome to use*
- 7. I think a chatbot is useful

(*reversed items)

Trust (5-point Likert scale)

- 1. I think the chatbot system is reliable
- 2. I think the chatbot is not likely to fail
- 3. I have trust in the chatbot technology
- 4. I think the chatbot is capable in meeting my needs
- 5. I would never trust a chatbot completely*
- 6. I think the chatbot delivers good work
- 7. I think the chatbot is credible
- 8. I am convinced that a chatbot can work alone, without assistance of people
- 9. I have faith in the skills of the chatbot
- 10. I trust the chatbot completely

(*reversed items)

Negative attitude towards robots (5-point Likert scale)

- 1. I would feel uneasy if robots really had emotions
- 2. Something bad might happen if robots developed into living beings
- 3. I would feel relaxed talking with robots*
- 4. I would feel uneasy if I was given a job where I had to use robots
- 5. If robots had emotions, I would be able to make friends with them*
- 6. I feel comfortable being with robots that have emotions*
- 7. The word 'robot' means nothing to me
- 8. I would feel nervous operating a robot in front of people
- 9. I would hate the idea that robots or artificial intelligence were making judgements about things
- 10. I would feel very nervous just standing in front of a robot
- 11. I feel that if I depend on robots too much, something bad might happen
- 12. I would feel paranoid talking with a robot
- 13. I am concerned that robots would be a bad influence on children
- 14. I feel that the future society will be dominated by robots

(*reversed items)

Empathy (5-point Likert scale)

- 1. I daydream and fantasize, with some regularity about things that might happen to me
- 2. I often have tender, concerned feelings for people less fortunate than me
- 3. I sometimes find it difficult to see things from the 'other guys' point of view*
- 4. Sometimes I don't feel very sorry for other people when they are having problems*

- 5. I really get involved with the feelings of the characters in a novel
- 6. In emergency situations, I feel apprehensive and ill-at-ease
- I am usually objective when I watch a movie or play, and I don't often get completely caught up in it*
- 8. I try to look at everybody's side of a disagreement before I make a decision
- 9. When I see someone being taken advantage of, I feel kind of protective towards them
- 10. I sometimes feel helpless when I am in the middle of a very emotional situation
- 11. I sometimes try to understand my friends better by imagining how things look from their perspective
- 12. Becoming extremely involved in a good book or movie is somewhat rare for me*
- 13. When I see someone get hurt, I tend to remain calm*
- 14. Other people's misfortunes do nu usually disturb me a great deal*
- 15. If I'm sure I'm right about something, I don't waste much time listening to other people's arguments*
- 16. After seeing a play or movie, I have felt as though I were one of the characters
- 17. Being in a tense emotional situation scares me
- 18. When is see someone being treated unfairly, I sometimes don't feel very much pity for them*
- 19. I am usually pretty effective in dealing with emergencies*
- 20. I am often quite touched by things I see happen
- 21. I believe that there are two sides to every question and try to look at them both
- 22. I would describe myself as a pretty soft-hearted person
- 23. When I watch a good movie, I can very easily put myself in the place of a leading character
- 24. I tend to lose control during emergencies
- 25. When I'm upset at someone, I usually try to 'put myself in his shoes' for a while
- 26. When I am reading an interesting story or novel, I imagine how I would feel if the events in the story were happening to me
- 27. When I see someone who badly needs help in an emergency, I go to pieces
- 28. Before criticizing somebody, I try to imagine how I would feel if I were in their place

(*reversed items)

Demographic information

- 1. What is your age?
- 2. What is your gender?
- 3. What is the highest educational level you have finished?

Appendix B: Chatbot conversation – Dutch script

Anthropomorphic characteristics:

Greeting Personality **Likes-dislikes** Conversational cues Emotional state Chatbot conversation - anthropomorphic (Dutch script) FADE IN: Int. Gesprek met een chatbot weergeven op mobiele telefoon CHATBOT: Hallo! Welkom bij pizza paleis, mijn naam is <mark>Jess </mark> wat kan ik voor je betekenen? USER: Ik heb honger CHATBOT: Gelukkig kan ik daarbij helpen. Wil je misschien een pizza bestellen? USER: Ja, graag 😊 CHATBOT: In wat voor pizza heb je trek? USER: Dat weet ik nog niet CHATBOT: Mijn favoriete pizza is de tonijn pizza, <mark>zeker een</mark> <mark>aanrader</mark>. En deze week is de hawaii pizza in de aanbieding. Als je het nog niet goed weet kun je de pizza's bekijken op de website: pizzapaleis.nl/pizzakeuzes. USER: Ik wil de peperoni pizza! CHATBOT: Welke peperoni pizza had je graag gewild? We hebben de keuze uit de volgende peperoni pizza's: - pizza peperoni en salami - pizza peperoni pikant

- pizza dubbel peperoni USER: Pizza dubbel peperoni CHATBOT: Hoe groot mag jouw pizza zijn: - Small (25 cm) - Medium (30 com) - Large (35 cm) USER: Medium CHATBOT: Wil je misschien nog wat anders bestellen? Of wil je de bestelling aanpassen? Bestelling wijzigen _ Product toevoegen _ _ Bestelling bevestigen USER: Ik wil graag een flesje coca cola toevoegen aan mijn bestelling 😳 CHATBOT: Oke! Geen probleem <mark>;)</mark> Ik heb het flesje coca cola toegevoegd aan je bestelling. Als je klaar bent met bestellen mag je op bestelling bevestigen drukken. Zo niet, kies dan uit de twee andere opties: Bestelling wijzigen _ _ Product toevoegen Bestelling bevestigen _ USER: Bestelling bevestigen CHATBOT: Wil je de bestelling laten bezorgen of hier afhalen? USER: Bezorgen, alsjeblieft

CHATBOT:

Oke, <mark>ik</mark> ga het voor je regelen. Wat is het adres van de bezorging?

USER: Geiserstraat 28a, Enschede

CHATBOT:

Ik heb de bestelling doorgegeven aan de bezorger! De kosten voor de bestelling zijn: 14,90. Hoe wil je betalen?

- Ideal
- Paypal
- Contant

```
USER:
```

Contant

CHATBOT: Oke, geregeld! De pizza wordt op dit moment gemaakt. De bezorger zal over ongeveer 27 minuten bij je zijn. Eet smakelijk!

Chatbot conversation - non-anthropomorphic (Dutch script)

FADE IN:

Int. Gesprek met een chatbot weergeven op mobiele telefoon

CHATBOT: Welkom bij pizza paleis, waarmee kan ik helpen?

USER: Ik heb honger

> CHATBOT: Oke, wil je een pizza bestellen?

USER: Ja, graag 😊

CHATBOT:

Welke pizza?

USER: Dat weet ik nog niet

```
CHATBOT:
bekijk de website voor de menukaart:
pizzapaleis.nl/pizzakeuzes.
     USER:
Ik wil de peperoni pizza!
     CHATBOT:
Welke peperoni pizza?
Selecteer keuze:
     - pizza peperoni en salami
     - pizza peperoni pikant
     - pizza dubbel peperoni
     USER:
Pizza dubbel peperoni
     CHATBOT:
Selecteer de grootte van de pizza:
     - Small (25 cm)
     - Medium (30 com)
     - Large (35 cm)
     USER:
Medium
     CHATBOT:
Wil je nog iets anders bestellen of wil je de
bestelling wijzigen?
          Bestelling wijzigen
     _
     _
         Product toevoegen
          Bestelling bevestigen
     -
     USER:
Ik wil graag een flesje coca cola toevoegen aan mijn
bestelling 😊
     CHATBOT:
Flesje coca cola toegevoegd aan bestelling
Wil je nog iets anders bestellen of wil je de
bestelling wijzigen?
     _
          Bestelling wijzigen
          Product toevoegen
     _
```

```
- Bestelling bevestigen
     USER:
Bestelling bevestigen
     CHATBOT:
Bestelling bezorgen of afhalen?
     USER:
Bezorgen, alsjeblieft
     CHATBOT:
Wat is het adres van bezorging?
     USER:
Geiserstraat 28a, Enschede
     CHATBOT:
Bestelling doorgegeven aan bezorger. De kosten zijn:
14,90. Hoe wil je betalen?
     _
         Ideal
     - Paypal
     - Contant
     USER:
Contant
     CHATBOT:
De bezorger zal arriveren in ongeveer 27 minuten.
```

```
Eet smakelijk!
```

Appendix C: Results pre-test 4 (Dutch)

Item	Opmerking
Introductie	
Introductietekst onderzoek	-
Introductietekst video/chatbot	Gebruik een ander lettertype en andere kleur bij de chatbot uitleg.
Video	
Video	Er zit geen geluid bij veranderen in: Het fragment heeft geen geluid. Geef een toelichting over de video, wat gebeurt er in de video?
	Video's in betere kwaliteit uploaden.
Vragen	
Anthropomorhism in conversational tone (Menselijkheid in gesprekstoon)	-
User satisfaction (Gebruikstevredenheid)	-
Anthropomorphism algemeen (Menselijkheid algemeen)	-
Likeability (Leuk, waard om van te houden)	-
Perceived intelligence (waargenomen intelligentie)	Item 2: Onwetend en geïnformeerd zijn vage begrippen, veranderen in: Heeft geen kennis van zaken – heeft kennis van zaken. Item 4: Dom veranderen in onverstandig. Item 5: Onintelligent veranderen in: Niet intelligent.

Intention to use & perceived usefulness (Intentie om te gebruiken & waargenomen nut)	-
Trust (vertrouwen)	Item 5: 'een chatbot' veranderen in 'de chatbot' Item 7: Typefout, 'je' veranderen in 'de'. Item 8: Onduidelijk. Veranderen in: Ik ben ervan overtuigd dat de chatbot zelfstandig zou kunnen functioneren (zonder hulp van mensen).
Negative attitude towards robots	Introductietekst: Deze vragenlijst gaat niet meer over de video. Begin daarom met de tekst: Hoe denk jij over robots? Item 8: Inkorten. Andere mensen veranderen in anderen. Item 9: Andere technologieën is een vaag begrip, dit begrip kan beter weggelaten worden.
Empathy	Introductietekst: Deze vragenlijst is heel lang. Laat aan de respondenten weten dat de enquête bijna voorbij is. Voeg toe: Bijna klaar! Je helpt mij ontzettend door deze laatste vragen nog even in te vullen. Item 15: Inkorten. Andere mensen veranderen in anderen. Item 16: Zinsformulering. Veranderen in: Tijdens het kijken van een film voelt het alsof ik een van de personages ben.
Demografische gegevens	-
Overige opmerkingen	Afsluitende tekst: typefout, na aanleiding van veranderen in: Naar aanleiding van.