A conversational agent for vulnerable care recipients: what role does embodiment have?

Research into the role of embodiment in conversational agents' evaluations by vulnerable care recipients in terms of user expectations, overall user experience and support in achieving hedonic and pragmatic goals

L. J. A. J. (Lauren) Verhoeven
Snr 2024280

Master's Thesis
Communication and Information Sciences
Specialization Communication & Cognition

Department Communication and Cognition
School of Humanities and Digital Sciences
Tilburg University, Tilburg

Supervisors: Dr. H. L. van Heijster & Dr. C. C. Liebrecht
Second Reader: Prof. Dr. M. L. Antheunis

July 26, 2023

Abstract

For vulnerable care recipients, receiving the right health and social care information and support has become increasingly difficult. For this problem, the use of virtual conversational agents (CAs) may offer a solution. A CA is a computer-system entity designed to interact with its users using natural language. Within the field of CAs, the use of embodiment is emerging. Embodiment is the implementation of virtual, animated, human-like characters in CAs, enabling CAs to not only engage through language, but also through humanly, non-verbal communication (e.g., body movements). The effects of CAs on and the possible influence of embodiment on CAs for vulnerable care recipients, are not explored yet. Therefore, this study focusses on embodiment's role in the design and application of CAs for vulnerable care recipients. The influence of both an ECA (Embodied Conversational Agent) and DCA (Disembodied Conversational Agent) on this target group's user experiences, are evaluated through the CAs' support in goal achievement of pragmatic goals (goals that include "doing", e.g., finding information) and hedonic goals (goals that include "being", e.g., feeling supported). The role of user's expectations regarding CAs on experience is taken into account as well. Qualitative semi-structured interviews were performed within a 2x2 (ECA/DCA vs. hedonic/pragmatic) within-subjects study design, meaning each participant interacted with both the ECA and DCA and on both scenarios (hedonic/pragmatic). Results showed that the role of embodiment on user experience was, unexpectedly, rather small and instead, the presented scenario (hedonic/pragmatic) was most impactful on the user's evaluations. Participants experienced a CA to be more suitable for the pragmatic scenario than for the hedonic scenario. Overall, there was no strong preference for the ECA or DCA. Future research suggestions and practical implications are made and limitations are discussed.

Keywords: vulnerable care recipients, embodiment, conversational agents, hedonic and pragmatic goals

Contents

Introduction

In recent times, the health care sector in the Netherlands has come under increasing pressure (WRR, 2021). Health care has become increasingly more expensive and the tightness of the labour market in this particular field is becoming a problem (Woittiez et al., 2018; WRR, 2021). Especially for vulnerable care recipients, such as people with intellectual disabilities, autism spectrum disorder (ASD) or mental health problems, it has become increasingly difficult to be supported by and meditated to the right health and social care (NZa, 2022). Besides, information regarding this needed health and social care is complex and hard to find, which makes it even harder for people in vulnerable positions to retrieve the right care and support in time (Nationale Ombudsman, 2018). In these difficulties, a virtual conversational assistant could offer a possible solution.

Virtual conversational assistance for health purposes is emerging and has been prominently studied the last years (Ter Stal et al., 2020). Virtual assistants (VAs) or conversational agents (CAs), which are different definitions for a similar concept, are often on artificial intelligence based-systems designed to interact with its users using natural language (textual and/or spoken) (Feine et al., 2019; Van Bussel et al., 2022). As CAs are able to operate at a large scale, at any time and at low cost (Curtis et al., 2021; Van Bussel et al., 2022) and as CA studies regarding health care purposes show promising results (e.g., Baptista et al., 2020; Fitzpatrick et al., 2017; Owens et al., 2019; Rabinowitz et al., 2022; Rhee et al., 2014), they offer an optimistic solution to people in vulnerable positions and solving the abovementioned problems in the health and social care sector. An example of a CA study that showed promising results for health is research by Owens et al. (2019), that showed that a CA called “iDecide”, delivering information about prostate cancer, increased knowledge about prostate cancer. Likewise, research on a diabetes information providing CA called “Laura” showed that this CA was perceived as, among other things, helpful and trustworthy (Baptista

et al., 2020). These studies show that CAs are capable of health related information provision and of supporting patients or clients.

However, studies that show the possibility of CAs to provide health related information and support, do not focus on vulnerable care recipients. Research gaps remain within the area of personalizing CAs' interactions with its users (Provoost et al., 2017). Adapting CAs to make them most suitable for this target group is not yet investigated properly, while it could be of key value in responding the target group's needs. Without investigation on this target group's needs, it would remain unclear how to best design and implement CAs for vulnerable care recipients. One of the few examples of CAs for more vulnerable user populations, is a CA for (mental) health purposes in research by Ahmad et al. (2022), which focused on the design of a personality-adaptive CA and pointed out that adapting CAs for people in vulnerable positions is challenging. Therefore, adapting CAs' design and functionalities to the needs of vulnerable users should be further investigated.

Another aspect on which CAs can differ, is level of embodiment. Within virtual conversational assistance, a distinction can be made between Embodied Conversational Agents (ECAs) and Disembodied Conversational Agents (DCAs). Both classify as computer systems using natural language processing to communicate to its target group (Araujo, 2018), but ECAs use embodiment (Kramer et al., 2020) and simulate properties of human conversations, such as (non)verbal behaviour (e.g., facial expressions) (Provoost et al., 2017). DCAs, on the other hand, only engage with the user through natural language (textual and/or spoken) (Araujo, 2018).

Research on the use of embodiment in CAs shows that there are potential benefits to it (Laranjo et al., 2018). Among the general population, embodiment has a positive effect on the user's CA experience (Song & Shin, 2022) and ECA's appearances can be evaluated by users in a way that is similar to evaluations of real humans' appearances (Schouten et al., 2022).

However, whether the use of embodiment also has potential benefits for vulnerable care recipients, has not been investigated yet. It is not clear yet whether using embodiment has a positive influence on vulnerable care recipients' assessments of CAs too. Besides, even within this target group, embodiment could be more beneficial to certain vulnerable care recipients than to others. Consequently, the comparison of using embodiment versus not using embodiment for vulnerable care recipients, needs to be further investigated.

In this study, the influence of embodiment will be evaluated along two dimensions, namely pragmatics and hedonics (Hassenzahl, 2007, 2018). *Pragmatics* refer to the perceived extent to which an interactive product can help achieve "do"-goals, e.g., finding information about a disease or condition. On the other hand, *hedonics* refer to the perceived extent to which an interactive product can help achieve "be"-goals, e.g., feeling supported by the product (Hassenzahl, 2007). These dimensions can also be used when looking at the possible influence of ECAs and DCAs. However, it is not explicitly investigated yet whether or how different CAs contribute to the achievement of these goals. Besides, the role of using embodiment in achieving these goals, or deliberately not, has not yet been researched either.

In addition, the outcome measures used in research on ECAs or DCAs for health purposes rarely focus on the user's expectations (Shechtman & Horowitz, 2003) or the influence these initial expectations have on the user's eventual experience. Although, user's experiences are largely influenced by the expectations the user has (Lohse, 2011; Luger & Sellen, 2016; Zamora, 2017). Therefore, if ECAs or DCAs could eventually partly fulfil the role of physical care givers through providing information and support, it is important to map what users' expectations are and how these expectations cohere with their actual experiences.

In conclusion, this thesis aims to contribute to understanding the role of embodiment and user expectations in CAs for vulnerable target groups through a qualitative user experience study, using semi-structured interviews with vulnerable care recipients. Through

expanding inclusive research on CAs and this specific target group, certain research gaps may be filled (Følstad et al., 2021). This thesis aims to answer the following research question and research sub question:

RQ: How does (dis)embodiment in a conversational agent (CA) for vulnerable care recipients contribute to user experience, in terms of the achievement of their pragmatic and hedonic goals?

Sub-RQ: How does (dis)embodiment contribute to the evaluation of a conversational agent (CA) by vulnerable care recipients, compared to the user expectations?

Theoretical framework

Conversational agents' definition, origin and impact on health-related purposes

A conversational agent (CA) is a software-based system designed to interact with its users using natural language (Feine et al., 2019), often simulating human conversation (Laranjo et al., 2018). CA is an encompassing term for software that interacts with its users through written or spoken natural language and includes systems like virtual assistants, which are based on vocal interaction, and chatbots, which are based on textual communication (Diederich et al., 2022). Despite the fact that chatbots in its most early form were introduced in research in 1966 (Weizenbaum, 1966), recent-years' machine learning and artificial intelligence (AI) resulted in the conversational agent as it is implemented into daily lives now (Song & Shin, 2022). CAs are modernly used in many different contexts, such as private life, education and health care (Diederich et al., 2022).

The latter, the use of CAs in health care, is a phenomenon that has been further developing in recent times and various studies have shown promising results. An example of a study that explored a CA in health care is research by Fitzpatrick et al. (2017), in which the effect of CAs on depression and anxiety were measured through the application of a cognitive-behavioral therapy-based CA called "Woebot". Participants that interacted with "Woebot" reduced their symptoms of depression and anxiety over time, whereas only symptoms of anxiety reduced for the information-only control group. Further, another example of successful application of CAs was found in a study by Rhee et al. (2014) about asthma self-management for adolescents. Participants indicated that the use of the CA, among other things, resulted in more awareness of asthma symptoms and triggers and a higher sense of control. Lastly, a third example of improved health through the use of CAs was shown in a study by Rabinowitz et al. (2022). This study focused on the assessment of a CA called "RehaBot" for people with severe traumatic brain injury. Focus groups in this study showed

that participants with this type of brain injury reported the CA as helpful, for example in serving as a memory aid.

Embodiment versus disembodiment

In the field of CAs, a distinction can be made between two subclasses of CAs; Embodied Conversational Agents (ECAs) and Disembodied Conversational Agents (DCAs). Both subclasses are software-based systems used to communicate with their users (Feine et al., 2019), but ECAs are extended by using embodiment in the form of virtual, animated, human-like characters (Araujo, 2018; Schouten et al., 2022). DCAs do not involve use of such character. Embodiment enables ECAs to not only engage through language, but also through more humanly, non-verbal communication, e.g., facial expressions and/or body movements (Araujo, 2018). Therefore, in short, embodiment can be seen as humanization of CAs.

Embodiment is a concept with many representations. The manifestation of embodiment can range from virtual, human-like characters to robots, and ECAs' communication ranges from text to other human modalities such as speech, gestures or facial expressions (Provoost et al., 2017). Additionally, embodiment can be expressed through many different ways. It includes the CA having a certain appearance, in the form of e.g., gender, age and/or ethnicity (Schouten et al., 2022), but also in the form of e.g., a two-dimensional display of the agent vs. a three-dimensional display of the agent, feature specification (socio-demographics and style), and/or the amount of realism that is used within the embodiment (Straßmann & Krämer, 2017).

Although embodiment includes many different ways and forms, the concept's expressions are unified by the idea that they have a positive effect on the user's CA experience (Song & Shin, 2022). People evaluate ECAs' appearances the same way as they evaluate human appearances (Schouten et al., 2022), which in turn provokes more sympathetic social behaviour in interaction with ECAs than during interaction with DCAs

(Luger & Sellen, 2016). In addition, embodiment can buffer for potential shortcomings in CAs (Go & Sundar, 2019). For instance, if CAs lack message interactivity and therefore evoke negative consequences for the user's experience (such as a negative attitude towards the CA's interface), the use of embodiment can compensate for that and counter these negative consequences (Go & Sundar, 2019). A literature study by Kramer et al. (2020) showed that while ECAs did not yet increase the user's health literacy, the use of ECAs did increase the user's motivation to apply measures that contribute to their health.

Additionally, for vulnerable care recipients, there is some evidence for a positive effect of ECAs in health care as well, even though the differences in effect between ECAs and DCAs are not yet studied extensively for this target group. A study by Bickmore et al. (2010) showed that interacting with ECAs for health purposes has positive effects on behaviour change and health education and also motivated why ECAs are especially convenient for vulnerable care recipients. Given the ability to simulate human verbal and non-verbal communication through including e.g., facial display or hand gestures, ECAs are suitable in conveying health information to vulnerable care recipients. Participants with inadequate health literacy, which is a characteristic that is closely related to vulnerable care recipients, had no problems in accepting and using the ECA from a study by Bickmore et al. (2010). This study even indicated that ECAs may be more admissible for inadequate health literates than for adequate health literates.

A possible explanation for the seemingly positive effect of ECAs over DCAs is the "Computers Are Social Actors" (CASA) paradigm, that is introduced by Nass et al. (1994) and supported by various other studies (Ahmad et al., 2022; Araujo, 2018; Curtis et al., 2021; Lee & Nass, 2003; Schouten et al., 2020; Song & Shin, 2022). The fundament of this paradigm is that humans respond socially to computers and that these social responses are produced naturally. Despite the fact that humans consciously know they are not interacting

with a human being, they respond to a computer's social signals (e.g., human-like characters) in an equal manner to the way they respond to social signals from other humans (e.g., facial expressions, body movements) (Nass et al., 1994).

Moreover, to further construe and expand the CASA paradigm and its supporting research, the "Media are Social Actors" framework was introduced as well (Lombard & Xu, 2021), in which the effects of social cues (e.g., face shapes, gestures, human-sounding voice, language use) are stressed. The MASA paradigm explains that some social cues might provoke different effects on the user's social perception than other social cues. This means that in provoking a social interaction with CAs, not only the amount of social cues matters, but mostly the selection of social cues (Lombard & Xu, 2021). The MASA paradigm also states that whether social cues provoke a social response or not is dependent on individual differences and context, making not yet existing research on the effects of those social cues on this specific vulnerable target group more interesting. As the user's e.g., personality (Lee et al., 2006) and age (Epley et al., 2007) influence someone's likeliness to ascribe and pick up social cues to and from CAs, their vulnerable position might influence this likeliness as well.

Nevertheless, research indicates that implementing CAs with certain social cues, such as visual human-like representations (Moon & Nass, 1996) or language use (Lombard & Xu, 2021), make the CA feel more human-like and human-likeness in turn contributes to effective human-computer interaction (Sundar, 2008). This designates that a more embodied and humanized CA is more preferable than a non-embodied CA (Ahmad et al., 2022; Song & Shin, 2022). However, while the idea of ECAs provoking more positive effects than DCAs is much supported, there is a hypothesis suggesting that there is a limit to adding embodiment elements and human-likeness to CAs: the "uncanny valley hypothesis" (Mori, 1970/2012). This hypothesis entails the idea that too realistic and/or too lifelike embodied representations of CAs provoke negative affective states (Song & Chin, 2022). It suggests that users feel

uncomfortable or sense eeriness when they interact with a CA that appears and acts like a real human (Mori, 1970/2012). According to this hypothesis, adding human-like features to a certain extent increases affinity with interactive entities like CAs. However, this applies up until a certain maximum of human-likeness. Beyond that maximum, affinity decreases, until it increases again for entities that are similar to human beings in appearance (e.g., hand and eye movements), but are far from hyper realistic in e.g., absolute size (Mori, 1970/2012). In creating an efficacious ECA, the uncanny valley should be avoided.

Concludingly, this indicates that applying embodiment in CAs will have a positive outcome for user evaluation's by vulnerable care recipients, under the condition that the CA's character is not hyper realistic and does not appear too human-like. This results in the following expectation for this study: interacting with an ECA will provoke a more positive user experience for vulnerable care recipients than interacting with a DCA.

Hedonic versus pragmatic attributes

People's experiences interacting with CAs, and with any other (interactive) product, can be assessed at different levels and the use of (dis)embodiment in CAs might influence each of these levels differently. According to Hassenzahl (2007, 2018), two levels at which interactive products can be assessed are *pragmatic* attributes and *hedonic* attributes. *Pragmatic attributes* refer to features of an interactive product that are responsible for fulfilling an individual's behavioural goals (Hassenzahl, 2007). Pragmatic attributes also cover what Hassenzahl (2007) refers to as *manipulation.* Manipulation is that what an interactive product can provoke in the user's environment, as a consequence of being useable and utile (Hassenzahl, 2007). These attributes refer to the product's ability to help in achieving "do"-goals (Hassenzahl, 2018). For example, if people are asked to "do" something, they require an instrument or tool to do so, e.g., lighting a candle by using a lighter. From a pragmatic perspective, the only goal for this lighter is that it can be used to

light the candle and that the user knows how to do so. In CAs, an example of a pragmatic attribute could be the CA's ability to provide the right information about health.

In contrast, *hedonic attributes* refer to the features of an interactive product that are influencing an individual's psychological well-being (Hassenzahl, 2007). These attributes refer to the product's ability to help achieving "be"-goals, such as "being supported". Hassenzahl (2007) emphasizes that hedonic attributes are covered through three sub-dimensions: *stimulation*, *identification* and *evocation*. In order to fulfil hedonic needs, a product could provide new opportunities and perceptions (stimulation), be part of the expression of someone's identity (identification) and/or represent memories or important values (evocation) (Hassenzahl, 2007). For instance, when looking at the example of the lighter and candle from a hedonic perspective, this lighter could not only be pragmatically valuable due to its functionality (i.e., lighting the candle), but also have hedonic value as it is a gift or heirloom from a family member. In CAs, an example of a hedonic attribute could be the CA's ability to help feel the user being competent.

In the case of CAs for vulnerable people, the user's experience can also be measured along these two different levels. Their wants and needs, such as finding information and feeling supported (Van Heijster et al., in preparation), are divisible across these two levels. It is not clear yet whether or how embodiment plays a role in the achievement of such goals through CAs. However, based on the CASA paradigm, it is expected that ECAs provoke a more social reaction than DCAs. As this paradigm suggests that social cues in computers such as CAs are provoking a social reaction (Ahmad et al., 2022; Song & Shin, 2022), the perceived social nature of ECAs could positively influence the achievement of hedonic needs. The achievement of hedonic needs requires a social aspect, that a product like an ECA might provide. Therefore, for this study, it is expected that ECAs have a greater ability to fulfil hedonic needs than DCAs have for vulnerable care recipients.

In comparison, the role of embodiment on the fulfilment of pragmatic needs might not be as influential. When CAs are used in the fulfilment of pragmatic needs and goals, they require mostly utility and usability (Hassenzahl, 2007); the manner in which these needs and goals are met, might not be important. If an ECA and a DCA are compared and if their functionality and contents are matching, they might be equally competent for meeting pragmatic needs and goals. As a result, for this study, it is expected that ECAs and DCAs are equally able to fulfil pragmatic needs for vulnerable care recipients.

User expectations, user experience and the interaction between them

In CA studies, the effect of user's expectations on user's experiences are not prominently taken into account (Shechtman & Horowitz, 2003). Instead, many studies' centres of attention are the development and design processes of CAs (e.g., Nguyen et al., 2021 Straßmann & Krämer, 2017; Ter Stal et al., 2020), in which the user is normally not included (Kramer et al., 2020). Other studies mainly focus on the CAs' efficacy in terms of opted behaviour or use-related outcomes (e.g., studies on nutrition focused mainly on whether eating patterns changed) (Kramer et al., 2020).

However, the user's experience is crucially influenced by the user's expectations (Lohse, 2011; Luger & Sellen, 2016; Olsson, 2014; Zamora, 2017). Currently, there often is a mismatch between the user's expectations of a CA and the user's actual experience. Users often feel like their expectations of CAs are not lived up to, which results in a more negative assessment of the experience (Luger & Sellen, 2016; Zamora, 2017). Moreover, the fact that users often are not included in the design process of CAs (Kramer et al., 2020) might facilitate the fact that their initial expectations differ from the eventual experience they have. Although, if the user's expectations are met, this positively influences the user's experience (Lohse, 2011).

In order to meet expectations, the user's expectations should first be mapped.

Therefore, as research on specifically vulnerable care recipients' expectations was not done before, exploration of their expectations of CAs is needed. Concludingly, user expectations of vulnerable care recipients will be briefly assessed in this study, to understand the potential role of those expectations on the user's experience. For this study, it is expected that positive expectations or positive previous experiences provoke a positive experience (Olsson, 2014). For negative expectations or negative previous experiences, a more negative user experience is expected for this study (Olsson, 2014).

Method

Design

To investigate the possible impact that (dis)embodiment in CAs has on user expectations and user experience, a qualitative user evaluation study was performed. The used research method was a semi-structured interview, with questions before and after the interaction with the CAs. Participants were presented with CAs in which embodiment (embodiment vs. disembodiment) and the scenario they are presented with (hedonic vs. pragmatic) were manipulated. This resulted in a 2 (embodiment vs. disembodiment) x 2 (scenario reflecting pragmatic attributes vs. scenario reflecting hedonic attributes) within-subjects design, where each participant is presented with both CAs and both scenarios. Through this type of design, participants were able to reflect on both embodiment and on the presented scenarios in particular and could compare both experiences with each other. This all was done in line with the qualitative nature of the study and to create a complete image of participants' view on embodiment and the scenarios.

Materials

Conversational agents

In order to answer the research question, two different CAs were created; one in which embodiment is included (ECA) and one in which embodiment is not included (DCA)[1]. The CAs were created together with Mobile Water © (https://mobilewater.nl/), a conversational technology design specialist. Ideas for the agent's possible features and appearance, (Dutch) pre-made text output were provided to Mobile Water by the author of this thesis, as well as the two scenarios, which were inspired by a related qualitative study on needs and interests of

[1] The ECA is accessible through the following link: https://noah.mobilewater.dev/ux/booking/SVA_EMBODY?start=INTRODUCTION&locale=NL&EMBODIMENT=TRUE and the DCA is accessible through this link: https://noah.mobilewater.dev/ux/booking/SVA_EMBODY?start=INTRODUCTION&locale=NL&EMBODIMENT=FALSE

vulnerable care recipients (Van Heijster et al., in preparation). Mobile Water implemented those ideas into the development and design of a virtual environment with and without an embodied character.

Designing the CAs' appearances. For the design of the character of the ECA, there were concrete suggestions of what different embodied features could be implemented into the design and what the design could potentially look like. Feature suggestions are presented in Table 1, together with a motivation based on literature on why these design features were proposed. This resulted in general ideas for the looks and features of the virtual environments and the ECA's character, which were communicated to Mobile Water (see Appendix A). However, important in this was that there was not just one way to design a CA, as there is no consensus on what an ECA should look like per se (Ter Stal et al., 2020). Instead, the design should rely on the task in question and the target group (Ring et al., 2014). Therefore, among others, the cultural background of participants, health care context and used interface (i.e., using the CA on desktop) were taken into account (see Table 1).

Table 1

Proposed design features for the ECA with motivations based on literature

Embodied design feature	Suggestion	Literature-based motivation for suggestion
Gender	Female	There is no consensus for a gender preference in ECAs (Ter Stal et al., 2020).
Race/ethnicity/cultural group	White	Research by Alsharbi and Richards (2017) showed that people have a preference (e.g., general liking) for a CA from the same cultural group. As all participants have a Dutch cultural background, it was expected effective to depict a large part of that cultural background.
Dimension	3D agent in a 2D-simulation	There are options for depicting an ECA in a 3D-space, for example through VR (Straßmann & Krämer, 2017). However, participants were interacting with the ECA via a computer screen, leaving only two

		options: a 2D agent or a 3D agent in a 2D-simulation. The latter is used most (Straßmann & Krämer, 2017) and was therefore suggested.
Appearance: Visual realism	Balance between high realism and not being too humanly realistic	Higher realism in CAs provokes more positive social interaction than lower realism (Yee et al., 2007). However, a CA's character being too realistic may provoke a negative affective state (Song & Shin, 2022). Therefore, a balance between high realism and low realism is suggested.
Appearance: Shading style (highly realistic vs. toon shaded) (Example of this difference in Figure 1)	Toon shaded	The use of *toon shading* (i.e., adding cartoonish shading) in CAs makes them being perceived as more likeable and caring, compared to highly realistic CAs (Ring et al., 2014).
Appearance: Agent's proportions (cartoon proportioned vs. human proportioned) (Example of this difference in Figure 2)	Balance between highly realistic and cartoonish	Cartoon proportioned agents are perceived as less appropriate than human proportioned agents in medical contexts. However, for social contexts, the opposite is true (Ring et al., 2014). As the context of this study can be seen as a combination of these contexts, balance between these two types of proportions was preferred.

Figure 1

Stimuli used in research by Ring et al. (2014), an example of highly realistic shading (left) and toon shading (right) in conversational agents

Figure 2

Stimuli used in research by Ring et al. (2014), an example of a human proportioned agent (left) and a cartoon proportioned agent (right)

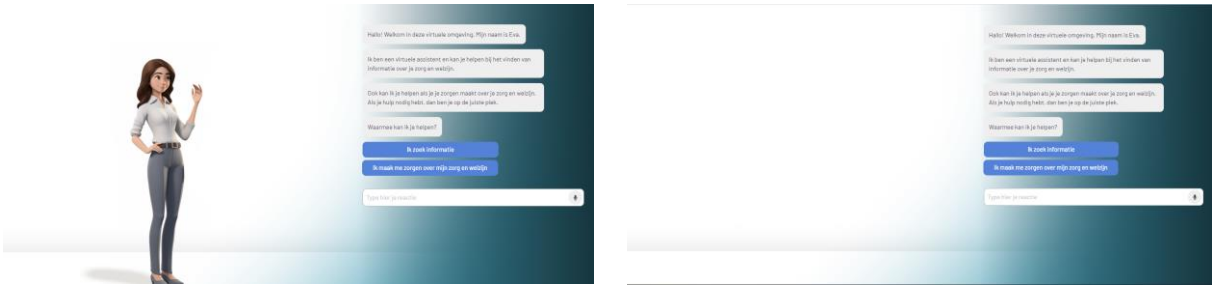
After proposing general ideas and features, the character that is depicted in Figure 3 was developed through AI prompting of the proposed character (see Appendix A). On the introduction screen of the ECA, the version of the character on the right (in Figure 3) was used, to imply the idea that the character was waving. After that, the version in Figure 3 which is the second from the left was used. The ECA's character was named "Eva", which is a Dutch name and also an abbreviation of embodied virtual assistant. The final virtual environments, both the embodied and disembodied version, are depicted in Figure 4. These virtual environments were based on an existing interface by Mobile Water, which was adapted to this particular study. To ensure that the two environments did not differ in any way except in absence or presence of embodiment, only that part is varying (see Figure 4).

Figure 3

Appearance of the ECA's virtual character named "Eva"

Figure 4

Screenshots of the virtual environments, embodied (left) and disembodied (right)



Designing the CAs' content. In terms of the CAs' dialogue and content, the CAs' pre-made input buttons, pre-made output and end-of-conversation infographics (for the infographics, see Appendix B) were created as well. Participants did not have the possibility to use natural language in the virtual environments. Instead, participants could click on pre-made input buttons (e.g., '*Ik wil informatie*') upon which the CA provided a pre-programmed output response. This was done to minimize the chances of different participants having completely different conversations and resulted in several paths that the participant could

walk through. Two of those paths were related to the given scenarios and the other paths were there to give a more complete impression of things that could be interacted upon with a CA. All fixed paths were pre-created and were elaborated in Microsoft Excel[2], after that, all paths were visualized on a Miro board.

At the end of each path, the participant had the opportunity to view and read an infographic about topics in the flow, such as e.g., how to apply for health care allowances. These infographics were used to provide the participant information and support. Besides, this method was also used to keep the interaction with the CA manageable. By using infographics, the conversation with the CA itself remained fairly short, but the opportunity to offer a large amount of information and support remained. The infographics are in Dutch and can be retrieved through Appendix B.

Scenarios reflecting hedonic and pragmatic goals

Besides the development of two different CAs, participants were also presented two different scenarios: one scenario reflecting a pragmatic goal and one scenario reflecting a hedonic goal (see Appendix C). As CAs can be assessed on the contribution to achieving both pragmatic as hedonic goals (Hassenzahl, 2007), it was important to evaluate how or to what extent CAs actually contribute to achieving these goals. By asking the participant to imagine a scenario where they are trying to achieve a pragmatic or hedonic goal (e.g., trying to get information), it could be demonstrated whether they feel like that goal is achieved through help from the CA.

The two scenarios that were used in this study are based on focus groups from prior research on what topics are concerning the target group and what problems they encounter (Van Heijster et al., in preparation). Therefore, the further exploring of these topics is

[2] The fixed chatbot flows in Excel can be retrieved via the following link: https://surfdrive.surf.nl/files/index.php/s/7rNBGlVcFX21XLN?path=%2F

relevant, as these are scenarios in which they benefit most from potential support. In case of the pragmatic scenario, participants were asked to imagine they are trying to seek information on their health care and had questions about their allowances (*uitkering* in Dutch), specifically about requesting allowances. As this is a task that entails *doing*, this is a scenario regarding pragmatic goals.

In case of the hedonic scenario, participants will be asked to imagine something has changed in allowances in general and they are not sure whether something is changing in their allowances. They were asked to imagine feeling a little worried about this and wanting to feel less worried through interacting with the CA. In this case, this task entails *feeling*, which makes it a scenario regarding hedonic goals. However, before reading the scenarios, prior explanation was given on the scenarios and it was emphasized that the scenarios were only hypothetical (see Appendix C). For each scenario, the participant exchanged a conversation in which obtaining the hedonic or pragmatic goal was the intended task.

Data collection

Data was collected through semi-structured interviews, with questions before and after the participants' interactions with the CAs. Important is that the language use was adjusted to vulnerable care recipients within the entire semi-structured interview in this study. For example, instead of CA, the term "chatbot(s)" was used. Different studies (Bartneck et al., 2009; Lohse, 2011; Zamora, 2017) that used questionnaires or semantic differentials as research methods, were used as inspiration for the questions in the semi-structured interview in this study. However, these research methods might be ambiguous or too abstract for vulnerable care recipients. This fact and the possible lower literacy of vulnerable care recipients were taken into account. This means that the language use and original formatting of questions (mostly semantic differential scales) in the studies used for inspiration (Bartneck et al., 2009; Lohse, 2011; Zamora, 2017) were adapted and changed into interview questions.

Additionally, as the interviews were semi-structured, the researcher asked additional questions where necessary. The reason for this was to further clarify the answers that were given and to get a complete sense of what the participants meant. The latter is facilitated by the flexible nature of semi-structured interviews (Kallio et al., 2016), which also enables space for participants to verbally express themselves in their own comfortable way. Therefore, semi-structured interviews as research method was most suitable for the vulnerable care recipients in this study.

User expectations and user experience

Prior to the participants interacting with the CAs, their expectations of such interaction were explored. The purpose of this was to later compare their initial expectations with their actual interaction. Participants' expectations were evaluated with three interview questions. The first question regarded their possible previous experience with CAs and the second question was about their view on what was good about CAs and less good about CAs. The third question regarded what they thought would make a CA better. These questions were based on research on user expectations by Zamora (2017).

After the interaction with the first CA, the participant was asked generally how they had experienced the interaction. After that, to further evaluate the user's experience, a document was used alongside interview questions to ask upon different CA aspects (see Appendix D). On this document, six different bars were depicted, where each bar represented semantic opposites that could describe their experience. Alongside this document, explanation was given about the bars and how participants could mark on the bar were they found the CA was fitting. Interview questions were then presented in the format of: 'Did you think the chatbot was more *x* or more *y*?', where *x* and *y* were the two opposites (e.g., *x* being "smart" and *y* being "dumb"). This question format was repeated 6 times, for each of the semantic opposites. All 6 questions can be found in Table 2.

The 6 questions used in this study are based on a semantic differential questionnaire by Lohse (2011), that includes 18 semantic differentials on user expectation and user evaluation. However, the formatting of semantic differentials has been changed to interview questions, visually supported by bars (see Appendix D). To not cognitively exhaust the vulnerable care recipients, only a selection of the 18 semantic differentials were converted to interview questions. Also, since Lohse (2011) states that: "it would probably be necessary to adapt the questionnaire for other scenarios and robots" (p. 486), it was figured that converting the questionnaire by Lohse (2011) into interview questions and removing differentials that are less relevant for this context would be a suitable choice. The original 18 semantic differentials by Lohse (2011) and the six extracted interview questions for this study can be found in Table 2.

After the interaction with the other CA, participants were presented with the same document with bars and were asked to mark where they found this CA to be fitting. However, the document was now paired with the question format: 'Did you find the chatbot more *x* or more *y* than the other one? And why is that?', where *x* and *y* again stand for the six pairs of semantic opposites (e.g., *x* being "smart" and *y* being "dumb"). Through this way, their two interactions could be compared with each other.

Table 2

Original semantic differentials by Lohse (2011), the English translation by Lohse (2011) and the six extracted interview questions.

Original items by Lohse (2011) in German	Translated items by Lohse (2011) in English	Interview questions in Dutch
lustig : ernst	funny : serious	*(removed)*
abwechslungsreich : langweilig	diversified : boring	*(removed)*
interessiert : desinteressiert	interested : indifferent	Vond je dat chatbot interesse toonde of geen interesse toonde?
gesprächig : ruhig	talkative : quiet	*(removed)*

freundlich : unfreundlich	friendly : unfriendly	*(removed)*
höflich : unhöflich	polite : impolite	Vond je de chatbot beleefd of onbeleefd is?
aktiv : passiv	active : passive	*(removed)*
kooperativ : unkooperativ	cooperative : uncooperative	Vond je de chatbot behulpzaam of onbehulpzaam?
aufmerksam : unaufmerksam	attentive : inattentive	Vond je dat de chatbot aandacht voor jou had of geen aandacht voor je had?
intelligent : dumm	intelligent : dumb	Vond je de chatbot slim of dom?
selbstständig : unselbstständig	autonomous : not autonomous	*(removed)*
unabhängig : abhängig	independent : dependent	*(removed)*
vorhersagbar : unvorhersagbar	predictable : unpredictable	*(removed)*
schnell : langsam	fast : slow	*(removed)*
lernfähig : lernunfähig	adaptive : not adaptive	*(removed)*
gehorsam : ungehorsam	obedient : disobedient	*(removed)*
praktisch : unpraktisch	practical : impractical	*(removed)*
nützlich : nutzlos	useful : useless	Vond je de chatbot nuttig of nutteloos?

Embodiment

Considering the application of (dis)embodiment was one of the manipulations in this study, it was important to explore whether an actual difference in the CAs was perceived. To evaluate whether the participants observed actual differences between the ECA and the DCA, each participant was asked four short questions that were indirectly regarded to embodiment (see Table 3). Two of those questions were more general questions, one was about the differences between the CAs and one was about general preference for either of the interactions. Two questions were based on semantic differentials in a questionnaire by Bartneck et al. (2009). Those were again converted to interview questions.

Table 3

Four questions regarding embodiment (partly based on Bartneck et al., 2009).

Interview questions for this study (translated to English)	Basis in Bartneck et al. (2009)

What differences did you see between the two chatbots?	No
Did you think one of the conversations was more like a real conversation than the other one? If so, why is that?	Yes, partly based on semantic differential (fake : natural)
Did you think one of the two was more like a conversation with a real human? Did you prefer that? Or not?	Yes, partly based on semantic differential (machine-like : human-like)
Did you prefer one of the two conversations in general? If so, what caused that?	No

Task success

In addition to the participants being presented with both the ECA or the DCA, each CA was also presented with either the scenario with a hedonic or pragmatic goal. To examine to what extent the CAs helped achieve the hedonic or pragmatic goals in the scenarios, task success was evaluated. Task success manifests itself in the CA contributing to more understanding in the pragmatic scenario and feeling better in the hedonic scenario. The questions regarding task success are included in Table 4.

Table 4

Questions regarding the effectiveness of the conversational agent contributing to the "do" goal in the pragmatic scenario and the "be" goal in the hedonic scenario

Scenario	Questions in Dutch, used in the interview	Questions translated to English
Pragmatic scenario	Weet je nog dat je je moest voorstellen dat je op zoek was naar informatie over je zorg en ondersteuning. Hoe vond je het om daarbij geholpen te worden door de chatbot? Vind je nu dat je goed hebt begrepen waar je naar toe moest met vragen over je uitkering en dat je meer informatie hebt ontvangen? Denk je dat de virtuele assistent daarbij heeft geholpen?	Do you remember that you had to imagine looking for information about your health care and support. How did you feel about being helped by the chatbot in this? Do you feel like you knew where to go to with questions regarding your allowances and like you received more information?

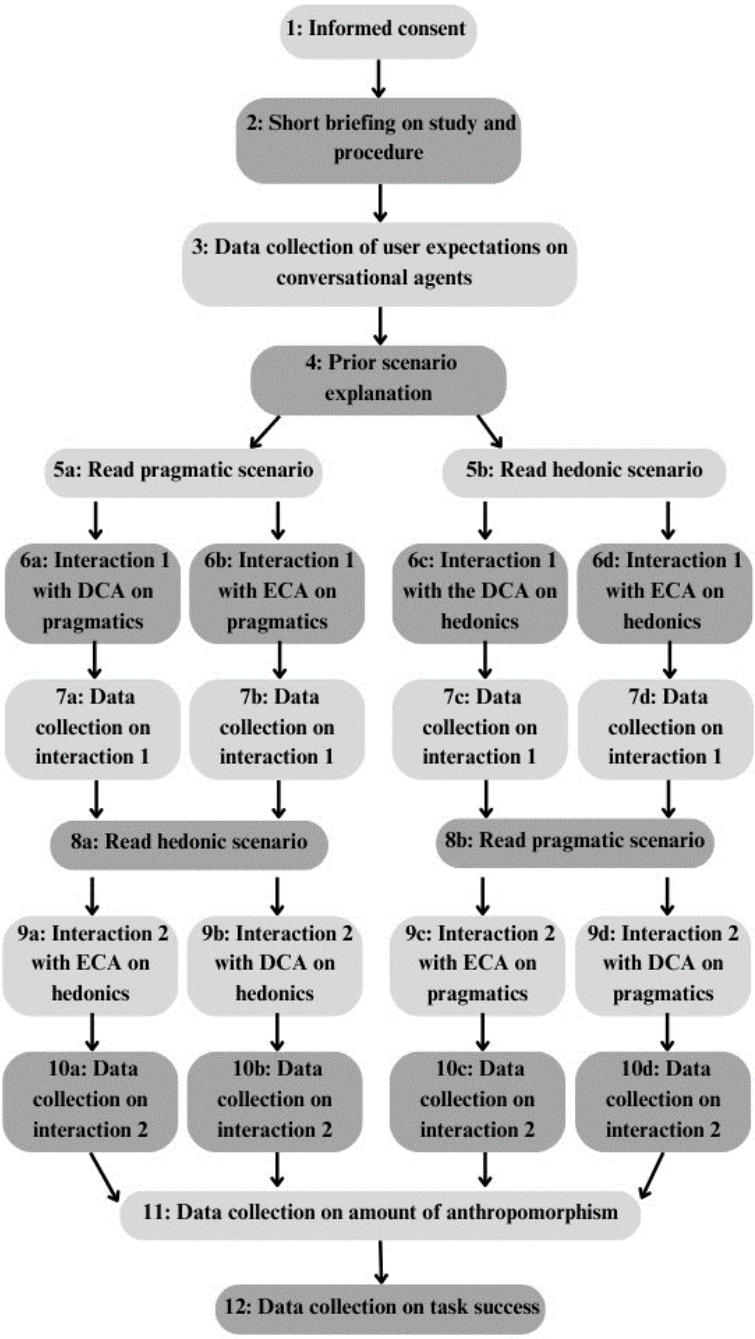
Hedonic scenario	Weet je nog dat je je moest voorstellen dat er iets was veranderd in je uitkeringen of zorg en ondersteuning en dat je je daar een beetje zorgen over maakte. Hoe vond je het om het daarover te hebben met de chatbot? Denk je dat je je na het gesprek met de chatbot minder zorgen zou maken?	Do you remember that you had to imagine something changed in your allowances or health care and support and that you felt a little worried about that. How did you feel about talking about this with the chatbot? Do you think you would be less worried after talking to a chatbot?

Procedure

The semi-structured interviews were all in Dutch and were partly carried out face-to-face and partly online via Google Meet. Before participants decided they wanted to participate, they all received an information bulletin via email with background information, to generate an idea of what was possibly expected from them and what their rights were (see Appendix E). Moreover, before the interviews were conducted, a pre-test took place with two representatives of vulnerable care recipients. One of these representatives walked through the full procedure and later tested both the CAs as well, the other representative walked through the full procedure only. The pre-test participants' opinions on the scenarios, the look and feel of the materials and the procedure were evaluated, but no adjustments needed to be made. The study procedure is schematically depicted in Figure 5.

Figure 5

Schematic overview of the study procedure



First, before participants started interacting with the created CAs, they were asked to read and sign the informed consent (see Appendix F). Since there were vulnerable people involved, the procedure and consent were also verbally explained by the researcher. Especially for this target group, it was important that participants were well aware of the

intention of the study and their rights. Therefore, participants were able to ask questions for possible clarification of any aspect of the study in case anything was unclear. In the second phase of this study, participants got a short briefing of what the study entailed and what the procedure would be.

In the third phase of this study, their expectations were evaluated with four interview questions. Then, in the fourth phase of this study, a scenario explanation was given to the participant (see Appendix C). As the scenarios are based on problems vulnerable care recipients encounter (Van Heijster et al., in preparation), it was likely that they were familiar with the scenarios. Through this prior scenario explanation, it was emphasized that their own situations did not have to be taken into account and that the scenarios could be seen as hypothetical. After that, the participant read either the pragmatic or the hedonic scenario, followed by an interaction on that scenario with either the ECA or DCA. This was then followed by the seventh phase of this study, were data collection on the first interaction was done through a general question and six questions regarding semantic opposites (together with the form with bars, see Appendix D).

Next, in phase eight, the participants read the other scenario, dependent on which one they read first. Then, they interacted with the other CA, again dependent on which on they had interacted with first. Then the interview continued, starting with a general question and the six questions regarding semantic opposites (through which the second interaction was compared to the first interaction), alongside the document with semantic opposites, presented through bars (see Appendix D). Both interactions with the CAs were variable to the extent that participants were able to choose from different buttons when answering the pre-made text output by the CA. After data collection on the second interaction, in the eleventh and twelfth phase of the study, data was collected regarding the amount of embodiment and regarding

task success (i.e., to what extent the CA has helped them achieve their hedonic or pragmatic goal).

Participants

For this study, a total of ten participants were interviewed. Six of these interviews were held online via Google Meet and four of the interviews were face-to-face. Before the interviews, the ten interviewees were randomly assigned to one of the four conditions of this study (first interaction: DCA on pragmatic, second interaction: ECA on hedonic/first interaction: ECA on pragmatic, second interaction: DCA on hedonic/first interaction: DCA on hedonic, second interaction: ECA on pragmatic/first interaction: ECA on hedonic, second interaction: DCA on pragmatic). Through this way, each condition was at least interacted with two times.

All ten interviewees were Dutch and two of them were female and eight of them were male. All participants were vulnerable care recipients. This target group included people who have a mental disorder, an intellectual disability and/or ASD. As this target group is fairly specific and it is important to approach people within this specific target group with care, it was quite challenging to recruit participants. Participants were acquired through purposive sampling and in collaboration with Stichting MEE Gelderse Poort (https://www.meegeldersepoort.nl/) and Radboud University Medical Centre (https://www.radboudumc.nl/en), which both have large access to and expertise regarding the target group. Given the fact that there were vulnerable people involved in this study, formal ethical approval was granted by the Research Ethics and Data Management Committee of Tilburg University.

Data analysis

To analyze the interviews, each interview was audio recorded and then transcribed with Microsoft's Office integrated transcribe function. Each interview transcript was then

checked and walked through by the writer of this thesis, to ensure that the produced transcript was complete and included all spoken utterances. After that, to analyze the transcripts, an Excel file was created to subdivide the interviewees' answers to the semi-structured interview questions. In this Excel file, eight different encodings were used to divide answers by interviewees. These seven encodings were: previous chatbot experiences or chatbot expectations (1), recognition and assessment of embodiment (2), task completion and chatbot understandability (3), virtual assistant's support in completion of hedonic or pragmatic goals (4), content suggestions (5), functional and technical suggestions (6), overall assessment (7), and lastly preference (8), for either of the interactions. These encodings were based on the interview questions and on any other emerging topics and therefore, each encoding represented a different topic. Each encoding got its own column in Excel, in which each row included one participant's view upon that particular topic and represented one participant. In these rows, their quotes about the different topics were included as well. Through this way, a complete overview of all utterances per topic, instead of per participant, was created.

Results

The results of these interviews are divided into subcategories, based on the interview questions and additional emerging categories.

Previous CA experiences and CA expectations

All participants had heard of the concept of chatbots. Only a small part of the participants had not interacted with CAs before. One of them expected CAs to be ‘robotic’. Many participants reported that they had previously interacted with or encountered CAs. In most cases, the CAs they used were produced for logistic or parcel companies or in the field of telecommunications. Besides, some participants mentioned they had previously participated in chatbot/CA research. However, each participant that had previously interacted with CAs also mentioned that their experiences were not very positive. Some of them described their experiences in terms as ‘deplorable’ or even ‘terrible’. Others were not as explicit in their judgements of previous experiences, but did mention that they preferred to ask their questions to a person or missed the opportunity to talk to a person. One participant said: *‘Often, people would like to talk to someone as well, for reassurance’*.

However, it did not seem like the user’s initial expectations or previous experiences later positively or negatively influenced the user’s experience per se. Some of the participants that had negative expectations, were not too positive about these interactions either. Nevertheless, there were also people with negative expectations, who did experience the interactions as positive. For example, one participant explained that he was previously annoyed at chatbots and said about his previous experiences: *‘I have experienced I prefer personal contact’*. Later, about one of his experiences in this study, he said: *‘This one* [CA] *was good’*.

Task completion and understandability

Participants experienced difficulties due to the fixed nature of the dialogues

Each participant interacted with both the ECA and DCA, where they interacted with one CA on one of the scenarios and with the other CA on the other scenario. For each scenario, there was an intended path within the interaction that would lead up to support in the scenario's goal and possibly one of the infographics. Within both the ECA and the DCA, it was possible to follow both paths that lead up to either support for the pragmatic scenario or the hedonic scenario. Some participants had some trouble making a distinction between those two paths and had some trouble figuring out which path was connected to either of the scenarios. For instance, when the scenario of seeking information was presented (pragmatic), some participants also, unpurposely, walked through the path of getting emotional support when feeling worried (hedonic) or the other way around.

Others had some trouble with completing the tasks as they could not decide for themselves what questions they wanted to ask. For some, this lead to confusion or frustration, and it seemed as if they felt like that option was missing and that the CA was not able to answer their personal questions. One participant stated: *'I think that I as a person should be able to determine the question'*. After attempting to explain that that was not possible yet in these CAs, it still was difficult for this person to separate the scenarios and personal life. Another participant said*: 'And if the question is not there, can I make it up then?'*, which reflected his confusion and comprehension problems.

Participants experienced difficulties due to the task goal

The cause of understandability problems did not seem to strongly relate to the presence or absence of embodiment. Instead, in terms of the intended goal of the task (hedonic vs. pragmatic), it did seem as if there was more general confusion in following the "right" intended path for the hedonic scenario in one specific area, since some participants

had difficulty with the three proposed emotions in this scenario. In the hedonic scenario, participants could interact on three possible emotions that were in line with the scenario of feeling a little worried (feeling worried and sad, feeling worried and angry or feeling worried and scared). They could receive support for these imagined worries and these imagined emotions, but some participants could not recognize themselves in any of those emotions or felt confused as neither of the emotions were included in the scenario description (see Appendix C).

User experiences and overall assessments

Overall user experiences, based on document with semantic opposites presented through bars

Overall, participants' user experiences with both CAs largely varied in terms of the six semantic opposites. Except for politeness, that was perceived high throughout each interaction by all participants, there were divergent evaluations of both CAs, regardless of embodiment. For perceived interest, for instance, some participants indicated that the ECA was more interested in them than the DCA. However, others did not report a difference between the ECA and DCA in terms of perceived interest. Instead of embodiment, perceived interest seemed to mainly rely on the nature of the task. Some participants indicated that the CA was perceived as more interested in the hedonic scenario. Moreover, many participants felt like perceived interest was not a characteristic that was attributable to a CA, but rather to humans only. A participant said: *'It is not about interest, he* [CA] *showed no interest in how I was doing'*. Another said: *'I cannot expect of a computer that it has personal interest in me'.* On the contrary, some did find in general that the CAs were interested in them, because it was asking them how they were feeling for example. Overall, neither the DCA or the ECA seemed to be strongly perceived as more interested.

For attentiveness, there were again many participants that felt like attentiveness was a

characteristic that could not be assigned to or was not fitting with CAs and throughout all interactions, regardless of embodiment, attentiveness was perceived as quite neutral. One participant stated: *'Attention to you, exactly in the middle* [of the bar]. *I do not think that* [being attentive] *is the point of the chatbot either'*. Another participant's view was in line with that: *'As it is not a person, it cannot really pay attention to me'*. However, another participant explained that the perceived lack of attention was not problematic: *'For me as a person, it did not have attention. But that is no problem. He did answer the question so that is fine. He is not distractable, so it pays full attention'*.

In terms of helpfulness and usefulness, participants were quite positive overall, regardless of embodiment, but there were some varying results too. It also seemed as if helpfulness and usefulness related to each other for many participants; if a CA was perceived as helpful, it was also perceived as useful and vice versa. One participant said: *'Completely helpful. … You just get an answer to your questions.'* However, one participant did see potential in the CAs usefulness, but was not very pleased yet: *'A chatbot could be useful if you would want to ask questions on strange times or if you want immediate answers, which is going on a lot currently. We don't want to wait for a live person. But reality is that you get very few good answers now* [from the CA]. *Or actually none yet.'* Many participants felt like the CAs did try to be helpful. One of them stated: *'He did do his best to help me seek* [for support]*, so it is programmed well. And certainly helpful in that sense'*. Another participant too stated: *'I assume that such a program, and I see that too, it does do its best'*. However, some participants did not find the CA helpful or useful at all. One participant expressed this through saying: *'I am not satisfied with this* [interaction]*'*.

Results regarding smartness were quite variating as well. Some participants mentioned a CA cannot be intelligent nor unintelligent, as it is a programmed computer. One participant expressed this by saying: *'I don't think it is smart or dumb. I think it is set up and*

programmed in such way to help you'. Nevertheless, others mentioned the CAs to be smart, for answering questions or showing empathy. Others did not assess the CAs as smart, as it could not answer all the possible health care questions there are or not really help them. About this, a participant said: *'It is kind of dumb when you make such a program* [the CA] *and that it does not contain all the … that not all answers are on there'*.

General evaluation of embodiment

Concerning embodiment, there was no strong preference for the embodied version or for the disembodied version. However, participants preferred the interaction with the ECA slightly more often, but only part of them specifically recalled this was due to the presence of the character. This preference was not any specific element of the character, mainly its presence in general. One participant said: *'I thought the figure provided a little more experience, let me put it that way. Without such a character, you know, I think it is such a standard piece of text'*. Later, this participant said: *'The figure from just now, it just gives the feeling that you are actually talking to someone'*. A few participants found all characteristics (perceived interest, politeness, helpfulness, attentiveness, smartness and usefulness) to be the same for both CAs and did not record major differences.

Role of scenarios in user experience

Some participants mentioned that they found differences in how the CAs were talking to them, but this mainly depended on the scenario that was covered by the CA and not on embodiment. Some participants found the CA to be brief, concise, clear and more extensive when they were trying to seek information about their health care in the pragmatic scenario, but felt like the CA was paying more attention to their feelings and what underlies their questions when they were talking about concerns in their health care in the hedonic scenario. This indicated that the interpretation and assessment of the CA was dependent on the hedonic or pragmatic scenario that it was presented with.

Instead of embodiment largely influencing how participants perceived the interactions, many participants mentioned the diverse nature of the both scenarios and that that was the most impactful on their assessment of the CA. This is well illustrated by a quote from one of the participants. This participant preferred the embodied version and interacted with it on the hedonic scenario, but mentioned that she was not sure how she would have assessed the other (disembodied) CA, if it was paired with the hedonic scenario. She said: '*The first topic* [hedonic] *obviously was that I was afraid and this* [second] *chat was purely about wanting information* [pragmatic]. *So, also, I don't really know how the second* [DCA] *would have answered this* [hedonic] *question, right?'*

Recognition and assessment of embodiment

Despite the fact that there was no strong preference for the embodied or disembodied interaction, a large part of the participants recognized the presence or absence of embodiment. Those that commented on the embodied character were either quite positive about its presence or had some remarks on it. For example, one participant commented: '*The figure was a little bit passive in the way she stood there'*. Another participant figured: *'If I may say so, then I think that what you showed, that figure, that you really have to let that figure talk, but also really let it speak out text'*.

In terms of deciding which interaction was more human-like, the majority of participants picked the embodied version, but not all participants reported that that was the result of the use of embodiment itself. In this, the scenario that was interacted upon seemed to be impactful again. One participant interacted with the ECA on the pragmatic scenario and found the conversation with this version more human-like than the other, because the ECA's content and dialogue were perceived as concrete and clear in information provision, not because of the presence of embodiment. Another participant interacted with the DCA on the hedonic scenario and found the conversation with the disembodied version more human-like,

as a result of asking about his feelings, not as a result of the absence of embodiment. However, some of them did assign more human-likeness to embodiment itself. For example, one participant said: '*That* [character] *makes it just a little friendlier or something'*. Thus, some participants found the CAs more human-like because of embodiment while the majority of participants found the CAs to be human-like because of its dialogue or content, regardless of application of embodiment or not.

Nevertheless, other participants were very sure in that they did not find the conversations human-like. They did not even want to call it a conversation. They stated, for example, the following: *'I have conversations with humans, but not with this'*. Besides, none of the participants really had the feeling that either of the conversation felt like it was with a real human. Some of them were very firm about this and answered the question regarding whether or not they found one of the interactions comparable to a conversation with a real human with: *'absolutely not'*.

CA's support in completion of hedonic or pragmatic goals: task success

Participants indicated that the use of CAs is more suitable for supporting the completion of pragmatic goals, than for hedonic goals. Many participants stated that interacting with the CA did not facilitate them feeling less worried, and therefore did not contribute to the completion of this hedonic goal. Firstly, they indicated that they would rather talk about such problems with a real person. This is well illustrated by the following statements by one of the participants: *'The purpose actually is to just get somewhere, let me put it that way, to talk to someone on the phone. So, I did not get that. So, I am just disappointed'*. Later, he said: *'Different for everyone, but if I am really in my feelings, then I find it terrible to do this'*. Another participant illustrated the same by saying: *'You know if this would happen* [hedonic scenario]*, then this* [interaction] *is not sufficient, then there has to be, well, that is what those tips said there, right? Then someone needs to talk to you. Then a*

chatbot is not enough.'

Secondly, some participants indicated that the dialogue and infographics in the interaction for completing the hedonic goal were not sufficient to complete those goals. The interaction was mainly focused on reducing possible worries that could be the result of changes in health care, while some participants preferred to interact on whether those concerns were justified or what was possibly going to change within their health care. A participant illustrated this by saying: *'You are concerned about the content, so you also want to know the content'*. Another participant described: *'I initially indicated that I was worried, that things were going to change and then they actually first led you to a website about dealing with anxiousness. That is like, well, that is not quite what I am looking for'*. In contrast, some participants mentioned that even though the tips that were given to deal with worries were not entirely suitable for themselves, they did see potential for others who might be in similar situations. For example, one participant said: *'Other people who are worried need to first be relaxed before getting to the content'.*

As for content, the content of the interaction for completing the pragmatic goal received a positive response, but there were also some remarks about this scenario by some participants. While many participants had the feeling that interacting with the CAs provided them with more information and that they felt helped by the CA in finding the right information, some participants had comments that regarded three categories. Firstly, some participants felt like the information that was provided could be found by themselves through searching on the internet and that the CA did not add value to this way of obtaining information. Secondly, some participants indicated that the infographics mostly covered one very specific topic, namely the *WLZ-indicatie,* which is a long-term care indication that people can apply for. In the infographics, it was explained how to apply for such an indication, as that was the described purpose of the pragmatic goal in this study. Some

participants felt like other long-term care indications that the Netherlands has to offer needed to be included as well. One person said: *'If my question was about the WLZ, then I was well helped, I think'*. Lastly, some participants felt like the CAs did not yet give a complete overview of all possible questions that vulnerable care recipients have and that they did miss the opportunity to get information about topics that applied to their own personal questions or situations.

Functional, technical and content-related suggestions

Besides the fact the participants evaluated interactions with the ECAs and DCAs, they also proposed some functional, technical and content-related suggestions for the CAs. These suggestions are valuable for the development of a CA that matches the desires of the target group. For example, some participants suggested a reading function, as they had some trouble with the amount of text that was presented by the CAs. Another participant suggested the ability to print the infographics that were given by the CAs. Many participants suggested the ability to be connected by telephone to a real person to ask further questions. Since talking to a real person was the only solution to obtain the hedonic goal in this study for some participants, a hand-over function like this was suggested multiple times.

Some content-related suggestions were the before-mentioned including of information about other Dutch long-term care indications (e.g., *WMO-indication*) for the pragmatic scenario and adjusting the presented emotions in the hedonic scenario. For the hedonic scenario, the CA asked participants whether they mostly felt angry, scared or sad when they worried about changes in their health care. Some participants mentioned that they could not really relate to any of those feelings and were confused about which option to choose as a result of that. However, they did not mention that other emotions needed to be included. It seemed as if they were having trouble with the specificity of these emotions; they rather felt worried and did not really want to connect a specific emotion to those worries.

Conclusion

The aim of this study was to explore the role of applying embodiment in conversational agents (CAs) for vulnerable care recipients, by qualitatively analyzing their user experiences overall and in terms of achievement of hedonic and pragmatic goals. To answer the research question, semi-structured interviews were performed. These interviews indicated that the role of applying embodiment is rather small; embodiment does not seem to play a role in the achievement of hedonic or pragmatic goals and overall user experience. Also, to answer the research sub-question, user's expectations did not have an one-to-one impact on user's experiences, as some participants with negative expectations had positive experiences.

With regard to overall assessments of the CAs, there was no strong preference for ECAs or DCAs. Instead, the evaluations of the CAs were influenced by the intended goal (either hedonic or pragmatic) and their dialogue and contents (either giving information or offering support). In general, there was a slight preference for using CAs for the obtainment of pragmatic goals (compared to hedonic goals), regardless of the use of embodiment. Evaluations showed that participants felt like using the CA for hedonic goals was not sufficient and that they preferred human-human interaction for hedonic goals.

This all is not exactly in line with the first expectation of this study, namely that ECAs provoke a more positive user experience than DCAs. Instead, part of the participants did indeed overall prefer using the ECA because of the presence of embodiment, while others did not or did not even recall the presence of embodiment. Secondly, it was expected that ECAs would be more suitable than DCAs for the completion of hedonic goals, but the results of this study indicate that that is not specifically the case. For pragmatic goals, however, ECAs and DCAs indeed were evaluated as equally able to support in completion of these goals, as was the third expectation.

Results indicated that the use of embodiment might provoke a positive user experience for some vulnerable people, but not for all of them. In conclusion, the type of goal (hedonic vs. pragmatic) is the largest predictor of a positive user experience with a CA for vulnerable care recipients. The use of embodiment could be beneficial to the overall evaluations of CA's for part of this target group, but did not seem to play a role in the achievement of goals. Future studies could use the current study as a starting point, as it provided broad insight into the yet to be further discovered field of CAs for vulnerable care recipients.

Discussion

Interpreting results and practical implications

While this study has revealed more insight into the unknown field of development and application of CAs in health and social care information and support for vulnerable care recipients, the results are not fully in line with expectations based on previous research. It was unexpected that embodiment did not play an evident role in overall assessment of the CAs or in the accomplishment of hedonic and pragmatic goals, while the nature of the goal itself (hedonic vs. pragmatic) seemed to be most influential on the user's experience with the CA.

A possible explanation for this result could be a possible insufficient emphasis on embodiment for participants. Some participants did not report on the differences in terms of embodiment and might not even have recognized them. However, it is not entirely clear whether or not they have not seen differences or whether they have not reported on them. If there would have been more emphasis on embodiment in the background briefing or during the interview, there might have been more responses regarding embodiment specifically. This could have given a more clear view of its influence on accomplishment of hedonic and pragmatic goals. Nevertheless, accentuating the background and the purpose of the study too much could have caused subject bias among participants. If all objectives of the study are clear, participants might (unconsciously) react to embodiment or answer questions in ways that they perceive are desirable for the study or researcher. Therefore, the right balance of emphasizing the fact that the study regards embodiment is required in future research on embodiment. Adding more questions about embodiment could be attributing to that balance.

In addition to emphasizing embodiment's role in the study, the expression of embodiment could also play a role in the explanation of the unexpected lack of a large influence of embodiment on user experience. In terms of embodiment's expression, this study made a distinction between using no embodiment and embodiment, meaning there was an

evident difference in use of embodiment. However, use of embodiment can be seen as a very broad spectrum of various aspects. There is not one singular way to use embodiment and there is no scientific consensus on what it exactly entails. As a result, use of embodiment is freely interpretable and this, however, might cause implications. While one shape or form of embodiment, for example the one in this study, may not be assessed positively by the entire target group, this does not have to be the case for every shape or form of embodiment in future research. For example, the addition of facial expressions, hand gestures or body movements could have all influenced the way the character and therefore the ECA would have been assessed. The use of those characteristics in other ECA studies, which are thus differing in expressions of embodiment, could possibly explain why embodiment's role in this study is less prominent.

Nevertheless, while the diverse possibilities for embodiment's manifestations make it difficult to design a "one-size-fits-all" CA for an entire target group, it could also be possible that the expression of embodiment might not be as important in CAs for this target group. The fact that the task scenarios and goals influenced user experience the most, leads to practical implications for the development of future CAs. Regardless of embodiment, the fixed nature of the created, fictional goals and scenarios could have been influential on participants' perceptions of CAs and their ability to provide information and support. In creating the hedonic and pragmatic goals and scenarios, the needs and goals of vulnerable care recipients have been taken into account (Van Heijster et al., in preparation) and the decision to formulate two fictional scenarios was done to maintain manageability for participants and similarity in their experiences. Some participants, however, implied that the study's scenarios and tasks were not completely representative for their own needs and that there should be the opportunity to ask questions or present problems apart from those fixed scenarios. This could have negatively influenced their experiences with the CA, in general and in the feeling of

accomplishing goals.

Thus, the results of this study imply that in future CA development and application, inclusion of functionalities through which participants can ask their own questions and (possibly AI-generated) tailored answers to those questions are needed. This would be in line with previous studies, that showed that conversational adaptation in CAs may be a tool to improve user experience (Kocaballi et al., 2019) and that adding more human-likeness to CAs is beneficial to a positive attitude towards them (Go & Sundar, 2019). Since all users have different conversational preferences, they might require different conversations in CAs (Kocaballi et al., 2019). Adapting the CAs answers to the user's demands in future CA application could reveal whether adding more human-likeness to CAs, visually but especially in terms of the conversation, could influence participants' beliefs about CAs capabilities of completing hedonic goals.

However, to nuance this, the other end of the spectrum, completely mimicking human-likeness and being on the verge of the uncanny valley, neither is beneficial to the assessment of CAs (Song & Shin, 2022). Besides, it is of ethical discussion whether you would want users, and especially users in vulnerable positions, to feel like they are having a conversation with a real human while actually interacting with a CA. When applying more conversational human-likeness and tailored answers, it remains important to maintain a manageable experience for this vulnerable target group and to keep in mind their vulnerable position.

Limitations and future directions for research

Although many valuable results were brought forward, there are some limitations to this study which bring along future directions for research. While the fixed nature of the scenarios induced practical implications, the manner in which support and information were offered in those scenarios could be adapted in future research. For this study, a combination of pre-made output and infographics were created to support and provide information, in order

for the interaction to be manageable, not too long and generally similar for each participant. However, especially for the hedonic scenario, this way of support provision and the dialogue itself were often reported as not sufficient to complete the hedonic goal of feeling supported and less worried. To enlarge CA's perceived competence in supporting the completion of hedonic goals, other support and information provision manners should be introduced. As many participants reported their needs to talk to a real person and reported a deficiency in human-likeness, in particular in the hedonic scenario, hybrid support and information provision by a combination of a CA and a human could have potential in the development and application of CAs for vulnerable care recipients. For example, in future research, the idea of a hand-over function from the CA to a real human could be further investigated.

Another limitation of methodological nature lies in some of the manner in which questions are asked. In future research, it might be of advantage to modify some questions used in this study. Asking participants about six CAs' characteristics through semantic opposites presented by bars (see Appendix D), which were based on quantitative semantic differentials, and creating a spectrum upon which the two interactions could be placed and compared, could be modified in future research. Firstly, for some participants, the way of asking was confusing; they did not understand that the bar represented a spectrum or found it confusing that there was no textual question above each bar. Secondly, presenting the questions in combination with bars also offered the opportunity to exactly place both interactions on the same place upon the spectrum. The bars offered the opportunity to pin down their thoughts on the CA before an attempt could be made to create a more comprehensive answer as to why they interpreted both the CAs in the same way. Participants could indeed have interpreted both interactions the exact same way, but placing the interactions on exactly the same place could also be the result of convenience or confusion. Lastly, the vulnerability of this target group also could have ensured that some of the

participants had some difficulties with assigning these characteristics, that were sometimes perceived as human-only characteristics, to the CAs. Concludingly, the conversion of these bars with opposites into verbal questions on these different aspects of CAs could perhaps be more suitable in the future.

Finally, a minor limitation in this study could lie in the fact that the diversity in the sample could have possibly been greater. In acquiring the sample, however, a caveat is that it is a complex process to acquire participants that are vulnerable care recipients and it is therefore of great value that ten participants were willing to participate in this study to provide in-depths insights into the application of CAs for this target group. Nevertheless, the large majority of participants' vulnerable position was related to ASD and the target group of vulnerable care recipients includes people with other vulnerable backgrounds as well (e.g., intellectual disabilities, mental health problems). While the inclusion of people with other vulnerable backgrounds could have been beneficial to this study, there still was a great variety between participants, creating an inclusive image and a wide range of valuable opinions either way. Every person's ASD, for instance, manifests itself differently and carries along different wants, needs and insights into the application of CAs. These different wants and needs may result in different wants and needs within CA interactions for each vulnerable care recipient. This creates space for challenging but also extremely important and valuable directions for future research in the field of creating an inclusive CA.

References

Ahmad, R., Siemon, D., Gnewuch, U., & Robra-Bissantz, S. (2022). Designing Personality-Adaptive Conversational Agents for Mental Health Care. *Information Systems Frontiers, 24,* 923-943. https://doi.org/10.1007/s10796-022-10254-9

Alsharbi, B., & Richards, D. (2017). Using Virtual Reality Technology to Improve Reality for Young People with Chronic Health Conditions. *Proceedings of the 9th International Conference on Computer and Automation Engineering, Sydney, Australia*, 11-15. https://doi.org/10.1145/3057039.3057080

Araujo, T. (2018). Living up to the chatbot hype: The influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions. *Computers in Human Behavior, 85,* 183-189. https://doi.org/10.1016/j.chb.2018.03.051

Baptista, S., Wadley, G., Bird, D., Olderburg, B., & Speight, J. (2020). Acceptability of an Embodied Conversational Agent for Type 2 Diabetes Self-Management Education and Support via a Smartphone App: Mixed Methods Study. *JMIR mHEalth and uHealth, 8*(7). https://doi.org/10.2196/17038

Bartneck, C., Kulic, D., Croft, E., & Zoghbi, S. (2009*).* Measuring Instruments for the Anthropomorphism, Animacy, Likeability, Perceived intelligence, and Perceived Safety of Robots. *International Journal of Social Robotics, 1*(1), 71-81. https://doi.org/10.1007/s12369-008-0001-3

Bickmore, T. W., Pfeifer, L.M., Byron, D., Forsythe, S., Henault, L. E., Jack, B. W., Silliman, R., & Paasche-Orlow, M. K. (2010). Usability of Conversational Agents by Patients With Inadequate Health Literacy: Evidence from Two Clinical Trials. *Journal of*

Health Communication, 15(2), 197-210. https://doi.org/10.1080/10810730.2010.499991

Curtis, R. G., Bartel, B., Ferguson, T., Blake, H. T., Northcott, C., Virgara, R., & Maher, C. A. (2021). Improving User Experience of Virtual Health Assistants: Scoping Review. *Journal of Medical Internet Research, 23*(12). https://doi.org/10.2196/31737

Diederich, S., Brendel, A. B., Morana, S., & Kolbe, L. (2022). On the Design of and Interaction with Conversational Agents: An Organizing and Assessing Review of Human-Computer Interaction Research. *Journal of the Association for Information Systems, 23*(1), 96-138. https://doi.org/10.17705/1jais.00724

Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review, 114*(4), 864-886. https://doi.org/10.1037/0033-295X.114.4.864

Feine, J., Gnewuch, U., Morana, S., & Maedche, A. (2019). A Taxonomy of Social Cues for Conversational Agents. *International Journal of Human-Computer Studies, 132,* 138-161. https://doi.org/10.1016/j.ijhcs.2019.07.009

Fitzpatrick, K. K., Darcy, A., & Vierhile, M. (2017). Delivering Cognitive Behavior Therapy to Young Adults With Symptoms of Depression and Anxiety Using a Fully Automated Conversational Agent (Woebot): A Randomized Controlled Trail. *JMIR Mental Health, 4*(2). https://doi.org/10.2196/mental.7785

Følstad, A., Araujo, T., Law, E. L.-C., Brandtzaeg, P. B., Papadopoulos, S., Reis, L., Baez, M., Laban, G., McAllister, P., Ischen, C., Wald, R., Catania, F., Meyer von Wolff, R., Hobert, S., & Luger, E. (2021). Future directions for chatbot research: an interdisciplinary research agenda. *Computing, 103*, 2915-2942. https://doi.org/10.1007/s00607-021-01016-7

Go, E., & Sundar, S. S. (2019). Humanizing chatbots: The effects of visual, identity and conversational cues on humanness perceptions. *Computers in Human Behavior*, *97*, 304-3016. https://doi.org/10.1016/j.chb.2019.01.020

Hassenzahl, M. (2007). The hedonic/pragmatic model of user experience. In E. Law., A. Vermeeren, M. Hassenzahl, & M. Blythe (Eds.), *Towards a UX Manifesto: Proceedings of the 21st British HCI Group Annual Conference, September 3-7, 2007* (pp. 10-14).

Hassenzahl, M. (2018). The Thing and I: Understanding the Relationship Between User and Product. In M. A. Blythe, & A. F. Monk (Eds.), *Funology 2: From Usability to Enjoyment* (pp. 301-313). Springer. https://doi.org/10.1007/978-3-319-68213-6_19

Kallio, H., Pietilä, A.-M., Johnson, M., & Kangasniemi, M. (2016). Systematic methodological review: developing a framework for a qualitative semi-structured interview guide. *Journal of Advanced Nursing*, *72*(12), 2954-2965. https://doi.org/10.1111/jan.13031

Kocabelli, A. B., Berkovsky, S., Quiroz, J. C., Laranjo, L., Tong, H. L. Rezazadegan, D., Briatore, A., Coiera, E. (2019). The Personalization of Conversational Agents in Health Care: Systematic Review. *Journal of Medical Internet Research*, *21*(11). https://doi.org/10.2196/15360

Kramer, L.L., ter Stal, S., Mulder, B. C., de Vet, E., & van Velsen, L. (2020). Developing Embodied Conversational Agents for Coaching People in a Healthy Lifestyle: Scoping Review. *Journal of Medical Internet Research*, *22*(2). https://doi.org/10.2196/14058

Laranjo, L., Dunn, A. G., Tong, H. L., Kocaballi, A. B., Chen, J., Bashir, R., Surian, D., Gallego, B., Magrabi, F., Lau, A. Y. S., & Coiera, E. (2018). Conversational agents in

healthcare: a systematic review. *Journal of the American Medical Informatics Association*, *25*(9), 1248-1258. https://doi.org/10.1093/jamia/ocy072

Lee, K. M., & Nass, C. (2003). Designing Social Presence of Social Actors in Human Computer Interaction. *Proceedings of the 2003 SIGCHI Conference of Human Factors in Computing Systems, Ft. Flauderdale, FL, USA, 5*(1), 289-296. https://doi.org/10.1145/642611.642662

Lee, K. M., Peng, W., Jin, S. A., & Yan, C. (2006). Can robots manifest personality? An empirical test of personality recognition, social responses, and social presence in human-robot interaction. *Journal of Communication*, *56*(4), 754-772. https://doi.org/10.1111/j.1460-2466.2006.00318.x

Lohse, M. (2011, July 31–August 3). *Bridging the gap between users' expectations and system evaluations* [Conference presentation]. 20th IEEE International Symposium on Robot and Human Interactive Communication, Atlanta, GA, USA. https://ieeexplore.ieee.org/document/6005252

Lombard, M., & Xu., K. (2021). Social Responses to Media Technologies in the 21st Century: The Media Are Social Actors Paradigm. *Human-Machine Communication*, *2*, 29-55. https://doi.org/10.30658/hmc.2.2

Luger, E., & Sellen, A. (2016). "Like Having a Really bad PA": The Gulf between User Expectation and Experience of Conversational Agents. *Proceedings of the 2016 SIGCHI Conference on Human Factors in Computing Systems, San Jose, CA, USA*, 5286-5297. https://doi.org/10.1145/2858036.2858288

Moon, Y., & Nass, C. (1996). How "Real" Are Computer Personalities?: Psychological Responses to Personality Types in Human-Computer Interaction. *Communication Research*, *23*(6), 651-674. https://doi.org/10.1177/009365096023006002

Mori, M. (2012). The Uncanny Valley (K. F. MacDorman, & N. Kageki,, Trans.). *IEEE Robotics and Automation, 19*(2), 98-100. https://doi.org/10.1109/MRA.2012.2192811 (Original work published 1970).

Nass, C., Steuer, J., & Tauber, E. R. (1994). Computers are Social Actors. In B. Adelson, S. Dumais, & J. Olson (Eds.), *Proceedings of the 1994 SIGCHI Conference on Human Factors in Computing Systems, Boston, MA, USA, April 24–28, 1994* (pp. 72-78). ACM. https://doi.org/10.1145/259963.260288

Nationale Ombudsman. (2018). *Zorgen voor Burgers: Onderzoek naar knelpunten die burgers ervaren bij de toegang tot zorg*, (2018/030). https://www.nationaleombudsman.nl/system/files/onderzoek/2018030%20Zorgen%20voor%20burgers.pdf

Nederlandse Zorgautoriteit (NZa). (2022, 13 December). *Zorg voor kwetsbare mensen steeds meer onder druk.* https://www.nza.nl/actueel/nieuws/2022/12/13/zorg-voor-kwetsbare-mensen-steeds-meer-onder-druk

Nguyen, T.-T., Sim, K., To You Kuen, A., O'donnell, R. R., Tee Lim, S., Wang, W., & Nguyen, H. D. (2021, July 12–14). *Designing AI-based Conversational Agent for Diabetes Care in a Multilingual Context* [Conference presentation]. Twenty-Fifth Pacific Asia Conference on Information Systems, Dubai, UAE. https://doi.org/10.48550/arXiv.2105.09490

Olsson, T. (2014). Layers of User Expectations of Future Technologies: an Early Framework. *CHI EA'14: CHI '14 Extended Abstracts on Human Factors in Computing Systems, Toronto, Canada*, 1957-1962. https://doi.org/10.1145/2559206.2581225

Owens, O. L., Felder, T., Tavakoli, A. S., Revels, A. A., Friedman, D. B., Hughes-Halbert, C., & Hébert, J. R. (2019). Evaluation of a Computer-Based Decision Aid for Promoting

Informed Prostate Cancer Screening Decisions Among African American Men: iDecide. *American Journal of Health Promotion, 33*(2), 267-278. https://doi.org/10.1177/0890117118786866

Provoost, S., Lau, H. M., Ruwaard, J., & Riper, H. (2017). Embodied Conversational Agents in Clinical Psychology: A Scoping Review. *Journal of Medical Internet Research, 19*(5). https://doi.org/10.2196/jmir.6553

Rabinowitz, A. R., Collier, G., Vaccaro, M., & Wingfield. R. (2022). Development of RehaBot – A Conversational Agent for Promoting Rewarding Activities in Users With Traumatic Brain Injury. *The Journal of Head Trauma Rehabilitation, 37*(3), 144-151. https://doi.org/10.1097/HTR.0000000000000770

Rhee, H., Allen, J., Mammen, J., & Swift, M. (2014). Mobile phone–based asthma self-management aid for adolescents (mASMAA): a feasibility study. *Patient Preference and Adherence, 8,* 63-72. https://doi.org/10.2147/PPA.S53504

Ring, L., Utami, D., & Bickmore, T. (2014). The Right Agent for the Job?: The effects of agent visual appearance on task domain. In T. Bickmore, S. Marsella, & C. Sidner (Eds.), *Intelligent Virtual Agents: 14th International Conference, IVA 2014, Boston, MA, USA, August 27–29, 2014, Proceedings* (pp. 374-384). Springer. https://doi.org/10.1007/978-3-319-09767-1_49

Schouten, D. G. M., Deneka, A. A., Theune, M., Neerincx, M. A., & Cremers, A. H. M. (2022). An embodied conversational agent coach to support societal participation learning by low-literate users. *Universal Access in the Information Society.* https://doi.org/10.1007/s10209-021-00865-5

Shechtman, N., & Horowitz, L. M. (2003). Media Inequality in Conversation: How People Behave Differently When Interacting with Computers and People. *Proceedings of the*

2003 SIGCHI Conference on Human Factors in Computing Systems, Ft. Lauderdale, FL, USA, 5(1), 281-288. https://doi.org/10.1145/642611.642661

Song, S. W., & Shin, M. (2022). Uncanny Valley Effects on Chatbot Trust, Purchase Intention, and Adoption Intention in the Context of E-Commerce: The Moderating Role of Avatar Familiarity. *International Journal of Human-Computer Interaction*, 1-16. https://doi.org/10.1080/10447318.2022.2121038

Straßmann, C., & Krämer, N. C. (2017). A Categorization of Virtual Agent Appearances and a Qualitative Study on Age-Related User Preferences. In J. Beskow, C. Peters, G. Castellano, C. O'Sullivan, I. Leite, & S. Kopp (Eds.), *Intelligent Virtual Agents: 17th International Conference, IVA 2017, Stockholm, Sweden, August 27–30 , 2017, Proceedings* (pp. 413-422). Springer. https://doi.org/10.1007/978-3-319-67401-8_51.

Sundar, S. S. (2008). The MAIN model: A heuristic approach to understanding technology effects on credibility. In M. J. Metzger & A. J. Flanagin (Eds.), *Digital media, youth, and credibility* (pp. 73-100). The MIT Press.

Ter Stal, S., Kramer, L. L., Tabak, M., Op den Akker, H., & Hermens, H. (2020). Design Features of Embodied Conversational Agents in eHealth: a Literature Review. *International Journal of Human-Computer Studies, 138*. https://doi.org/10.1016/j.ijhcs.2020.102409

Van Bussel, M. J. P., Odekerken-Schröder, G. J., Ou, C., Swart, R. R., & Jacobs, M. J. G. (2022). Analyzing the determinants to accept a virtual assistant and use cases among cancer patients: a mixed methods study. *BMC Health Services Research, 22*(1), 890-913. https://doi.org/10.1186/s12913-022-08189-7

Van Heijster, H. L., Liebrecht, C. C., Bol, N., Van Calis, J., Van der Cruijsen, A., Dijkgraaf, M., Feitz, W., Muller, M., Tromp, E., Bevelander, K., & Antheunis, M. L. (2023).

Inclusive design of a sensitive virtual assistant to support people in vulnerable positions in their access to care. [Manuscript in preparation]. Department of Communication and Cognition, Tilburg University.

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. https://doi.org/10.1145/365153.365168

Wetenschappelijke Raad voor het Regeringsbeleid (WRR). (2021, 15 September). *Zorg onder druk: keuzes overheid noodzakelijk.* https://www.wrr.nl/actueel/nieuws/2021/09/15/zorg-onder-druk-keuzes-overheid-noodzakelijk

Woittiez, I., Eggink, E., Putman, L., & Bas, M. (2018). *An international comparison of care for people with intellectual disabilities: An exploration*. Sociaal en Cultureel Planbureau. https://repository.scp.nl/bitstream/handle/publications/371/An%20internati onal%20comparisoc%20of%20care%20for%20people%20with%20intellectual%20dis abillities.pdf?sequence=1

Yee. N., Bailenson, J. N., & Rickertsen K. (2007). A Meta-Analysis of the Impact of the Inclusion and Realism of Human-Like Faces on User Experiences in Interfaces. *Proceedings of the 2007 SIGCHI Conference on Human Factors in Computing Systems, San Jose, CA, USA,* 1-10. https://doi.org/10.1145/1240624.1240626

Zamora, J. (2017). I'm Sorry, Dave, I'm Afraid I Can't Do That: Chatbot Perception and Expectations. *Proceedings of the 5th International Conference on Human Agent Interaction, Bielefeld, Germany,* 253-260. https://doi.org/10.1145/3125739.3125766

Appendix A

Proposals for the ECA

Figure A1

General idea of what the virtual environment with the ECA could look like, created with Canva Pro

Figure A2

Proposed inspiration for Mobile Water © for the design of an embodied conversational agent, created with Canva Pro

Appendix B

Infographics

Figure B1 depicts an example of one of the end-of-conversation infographics. Access to all infographics is provided through the following link:

https://surfdrive.surf.nl/files/index.php/s/7rNBGlVcFX21XLN?path=%2F

Figure B1

Example of one of the end-of-conversation infographics

Appendix C

Hedonic and pragmatic scenario and prior explanation

Prior scenario explanation

Welkom bij dit onderzoek!

Het onderzoek gaat over een chatbot die mensen zou kunnen helpen die vragen hebben over zorg. Een chatbot is een computerprogramma dat met iemand kan praten door te chatten. Veel mensen vinden het moeilijk om te weten waar ze informatie kunnen krijgen over zorg. Een chatbot kan daar bij helpen.

Voor dit onderzoek ga je twee keer kort met een chatbot praten. Je chat over twee onderwerpen. Wat de onderwerpen zijn, lees je in deze brief.

De onderwerpen gaan over uitkeringen. Misschien heb jij daar ook wel eens vragen over gehad. De chatbot stelt je straks een paar vragen. Je hoeft je antwoorden niet zelf te bedenken. Je kunt steeds kiezen uit twee of drie antwoorden.

Scenario reflecting pragmatic goal

Onderwerp A:

Stel je voor dat je op zoek bent naar informatie over je zorg en ondersteuning. Je hebt wat vragen over je uitkering en je wil er graag meer over weten. Je wil vooral graag weten hoe je de uitkering moet aanvragen. Je probeert hier achter te komen door met de chatbot te praten.

Scenario reflecting hedonic goal

Onderwerp B:

Stel je voor dat er iets veranderd waardoor zorg-uitkeringen anders worden. Het is nog niet duidelijk of jouw zorg en ondersteuning hierdoor gaan veranderen. Je maakt je daar een beetje zorgen over. Je gaat met de chatbot praten, misschien kan die je geruststellen.

Appendix D

Document with semantic opposites presented through bars

Toont interesse

Toont geen interesse

Beleefd

Onbeleefd

Behulpzaam

Onbehulpzaam

Aandacht
heeft voor jou

Geen aandacht
heeft voor jou

Slim

Dom

Nuttig

Nutteloos

Appendix E

Information bulletin

Chatbot voor vragen over zorg

Hallo, ik ben Lauren Verhoeven.
Ik studeer aan de Universiteit van Tilburg en doe hier onderzoek.

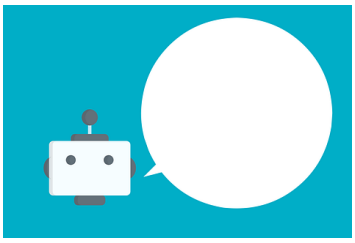
Wil je mij helpen met onderzoek?

Deze brief helpt je om te kiezen of je mee wil doen aan dit onderzoek of niet. Je kunt ook je begeleider of coach om hulp vragen bij het kiezen, of als je iets niet snapt. Heb je vragen? Dan kun je altijd mailen of bellen naar Lauren: xxxxx@tilburguniversity.edu, 06xxxxxxxx

Waar gaat dit onderzoek over?

Veel mensen vinden het moeilijk om weten waar ze informatie kunnen krijgen over zorg, bijvoorbeeld van de gemeente. Maar eigenlijk zou deze zorg of ondersteuning fijn voor ze zijn. Ze hebben bijvoorbeeld autisme, een verstandelijke beperking of iets aan hun hersenen. Ze hebben vragen over hun zorg waar niet altijd meteen antwoord op te vinden is. Een chatbot kan bij deze vragen helpen.

Wat is een chatbot?



Een chatbot is een computerprogramma dat met iemand kan praten door te chatten. Een chatbot kan zo worden gemaakt dat hij zich kan aanpassen aan degene met wie hij praat. Hij kan daarom sneller helpen bij het vinden van informatie over zorg en ondersteuning, bijvoorbeeld van de gemeente. Dit onderzoek gaat over de hulp van een chatbot bij speciale vragen over de zorg.

Wat gaan we doen?

Als je mee wil doen aan het onderzoek, dan maak ik een afspraak over wanneer je tijd hebt om mee te doen. Ik bespreek met jou waar jij het liefst wil afspreken. Als je het spannend vindt om alleen te gaan, dan mag je begeleider er ook bij zijn.

In het onderzoek praat je even met een chatbot over twee verschillende onderwerpen. Je kunt de onderwerpen voor het onderzoek alvast lezen als je dat fijn vindt. Tijdens het onderzoek mag je het onderwerp nog een keer lezen. De onderzoeker kan je daar ook bij helpen. Voordat je met de chatbot praat, stel ik je alvast een paar vragen over wat je verwacht van een chatbot. Daarna praat je even met de chatbot. De chatbot stelt je wat vragen en jij kunt antwoorden door met de computermuis op knoppen te klikken. Als je klaar

bent met praten met de chatbot, stel ik je een paar vragen over wat je ervan vond. Als je het goed vindt, maak ik een geluidsopname van het gesprek. Het duurt maximaal een uur.

Doe je mee met ons onderzoek?

Je helpt ons als je meedoet met dit onderzoek. Doordat je mee doet, weten we hoe we een chatbot kunnen maken die antwoorden weet op zorgvragen. Je mag altijd zelf kiezen of je meedoet. Er zijn geen goede of foute antwoorden, ik ben benieuwd naar wat jij van de chatbot vindt. Meedoen heeft geen nadelen voor jou.

Wat zijn de afspraken als je meedoet?

We overleggen samen wanneer je meedoet. Als je het fijn vindt om ook te overleggen met je begeleider of coach erbij, dan mag dat.

Als je meedoet, maar toch opeens wil stoppen, dan mag dat. Je hoeft daar dan geen reden voor te geven als je dat niet wil.

Als je wil meedoen en het goed vindt dat ik een geluidsopname van het gesprek maak, dan geef je toestemming door je handtekening te zetten op een toestemmingsformulier.

We bewaren de geluidsopname op een veilige plek, op de computers van Tilburg Universiteit. Hier maken we een verslag van. Daarna worden de geluidsopnames verwijderd. Jouw naam staat niet in het verslag, dat noemen we gepseudonimiseerd. Je mag het verslag van jouw gesprek inzien en aanpassen als je dat wil. Dit kan tot het onderzoek is afgelopen (augustus 2023). De verslagen bewaren we 10 jaar op de computers van Tilburg Universiteit.

De antwoorden die je aan de chatbot geeft worden niet bewaard en na afloop van het gesprek gewist.

Het kan zijn dat je een klacht hebt over het onderzoek. Als je dit liever niet tegen de onderzoeker wil zeggen, kun je naar dit e-mailadres mailen: tshd.redc@tilburguniversity.edu. Hier kun je hulp bij vragen van je begeleider.

Dit onderzoek is goedgekeurd door een groep mensen die kijkt of onderzoeken goed zijn. Dit is de REDC van Tilburg Universiteit.

Bedankt voor het lezen en misschien tot snel!

Appendix F

Informed consent form

Chatbot voor vragen over zorg

Dit formulier vul je in als je mee wil doen aan het onderzoek naar de chatbot. Je kunt iemand om hulp vragen om samen met je de tekst te lezen en je handtekening te zetten.

Afspraken

Meedoen aan het onderzoek betekent dat je vragen van de onderzoeker beantwoordt en een gesprekje hebt met een chatbot.

Tijdens de gesprekken maken we een geluidsopname. Zo kunnen we het gesprek terugluisteren en opschrijven in een verslag wat er is gezegd. Daarna wordt de geluidsopname verwijderd. De antwoorden die je aan de chatbot geeft worden niet bewaard en na afloop van het gesprek met de chatbot gewist.

We bewaren de geluidsopname tijdelijk op de computers van de Universiteit van Tilburg. Dit is een veilige plek. Nadat we het verslag hebben gemaakt, verwijderen we de geluidsopname. Alleen onderzoekers kunnen bij de geluidsopname en het verslag.

Je mag stoppen als je niet meer mee wilt doen.
Je hoeft niet te vertellen waarom je stopt.
Als je ergens geen antwoord op wil geven, dan hoeft dat niet. Er zijn geen goede of foute antwoorden.

De onderzoeker maakt een verslag van het gesprek. Jouw naam staat niet in het verslag (anoniem). Alles wat je zegt, kunnen alleen de onderzoekers van het project over chatbots zien en gebruiken.

Vink de hokjes aan als ze kloppen:

☐ Ik heb informatie gekregen over het onderzoek.

☐ Ik heb geen vragen meer over het onderzoek.

☐ Ik ben het eens met de afspraken.

☐ Ik wil meedoen aan het onderzoek.

Handtekeningen

Deelnemer

Datum: ……-……-……

Naam: ……

Handtekening:

Onderzoeker

Datum: ……-……-……

Naam: ……

Handtekening:

Als je een vraag hebt over dit onderzoek kun je mailen of bellen naar Lauren: xxxxxx@tilburguniversity.edu, 06xxxxxxxx

Dat mag ook als je na het onderzoek nog vragen hebt.