

Who Do You Prefer? The Effect of Age, Gender and Role on Users' First Impressions of Embodied Conversational Agents in eHealth

Silke ter Stal ^{a,b}, Monique Tabak ^{a,b}, Harm op den Akker ^{a,b}, Tessa Beinema ^{a,b}, and Hermie Hermens ^{a,b}

^aeHealth Group, Roessingh Research and Development, Enschede, The Netherlands; ^bMathematics and Computer Science, University of Twente, Enschede, The Netherlands

ABSTRACT

Embodied conversational agents may be used to engage users in adopting eHealth applications. The aim of this research is to investigate which design features establish a positive first impression of an agent in this context. A set of eight static agent images, different in age, gender and role, were subjected to testing in an online questionnaire. Respondents ($n = 155$) selected their preferred design and rated the characteristics – friendliness, expertise, reliability, involvement and authority – and the likeliness of following the agent's advice for all designs. In addition, focus groups ($n = 13$) were conducted for detailed understandings supporting these impressions. Our results show that, for both a general and elderly population, (1) people seem to prefer images of young, female agents over old, male agents, (2) the (a) age, (b) gender and (c) role of the agent image affect the perception of the agent's characteristics and the likeliness of following the agent's advice, and that (3) both the general and elderly population prefer an agent image that is similar in (a) age and (b) gender. A next step would be to investigate how the characteristics of the agent designs are perceived after interaction with the agent.

1. Introduction

The use of eHealth applications is being widely investigated as a way to relieve the burden on the healthcare sector caused by the aging society. Such applications are used for accessing relevant health information, improving the quality of care, reducing healthcare errors, increasing collaboration and encouraging the adoption of healthy behaviors (Krebs & Neuhauser, 2010). eHealth applications can be used in daily life, often under the supervision of a healthcare professional. For example, an application can provide users with video instructions on how to perform particular physical exercises via an online platform. The healthcare professional is not actually present, but supervises the exercising process at a distance.

For many eHealth applications, actual use typically decreases rapidly after several weeks (Nijland, 2011). One challenge for the design of eHealth applications is to keep the user engaged without including face-to-face interaction with the healthcare professional. Most of the existing eHealth applications provide advice in the form of plain text or via a text-based question-answer module (Kaptein, De Ruyter, Markopoulos, & Aarts, 2012). However, face-to-face interaction remains one of the best ways to communicate health information; it incorporates grounding – dynamically assessing the other person's level of understanding and repeating or elaborating on information when necessary (Clark & Brennan, 1991). In addition, face-to-face interaction makes use of verbal and non-verbal behavior to elicit trust, better communication and satisfaction (Bickmore, Pfeifer, & Jack, 2009).

1.1. Toward ECAs for eHealth

Since face-to-face interaction seems a good approach to elicit user engagement and stimulate actual use of eHealth applications, the use of embodied conversational agents (ECAs) is investigated. ECAs are *more or less autonomous and intelligent software entities with an embodiment used to communicate with the user* (Ruttkey, Dormann, & Noot, 2004). By interacting with the user face-to-face, ECAs can build trust and rapport, leading to companionship and long-term, continual use (Vardoulakis, Ring, Barry, Sidner, & Bickmore, 2012).

In order to reach engagement and actual use of the underlying application, it is important to adapt the characteristics of the ECA to the health context. Current research does not show what agent characteristics are especially important. But, existing studies explored characteristics, such as *friendliness* in relation to likability of the agent and engagement (Cafaro et al., 2012), and show that characteristics, such as *expertise, reliability, involvement* and *authority* are important aspects in patient-caregivers relations (Paap, Schrier, & Dijkstra, 2018; Pearson & Raeke, 2000; Ridd, Shaw, Lewis, & Salisbury, 2009), persuasive systems (Oinas-Kukkonen & Harjumaa, 2009), and eHealth applications in particular (Fogg, 2011; Mohr, Cuijpers, & Lehman, 2011; van Velsen et al., 2016). Therefore, these characteristics might also important be relevant for an ECA in the health context, to eventually, *follow the agent's advice*.

1.2. Engagement starts with a good first impression

In human-human interaction, first impressions of other humans are important (Bar, Neta, & Linz, 2006). Humans are excellent at judging personality traits and complex social characteristics such as dominance, hierarchy, warmth, and threat (Bar et al., 2006), but also in judging a person's skills, sexual orientation, political views and attitudes toward others (Cafaro, Vilhjálmsón, & Bickmore, 2016). Our first impressions judge the personality of someone to the extent of allowing us to anticipate the future behavior of that person (Vartanian et al., 2012). Therefore, they affect the development of relationships. Our first impressions can be shaped by both the static characteristics (Cafaro et al., 2016) and the dynamic characteristics (Bergmann, Eyssel, & Kopp, 2012; Cafaro et al., 2016) of a person. *Static characteristics often relate to a person's visual appearance, whereas dynamic characteristics include a person's verbal and non-verbal behaviors.*

Just as in human-human interaction, first impressions are important in human-agent interaction. According to Bergmann et al. (2012), there is evidence that both static and dynamic characteristics are cues of major importance for how humans evaluate ECAs.

1.3. Agent design and perception of agent characteristics

Several research has been performed on agent design features, among which the agent's age, gender and role in particular. First, research indicates that people prefer an agent of a particular age. Some research indicates that people prefer to interact with a young character (Cowell & Stanney, 2003; Rosenberg-Kima, Baylor, Plant, & Doerr, 2008), whereas other research indicates that people tend to prefer an agent of their own age (Alsharbi & Richards, 2017; Lee, Xiao, & Wells, 2018). In addition, the agent's age affects the users' perception of particular agent characteristics, such as trustworthiness (Lee et al., 2018).

Second, several research has been performed on the agent's gender. Research shows both a preference for female (Alsharbi & Richards, 2017) and male agents (Kim, Baylor, & Shen, 2007) and indicates that people prefer an agent of the same gender as themselves (Bailenson, Blascovich, & Guadagno, 2008; Guadagno, Blascovich, Bailenson, & McCall, 2007). Furthermore, on the one hand, research suggests that for particular tasks people prefer agents that are associated with these tasks conform gender stereotypes (Forlizzi, Zimmerman, Mancuso, & Kwak, 2007; Zimmerman, Ayoob, Forlizzi, & McQuaid, 2005). On the other hand, research also suggests the opposite (Baylor & Kim, 2004; Rosenberg-Kima et al., 2008). In addition, the agent's gender affects the users' perception of the agent. Males are perceived as more powerful (Nunamaker, Derrick, Elkins, Burgoon, & Patton, 2011), more knowledgeable (Baylor & Kim, 2004), more intelligent (Baylor & Kim, 2004) and are seen as better facilitated learners (Kim et al., 2007), whereas females are perceived as more likable (Nunamaker et al., 2011), but are also more verbally abused (Silvervarg, Raukola, Haake, & Gulz, 2012).

Third, research shows that the agent's role, such as the agent being an expert or a mentor, affects how its characteristics are

perceived. The agent's role affects the users' perception of the agent's level of expertise and intelligence (Baylor & Kim, 2004; Nguyen & Masthoff, 2009; Veletsianos, 2010) and how motivating the agent is (Kim, Baylor, & Reed, 2003). In particular, one study indicates that agents dressed to fit their role are perceived to be more professional, trustworthy, reassuring, and more persuasive compared to agents whose appearance is not role-appropriate (Parmar, Olafsson, Utami, & Bickmore, 2018).

In general, several studies suggest that people prefer agents who are similar to themselves (Bailenson et al., 2008; Guadagno et al., 2007; Rosenberg-Kima et al., 2008; Wissen, Vinkers, & Halteren, 2016).

Although this research shows that the agent design features age, gender and role affect the users' perception of ECAs, *research shows mixed results when it comes to what age, gender and role to implement.* In addition, we do not know whether these results can also be translated to agents in the health context and how these design features affect the users' perception of agent characteristics relevant for the health context in particular.

1.4. Research objectives

This research aims to investigate how ECAs in the eHealth domain can be designed to trigger positive first impressions. The research takes a *zero acquaintance approach* (Vartanian et al., 2012), meaning that *the design of ECAs will be evaluated based on solely a static agent image, without implementing actual interaction between the agent and the user.* We are specifically interested in how the design features age, gender and role affect the first impressions of the agent's characteristics *friendliness, expertise, reliability, authority and involvement, the likeliness of following the agent's advice,* and whether the design features relate to the characteristics of the respondents. The hypotheses to be tested by this research can be seen in Table 1.

2. Method

We aim to determine which static agent design features are important to establish a positive first impression, in particular for an elderly population (≥ 55 years). We combined both quantitative and qualitative research methods. First, we

Table 1. Hypotheses to be tested by this research.

Nr.	Description
H1.1	The age of a static agent image affects the respondents' perceptions on at least one of the five characteristics – friendliness, expertise, reliability, authority and involvement – or the likeliness of following the agent's advice
H1.2	The gender of a static agent image affects the respondents' perceptions on at least one of the five characteristics – friendliness, expertise, reliability, authority and involvement – or the likeliness of following advice
H1.3	The role of a static agent image affects the respondents' perceptions on at least one of the five characteristics – friendliness, expertise, reliability, authority and involvement – or the likeliness of following advice
H2.1	Respondents prefer an image of an agent of similar age as themselves
H2.2	Respondents prefer an image of an agent of the same gender as themselves
H2.3	Respondents prefer an image of an agent having the same level of health expertise as themselves

investigated the impressions of a general and elderly population using an online questionnaire. Next, we investigated impressions of an elderly population by means of focus groups.

2.1. First impressions general and elderly population

2.1.1. Participants

Respondents to the questionnaire should be fluent in the Dutch or the English language. No other inclusion or exclusion criteria were set. We recruited the respondents via a Dutch panel of adults that indicated they were interested in participating in research on eHealth and through snowball sampling via social media and personal connections. The questionnaire was accessible via a public link of the survey program Qualtrics and available for two months, in July and August 2018.

2.1.2. Agent design characteristics

To be able to compare the agent designs, the style of the agent designs was kept constant. The agent images differed on three features: the agent's *age* (young or old), the agent's *gender* (male or female) and the agent's *role* (expert – having a high level of health expertise – or peer – having a low level of health expertise). Table 2 provides an overview of the features. Combinations of all variations were tested, leading to a set of

Table 2. The agents subjected to test in the study differed in three features: age, gender and role, resulting in a set of eight agent permutations.

Design Feature	Value 1	Value 2
Age	Young	Old
Gender	Male	Female
Role	Expert	Peer
(Level of health expertise)	(high level of health expertise)	(low level of health expertise)

eight agent permutations. The individual agent designs are shown in Figure 1.

2.1.3. Measurements

The following data were collected via the online questionnaire:

- Characteristics of respondents (age, gender, health literacy).
- Preferred agent design at first glance.
- For each agent design: likeliness of following the agent's advice.
- For each agent design: ratings of importance of five agent characteristics: friendliness, trustworthiness, involvement, expertise and authority.

2.1.4. Procedure

Figure 2 shows the flow of the online questionnaire. In the introduction to the questionnaire, the context of the research was explained and informed consent was obtained. The questionnaire itself consisted of three parts. The first part consisted of questions on the characteristics of the respondent. In the second part, the eight agent designs were shown to the respondent simultaneously. Then, the respondent selected one of the designs as his or her preferred design at first glance. The position of the various agents on the screen was randomized to avoid any bias. In addition, the respondent had the opportunity to state the rationale behind his or her preference in a text box. In the final part of the questionnaire, each agent was shown individually. For each agent, the respondent rated the likeliness of following the advice of the agent on a 7-point Likert scale. Also, he or she was asked to rate the five characteristics of the agent on a 7-point Likert scale. The order in which the individual agent designs were shown was randomized. In total, the questionnaire took around 15 minutes.



Figure 1. The eight agent designs subjected to testing.

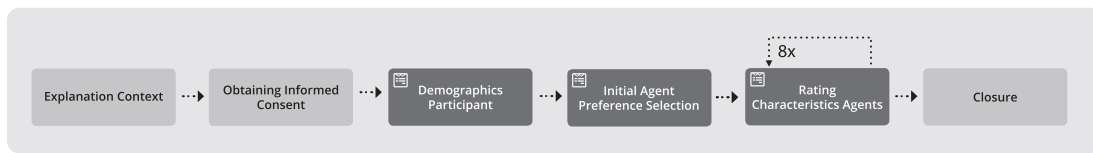


Figure 2. Flow of the online questionnaire.

2.1.5. Data analysis

The questionnaire data was exported from Qualtrics to Excel. The respondents' age was treated as a continuous variable, whereas all other respondents' characteristics were treated as categorical variables and responses on Likert scale questions as discrete variables. As a next step, for all categorical respondents' characteristics, the respondents were divided into two new categorical variables per characteristic. The answers to the open question requesting an explanation for the selected agent design, were coded by two independent researchers (StS, TB) and finally coded using the final coding scheme that both researchers agreed upon.

As a next step, the Excel file was imported into SPSS 25 statistics program to perform statistical analyses. For all relations between two categorical and/or discrete variables, a Chi-square test or Fisher's Exact test was conducted as appropriate. For all relations with a continuous variable, an independent-samples or paired-samples t-test was conducted. All tests were performed using a 95% Confidence Interval, and included respondents that filled out the questions regarding the variables necessary for the particular tests.

2.2. Attitude elderly population toward agent characteristics

2.2.1. Participants

Participants in the focus groups should be aged 55 years or above and fluent in the Dutch language. We recruited the respondents via a Dutch panel of adults that indicated they were interested in participating in research on eHealth. The focus groups were performed in July 2018.

2.2.2. Measurements

Two identical focus groups were conducted. The following data, quantitative and qualitative respectively, were collected in the focus groups:

- Ratings of importance for a set of twenty predefined agent characteristics (hair color, skin color, clothing, gender, age, voice, language usage, humor, intelligence, reliability, cultural background, political preferences,

posture, role, shape, friendliness, expertise, authority, involvement and hobbies)

- Explanation of the rating of the importance of all agent characteristics.

2.2.3. Procedure

Figure 3 shows the outline of the focus group process. The focus group interactions were audio recorded. All participants signed informed consent at the beginning of the focus group. Two researchers not involved in the study led the focus groups. The focus groups started with an introduction, acquainting the participants with each other, the purpose of the research and the goal of the meeting. Then, the participants were familiarized with the concept of virtual coaches using some examples of virtual characters and an example scenario in a non-health context. After the introduction, the required functionalities of virtual coaches were discussed. The discussion was aligned to three scenarios: a virtual coach for (1) physical activity; (2) nutrition; and (3) social activity. Next, the participants individually performed a card-sorting task. Participants received a set of cards with twenty key characteristics of agents and a few blank cards. In addition, they received a sheet with the title: "My ideal coach is ..." with two columns below, labeled "important" (left) and "less important" (right). Participants were asked to place the cards of characteristics they perceived as important at the left and the cards of characteristics they believed were less important at the right. After the card-sorting task, each participant explained which characteristics he or she believed were important and why. Participants were then encouraged to respond to each other in a general discussion.

2.2.4. Data analysis

The attitudes of the participants toward the functionalities and characteristics of virtual coaches were thematically analyzed. The themes were coded using ATLAS.ti 8, based on the steps proposed by Pope and Mays (2006). One researcher (StS) created a first coding scheme and labeled all the data accordingly. A second researcher (TB) used the coding scheme to code a subset of the data. Disagreements between



Figure 3. Outline of the focus group process.

the first and second researcher were discussed and overcome, leading to an updated coding scheme. The first researcher used that updated coding scheme to re-code all data entries and the second researcher then independently re-coded a new subset. Again, disagreements between the two researchers were discussed and overcome, leading to the final coding scheme used by the first researcher to re-code all data one final time.

In addition, the sheets of the card-sorting task were digitized. The importance of a characteristic was labeled as either “important” (left side of the sheet), “neutral” (center of the sheet) or “less important” (right side of the sheet). For each characteristic, the frequency of each label was counted.

3. Results

3.1. First impressions general and elderly population

3.1.1. Descriptives

In total, 155 people participated in the online questionnaire, of which 115 people filled out the complete questionnaire. The age of the general population that filled out the complete questionnaire ranged from 17 to 87 years ($M = 51.36$ years, $SD = 20.71$, 22 unknown) and 69 were female and 67 were male (19 unknown). The elderly population that filled out the complete questionnaire consisted of 66 participants, of which the age ranged from 55 to 87 years ($M = 67.85$ years, $SD = 6.97$ years, 1 unknown) and 30 were female and 35 were male (1 unknown).

3.1.1.1. Preference agent designs at first glance. Figure 4 shows the frequencies of the preference for the agent designs at first glance. Overall, the young female peer agent was selected most (32 times, 24.61%). Also, the young female doctor agent was selected often (23 times, 17.69%). The old male doctor agent and old male peer agent were selected 6

times (5.00%) and 5 times (4.00%) respectively. Overall, these results indicate a high preference for the young female agents and a low preference for the old male agents. When looking at the elderly population in particular, we see that, similar to the general population, the young, female peer is selected the most and the old, male peer the least. However, the ranking of the agents in between shows some differences compared to the general population.

Arguments supporting the agent preferences provided in the open question within the questionnaire varied substantially. Firstly, respondents indicated they preferred a certain agent because of its appearance: its age (23 times), its appearance in general (9 times), its clothing (10 times), its gender (6 times), or because they thought the agent looked beautiful (3 times) or healthy (1 time). Secondly, respondents provided arguments related to the perceived personality of the agent, such as the agent looking friendly (5 times), sympathetic (4 times), accessible (4 times), professional (4 times), reliable (3 times), credible (1 time), happy (1 time), calm (1 time) and authoritative (1 time). Others selected the agent because of its role (10 times) or expertise (8 times). Finally, some respondents indicated they could identify with the agent (8 times).

3.1.2. Comparison perceived characteristics agents designs

For each agent characteristic (friendliness, expertise, reliability, authority and involvement) and the likeliness of following advice, the mean ratings were compared for both agent feature categories (i.e. young agents vs old agents, female agents vs male agents and peer agents vs expert agents), for both the general and elderly population. First, except for likeliness of following advice, expertise and reliability for the elderly population, the mean ratings of the characteristics of the images of the young and old agents are significantly different (see Table 3). Second, except for likeliness of following advice and reliability for both the general and elderly population, and friendliness, expertise and

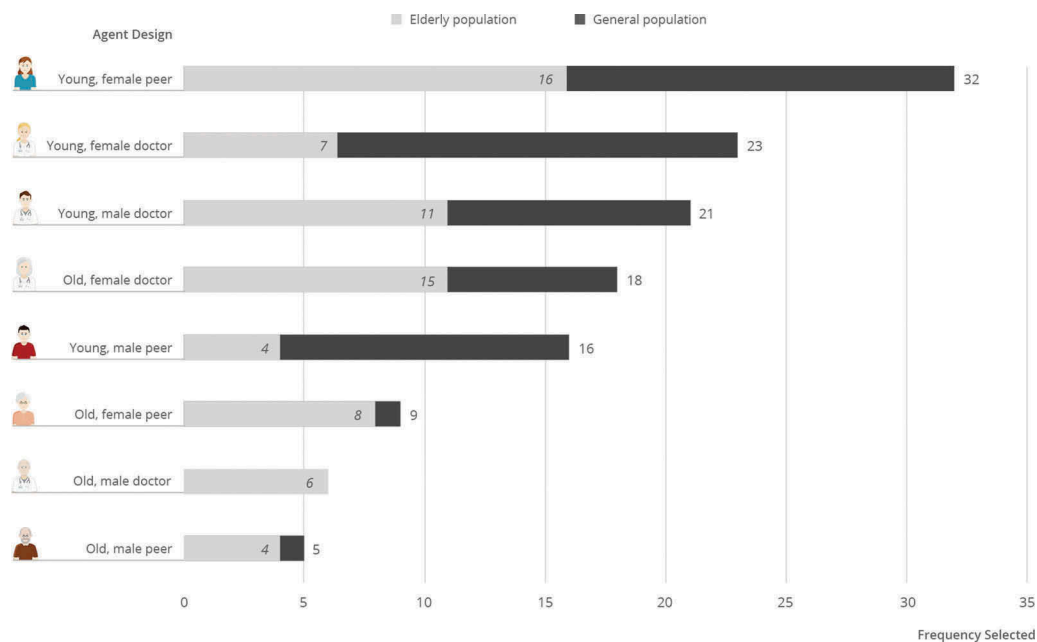


Figure 4. Frequencies of the agent designs preferred at first glance.

Table 3. Results of the paired-samples t-tests comparing the mean ratings of the characteristics of the images of young and old agents for both the general population (n = 117) and the elderly population (n = 66). The range of the mean ratings is from 1 (strongly disagree) to 7 (strongly agree). The relations for which the p-values are shown in bold are statistically significant. Except for likeliness of following advice, expertise and reliability for the elderly population, the mean ratings of the characteristics of the images of the young and old agents are significantly different. The cells colored in gray show similar results for both the general and elderly population.

	General Population				Elderly Population			
	Young, M(SD)	Old, M(SD)	p	Cohen's d	Young, M(SD)	Old, M(SD)	p	Cohen's d
Likeliness of following advice	4.78(1.127)	4.35(1.150)	<.001	0.471	4.71(1.298)	4.54(1.216)	.095	–
Friendliness	5.50(0.880)	4.84(1.038)	<.001	0.769	5.25(0.971)	4.67(1.095)	<.001	0.654
Expertise	4.79(0.960)	4.62(0.960)	.038	0.191	4.61(1.048)	4.61(1.030)	1.000	–
Reliability	4.97(0.938)	4.83(0.935)	.046	0.186	4.69(1.020)	4.73(0.983)	.643	–
Authority	3.59(1.155)	4.19(0.935)	<.001	–0.517	3.53(1.215)	4.04(0.962)	<.001	0.486
Involvement	4.79(1.099)	4.50(1.019)	.001	0.314	4.75(0.990)	4.53(1.056)	.023	0.286

involvement for the elderly population, the mean ratings of the characteristics of the images of the female and male agents are significantly different (Table 4). Third, except for friendliness for both the general and elderly population, and involvement for the general population, the mean ratings of the characteristics of the images of the peer and expert agents are significantly different (see Table 5). Table 6 shows a summary of the results, the results support all three hypotheses H1.1, H1.2 and H1.3, for both the general and elderly population.

3.1.3. Relation characteristics respondent and features preferred agent design

We tested the relation of the age, gender and health literacy of the respondents with the three agent features – age, gender and role – respectively. The results of these tests are shown in Table 7–9. A summary of all results can be seen in Table 10. H2.1 and H2.2 are supported, whereas H2.3 is rejected.

3.1.3.1. Relation age preferred agent and age respondent.

Table 7 shows that there was a significant difference between the mean age of the respondents' that preferred an image of a young agent (M = 47.52 years, SD = 21.925) and the mean age of the respondents' that preferred an image of an old agent (M = 64.46 years, SD = 13.312). *Younger respondents are more likely to select an image of a young agent than older respondents.*

3.1.3.2. Relation gender preferred agent and gender respondent.

Table 8 shows that there was a significant relation between the gender of the respondents (male or female) and the gender of the selected agent design for both the general population and the elderly population. *Male respondents are more likely to select an image of a male agent.*

3.1.3.3. Relation role preferred agent and health literacy respondent.

No significant effect of the role of health literacy

Table 4. Results of the paired-samples t-tests comparing the mean ratings of the characteristics of the images of the female and male agents for both the general population (n = 117) and the elderly population (n = 66). The range of the mean ratings is from 1 (strongly disagree) to 7 (strongly agree). The relations for which the p-values are shown in bold are statistically significant. Except for likeliness of following advice and reliability for both the general and elderly population, and friendliness, expertise and involvement for the elderly population, the mean ratings of the characteristics of the images of the female and male agents are significantly different. The cells colored in gray show similar results for both the general and elderly population.

	General Population				Elderly Population			
	Female, M(SD)	Male, M(SD)	p	Cohen's d	Female, M(SD)	Male, M(SD)	p	Cohen's d
Likeliness of following advice	4.64(1.150)	4.49(1.163)	.063	–	4.67(1.237)	4.57(1.262)	.284	–
Friendliness	5.28(0.936)	5.06(0.882)	<.001	0.113	5.02(0.993)	4.89(0.961)	.086	–
Expertise	4.63(0.891)	4.77(0.898)	.005	–0.069	4.59(0.998)	4.63(0.982)	.504	–
Reliability	4.93(0.919)	4.88(0.881)	.354	–	4.73(0.994)	4.69(0.932)	.483	–
Authority	3.82(0.911)	3.96(0.956)	.019	–0.071	3.69(0.973)	3.88(1.032)	.014	–0.310
Involvement	4.74(1.023)	4.54(0.980)	<.001	0.100	4.71(0.969)	4.58(1.010)	.053	–

Table 5. Results of the paired-samples t-tests comparing the mean ratings of the characteristics of the images of the peer and expert agents for both the general population (n = 117) and the elderly population (n = 66). The range of the mean ratings is from 1 (strongly disagree) to 7 (strongly agree). The relations for which the p-values are shown in bold are statistically significant. Except for friendliness for both the general and elderly population, and involvement for the general population, the mean ratings of the characteristics of the images of the peer and expert agents are significantly different. The cells colored in gray show similar results for both the general and elderly population.

	General Population				Elderly Population			
	Peers, M(SD)	Experts, M(SD)	p	Cohen's d	Peers, M(SD)	Experts, M(SD)	p	Cohen's d
Likeliness of following advice	3.69(1.353)	5.16(1.076)	<.001	–0.994	4.10(1.508)	5.15(1.199)	<.001	–0.789
Friendliness	5.20(1.047)	5.14(0.873)	.464	–	4.94(1.107)	4.98(0.946)	.700	–
Expertise	4.04(1.167)	5.37(0.974)	<.001	–1.027	4.17(1.180)	5.05(0.979)	<.001	–0.880
Reliability	4.55(1.061)	5.25(1.061)	<.001	–0.697	4.44(1.139)	4.99(0.908)	<.001	–0.630
Authority	3.44(1.003)	4.34(1.067)	<.001	–0.816	3.51(1.041)	4.06(1.086)	<.001	–0.595
Involvement	4.57(1.131)	4.72(0.992)	.077	–	4.52(1.136)	4.77(0.964)	.026	–0.281

Table 6. Results supported the hypotheses H1.1, H1.2 and H1.3.

Hypothesis	General Population		Elderly Population	
	Outcome	Explanation	Outcome	Explanation
H1.1 – Age affects perceptions	Supported	– Images of young agents are rated higher on (a) <i>likeliness of following advice</i> , (b) <i>friendliness</i> , (c) <i>expertise</i> , (d) <i>reliability</i> and (e) <i>involvement</i> compared to images of old agents – Images of old agents are rated higher on (f) <i>authority</i> compared to images of young agents	Supported	– Images of young agents are rated higher on (b) <i>friendliness</i> and (e) <i>involvement</i> compared to images of old agents. – Images of old agents are rated higher on (f) <i>authority</i> compared to images of young agents.
H1.2 – Gender affects perceptions	Supported	– Images of female agents are rated higher on (b) <i>friendliness</i> and (e) <i>involvement</i> than images of male agents – Images of male agents are rated higher on (c) <i>expertise</i> and (f) <i>authority</i> than images of female agents	Supported	– Images of male agents are rated higher on (f) <i>authority</i> than images of female agents.
H1.3 – Role affects perceptions	Supported	– Images of expert agents are rated higher on (a) <i>likeliness of following advice</i> , (c) <i>expertise</i> , (d) <i>reliability</i> and (f) <i>authority</i> than images of peer agents	Supported	– Images of expert agents are rated higher on (a) <i>likeliness of following advice</i> , (c) <i>expertise</i> , (d) <i>reliability</i> , (e) <i>involvement</i> and (f) <i>authority</i> than images of peer agents.

of the respondent on the role of the preferred agent was found, neither for the general and elderly population.

3.2. Attitude elderly population toward agent characteristics

3.2.1. Descriptives

Thirteen people ($n = 13$) participated in the focus groups. The age of the participants ranged from 58 to 81 ($M = 71.23$ years, $SD = 6.82$). Five males and eight females participated. In the first focus group, four males and four females ($n = 8$, $M = 73.13$ years, ages ranged from 58 to 81 years) participated. In the second focus group, one male and four

females participated ($n = 5$, $M = 68.20$ years, ages ranged from 61 to 75 years).

3.2.2. Frequencies agent characteristics card sorting task

In the focus groups, elderly provided their opinion on agent features important for an agent in the health domain.

Table 7. Results of the independent-samples t-test testing the relation between the age of the respondent and the age of the agent design preferred by the respondent for the general population ($n = 123$): the relation is statistically significant with a large effect size.

Age Respondent	Age Preferred Agent			Cohen's <i>d</i>
	Young, <i>M(SD)</i>	Old, <i>M(SD)</i>	<i>p</i>	
	47.52(21.925)	64.46(13.312)	<.001	0.934

Table 8. Results of the Chi-square tests testing the relation between the gender of the respondent and gender of the agent design preferred by the respondent for both the general population ($n = 126$) and the elderly population ($n = 71$): both relations are statistically significant, having a small effect size for both populations.

		Gender Preferred Agent			<i>p</i>	Phi ϕ
		Female, <i>n(%)</i>	Male, <i>n(%)</i>			
General Population	Female	48(73.8)	17(26.2)			0.238
	Male	31(50.8)	30(49.2)			
Elderly Population	Female	27(79.4)	7(20.6)			0.293
	Male	19(51.4)	18(48.6)			

Table 9. Results of the Fisher's exact tests testing the relations between the health literacy of the respondent and the role of the agent design preferred by the respondent for both the general population ($n = 126$) and the elderly population ($n = 71$): both relations are not statistically significant.

		Role Preferred Agent			<i>p</i>
		Peer, <i>n(%)</i>	Expert, <i>n(%)</i>		
General Population	Low literate	3(42.9)	4(57.1)		1.000
	Moderate or high literate	56(47.1)	63(52.9)		
Elderly Population	Low literate	2(66.7)	1(33.3)		.585
	Moderate or high literate	30(44.1)	38(55.9)		

Table 10. Results supported the hypotheses H2.1, H2.2 and rejected H2.3.

Hypothesis	General Population		Elderly Population	
	Outcome	Explanation	Outcome	Explanation
H2.1 – Preference image of similar age	Supported	– Younger respondents were more likely to select an image of a young agent than older respondents	n.a.	
H2.2 – Preference image of same gender	Supported	– Male respondents are more likely to select an image of a male agent	Supported	– Male respondents are more likely to select an image of a male agent
H2.3 – Preference image of similar health expertise	Rejected		Rejected	

Whereas this research focuses on static design features (features related to a person's visual appearance) participants were not restricted to provide their opinion on static design features; they could also provide information on dynamic design features (a person's verbal and non-verbal behaviors).

Figure 5 shows the frequencies of the ratings of importance for agent characteristics of the card sorting task in the focus groups. All participants (13, 100%) rated humor, friendliness, expertise, involvement and reliability as important. In addition, most participants (11, 85%) considered the agent's role important. The same applies to the agent's language usage (10, 77%), intelligence (10, 77%) and voice (9, 69%). Shape and hobbies were rated quite neutral – the characteristics were considered important by six (46%) and five participants (38%) respectively. Overall, few participants rated the agent's age (3, 23%), authority (3, 23%), clothing (2, 15%), gender (1, 8%) political preference (1, 8%), posture (1, 8%), hair color (1, 8%) and skin color (1, 8%) as important.

3.2.3. Attitude toward design features

Figure 5 shows that, overall, the design features *age* and *gender* subjected to test in the online questionnaire were not considered important. On the contrary, the agent's *role* was considered important by almost all participants. Regarding the agent's *age*, two participants indicated they prefer a young agent, since they are themselves elderly. As one of these participants said: "I believe that, if the agent in front of me is around 25 or 30 years old, this agent will be very refreshing and will think differently from someone of my age. For example, it sees everything much more royal". This finding contradicts the findings of the online questionnaire, in which elderly preferred elderly agents. About the agent's *role*, several participants indicated they could see various roles for the agent, depending on the context. In certain situations they prefer an expert; in other situations a peer. One participant said: "That the agent is an expert in the subject he is coaching on". Another participant specifically indicated the importance of expertise in the medical field: "It depends on the subject. When the conversation is on medical subjects, then the agent should be an expert, to be able to provide correct answers". Conversely, another participant indicated the importance of the agent as a peer: "Sometimes one prefers

someone who is half peer, half tutor, instead of someone who is telling you what to do".

3.2.4. Attitude toward agent characteristics

Figure 5 shows that almost all agent characteristics used in the online questionnaire – *friendliness*, *expertise*, *reliability* and *involvement* – were considered important. *Authority* was considered less important. One female participant said "What I also wrote is (...) positive conformations. In contrast to saying: 'Oh, well, it went completely wrong', that would be horrible." aiming for the agent to be *friendly*. During the focus group discussion, some participants related *expertise* to *intelligence* and *authority*. Participants had difficulties with the meaning of the word *authority*. Some understood *authority* as "being an expert in a certain subject", whereas others understood it as "being authoritative". Overall, participants did not want their coach to be authoritative: "Well, the agent should not be like, 'everything I say is good, because I studied for it'", but some mixed opinions existed. Some participants that indicated in some situations an authoritative coach could be useful: "An authority, yes, I could benefit from an authoritative figure now and then". On the contrary, the advice given by the agent should be trustworthy; as one participant described it: "Expertise: that I can be 100% confident that when someone tells or explains something to me, this is researched". Another participant explained "that the agent is reliable, you know, that one can blindly follow it at some point".

4. Discussion

The aim of this research was to determine which static agent design features are important to establish a positive first impression in the eHealth domain. We specifically researched the influence of the agent's age, gender and role. We found that, for both a general and elderly population, (1) *people seem to prefer images of young, female agents over images of old, male agents*, that (2) *the agent's (a) age, (b) gender and (c) role affect the perception of the characteristics of the agent image and the likeliness of following the agent's advice* and that (3) *people show a preference for an image of an agent that is similar in (a) age and (b) gender*.

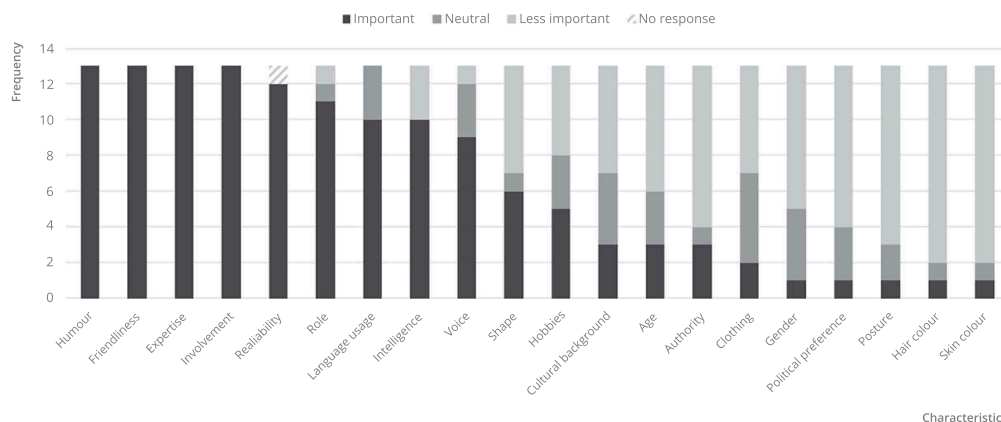


Figure 5. Results of card sorting task in focus groups. Participants rated each characteristic as either important, neutral or less important.

4.1. Overall: Preference for young females

Our results indicate that, overall, respondents preferred an image of the young, female peer agent, followed by the image of a young, female doctor agent. The preference for a female agent design could be the result of the health coaching task being associated with female gender stereotypes (Forlizzi et al., 2007). The young, female peer agent design being preferred above the design of a female doctor agent, could be explained by the suggested role for the agent – a coach – at the introduction of the questionnaire. In addition, images of old male agents were preferred least. The preference for images of young agents could be a result of these agents being perceived as having more expertise, as explained in the next section.

4.2. The age, gender and role of the agent image affect the perceived agent characteristics and likeliness of following advice

Differences in perception of the agent designs that differed in age, gender and role exist. Firstly, young agents were rated higher on friendliness and involvement compared to old agents by both the general and elderly population and were rated higher on expertise and reliability by the general population, whereas the old agents were seen as more authoritative by both the general and elderly population. These different perceptions might be explained by The Stereotype Content Model (Fiske, Cuddy, Glick, & Xu, 2002). The model defines two fundamental dimensions of social perception, namely warmth (communion) and competence (agency). Older people are perceived as low in competence, whereas young people are perceived as high in competence (Andreoletti, Leszczynski, & Disch, 2015; Cuddy, Fiske, & Glick, 2008), explaining why the young agents were rated higher on expertise. This could relate to the higher ratings of reliability. Also, older people are perceived as high in warmth and young people as low in warmth (Andreoletti et al., 2015; Cuddy et al., 2008), which seems not to apply to the agents in our research. Younger agents were namely seen as more friendly and involved, mapping onto the warmth dimension, than older agents. What could explain that the younger agents were rated higher on friendliness and involvement is the childlike qualities of their faces. A study by Berry and McArthur (1986) explains that adults with immature facial qualities are more likely to be perceived as having childlike psychological attributes, which identify infants' helplessness. Adults with various childlike facial qualities are perceived to afford more warmth and submission, which could explain why the younger agents were seen as more friendly. On the contrary, authority tends to be the opposite of submission, which could explain why the younger agents were perceived as less authoritative than the older agents.

Secondly, female agents were perceived as more friendly and involved by the general population (not by the elderly population), whereas male agents were perceived as more authoritative by both the general and elderly population and perceived as having more expertise by the general population. These different perceptions of female and male agents might

again be explained by The Stereotype Content Model. Men are perceived as high in competence and low in warmth – being envious –, whereas traditional women (housewives) are perceived as low in competence and high in warmth – being paternalistic (Cuddy et al., 2008; Fiske et al., 2002). Similarly, traits implying hardness (e.g. dominating) are assigned to men and soft traits (e.g. being submissive) are assigned to females (Ashmore & Tuma, 1980) and men are perceived as more assertive, having a higher self-esteem, whereas females are seen as more trustworthy and tender-mindedness (Feingold, 1994). Referring back to the Stereotype Content Model, expertise and authority might relate to the competence dimension, whereas friendliness and involvement might relate to the warmth dimension, explaining the different perceptions of the characteristics of the female and male agents.

Thirdly, respondents of both the general and elderly population rated expert agents higher on expertise, reliability and authority than peer agents. In addition, the elderly population rated the expert agents also higher on involvement. These results might relate to the results of research by Veletsianos (2010). Veletsianos researched whether learners stereotype a pedagogical agent as being knowledgeable or not knowledgeable based on its appearance (either having the appearance of a scientist or an artist) and the type of tutorial to be presented (nanotechnology or punk rock). Indeed, the agents were stereotyped on the basis of their appearance. Similarly, our expert agents (wearing a doctor's coat) could have been stereotyped as being more knowledgeable – having more expertise in the health domain – than the peer agents. The perception of experts having more expertise could have affected the perceived level of authority, reliability and involvement of the agents.

In addition to rating the agents' characteristics in the online questionnaire, the respondents indicated how likely they are to follow the advice of the agents. Respondents of the general population indicated they are more likely to follow the advice of young and expert agents, whereas no difference between female and male agents existed. The elderly population indicated only to be more likely to follow advice of experts. This finding might relate to young and expert agents being rated higher on expertise and reliability. As a next step, these agents could provide health coaching. The user's likeliness of following advice could be an important factor when providing health coaching. When users indicate they are more likely of following the agent's advice, their actual behavior might be positively affected: they might follow the agent's advice more often. Using a young, expert agent might thus be a good start for agents that provide coaching advice.

4.3. The preferred agent design relates to the respondent's age and gender

Relations between the respondents' characteristics and design features of the preferred agent exist. For both the general and elderly population, young respondents are more likely to select an image of a young agent and males are more likely to select an image of a male agent. This result could be explained by the tendency to prefer agents that look similar

to ourselves, (Bailenson et al., 2008; Baylor, 2009; Baylor & Kim, 2004; Guadagno et al., 2007; Kim et al., 2007). Finally, the focus groups showed that elderly believe that dynamic design features of the agent, such as humor, friendliness, expertise, involvement and reliability are more important than the agent's static design features, among which age, gender and role. However, the outcomes of the online questionnaire indicate that the static design features do affect the perception of the agent. In addition, some participants in the focus group indicated they prefer a young agent, which contradicts the findings of the questionnaire.

4.4. Limitations research

Our results show that the perceptions of characteristics of agent designs relate to the implementation of design features of these agents. However, to value differences in perception of specific characteristics among agents, it would be relevant to know how these characteristics are valued in the first place. By researching the "absolute importance" of the agent characteristics, we could estimate the relative importance of the different characteristics to each other and better value the differences in these perceptions.

In addition, we compared the perceptions of both a general and elderly population. Although, the sample sizes of both groups were quite large, the sample size of the elderly population was smaller than the sample size of the general population, which might have caused differences in results between both groups.

Another factor that might have influenced the results of our research is that the agents that differed on one of the three design features (either age, role or gender) indirectly differed on multiple design features. Design features, such as hair color or the presence of glasses, were not constant over the designs, and, therefore, might have affected the frequencies of the preferred agent designs and the ratings of the agents' characteristics. Instead of, for example, the old female peer agents representing the young female peer agents at a later stage in their life, the agents were designed as eight individuals to prevent any bias in the results by respondents linking certain agents to each other.

This research solely focused on the first impression of static images representing agents. Future work will include research on whether the differences in first impressions are still present after people actually interact with the agents.

4.5. Implications research

Although this research focused on first impressions of static agent images – images representing agents, but without interaction between the agent and the user –, the results of our research might be translated to embodied conversational agents, or just conversational agents, such as chat bots. What can be learned from this research is that:

- *The agent design features – age, gender and role – affect the perception of the characteristics of a static agent image at first glance*; existing research confirms this finding (Bergmann et al., 2012). Therefore, considering

these design features when developing an (embodied) conversational agent can be beneficial.

- In addition, particular user characteristics, such as age and gender, affect the first impressions we have of the characteristics of a static agent image; *personalization of agent designs matters*. Adapting the agent design to the user could optimize the user's first impression and result in a positive start of the interaction with an (embodied) conversational agent.

Although this study only focused on eight different static agents designs differing on three design features and designed for a health care context, our results may be generalized toward other contexts.

5. Conclusion

The static design of an ECA does affect people's first impression of the agent image. Both a general and elderly population seem to prefer images of young, female agents over old male agents. Furthermore, the agent's design features affect people's first impressions of the characteristics friendliness, expertise, reliability, involvement and authority of the agent image and the likeliness to follow the agent's advice. The majority of the characteristics of these different agent images are perceived similar by the general and elderly population, whereas also some differences in perception exist. Adapting the design of an agent to the user might thus be beneficial. People seem to perceive the level of friendliness, expertise, reliability, involvement and authority of the agent images that differ in age, gender and role differently. In addition, relations between people's characteristics – age and gender – and design features of their preferred agent design exist. In addition, we have shown that the design features age, gender and role influence people's first impression of the agent. These findings provide a motivation to test the effect of other design features on people's first impressions as well. As an example, the implementation of humor, friendliness, expertise, involvement and reliability could be tested, since these characteristics were seen as important by the elderly in the focus groups.

These findings indicate that carefully designing the appearance of the agent is of value for establishing user engagement in eHealth applications. Adapting the agent design to the personal characteristics of the user influences the user's first impression of the agent, of which the interaction between the user and agent could benefit later on. Whereas our research focused on the eHealth domain, the findings might also be applicable to other contexts, such as education.

Funding

This work was supported by the European Commission's Horizon 2020 Research and Innovation Programme projects GOAL [Grant Agreement Number 731656] and Council of Coaches [Grant Agreement Number 769553].

ORCID

Silke ter Stal  <http://orcid.org/0000-0001-9458-717X>
 Monique Tabak  <http://orcid.org/0000-0001-5082-1112>
 Harm op den Akker  <http://orcid.org/0000-0001-6312-6063>

Tessa Beinema  <http://orcid.org/0000-0003-3513-0641>
 Hermie Hermens  <http://orcid.org/0000-0002-3065-3876>

References

- Alsharbi, B., & Richards, D. (2017). Using virtual reality technology to improve reality for young people with chronic health conditions. In *Proceedings of the 9th international conference on computer and automation engineering* (pp. 11–15). ACM. doi:10.1145/3057039.3057080
- Andreoletti, C., Leszczynski, J. P., & Disch, W. B. (2015). Gender, race, and age: The content of compound stereotypes across the life span. *International Journal of Aging and Human Development*, 81(1–2), 27–53. doi:10.1177/0091415015616395
- Ashmore, R. D., & Tumia, M. L. (1980). Sex stereotypes and implicit personality theory. I. A personality description approach to the assessment of sex stereotypes. *Sex Roles*, 6(4), 501–518. doi:10.1007/BF00287882
- Bailenson, J. N., Blascovich, J., & Guadagno, R. E. (2008, November). Self-representations in immersive virtual environments. *Journal of Applied Social Psychology*, 38(11), 2673–2690. doi:10.1111/j.1559-1816.2008.00409.x
- Bar, M., Neta, M., & Linz, H. (2006). Very first impressions. *Emotion*, 6(2), 269–278. doi:10.1037/1528-3542.6.2.269
- Baylor, A. L. (2009). Promoting motivation with virtual agents and avatars: Role of visual presence and appearance. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1535), 3559–3565. doi:10.1098/rstb.2009.0148
- Baylor, A. L., & Kim, Y. (2004). Pedagogical agent design: The impact of agent realism, gender, ethnicity, and instructional role. In *International conference on intelligent tutoring systems* (pp. 592–603). Berlin, Heidelberg: Springer. doi:10.1007/978-3-540-30139-4\56
- Bergmann, K., Eyssel, F., & Kopp, S. (2012). A second chance to make a first impression? How appearance and nonverbal behavior affect perceived warmth and competence of virtual agents over time. In *International conference on intelligent virtual agents* (pp. 126–138). Berlin, Heidelberg: Springer. doi:10.1007/978-3-642-33197-8\13
- Berry, D. S., & McArthur, L. Z. (1986). Perceiving character in faces. The impact of age-related craniofacial changes on social perception. *Psychological Bulletin*, 100(1), 3–18. doi:10.1037/0033-2909.100.1.3
- Bickmore, T. W., Pfeifer, L. M., & Jack, B. W. (2009). Taking the time to care: Empowering low health literacy hospital patients with virtual nurse agents. In *27th international conference on human factors in computing systems* (pp. 1265–1274). ACM Press. doi:10.1145/1518701.1518891
- Cafaro, A., Vilhjálmsón, H. H., & Bickmore, T. (2016). First impressions in human-agent virtual encounters. *ACM Transactions on Computer-Human Interaction*, 23(4), 1–40. doi:10.1145/2940325
- Cafaro, A., Vilhjálmsón, H. H., Bickmore, T., Heylen, D., Jóhannsdóttir, K. R., & Valgarðsson, G. S. (2012). First impressions: Users' judgments of virtual agents' personality and interpersonal attitude in first encounters. In *International conference on intelligent virtual agents* (pp. 67–80). doi:10.1007/978-3-642-33197-8\7
- Clark, H. H., & Brennan, E. S. (1991). Grounding in communication. *Perspectives on Socially Shared Cognition*, 13, 222–233. doi:10.1037/10096-006
- Cowell, A., & Stanney, K. (2003). Embodiment and interaction guidelines for designing credible, trustworthy embodied conversational agents. In *International workshop on intelligent virtual agents* (pp. 301–309). Berlin, Heidelberg: Springer. doi:10.1007/978-3-540-39396-2\50
- Cuddy, A. J., Fiske, S. T., & Glick, P. (2008). Warmth and competence as universal dimensions of social perception: The stereotype content model and the BIAS map. *Advances in Experimental Social Psychology*, 40(07), 61–149. doi:10.1016/S0065-2601(07)00002-0
- Feingold, A. (1994). Gender differences in personality. *Psychological Bulletin*, 116(3), 429–456. doi:10.1176/ajp.152.4.579
- Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, 82(6), 878–902. doi:10.4324/9781315187280
- Fogg, B. (2011). Computers as persuasive social actors. *Persuasive Technology*, 89–120. doi:10.1016/B978-155860643-2/50007-X.
- Forlizzi, J., Zimmerman, J., Mancuso, V., & Kwak, S. (2007). How interface agents affect interaction between humans and computers. In *Proceedings of the 2007 Conference on Designing Pleasurable Products and Interfaces* (pp. 209–221). doi:10.1145/1314161.1314180
- Guadagno, R., Blascovich, J., Bailenson, J., & McCall, C. (2007). Virtual humans and persuasion: The effects of agency and behavioral realism. *Media Psychology*, 10, 1–22. doi:10.108/15213260701300865
- Kaptein, M., De Ruyter, B., Markopoulos, P., & Aarts, E. (2012). Adaptive persuasive systems. *ACM Transactions on Interactive Intelligent Systems*, 2(2), 1–25. doi:10.1145/2209310.2209313
- Kim, Y., Baylor, A. L., & Reed, G. (2003). The impact of image and voice of pedagogical agents. In *E-learn 2003 - world conference on e-learning in corporate, government, healthcare, and higher education* (pp. 2237–2240). Berlin, Heidelberg: Springer. doi:10.1007/978-3-540-30139-4\56
- Kim, Y., Baylor, A. L., & Shen, E. (2007). Pedagogical agents as learning companions: The impact of agent emotion and gender. *Journal of Computer Assisted Learning*, 23(3), 220–234. doi:10.1111/j.1365-2729.2006.00210.x
- Kreps, G. L., & Neuhauser, L. (2010, March). New directions in eHealth communication: Opportunities and challenges. *Patient Education and Counseling*, 78(3), 329–336. doi:10.1016/j.pec.2010.01.013
- Lee, Y.-H., Xiao, M., & Wells, R. H. (2018). The effects of avatars' age on older adults' self-disclosure and trust. *Cyberpsychology, Behavior, and Social Networking*, 21(3), 173–178. doi:10.1089/cyber.2017.0451
- Mohr, D. C., Cuijpers, P., & Lehman, K. (2011). Supportive accountability: A model for providing human support to enhance adherence to eHealth interventions. *Journal of Medical Internet Research*, 13(1), 1–11. doi:10.2196/jmir.1602
- Nguyen, H., & Masthoff, J. (2009). Designing empathic computers: The effect of multi-modal empathic feedback using animated agent. In *Proceedings of the 4th international conference on persuasive technology*. ACM. doi:10.1145/1541948.1541958
- Nijland, N. (2011). *Grounding eHealth* (Doctoral dissertation). University of Twente, Enschede, The Netherlands. doi:10.3990/1.9789036531337
- Nunamaker, J. F., Derrick, D. C., Elkins, A. C., Burgoon, J. K., & Patton, M. W. (2011). Embodied conversational agent-nased kiosk for automated interviewing. *Journal of Management Information Systems*, 28(1), 17–48. doi:10.2753/mis0742-1222280102
- Oinas-Kukkonen, H., & Harjuma, M. (2009). Persuasive systems design: Key issues, process model, and system features. *Persuasive Systems Design*, 24(28), 485–500. doi:10.17705/1cais.02428
- Paap, D., Schrier, E., & Dijkstra, P. U. (2018). Development and validation of the working alliance inventory Dutch version for use in rehabilitation setting. *Physiotherapy Theory and Practice*, 1–12. doi:10.1080/09593985.2018.1471112
- Parmar, D., Olafsson, S., Utami, D., & Bickmore, T. (2018). Looking the part: The effect of attire and setting on perceptions of a virtual health counselor. In *Proceedings of the 18th international conference on intelligent virtual agents* (pp. 301–306). ACM. doi:10.1145/3267851.3267915
- Pearson, S. D., & Raeke, L. H. (2000). Patients' trust in physicians many theories, few measures, and little data. *Journal of General Internal Medicine*, 15(7), 509–513. doi:10.1046/j.1525-1497.2000.11002.x
- Pope, C., & Mays, N. (2006). Qualitative methods in health research. In *Qualitative research in health care* (pp. 1–11). doi:10.1002/9780470750841.ch1
- Ridd, M., Shaw, A., Lewis, G., & Salisbury, C. (2009). The patient-doctor relationship: A synthesis of the qualitative literature on patients' perspectives. *British Journal of General Practice*, 59(561), 268–275. doi:10.3399/bjgp09X420248
- Rosenberg-Kima, R. B., Baylor, A. L., Plant, E. A., & Doerr, C. E. (2008). Interface agents as social models for female students: The effects of agent visual presence and appearance on female students' attitudes and beliefs. *Computers in Human Behavior*, 24(6), 2741–2756. doi:10.1016/j.chb.2008.03.017
- Ruttkay, Z., Dormann, C., & Noot, H. (2004). Embodied conversational agents on a common ground: A framework for design and evaluation.

- From Brows to Trust: Evaluating Embodied Conversational Agents*, 27–66. doi:10.1007/1-4020-2730-3{_}2
- Silvervarg, A., Raukola, K., Haake, M., & Gulz, A. (2012). The effect of visual gender on abuse in conversation with ECAs. In *International conference on intelligent virtual agents* (pp. 153–160). Berlin, Heidelberg: Springer. doi:10.1007/978-3-642-33197-8{_}16
- van Velsen, L., Wildevuur, S., Flierman, I., van Schooten, B., Tabak, M., & Hermens, H. (2016). Trust in telemedicine portals for rehabilitation care: An exploratory focus group study with patients and healthcare professionals. *BMC Medical Informatics and Decision Making*, 16(11), 1–12. doi:10.1186/s12911-016-0250-2
- Vardoulakis, L. P., Ring, L., Barry, B., Sidner, C. L., & Bickmore, T. (2012). Designing relational agents as long term social companions for older adults. In *International conference on intelligent virtual agents* (pp. 289–302). Berlin, Heidelberg: Springer. doi:10.1007/978-3-642-33197-8{_}30
- Vartanian, O., Stewart, K., Mandel, D. R., Pavlovic, N., McLellan, L., & Taylor, P. J. (2012). Personality assessment and behavioral prediction at first impression. *Personality and Individual Differences*, 52(3), 250–254. doi:10.1016/j.paid.2011.05.024
- Veletsianos, G. (2010). Contextually relevant pedagogical agents: Visual appearance, stereotypes, and first impressions and their impact on learning. *Computers and Education*, 55(2), 576–585. doi:10.1016/j.compedu.2010.02.019
- Wissen, A. V., Vinkers, C., & Halteren, A. V. (2016). Developing a virtual coach for chronic patients: A user study on the impact of similarity, familiarity and realism. In *International conference on persuasive technology* (pp. 263–275). Cham: Springer. doi:10.1007/978-3-319-31510-2{_}23
- Zimmerman, J., Ayoob, E., Forlizzi, J., & McQuaid, M. (2005). Putting a face on embodied interface agents. In *Designing pleasurable products and interfaces conference proceedings* (pp. 233–248). doi:10.1184/R1/6470366.v1

About the Authors

Silke ter Stal is a junior researcher and PhD candidate at Roessingh Research and Development. She has a background in Creative Technology and Human Media Interaction at the University of Twente. Silke currently researches the design of embodied conversational agents for eHealth in several European and national research projects.

Monique Tabak is Associate professor at the University of Twente and Roessingh Research and Development, focusing her research on Personalized eHealth technology for rehabilitation. She has a background in Biomedical Engineering and obtained her PhD degree on promoting physical activity behavior in COPD. She participated in >20 (inter)national research projects.

Harm op den Akker is a senior researcher at Roessingh Research and Development. Harm has a background in computer science and a PhD degree from the University of Twente. Harm is currently coordinating the European Project “Council of Coaches” and has worked in various European and national research before that.

Tessa Beinema is a junior researcher at Roessingh Research and Development with a background in artificial intelligence. She currently researches how to select personalized coaching topics that are to be discussed by virtual coaches in eHealth applications.

Hermie Hermens is professor Telemedicine at the University Twente, director Personalized eHealth Technology Program and director Telemedicine at Roessingh Research and Development. His work focusing on remote monitoring and smart coaching technology has been cited over 19000 times (H-index 63). He participated in over 25 European projects, coordinated 4 of them.