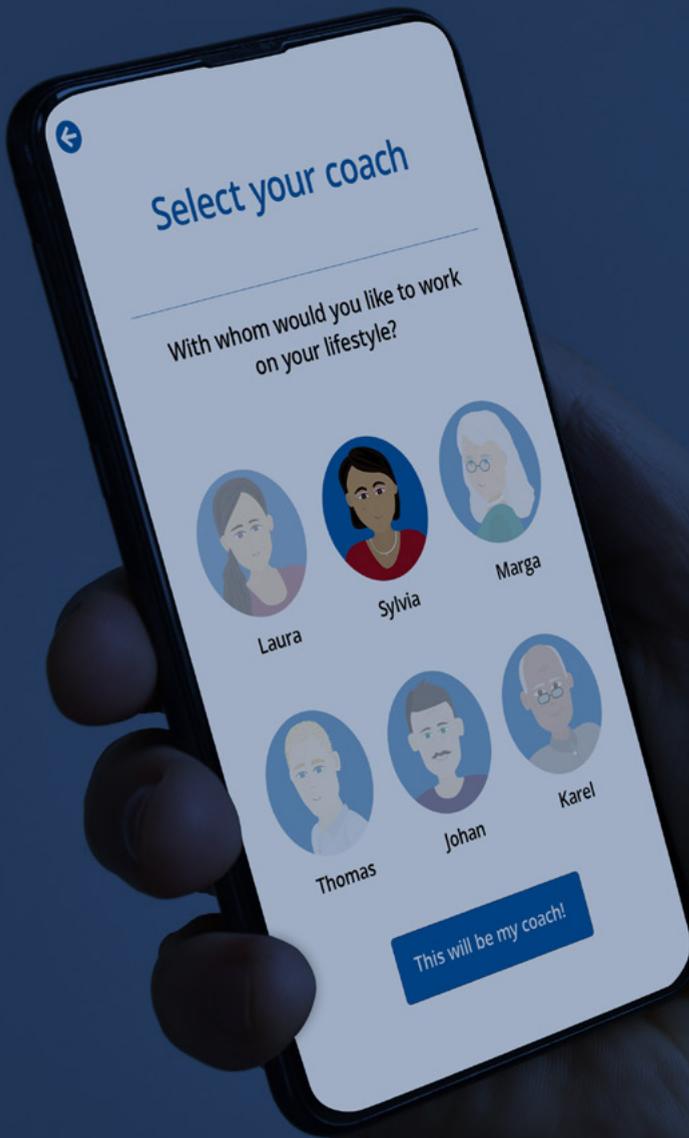


Look Who's Talking

*Appearance of Embodied
Conversational Agents in eHealth*



LOOK WHO'S TALKING

APPEARANCE OF EMBODIED CONVERSATIONAL AGENTS IN EHEALTH

Silke ter Stal

The publication of this thesis was financially supported by:



Cover design: Silke ter Stal

Printed by: Ipskamp Printing

Lay-out: Silke ter Stal

ISBN: 978-90-365-5126-7

DOI: 10.3990/1.9789036551267

© **Silke ter Stal, Enschede, the Netherlands, 2021**

All rights reserved. No parts of this thesis may be reproduced, stored in a retrieval system or transmitted in any form or by any means without permission of the author. Alle rechten voorbehouden. Niets uit deze uitgave mag worden vermenigvuldigd, in enige vorm of op enige wijze, zonder voorafgaande schriftelijke toestemming van de auteur.

LOOK WHO'S TALKING

APPEARANCE OF EMBODIED CONVERSATIONAL AGENTS IN EHEALTH

DISSERTATION

to obtain
the degree of doctor at the University of Twente,
on the authority of the rector magnificus,
prof. dr. ir. A. Veldkamp,
on account of the decision of the Doctorate Board,
to be publicly defended
on Wednesday the 31st March 2021 at 14.45 hours

by

Silke ter Stal
born on the 31st of May, 1992
in Enschede, the Netherlands

THIS DISSERTATION HAS BEEN APPROVED BY

Supervisor

prof. dr. ir. H. J. Hermens

Co-supervisor

dr. ir. M. Tabak

GRADUATION COMMITTEE

Chairman/secretary

prof. dr. J. N. Kok | *University of Twente*

Supervisor

prof. dr. ir. H. J. Hermens | *University of Twente, Roessingh Research and Development*

Co-supervisor

dr. ir. M. Tabak | *University of Twente, Roessingh Research and Development*

Committee Members

prof. dr. ir. G. D. S. Ludden | *University of Twente*

prof. dr. D. K. J. Heylen | *University of Twente*

prof. dr. ir. J. F. M. Masthoff | *Utrecht University*

prof. dr. C. Pelachaud | *Sorbonne University*

Welcome! Nice to meet you, my name is Sylvia. I will be guiding you throughout this thesis. I will provide you with a quick summary at the beginning of every chapter. I must say, I found it very interesting. I hope you enjoy reading it too. See you soon!



CONTENTS

01	General Introduction	9
02	Design Features of Embodied Conversational Agents (ECAs) in eHealth: a Literature Review	23
03	Effects of ECA Age, Gender and Role on Users' Impressions at First Glance	43
04	Effects of ECA Age and Gender on Users' Impressions after Short Interaction	65
05	Effects of ECA Emotion on Users' Perceptions of Rapport after Short Interaction	85
06	ECA Appearance in a Long-term, Daily Life Setting for Self-management	101
07	Effects of ECA Age and Gender in a Multi-Agent Application in a Long-term, Daily Life Setting	123
08	State-of-the-Art and Design Strategies for ECA Appearance	139
09	General Discussion	169
10	Appendices References Summary Samenvatting Dankwoord About the Author Progress Range	181

01

General Introduction

THE FUTURE OF HEALTH CARE: FROM CURE TO CARE TO COACHING

In 2050, the number of persons aged 60 or above worldwide is expected to be double compared to 2017 (*U.N. Department of Economic and Social Affairs, 2017*). Although we live longer, we do not necessarily become healthier. As an example, it is expected that by 2030, nearly 40% of the Dutch population – 7 million people – will have a chronic condition, such as heart disease, cancer and diabetes (*RIVM - National Institute for Public Health and the Environment, 2014*). This percentage is similar in other European countries. In addition, ageing is associated with an increased risk of multi-morbidity: experiencing more than one chronic condition at the same time (*World Health Organization, 2015*). Multi-morbidity further increases the patient burden, mortality and health care costs (*World Health Organization, 2015; van Boven, 2017; Chen et al., 2017*), adds to the complexity of care and impacts health care consumption (*Vanfleteren et al., 2016*). Importantly, chronic diseases are linked to major behavioural risk factors, such as unhealthy diet, physical inactivity and tobacco use (*World Health Organization, 2015*). Thus, the focus of our health care has to shift from curing acute complaints to (secondary) prevention of complaints via long-term care and lifestyle coaching.

Via this long-term care, health complaints can be prevented (to worsen) by supporting people in having a healthy lifestyle. Aspects of a healthy lifestyle are, for example, physical activity and healthy nutrition. Physical activity, defined as ‘any bodily movement produced by skeletal muscles that requires energy expenditure’ (*World Health Organization, 2010*), includes activities undertaken while working, playing, carrying out household chores, travelling, and engaging in recreational pursuits. Physical activity improves a person’s physical health (e.g. it maintains muscle strength), cognition (e.g. it reduces anxiety and depression), and social health (e.g. it increases community involvement). Moreover, being physically active reduces the risk of, for example, heart diseases, diabetes and stroke (*World Health Organization, 2015*). Healthy nutrition consists of a balanced calorie intake and expenditure and a limited salt and free sugar intake (*World Health Organization, 2020*). Like physical activity, healthy nutrition reduces the risk of many chronic diseases, including diabetes, heart disease, stroke and cancer (*World Health Organization, 2020*).

Several initiatives exist to support people in developing and maintaining a healthy lifestyle. For example, in the Netherlands, combined lifestyle intervention (CLI) programs are offered for overweight people showing positive effects of CLI on participants’ lifestyles (*Preller et al., 2011*). A CLI program consists of advice and support for adopting a healthier diet, adopting eating habits, getting more physical exercise and achieving behavioural change (*Zorginstituut Nederland, n.d.*). It is offered by a multidisciplinary team of health care providers, involving lifestyle advisors, physiotherapists and dieticians.

Yet, insufficient health care professionals will be available to provide this personalised coaching in the future. To be able to keep helping all persons in need for care, alternative ways to provide coaching are investigated. A solution widely investigated are eHealth applications (*Craig & Patterson, 2005; Heinzelmann et al., 2005; Kreps & Neuhauser, 2010; World Health Organization, 2012*). eHealth, often also referred to as telemedicine or tele-health (*Fatehi & Wootton, 2012*), is defined in many different ways. *Pagliari et al.* updated the definition by *Eysenbach* and define eHealth as:



e-health is an emerging field of medical informatics, referring to the organisation and delivery of health services and information using the Internet and related technologies. In a broader sense, the term characterises not only a technical development, but also a new way of working, an attitude, and a commitment for networked, global thinking, to improve health care locally, regionally, and worldwide by using information and communication technology.

– Pagliari et al. (adapted from Eysenbach) –

The use of eHealth for lifestyle coaching is getting more and more common. But, for eHealth applications to be effective, use of the applications is crucial. In that sense, engagement with an eHealth application is necessary; those who are more engaged are significantly less likely to stop using it (Yardley et al., 2016; Kohl et al., 2013; Scherer et al., 2017; Crutzen et al., 2011). Yet, in the beginning, users often show an increased interest in interacting with the technology, since it is new, but this interest usually fades after two weeks of use (Nijland, 2011).

To fill in the gap of long-term interactions, virtual health assistants arise. Virtual assistants are available 24/7 and can promote engagement over hundreds, if not thousands, of interactions. From voice assistants controlling our smart homes, such as Amazon's Alexa¹ and Apple's Siri², till chatbots and robots; virtual assistants start playing the role of counsellors, coaches and educators in our daily lives. These virtual assistants also arise in health care. Examples of chatbots in health care are Lark³ – a chatbot that supports chronic disease management – and Woebot⁴ – an intelligent and emphatic chatbot that delivers mental health therapy programs. Examples of robots used in health care are Pepper⁵ – A child-size robot with tablet display, not specifically designed for health care, but used, for example, for health data acquisition among older adults (Boumans et al., 2019) – and Tessa⁶ – a tabletop model robot in the shape of a flower pot that supports cognitive impaired with auditory reminders. Within fifteen years, will we all have such a coach accompanying us to support our health and well-being?

1 <https://developer.amazon.com/alexa>

2 <https://www.apple.com/siri>

3 <https://www.lark.com>

4 <https://woebothealth.com>

5 <https://www.softbankrobotics.com/emea/en/pepper>

6 <https://www.tinybots.nl>



May 15th, 2035. It is 7.30h, you are still asleep. On your bedside table is your virtual coach Sylvia – a 3D, holographic virtual character. While you are still enjoying your sleep, Sylvia gently opens the curtains in your room, allowing the first sunlight of the day to enter. She wakes you up softly:



Good morning. It is 7.30h, time to wake up.

After turning around once more, you get out of bed. You greet Sylvia:

Good morning, Sylvia. What's the weather going to be today?



I just checked for you. It will be a beautiful spring day, with a bright blue sky and a temperature of around 20 degrees Celsius.

You smile, then take a refreshing shower and put on a jeans and a polo. Then, you are headed to the kitchen, where Sylvia already awaits you at the counter:



How are you feeling today?

Thank you for asking, sadly, I am having a slight headache.



That is sad to hear, I can imagine you have had better mornings.

You do not really have an appetite this morning, but Sylvia reminds you of how important it is to eat breakfast:

You are probably not that hungry, but you might want to eat something small. What about adding a piece of fruit to your breakfast? There are still two of those lovely kiwis that you like in the refrigerator.



You make yourself a cup of tea, a cracker and grab one of the kiwis. You recall you have to take your medication as well. Sylvia had a similar thought:

Oh, and let me remind you that it is time for your medication.



You thank Sylvia, take your medication and start eating your breakfast. During breakfast, you ask Sylvia whether there is anything on your schedule today.

Sylvia, what is on my schedule today?

You do not have anything on your schedule today. Shall we go for a walk and enjoy the lovely weather? A bit of fresh air might help to soothe your headache.



It did not cross your mind, but you accept Sylvia's suggestion:

Actually, that might be a good idea.

You make a quick stop at the toilet, take your jacket and go outside. Sylvia pops-up at your smart watch:



So, where are we going?

I was thinking we could make a small round through the park.



Excellent choice. Let's go there. I have selected some nice routes for you.

You look at the routes Sylvia suggested at your smart watch and select one of them. You hear the birds chirping, you smell the freshly mown grass and you feel the warmth of the sun on your skin. There are quite some other people in the park. You friendly greet them. You forgot about your headache. After twenty minutes, you feel the need for a small stop. You apologise towards Sylvia:

Sylvia, sorry, I have to sit down for a minute.



No need to apologise. I can use some rest too.

You sit down on a wooden bench, observing the playing children, fanatic joggers and chatting families.

What a beautiful day, isn't it?



Definitely.

PROMOTING ENGAGEMENT BY EMBODIED CONVERSATIONAL AGENTS

Sylvia is a virtual health coach. She is not a chatbot, voice assistant or robot, but an Embodied Conversational Agent (ECA) – also known as a virtual human, animated character, intelligent virtual agent or relational agent. ECAs are defined as more or less autonomous and intelligent software entities with an embodiment used to communicate with the user (*Ruttkay et al., 2004*). ECAs go beyond chatbots and voice assistants that mainly communicate with the user via text or speech respectively, by having an additional layer for communication: their embodiment. Communication via embodiment involves, for example, facial expressions and hand and body gestures. In this way, ECAs are similar to robots, only having a virtual embodiment instead of physical one. Two examples of an ECA are Laura and Molly (*Figure 01.1*). Laura is an early ECA, a virtual exercise advisor that promotes walking behaviour, created by the Relational Agents Group (*Bickmore et al., 2005*). A more recent and commercial ECA is Molly, created by Sensely⁷. Molly is a virtual assistant that can perform symptom assessment, provide health information and support users having chronic conditions, such as chronic heart failure and diabetes.

ECAs could contribute to user engagement in several ways. First, ECAs can contribute to engagement by creating user experiences that are more fun, absorbing and intrinsically enjoyable, since ECAs are a form of interactive multi-media (*Lefebvre et al., 2010; Yardley et al., 2016*). Second, ECAs can contribute to engagement by providing users with social support, which is one of most important persuasive drivers in eHealth (*Kelders et al., 2012*). ECAs can provide this social support by building trust and rapport (e.g. mutual understanding) with the user, leading to companionship. Such a bonding companionship could help to maintain a user's engagement for applications offering long-term care and coaching (*Kelders et al., 2012*).

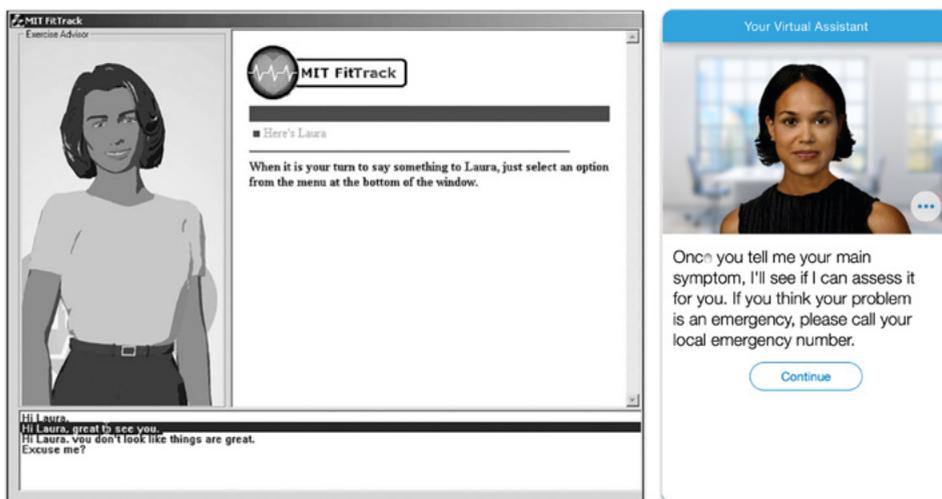


Figure 01.1 – Two ECA examples. Left: Laura, an early ECA for walking promotion (picture from *Bickmore et al., 2005*). Right: Molly, an ECA performing symptom assessment (picture from Sensely).

7 <https://www.sensely.com>

In literature, no clear definition of engagement exist. In the context of interaction with technology, *O'Brien and Toms (2008)* define engagement as a quality of user experience characterised by attributes of challenge, positive affect, endurance, aesthetic and sensory appeal, attention, feedback, variety/novelty, interactivity, and perceived user control.

In this thesis, we are interested in engagement with technology in the context of personalised long-term care, in particular, support for the adoption of healthy behaviours. In the context of such digital behaviour change interventions (DBCIs), *Yardley et al. (2016)* differentiate two types of engagement: they differentiate engagement with the DBCI from engagement with behaviour change itself. *Cole-Lewis et al. (2019)* make a similar distinction. They distinguish health behaviour engagement, referred to as 'Big E', from DBCI engagement, referred to as 'Little e'. In addition, *Cole-Lewis et al.* go one step further than *Yardley et al.* by splitting DBCI engagement further into two subclasses: A) user interactions with features of the intervention designed to encourage frequency of use (i.e. simple login, games, and social interactions) and make the user experience appealing, and B) user interactions with behaviour change intervention components (i.e. behaviour change techniques). We took this characterisation of engagement with DBCIs by *Cole-Lewis et al. (2019)* as a guideline for engagement described in this thesis. We positioned an ECA that supports users in adopting healthy behaviours as a feature to encourage use and user experience (subclass A), as seen in **Figure 01.2**.

Furthermore, **Figure 01.2** visualises phases of engagement in the context of DBCIs. As explained by *O'Brien and Toms (2008)*, engagement is a process consisting of four stages, namely the **point of engagement, the period of sustained engagement, dis-engagement, and re-engagement**. *Yardley et al.* specify similar phases of engagement with DBCIs specifically. They illustrate the following four phases: 1) engagement with the DBCI only (preparation for behaviour change), 2) engagement with behaviour change, mediated by DBCI, 3) DBCI usage no longer required for maintenance of engagement with behaviour change and 4) re-engagement with DBCI if needed (problem solving, relapse management). For this thesis, we slightly adapted this specification of phases of engagement by *Yardley et al. (2016)*. We added a t_0 , which represents the users' perception at first glance, as first impressions of an ECA are an important determiner for whether a person continues interacting with an ECA (*Bergmann et al., 2012*). For clarity, we named phases 1 to 4 short-term, long-term, maintenance and relapse respectively.

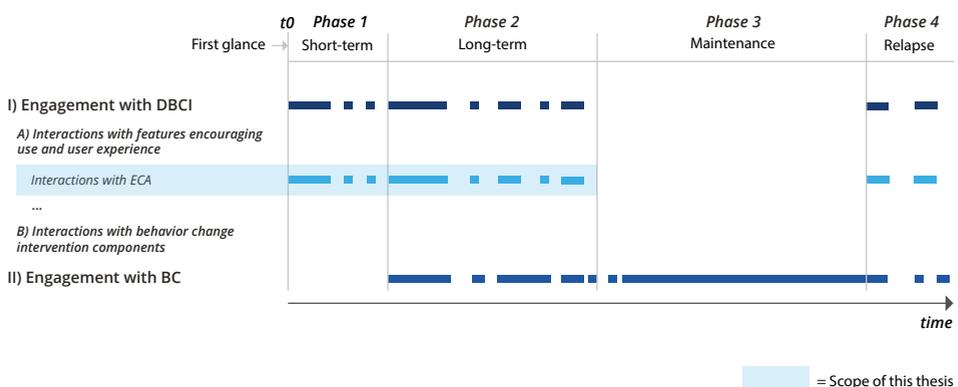


Figure 01.2 – Our illustration of the phases of engagement, adapted from Yardley et al. We position interactions with an ECA as a feature to encourage use and user experience to promote engagement with the DBCI. We define one point in time (first glance) and named the four phases: short-term, long-term, maintenance and relapse.

APPEARANCE OF ECAS IN EHEALTH: A LACK OF GUIDELINES

Incorporating ECAs into eHealth applications could be a means to promote user engagement. But how should we design these ECAs to promote this user engagement? The design of an ECA incorporates many different aspects, such as design of the content of the ECA's messages and the way of interacting with the ECA. Another aspect that should be designed is an ECA's appearance. As in human-human interaction, an ECA's appearance affects the impressions we have of this ECA (*Bergmann et al., 2012; Kelley, 1950*). When we interact with another person, or ECA, for the first time, we immediately form initial ideas about the other. For example, within milliseconds, we judge how friendly or competent someone is. Even one step further, people that have a positive first impression of a person, tend to interact more with that person (*Kelley, 1950*). This might also be true for human-agent interaction (*Bergmann et al., 2012*). Thus, if we want to establish and maintain user engagement with eHealth applications over time, by stimulating users to continue interacting with an ECA, we have to design an ECA's appearance such that it positively affects users' impressions of this ECA. Yet, how do we design an ECA's appearance to reach this?

Some research has been performed on developing a set of ECA design features. For example, *Ruttkey et al.* developed a taxonomy for relevant design and evaluation aspects of ECAs (*Ruttkey et al., 2004*). They identify the following design aspects as part of their taxonomy: an ECA's embodiment – its physical appearances (its looks, speech and/or textual output, hand and body gestures and facial and gaze expressions), mental capacities (its social role, personality, user model, natural language generator and dialogue manager) and the application interface (including background knowledge processing). In addition, *Straßmann and Krämer* identify design features related to the ECA's appearance as applied in prior ECA research (not restricted to a particular domain or effect on particular outcome measures). They categorise the variables: embodiment vs no embodiment, species, realism, 2D vs 3D and feature specification (socio-demographic and styling). Although these taxonomies define ECA design features, they do not show how these ECA design features should actually be designed. Some ECA design guidelines exist in other contexts than eHealth, such as the guidelines for pedagogical agents by *Veletsianos et al.*. They propose a three-tier framework of 15 research-based guidelines, focusing on 1) the user interaction – an ECA should be attentive and sensitive to the learner's needs and wants –, 2) the message – an ECA should consider intricacies of its message –, and 3) the agent's characteristics – an ECA should display socially appropriate demeanour, posture and representation. Yet, these guideline focus little on the ECA's appearance specifically. We conclude that little is known about how an ECA's appearance affects user engagement: a set of design guidelines for an ECA's appearance in eHealth is missing (*Figure 01.3*).

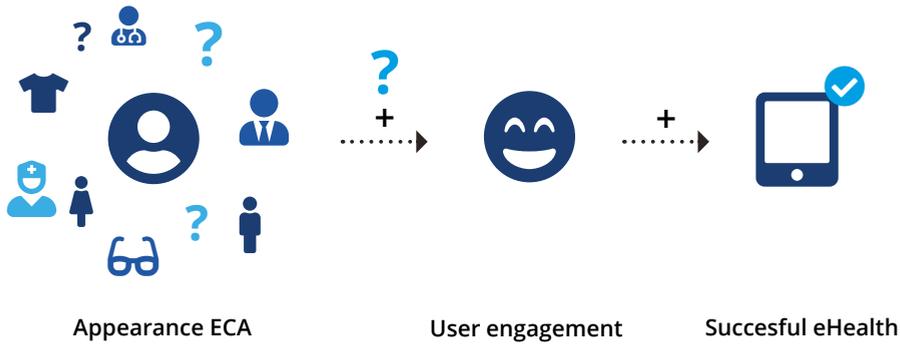


Figure 01.3 – The objective of this thesis: how to design the appearance of an ECA in eHealth to contribute to user engagement – a key to successful eHealth.

THIS THESIS



Main objective

How to design the appearance of an ECA in eHealth to promote user engagement?

By researching how this ECA appearance should be designed to trigger positive impressions in the different phases of engagement, we can stimulate users to continue to a next phase of interaction and eventually reach long-term engagement. Therefore, the sub objectives of this thesis focus on an ECA's appearance in different phases of interaction.



Subobjectives

1. How to design the appearance of an ECA in eHealth for positive impressions at first glance?
2. How to design the appearance of an ECA in eHealth for positive impressions after short interaction?
3. How to design the appearance of an ECA in eHealth for positive impressions after long-term interaction?

Outline Thesis

Figure 01.4 shows how the different chapters of this thesis contribute to answering the main objective.



Figure 01.4 – Outline of this thesis.

In **chapter 2**, we start with a literature review identifying the researched design features for ECAs in eHealth, the outcome variables that were used to measure the effect of these design features and what the found effects for each variable were. Outcomes of this literature review were used to determine the design features to be investigated in the experimental research.

In the experimental research, we first explored users' perceptions of different ECA designs at first glance. **Chapter 3** describes a study in which we investigated the effect of the ECA design features age, gender and role on users' first impressions of ECAs, to gain more insight into what ECA appearance triggers positive impressions at first glance.

Next, we investigated how an ECA appearance can contribute to positive impressions of an ECA after short interaction. **Chapter 4** describes a study that follows from findings of the study presented

in **chapter 3**. This study focuses on the ECA design features age and gender by comparing users' perceptions after short interaction with a young female ECA and an old male ECA for performing health questionnaires. **Chapter 5** describes another study on short interaction. This study focused on the ECA design feature emotion and compares emotion in an ECA's textual expressions and facial expressions.

Eventually, ECAs for eHealth will be used in a long-term, daily life setting. Therefore, **chapter 6 and 7** describe studies in which we evaluated ECA designs in such a long-term, daily life setting, to gain insight into what ECA appearance positively affects users' impressions of an ECA in this setting. The study described in **chapter 6** evaluated the design of an ECA implemented in an eHealth self-management intervention for patients with both Chronic Obstructive Pulmonary Disease (COPD) and Chronic Heart Failure (CHF). Whereas the majority of eHealth applications implement just a single ECA, a new research area focuses on using multiple ECAs providing holistic coaching. Therefore, **chapter 7** focused on ECA design in a multi-agent eHealth application during daily life. In this chapter we specifically researched users' perceptions of the ECA's age and gender.

Finally, we conclude this thesis with a synthesis of the literature review and experimental research in **chapter 8**. We created a set of design strategies for an ECA's appearance in eHealth and showed how one could design an ECA based on these strategies in an implementation example of a mobile physical activity coach.

02

Design Features of Embodied Conversational Agents (ECAs) in eHealth: a Literature Review

BASED ON:

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, 102409. [doi:10.1016/j.ijhcs.2020.102409](https://doi.org/10.1016/j.ijhcs.2020.102409)



Abstract

Embodied conversational agents (ECAs) are gaining interest to elicit user engagement and stimulate actual use of eHealth applications. In this literature review, we identify the researched design features for ECAs in eHealth, the outcome variables that were used to measure the effect of these design features and what the found effects for each variable were. Searches were performed in Scopus, ACM Digital Library, PsychINFO, Pubmed and IEEE Xplore Digital Library, resulting in 1284 identified articles of which 33 articles were included. The agents speech and/or textual output and its facial and gaze expressions were the most common design features. Little research was performed on the agent's looks. The measured effect of these design features was often on the perception of the agent's and user's characteristics, relation with the agent, system usage, intention to use, usability and behaviour change. Results show that emotion and relational behaviour seem to positively affect the perception of the agents characteristics and that relational behaviour also seems to positively affect the relation with the agent, usability and intention to use. However, these design features do not necessarily lead to behaviour change. This review showed that consensus on design features of ECAs in eHealth is far from established. Follow-up research should include more research on the effects of all design features, especially research on the effects in a longterm, daily life setting, and replication of studies on the effects of design features performed in other contexts than eHealth.



To relieve the burden on the healthcare sector caused by the ageing society, the use of eHealth applications is being widely investigated. These applications can be used in establishing a user's behaviour change in daily life either under the supervision of a healthcare professional, or in stand-alone use to promote self-management. Although they seem promising, many eHealth applications face the problem of actual use rapidly decreasing after several weeks (*Nijland, 2011*). Often, existing eHealth applications provide advice in the form of plain text or via a text-based question-answer module (*Kaptein et al., 2012*). Face-to-face interaction remains one of the best ways to communicate health information; it incorporates grounding – dynamically assessing the other persons level of understanding and repeating or elaborating on information when necessary (*Clark and Brennan, 1991*). In addition, face-to-face interaction elicits trust, better communication and satisfaction via both verbal and non-verbal behaviour (*Bickmore et al., 2009b*).

Face-to-face interaction seems to be a possibility to elicit user engagement and stimulate actual use of eHealth applications. Therefore, the use of embodied conversational agents (ECAs) is gaining interest as an alternative means. ECAs are more or less autonomous and intelligent software entities with an embodiment used to communicate with the user (*Ruttkay et al., 2004*). By interacting with the user, ECAs can build trust and rapport, leading to companionship and long-term, continual use (*Vardoulakis et al., 2012*).

ECAs in eHealth: a Lack of Design Guidelines

Although research indicates that incorporating ECAs into eHealth applications could elicit user engagement, little is known about how these agents should be designed in order to accomplish this engagement. Some research on the agent's design has been performed, but no design guidelines exist. A taxonomy of the different design features of ECAs can be essential to establish a common ground for developing design guidelines. *Ruttkay et al.* (*Ruttkay et al., 2004*) created a taxonomy of relevant design and evaluation aspects of ECAs. They distinguish the agent's embodiment (its looks, speech and/or textual output, hand and body gestures and facial and gaze expressions), mental capacities (its social role, personality, user model, natural language generator and dialogue manager) and the application interface (including background knowledge processing). In addition, *Straßmann and Krämer* (*Straßmann and Krämer, 2017*) identify design features related to the agent's appearance. They categorise the variables: embodiment vs no embodiment, species, realism, 2D vs 3D and feature specification (socio-demographic and styling).

Despite the attempts to create a taxonomy of design features, little is known about how these features should actually be designed. Some agent design guidelines exist, such as the design guidelines for pedagogical agents by *Veletsianos et al.* (*Veletsianos et al., 2009*), but these guidelines do not focus on eHealth. Many studies on agent design features with respect to eHealth explore a single design feature (such as an agent's culture background (*Zhou et al., 2017*) and body shape (*van Vugt et al., 2006*)). Findings of these studies provide some direction for the design of an ECA, but were not translated into actual guidelines. Therefore, we conclude that no design guidelines for ECAs in eHealth exist.

A literature review of research on design features for ECAs in eHealth can, therefore, be a valuable input for the development of these guidelines. Such a literature review could provide insight into

how often particular design features have been researched and could draw general conclusions on the effects of particular design features flowing from results of multiple studies. Several literature reviews on conversational agents in eHealth have been performed. However, they either focus on conversational agents in general, not on ECAs specifically (*Laranjo et al., 2018; Rist et al., 2004*), are not up-to-date (*Rist et al., 2004*), focus on a broader context than health (*Scholten et al., 2017*) or focus on a subarea of health, such as clinical psychology (*Kramer et al., 2019; Provoost et al., 2017; Rist et al., 2004*). In addition, all of the reviews focus on technological and clinical possibilities. Although they sometimes include a description of the ECA designs used, they do not present effects of particular design features. Thus, a structured literature review of the available studies on particular design features, including a general conclusion with respect to the researched effect of the design features, is missing.



Research objectives

The main goal of this literature review is to identify the researched design features for ECAs in eHealth. The sub goals of this research are to identify the outcome variables used to measure the effect of these design features and to identify what the found effects for each variable were.

METHOD

Search Strategy

Searches were performed in November 2018 in the electronic databases of Scopus, ACM Digital Library, PsychINFO, PubMed and IEEE Xplore Digital Library, as discussed and agreed upon by three researchers: the first, third and fourth author. The searches were restricted to queries containing terms related to (1) embodied conversational agent and (2) eHealth. The list of search terms was composed after several iterations and refinement by the first, third and fourth author. The final list of search terms can be seen in **Table 02.1**.

The searches were performed on titles and abstracts and were not restricted on publication date. For databases that allowed to, Scopus and Pubmed, the language was limited to English and Dutch. In addition, we limited the searches on Scopus to the subject areas Computer Science, Medicine, Mathematics, Social Sciences, Engineering, Psychology, Health Professions, Neuroscience, Nursing, Arts and Humanities and Decision Sciences and the document type Conference Paper, Article, Book Chapter and Book. Again, these limitations were discussed and agreed upon by three researchers: the first, third and fourth author. The final database searches were performed by one researcher (StS).

Table 02.1 – Terms used for the database searches. For databases that do not allow the use of the asterisk (*), the asterisks were removed.

Term [Embodied Conversational Agent]		Term [eHealth]
<p>"virtual agent*" OR "conversational agent*" OR "virtual * coach *" OR "digital coach*" OR "counsel* agent*" OR "virtual counsel*" OR "virtual advisor*" OR "motivational agent*" OR "virtual human*" OR "animated character*" OR "virtual character*" OR "relational agent*" OR "social agent*" OR "interface agent*" OR "interface character*"</p>	AND	<p>e-health OR ehealth OR tele-medicine OR telemedicine OR tele-health OR telehealth OR m-health OR mhealth OR health* OR wellbeing OR e-coaching OR ecoaching OR medic*</p>

Screening Strategy

From the articles identified by the database searches, the duplicates were removed by the first author. Then, two researchers, the first and second author, performed the title, abstract and full text screening independently. The inclusion and exclusion criteria used for the screenings were discussed and agreed upon by the first, third and fourth author and can be seen in [Table 02.2](#) and [Table 02.3](#). The taxonomy of design features used for exclusion criterion E4 was created by combining the categories identified by *Ruttkey et al.* and *Straßmann and Krämer* ([Table 02.3](#)). After each screening, the researchers discussed disagreements until they reached consensus. For the full-text screening, a third researcher, the third author, screened the texts for which the other two researchers had difficulties in reaching consensus. Finally, forward-snowballing was used to screen the references in the included articles using the same technique as used for the screening of the database searches, consisting of a title screening, abstract screening and full-text screening. Duplicates and articles that were already selected for the review through the screening of the database searches, were removed in a pre-processing stage.

Table 02.2 – Inclusion criteria used for the article screenings.

Inclusion Criteria
I1 – The article is written in English or Dutch
I2 – The article is a journal article, conference paper or book (chapter)

Article Reviews and Synthesis

Two review tables were created. The first table, [Table A.1 \(Appendix A\)](#), lists general information about the articles found: the goal of the application in which the ECA was implemented (either in the context of alcohol consumption, mental health, nutrition, physical activity, medical treatment or other) and characteristics of the participants in the research (the age group: adults, children or elderly; education: low, at least some college, students and university; and cultural background: Asian, African

Table 02.3 – Exclusion criteria used for the article screenings.

Exclusion Criteria	Explanation
<i>E1 – The article does not report on primary data</i>	The article is a review article
<i>E2 – The virtual agent is not an embodied conversational agent</i>	Embodied conversational agents are more or less autonomous and intelligent software entities with and embodiment used to communicate with the user (<i>Ruttkay et al., 2004</i>)
<i>E3 – The virtual agent is not applied in a health context</i>	Health is a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity (<i>WHO, 1946</i>)
<i>E4 – The article does not research a design feature of the virtual agent</i>	Design features: <ul style="list-style-type: none"> ➤ Looks (<i>Straßmann & Krämer, 2017</i>): <ul style="list-style-type: none"> ➤ Species (e.g. human, animal, robots, objects, and mystical creatures) ➤ Realism (e.g. stylization, resolution and detailedness) ➤ 2D/3D ➤ Feature specification: Socio-demographic (e.g. gender, ethnicity, race) and Styling (hair, make-up) ➤ Speech and/or textual output (<i>Ruttkay et al., 2004</i>)Hand and body gestures (<i>Ruttkay et al., 2004</i>) ➤ Facial and gaze expressions (<i>Ruttkay et al., 2004</i>)
<i>E5 – The article does not provide any outcomes on the effect of or opinions of users on a design feature of the virtual agent</i>	
<i>E6 – There is no full-text available</i>	

American, Caucasian and Hispanic). In addition, the evaluation of each study was classified as one of the four evaluation stages of DeChant, according to the renewed framework for the evaluation of telemedicine by *Jansen Kosterink et al. (Jansen Kosterink et al., 2016)*. Evaluations were classified as either being in:

- Stage I: technical efficacy – focus on the feasibility and usability of the technology.
- Stage II: specific system objectives – gaining an initial idea about the potential added value for clinical practice and possible working mechanism.
- Stage III: system analysis – technology evaluated in the way they will be implemented in daily clinical practice.
- Stage IV: external validity – elaboration of the adoption as addressed in stage III.

Furthermore, each study was classified as either experimental (meaning the researcher allocates subjects to an intervention or exposure group), observational analytic (the researcher simply measures the exposure or treatments of the groups) or as a survey or qualitative study.

The structure of the second review table, **Table A.2 (Appendix A)**, was agreed upon by two researchers (StS and MT). For each article, the table presents the category of the design feature(s) researched, the design feature(s) researched and the outcome variable(s). The design feature categories were the same categories as used in exclusion criterion E4 (see **Table 02.3**). The categories of the outcome variables were designed retrospectively by thematic analysis of all outcome variables found in the articles. In addition, the table displays the method and the results of the research with respect to the design feature. The articles are grouped on design feature and sorted alphabetically within this category.

RESULTS

From the 1284 articles identified by the database searches, 23 articles were included in the review. In addition, 10 articles were included via the snowballing method, resulting in **33 articles included in the review**. **Figure 02.1** shows the flow diagram of the database searches and article screenings.

Table A.1 (Appendix A) lists general information about the articles found. The included studies were published between **2001 and 2018**. Most of the ECAs were developed in the context of **physical activity** (thirteen ECAs (*Bickmore et al., 2005a; 2009a; 2010; Bickmore and Picard, 2004; 2005; Forlizzi et al., 2007; Frost et al., 2012; Nguyen and Masthoff, 2007; Olafsson et al., 2017; Schmeil and Suggs, 2014; van Wissen et al., 2016; Yin et al., 2010; Zhou et al., 2017*)), **medical treatment** (eight ECAs (*Forlizzi et al., 2007; Parmar et al., 2018; Ring et al., 2014; Robertson et al., 2015; Silverman et al., 2001; Skalski et al., 2007; van Wissen et al., 2016; Zhou et al., 2014*)), **mental health** (six ECAs (*Alsharbi and Richards, 2017; Bickmore and Schulman, 2007; Grillon and Thalmann, 2008; Kang and Gratch, 2011; Nguyen and Masthoff, 2009; Tielman et al., 2017*)) and **nutrition** (four ECAs (*Creed and Beale, 2012; Creed et al., 2015; Olafsson et al., 2017; Schmeil and Suggs, 2014*)). Just a few articles describe ECAs in the context of **alcohol consumption** (3 articles (*Amini et al., 2014; 2013; Lisetti et al., 2013*)) or **other topics** (three articles (*Bickmore and Ring, 2010; Malhotra et al., 2016; van Vugt et al., 2006*)). The amount of participants differed from **11 to 764** ($M = 91, SD = 147$). Most studies included both **male and female participants**. Three studies focused on children (*Alsharbi and Richards, 2017; Frost et al., 2012; Zhou et al., 2017*), two on elderly (*Malhotra et al., 2016; van Wissen et al., 2016*) and the rest on **adults**. Of the articles that reported on the participants' education, **most participants were students** (*Amini et al., 2013; Bickmore and Schulman, 2007; Bickmore and Picard, 2004; Creed et al., 2015; Lisetti et al., 2013; Nguyen and Masthoff, 2007; 2009; Olafsson et al., 2017; Skalski et al., 2007; Tielman et al., 2017; van Vugt et al., 2006*), had a university degree (*Creed et al., 2015; Nguyen and Masthoff, 2007; 2009; Tielman et al., 2017; van Wissen et al., 2016; Zhou et al., 2017*) or had at least some college (*Bickmore and Ring, 2010; Bickmore et al., 2009a; 2010; Robertson et al., 2015; Silverman et al., 2001; van Wissen et al., 2016; Zhou et al., 2017*). Just one study particularly focused on lower-educates (*Robertson et al., 2015*). Of the articles that reported on the participants' cultural background, participants were **Caucasian** (thirteen articles (*Alsharbi and Richards, 2017; Amini et al., 2013; Bickmore et al., 2009a; 2010; Creed et al., 2015; Frost et al., 2012; Olafsson et al., 2017; Robertson et al., 2015; Schmeil and Suggs, 2014; Tielman et al., 2017; van Wissen et al., 2016; Yin et al., 2010; Zhou et al., 2014*)), **Afro American** (seven articles (*Amini et al., 2013; Bickmore and Ring, 2010; Bickmore et al., 2009a; 2010; Olafsson et al., 2017; Robertson et al., 2015; Zhou et al., 2014*)), **Hispanic** (three articles (*Amini et al., 2013; Yin et al., 2010; Zhou et al., 2014*)) and **Asian** (three articles (*Amini et al., 2013; Olafsson et al., 2017; Zhou et al., 2017*)). Of the evaluations performed, **no evaluation was in stage IV, technical efficacy**. **Just one article re-**

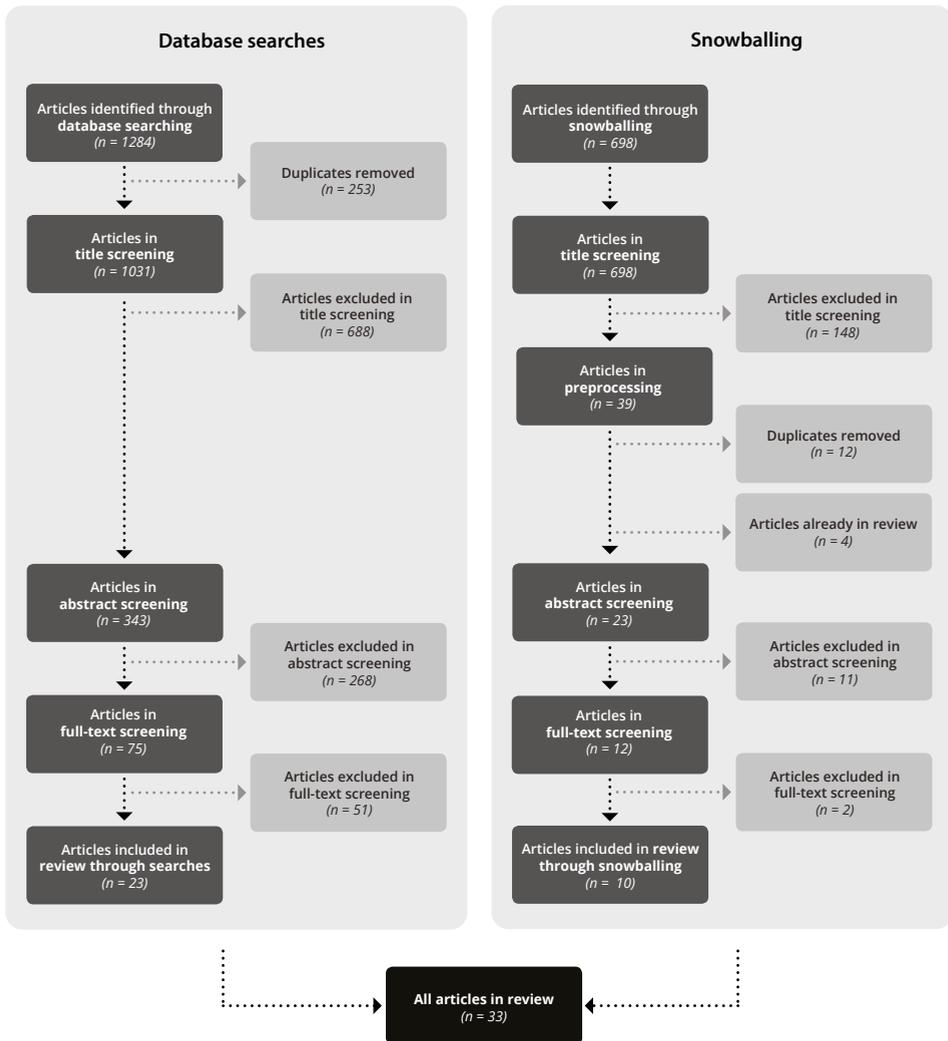


Figure 02.1 – Flow diagram of the database searches and article screenings.

ports on stage III, system analysis (Zhou et al., 2014). Some articles performed evaluations in stage II, specific system objectives (Bickmore et al., 2005a; 2009a; 2010; Bickmore and Picard, 2004; 2005; Creed and Beale, 2012; Creed et al., 2015; Nguyen and Masthoff, 2009; Schmeil and Suggs, 2014; Skalski et al., 2007; Tielman et al., 2017; Yin et al., 2010; Zhou et al., 2017). However, the majority of the articles report on evaluations in the stage I, technical efficacy. In addition, no article described an observational analytic study and few articles describe qualitative studies (two articles (Nguyen and Masthoff, 2007; Robertson et al., 2015)) and survey studies (four articles (Alsharbi and Richards, 2017; Forlizzi et al., 2007; Nguyen and Masthoff, 2007; Parmar et al., 2018)). The majority of the studies performed were experimental studies; they compared multiple variants of a particular design feature.



Figure 02.2 – Frequency of design features and outcome variables in the articles found. The width of the bubble corresponds to the number of articles that research a particular outcome variable for a particular design feature category.

The Design Features and Outcome Variables Researched

Table A.2 (Appendix A) provides information about the design features researched, corresponding outcome variables and results for each article in the review. All articles were grouped on design feature category. In addition, **Figure 02.2** provides an overview of the frequencies of the design feature categories and outcome variables identified in articles included in the review. Some articles research design features in multiple categories. **Most of the research is performed on the categories speech and/or textual output and facial and gaze expressions.** The categories species and 2D/3D are researched the least.

The thematic analysis of the outcome variables resulted in the following categories: **usage, intention to (continue) using, (intention towards) behaviour change, usability and user experience, agent characteristics (e.g. demographics, personality, styling), relation with agent, user characteristics and other.** The majority of the articles provide outcomes regarding the users' perception of the agent characteristics (Alsharbi and Richards, 2017; Amini et al., 2014; 2013; Bickmore and Ring, 2010; Bickmore and Schulman, 2007; Bickmore et al., 2009a; 2010; 2005b; Bickmore and Picard, 2005; Creed and Beale, 2012; Forlizzi et al., 2007; Grillon and Thalmann, 2008; Lisetti et al., 2013; Malhotra et al., 2016; Nguyen and Masthoff, 2007; 2009; Olafsson et al., 2017; Parmar et al., 2018; Ring et al., 2014; Robertson et al., 2015; Silverman et al., 2001; Skalski et al., 2007; Tielman et al., 2017; van Vugt et al., 2006; van Wissen et al., 2016; Yin et al., 2010; Zhou et al., 2014; 2017). In addition, many articles report on the users' perception of the relation with the agent (Alsharbi and Richards, 2017; Amini et al., 2014; 2013; Bickmore et al., 2005b; Bickmore and Picard, 2004; 2005; Creed et al., 2015; Kang and Gratch, 2011; Lisetti et al., 2013; Olafsson et al., 2017; Parmar et al., 2018; Skalski et al., 2007; van Vugt et al., 2006; Zhou et al., 2014), usability and user experience (Amini et al., 2014; 2013; Bickmore and Ring, 2010; Bickmore and Schulman, 2007; Bickmore et al., 2009a; 2010; 2005b; Lisetti et al., 2013; Nguyen and Masthoff, 2009; Olafsson et al., 2017; Ring et al., 2014; Silverman et al., 2001; Tielman et al., 2017; van Wissen et al., 2016; Zhou et al., 2014; 2017), intention to use (Amini et al., 2014; 2013; Bickmore and Schulman, 2007; Bickmore et al., 2010; 2005b; Bickmore and Picard, 2004; 2005; Creed and Beale, 2012; Lisetti et al., 2013; Olafsson et al., 2017; Parmar et al., 2018; Ring et al., 2014; Schmeil and Suggs, 2014; van Vugt et al., 2006; van Wissen et al., 2016; Zhou et al., 2014) and system usage (Bickmore et al., 2009a; 2010;

2005b; Bickmore and Picard, 2005; Creed and Beale, 2012; Creed et al., 2015; Olafsson et al., 2017). Also, many articles already provide information about the effect of the design of a particular feature on users' (intention towards) behaviour change (Bickmore and Schulman, 2007; Bickmore et al., 2010; 2005b; Bickmore and Picard, 2005; Creed et al., 2015; Kang and Gratch, 2011; Nguyen and Masthoff, 2009; Parmar et al., 2018; Schmeil and Suggs, 2014; Silverman et al., 2001; Skalski et al., 2007; Tielman et al., 2017; Yin et al., 2010; Zhou et al., 2017). Just a few articles provide results related to the users' perception of their own characteristics (Creed et al., 2015; Frost et al., 2012; Tielman et al., 2017; van Vugt et al., 2006; Yin et al., 2010; Zhou et al., 2014) or report on other outcome variables (Bickmore and Ring, 2010; Bickmore and Picard, 2005; Creed and Beale, 2012; Nguyen and Masthoff, 2007; Olafsson et al., 2017; Tielman et al., 2017; Zhou et al., 2014).

Looking at the relation between the design feature categories and outcome variables specifically, we see that for realism, species and 2D/ 3D just a few outcome variables are researched, whereas for the other categories, almost all outcome variables have been researched.

In the remainder of this section, **the research and outcomes are grouped by the design features categories**. We start with describing research related to speech and/or textual output, facial and gaze expressions and hand and body gestures, followed by research on the agent's looks.

Speech and/or Textual Output, Facial and Gaze Expressions and Hand and Body Gestures

Table 02.4 provides a summary of the effects found for the different outcome variables with respect to design features in the categories speech and/or textual output, facial and gaze expressions and hand and body gestures.

First, **some articles provide research on an agent's emotion**. Compared to agents not showing emotion, agents showing emotion are rated higher on several characteristics, such as likeability and believability (Creed and Beale, 2012; Creed et al., 2015; Silverman et al., 2001) and resulted in higher usability (Silverman et al., 2001) and intention to use (Creed and Beale, 2012). However, no clear consensus exist for emotional agents triggering behaviour change; one study found that users interacting with an emotional agent showed a larger behaviour change than users interacting with a non-emotional agent (Silverman et al., 2001). Another study found the opposite: users interacting with a non-emotional agent showed a larger behaviour change than users interacting with an emotional agent. But, on other behaviour variables, they did not find any differences (Creed et al., 2015). It should be noted that the two studies offered different application goals: change in awareness on heart attack scenarios and change in food intake.

Second, **some articles provide research on an agent's relational, empathic behaviour**. First, relational agents are liked more: they score higher on characteristics, such as likeability, perceived caring, trustworthiness and enjoyment (Amini et al., 2014; 2013; Bickmore et al., 2005a; Bickmore and Picard, 2004; 2005; Lisetti et al., 2013; Nguyen and Masthoff, 2009). In addition, relational behaviour positively affects the users' relation with the agent (Amini et al., 2014; 2013; Bickmore et al., 2005a; Bickmore and Picard, 2004; 2005; Lisetti et al., 2013). Lastly, the use of relational agents leads to higher usability (Amini et al., 2014; 2013; Bickmore and Schulman, 2007; Lisetti et al., 2013; Nguyen and Masthoff, 2009) and intention to use (Amini et al., 2014; 2013; Bickmore et al., 2005a; Bickmore and Picard, 2004; 2005; Lisetti et al., 2013). However, with respect to behaviour change, literature presents mixed results; some studies did not find any effect (Bickmore et al., 2005a; Bickmore and Picard, 2004; 2005; Nguyen and Masthoff, 2009), whereas just one article provided some positive results related to an agents rela-

tional behaviour (*Bickmore and Schulman, 2007*). Though, the studies researched applications with different goals. *Bickmore and Schulman*, finding a positive effect of relational behaviour, tested the effect of relation behaviour on mood, whereas the majority of the other studies, not finding any effects, focused on physical activity. A last note with respect to an agent's relational behaviour: as described by *Nguyen and Masthoff*, people seem not to care or expect whether a system could understand and care for their feelings, but when a system is represented by a human-like agent, its lack of empathy could lead to negative user experience and worsen the user's attitude towards the system.

In addition, **some articles research an agent providing personal information**. High self-disclosure positively affects the user's (intention towards) behaviour change and its relationship with the agent (*Kang and Gratch, 2011*), whereas stories told in first person result in high system usage and usability (*Bickmore et al., 2009a*).

Furthermore, **some research on variability in an agent's behaviour has been performed**. Variability in an agent's behaviour positively affects system usage (*Bickmore et al., 2010*) and intention to use (*Bickmore et al., 2010*), but, with respect to behaviour change, non-variable behaviour is preferred over variable behaviour (*Bickmore et al., 2010*). When varying the behaviour of an agent, changing its behaviour with respect to human eye contact behaviour seems to be better than randomly changing its behaviour, since an agent changing its behaviour with respect to human eye contact behaviour is perceived to be more normal and realistic (*Grillon and Thalmann, 2008*).

Some last remarks, based on research presented in single articles. First, allowing users to control an agent's prosody (the stress and intonation patterns of an utterance) and facial expressions when the agent's task is to retell a story results in high satisfaction (*Bickmore et al., 2010*). Second, users rate their characteristics (e.g. intrinsic motivation and self-efficacy) higher after interaction with an interactive coach than after interaction with a non-interactive coach and higher after interaction with a moving coach than after interaction with a non-moving coach (*Frost et al., 2012*). Furthermore, adding rap music to a dialogue positively affects engagement and the user's relation with the agent, whereas rap music reduces trust in an agent (*Olafsson et al., 2017*). The presence or absence of rap music did not influence the perception of the agent's characteristics (e.g. naturalness, knowledge-ability, perceived similarity and liking), system usage, intention to (continue) using the agent and the systems usability. Also, presenting psycho-education via text results in higher task adherence than when an agent provides the psycho-education verbally (*Tielman et al., 2017*), since psycho-education in text was better recollected. Finally, linguistic tailoring had no effect with respect to persuasion in the context of behaviour change (*Yin et al., 2010*).

Looks

Other research identified in the review focuses on the agent's looks. **Table 02.4** provides a summary of the effects found for the different outcome variables with respect to the agent's looks. Research has been performed on the subcategories species, realism, styling and socio-demographics. **No article in the review presented research on effects of agents in either 2D or 3D.**

Just one article researched the agent's species and a few the agent's realism. Research shows mixed results with respect to the best rendering style. Although stylised agents are rated positively on characteristics such as friendliness (*Ring et al., 2014; Robertson et al., 2015*), several studies indicate that human agents are preferred over abstract, and stylised (cartoon-like) agents (*Forlizzi et al., 2007; Ring et al., 2014; Robertson et al., 2015; van Wissen et al., 2016*). However, the application goal,

Table 02.4 – Summary of the effects of the design features on the outcome variables, either a positive effect (+), negative effect (-), no effect (0) or an effect that depends on the context (~). For every row, symbols having the same number in superscript are researched within the same study.

	Design feature	Usage	Intention to (continue) using	(Intention towards) behaviour change
Speech/Text, Output, Facial & Gaze Exp., Hand & Body Gest.	Emotion (vs. no)		+ ²	- ³ 0 ³ + ¹
	Relational, empathic behaviour (vs. no)		+ ¹ + ² + ³ + ⁴ + ⁵ + ⁶	0 ³ 0 ⁴ 0 ⁵ 0 ⁷ + ⁸
	Providing personal information (vs. no)	+ ²		+ ¹
	Variable behaviour (vs. no)	+ ¹	+ ¹	- ¹ ~ ²
	User control prosody & facial expressions (vs. no)			
	Interactivity (vs. no)			
	Rap (vs. no)			
	Agent message: in text (vs. verbally)	+ ¹		+ ¹
	Linguistic tailoring (vs. no)			0 ¹
Looks	Rendering style: human-like (vs. cartoon)		0 ⁴	
	Clothing: professional (vs. casual)		+ ¹	
	Body shape: slim (vs. fat)		- ⁴	- ⁴ + ¹ + ² + ³
	Gender: female (vs. male)			
	Age: young (vs. old)			
	Cultural tailoring (vs. no)			- ² - ⁴
	Role: friend (vs. professional)			

	Usability and user experience	Agent characteristics	Relation with agent	User characteristics	Preference
	+ ¹	+ ¹ + ² + ³			
	+ ¹ + ² + ⁶ + ⁷ + ⁸	+ ¹ + ² + ³ + ⁴ + ⁵ + ⁶ + ⁷	+ ¹ + ² + ³ + ⁴ + ⁵ + ⁶		
	+ ²		+ ¹		
	+ ¹				
				+ ¹	
	0 ¹ + ¹	- ¹ 0 ¹	+ ¹		
		- ¹ - ² + ¹ + ² + ³ + ⁴			
		+ ¹	+ ¹		
		- ⁴	- ⁴		
					+ ¹ ~ ²
					+ ¹ ~ ²
		+ ¹ + ²	- ³		
					+ ¹ ~ ²

participants and study methods of the studies varied a lot. With respect to intention to use, no effect of rendering style was found (*van Wissen et al., 2016*).

Second, **some articles research the agent's clothing and body shape**. A professional looking agent, dressed in a white coat and wearing a stethoscope, is positively rated on its characteristics (e.g. credibility, trustworthiness, reassurance, caring and friendliness), relation with the user and intention to use (*Parmar et al., 2018*) compared to a casually dressed agent. Regarding the agent's body shape, literature shows mixed results. With respect to behaviour change, some research shows a preference for attractive agents above unattractive agents (*Nguyen and Masthoff, 2007; Schmeil and Suggs, 2014; Skalski et al., 2007*), whereas other research shows a preference for non-ideal, fatter characters above ideal, slim characters (*van Vugt et al., 2006*). Also, with respect to the perception of the agent's characteristics, relation with the agent and intention to use, results show positive effects for non-ideal body shapes (*van Vugt et al., 2006*). Although the studies show different results, the target users and application goal were similar.

Lastly, **some articles research the agent's demographics**. Literature does not show a clear consensus when it comes to preference for a particular gender. Some research indicates a preference for female agents (*Alsharbi and Richards, 2017*), whereas other research shows that the preferred gender depends on the task of the agent (*Forlizzi et al., 2007*). However, the studies differed in target group (children vs. adults) and application goal (providing medical advice or physical activity training vs. treatment of anxiety and post-traumatic stress disorder). In addition, no clear consensus exists on the age of the agent; some research suggests that young agents are preferred over old agents (*van Wissen et al., 2016*), whereas other research suggests that users prefer agents of the same age or older (*Alsharbi and Richards, 2017*). Again, the studies differed in target group (older adults vs. children) and application goal (increase physical activity and medication vs. treatment of anxiety and post-traumatic stress disorder). Also, some research indicates an agent having the same cultural background as the user is more positively rated on its characteristics (e.g. perception of caring, general liking) (*Alsharbi and Richards, 2017; Yin et al., 2010*) and its relation with the user compared to an agent with a different cultural background (*Zhou et al., 2014*), whereas, with respect to behaviour change, agents with a different cultural background could be beneficial (*Yin et al., 2010; Zhou et al., 2017*). However, the studies targeted different users and researched applications with different goals. In addition, the studies were in different stages of change (ranging from I to III). Furthermore, some research indicates a preference for an agent that acts as a friend (*Alsharbi and Richards, 2017*), whereas other research indicates that the preferred agent role relates to the agent's task (*Nguyen and Masthoff, 2009*). Though, the studies focused on a different target group (children vs. adults) and application goal (treatment of anxiety and post-traumatic stress disorder vs. mood manipulation).

Two final remarks with respect to the agent's looks. First, several studies stress the importance of aligning the agent's looks to the looks of the user (*Malhotra et al., 2016; Robertson et al., 2015*). Similarity with the agent seems to influence the perception of the characteristics of and preferences for particular agents (*Zhou et al., 2014*). It seems that some users prefer agents that are similar to themselves, for example, in age (*Alsharbi and Richards, 2017*), body shape (*van Vugt et al., 2006*) and cultural background (*Alsharbi and Richards, 2017; Zhou et al., 2017*). Second, literature indicates that preference for particular agents and perception of their personalities depend on the task of the agent (e.g. providing medical advice, encouraging to perform physical activity) (*Nguyen and Masthoff, 2007; Ring et al., 2014*). It seems that people apply human stereotypes to agents, and therefore, for example, have preferences for a particular gender for a particular task (e.g. male agents are preferred for the role of athletic trainer) (*Forlizzi et al., 2007*).

DISCUSSION

In this chapter, we provided a state of the art of design features for ECAs in eHealth, showing a field that is immature and without consensus on effective design features. Emotion and relational behaviour seem to have positive effects, but do not necessarily lead to behaviour change in the context of eHealth. In the remainder of this section, we discuss the design features researched, the outcome variables on which the effect of these design features were researched and what the measured effects were.

Design Features for ECAs in eHealth

The included articles show that most of the research focused on speech and/or textual output, gaze and facial expressions and hand and body gestures, and not on an agent's looks. Therefore, we see an opportunity for future work on the agent's looks. We are not aware of a literature review identifying design features for ECAs in other contexts to compare our results with. However, articles that research speech and/or textual output, facial expressions and hand and body gestures (*Acosta and Ward, 2011; Berry et al., 2005; Kim et al., 2007; Lee et al., 2007; Pelachaud, 2009; von der Pütten et al., 2009*) or the agent's looks (*Baylor and Kim, 2004; Cowell and Stanney, 2003; Guadagno et al., 2007; Khan and Angeli, 2009; Khan and Sutcliffe, 2014; Kim et al., 2003; 2007; Lee et al., 2018; Rosenberg-Kima et al., 2008; Straßmann and Krämer, 2017; Veletsianos, 2010*) in general, or in another context than eHealth, do exist. Therefore, we believe that in other contexts similar design features might have been researched.

Outcome Variables for Measuring the Effect of Design Features

The measured effect of the design features for ECAs in eHealth was often on the perception of the agent's and users characteristics, relation with the agent, system usage, intention to use, usability and behaviour change. Again, although we are not aware of a literature review identifying outcome variables used to evaluate ECAs in other contexts, we do see similar outcome variables researched in other contexts by individual articles. For example, research has been performed on the perception of the agent's characteristics (*Kim et al., 2007; Pelachaud, 2009*), user's characteristics (*Kim et al., 2007*), relation with the agent (*Acosta and Ward, 2011; Lee et al., 2007; von der Pütten et al., 2009*) and behaviour change (*Berry et al., 2005; Kim et al., 2007; Lee et al., 2007*). Therefore, we believe that outcome variables measured in other contexts might be similar to the outcome variables researched in the health context.

Measured Effects of the Design Features

Effects of an ECA's Speech and/or Textual Output, Facial and Gaze Expressions and Hand and Body Gestures

Existing literature shows some consensus on the effects of an ECAs emotion and relational behaviour, but no consensus on the effects of other design features with respect to an agent's speech and/or textual output, facial and gaze expressions and hand and body gestures. Research on design features for ECAs in other contexts supports our findings.

First, some positive effects of **an agent's emotion** were found in research in other contexts, such as e-learning. Research shows that an agent's positive emotions positively affect the users perception of an agent's characteristics, such as the ability to facilitate learning (*Kim et al., 2007*) and that the way emotion is implemented might affect the function of emotion (e.g. attract the user's attention, persuade the user) (*Pelachaud, 2009*). Whereas our literature review does not show a clear effect of

emotion on behaviour change, existing research in other contexts shows some positive effects, such as an agent's emotion increasing the user's interest in learning (Kim et al., 2007) and increasing the user's cognitive performance (Berry et al., 2005).

Second, research in other contexts shows positive results for the implementation of **relational behaviour**. For example, research on pedagogical and co-learner agent's shows that an agent's empathy positively impacts students self-efficacy (Kim et al., 2007), adapting the agent's speech to the emotional state of the user results in higher rapport (Acosta and Ward, 2011), that trust in the agent is higher for a caring co-learner compared to a non-caring co-learner agent (Lee et al., 2007) and that a higher mutual awareness when increasing the agent's behavioural realism (von der Pütten et al., 2009). Whereas the articles in the review mainly show no effect of relational behaviour on behaviour change, some articles in other contexts do show a positive effect of relational behaviour on behaviour change. For example, in the context of e-learning, research shows an increased learner interest (Kim et al., 2007) and increased learning (Lee et al., 2007) when implementing relational behaviour.

For **other design features** with respect to speech and/or textual output, facial and gaze expressions and hand and body gestures, results show no clear consensus. Either few studies have been performed on these features, which makes it difficult to generalize the results, or results show contradictory effects. Differences might be caused by the studies involving different target groups, ranging from children to older adults, or the applications having too different goals, ranging from mood manipulation to increase of physical activity. Therefore, more research is needed:



Recommendation

We recommend to perform research on the effect of design features regarding the agent's speech and/or textual output, facial and gaze expressions and hand and body gestures on the same outcome variables and with a similar target group and application goal.

Effects of an ECA's Looks

Until now, just a few studies have been carried out on the agent's looks in the context of eHealth. No consensus exists on the agent's species and rendering style (also shown by research on rendering style in other fields, such as by McDonnell et al., 2012; Zell et al., 2015; Zibrek et al., 2018), and no research has yet been performed on the effects of agents in either 2D or 3D. Also, different opinions on the most appropriate agent demographics exist. Our review does not show consensus with respect to the preferred gender, age, role and cultural background. Also, an agent's clothing and body shape seem to be factors to take into account when it comes to creating a positive perception of the agent. These results are in line with research in other contexts, for example in the context of e-learning, showing mixed results on the agent's rendering style (Straßmann and Krämer, 2017), gender (Baylor and Kim, 2004; Cowell and Stanney, 2003; Guadagno et al., 2007; Kim et al., 2007; Rosenberg-Kima et al., 2008), age (Cowell and Stanney, 2003; Lee et al., 2018), role (Baylor and Kim, 2004; Kim et al., 2003), cultural background (Baylor and Kim, 2004; Cowell and Stanney, 2003) and the agent's clothing and

body shape (Baylor and Kim, 2004; Khan and Angeli, 2009; Khan and Sutcliffe, 2014; Veletsianos, 2010). What should be noted is that some articles included in the review stress the importance of aligning the agent's looks, especially its demographics age and gender, to the looks of the user. Research in other contexts, such as e-learning, supports this note (Baylor, 2009; Guadagno et al., 2007; Gulz and Haake, 2005; Lee et al., 2018; Rosenberg-Kima et al., 2008; Straßmann and Krämer, 2017; Veletsianos, 2010). Thus, it seems important to personalise the agent's looks:



Recommendation

We see opportunities for future work researching the effect of design features regarding the agent's looks in relation to the characteristics of the user.

Transferring Effects to a Long-term, Daily Life Setting

The research area on the design of ECAs in eHealth is relatively immature. Most of the articles describe research in stage I or II of the renewed framework of evaluation for telemedicine (Jansen Koosterink et al., 2016). Therefore, we learned about the effects of the agent's design features in a lab setting, but do not yet know how these effects translate to a daily life setting for which the ECAs are designed. In addition, in most lab studies users interacted with the agent's for a short period of time, which is different from the long-term interaction for which the majority of the ECAs are designed. Therefore, we should be careful with interpreting the results of this literature review:



Recommendation

We see an opportunity for replicating the studies identified in this literature review that were in stage I or stage II of the renewed framework of evaluation for telemedicine, in a long-term, daily life setting (stage III or IV).

Applying Design Guidelines from Other Contexts

As indicated above, the findings of our review are in line with results of general research on design features for ECAs or research on design features for ECAs in another context. Therefore, we believe that general guidelines, like guidelines for designing personalities for social agent's by Dryer (Dryer, 1999) and design guidelines for other contexts, such as guidelines of the Enhancing Agent Learner Interactions (EnALI) Framework (Veletsianos et al., 2009), might be applicable to agent's in eHealth as well. However, research also indicates that the user's perception of ECAs depends on the agent's task (Baylor, 2009). Whereas the task of agent's in eHealth might be similar to agent's in other contexts, such as being informative, like embodied chatbots on commercial websites, agent's used in muse-

ums or pedagogical agent's, the task of the agent might differ from the task of these agent's when supporting behaviour change in the health domain. Therefore, we should research whether results of other contexts are still applicable to eHealth:



Recommendation

We suggest to repeat studies focusing on the effects of design features in a general or other context in the eHealth context.

Strengths and Limitations Review

The strength of this review was that it focused on design features for ECAs in a specific context, eHealth, since preferences for agent designs might be context-dependent. However, this narrow focus could also be a limitation of the research. General research – research not restricted to a particular context – on agent design features might still be applicable to a health context. In addition, research regarding design features in other contexts, such as research on design features for pedagogical agent's, is not included in the review. Some of the results might still be applicable to the eHealth context.

CONCLUSION

This literature review identified (1) the **researched design features** for ECAs in eHealth, (2) the **outcome variables** used to measure the effect of these design features and (3) the **found effects** for each variable. Results show that the agent's speech and/or textual output and its facial and gaze expressions were the most common design features, whereas little research was performed on the agent's looks. The measured effect of these design features was often on the perception of the agent's and user's characteristics, relation with the agent, system usage, intention to use, usability and behaviour change. With respect to the effects found, consensus on design features of ECAs in eHealth is far from established. Solely, emotion and relational behaviour seem to positively affect the perception of the agent's characteristics and that relational behaviour also seems to positively affect the relation with the agent, usability and intention to use. However, these design features do not necessarily lead to behaviour change. The research area of ECAs in eHealth is immature, therefore, we see four opportunities for future work: (1) more research on the agent's speech and/or textual output, facial and gaze expressions and hand and body gestures, (2) research on the agent's looks, (3) for all categories: evaluations in a long-term, daily life setting, and (4) replication of studies regarding design features performed in other contexts than eHealth. By performing research in these four areas, we can work towards a set of design guidelines for ECAs in the eHealth domain.

Acknowledgements – This work was supported by the European Commissions Horizon 2020 Research and Innovation Programme project Council of Coaches (Grant Agreement Number 769553).

03

Effects of ECA Age, Gender and Role on Users' Impressions at First Glance

BASED ON:

ter Stal S., Tabak, M., op den Akker, H., Beinema, T., Hermens, H. (2019). Who do you prefer? The effect of age, gender and role on users' first impressions of embodied conversational agents in eHealth. *International Journal of Human-Computer Interaction*, 36(9), 881–892. [doi:10.1080/10447318.2019.1699744](https://doi.org/10.1080/10447318.2019.1699744)



Abstract

Embodied conversational agents may be used to engage users in adopting eHealth applications. The aim of this chapter 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, and that (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. Furthermore, our results show that (3) females seem to prefer images of agents similar in gender (i.e. female agents), whereas males do not seem to have a clear preference for a particular gender, and that (4) younger people seem to have a preference for images of agents similar in age (i.e. young agents), whereas older people do not seem to have a clear preference for a particular age. A next step would be to investigate how the characteristics of the agent designs are perceived after interaction with the agent.



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 behaviours (Kreps & 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 behaviour to elicit trust, better communication and satisfaction (Bickmore, Pfeifer, & Jack, 2009).

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 (Ruttkay, 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 likeability 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 be important to be relevant for an ECA in the health context, to eventually, **follow the agent’s advice**.

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 behaviour 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, Eyszel, & 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 behaviours.**

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.

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, 2016; 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, 2016) 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, Ayoub, 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 likeable (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; 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.



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 03.1**.

Table 03.1 – 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

METHOD

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

First Impressions General and Elderly Population

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.

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 03.2** provides an overview of the features. Combinations of all variations were tested, leading to a set of eight agent permutations. The individual agent designs are shown in **Figure 03.1**.

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.

Table 03.2 – 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 (level of health expertise)	Expert (high level of health expertise)	Peer (low level of health expertise)



Figure 03.1 – The eight agent designs subjected to testing.

Procedure

Figure 03.2 shows the flow of the online questionnaire. The questionnaire itself can be found in **Appendix B**. 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. Next, participants were informed that they would get to see a set of agent designs, and would be asked about their preferences for these agent designs when employed as a coach to gain a healthy lifestyle. 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 third 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.

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

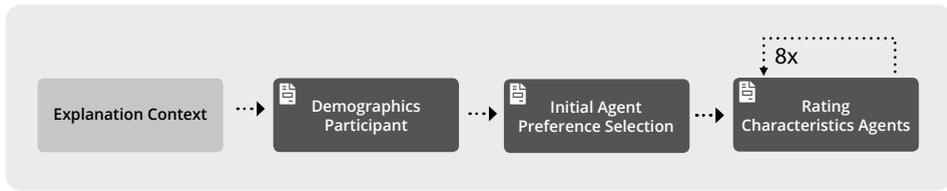


Figure 03.2 – Flow of the online questionnaire.

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.

Attitude Elderly Population toward Agent Characteristics

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 (mostly living in Twente, the Netherlands). The focus groups were performed in Dutch, in July 2018.

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.

Procedure

Figure 03.3 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).



Figure 03.3 – Outline of the focus group process.

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.

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 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.

RESULTS

First Impressions General and Elderly Population

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).

Preference Agent Designs at First Glance

Figure 03.4 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

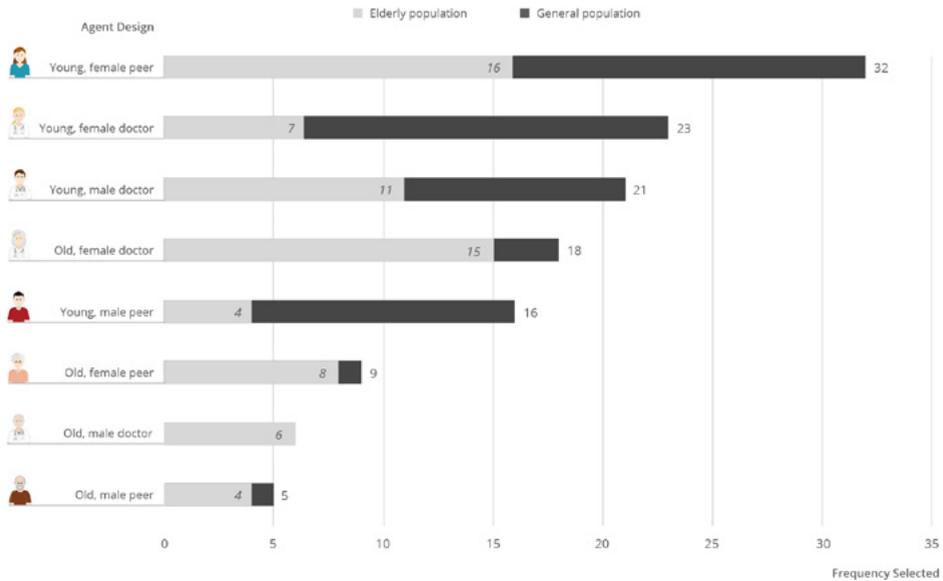


Figure 03.4 – Frequencies of the agent designs preferred at first glance.

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**).

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 03.3](#)). Second, 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 ([Table 03.4](#)). 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.

Table 03.3 – 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 (n = 117) and elderly population (n = 66).*

	General Population				Elderly Population			
	Young, M(SD)	Old, M(SD)	p	Coh. d	Young, M(SD)	Old, M(SD)	p	Coh. 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

Table 03.4 – 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 (n = 117) and elderly population (n = 66).*

	General Population				Elderly Population			
	Female, M(SD)	Male, M(SD)	p	Coh. d	Female, M(SD)	Male, M(SD)	p	Coh. 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 03.5 – 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 (n = 117) and elderly population (n = 66).*

	General Population				Elderly Population			
	Peers, M(SD)	Experts, M(SD)	p	Coh. d	Peers, M(SD)	Experts, M(SD)	p	Coh. 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

*Ratings from 1 (strongly disagree) to 7 (strongly agree). Relations for which p-values are shown in bold are statistically significant. Cells colored in gray show similar results for general and elderly population.

Table 03.6 – Results supported the hypotheses H1.1, H1.2 and H1.3.

<i>Hypothesis</i>	General Population		Elderly Population	
	<i>Outcome</i>	<i>Explanation</i>	<i>Outcome</i>	<i>Explanation</i>
<i>H1.1 – Age affects perceptions</i>	Supported	<ul style="list-style-type: none"> ➤ Images of young agents are rated higher on (a) likeliness of following advice, (b) friendliness, (c) expertise, (d) reliability and (e) involvement compared to images of old agents ➤ Images of old agents are rated higher on (f) authority compared to images of young agents 	Supported	<ul style="list-style-type: none"> ➤ Images of young agents are rated higher on (b) friendliness and (e) involvement compared to images of old agents. ➤ Images of old agents are rated higher on (f) authority compared to images of young agents.
<i>H1.2 – Gender affects perceptions</i>	Supported	<ul style="list-style-type: none"> ➤ Images of female agents are rated higher on (b) friendliness and (e) involvement than images of male agents ➤ Images of male agents are rated higher on (c) expertise and (f) authority than images of female agents 	Supported	<ul style="list-style-type: none"> ➤ Images of male agents are rated higher on (f) authority than images of female agents.
<i>H1.3 – Role affects perceptions</i>	Supported	<ul style="list-style-type: none"> ➤ Images of expert agents are rated higher on (a) likeliness of following advice, (c) expertise, (d) reliability and (f) authority than images of peer agents 	Supported	<ul style="list-style-type: none"> ➤ Images of expert agents are rated higher on (a) likeliness of following advice, (c) expertise, (d) reliability, (e) involvement and (f) authority than images of peer agents.

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 **Table 03.7**, **Table 03.8** and **Table 03.9**. A summary of all results can be seen in **Table 03.10**. H2.1 and H2.2 are supported, whereas H2.3 is rejected.

Relation Age Preferred Agent and Age Respondent

Table 03.7 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.

Relation Gender Preferred Agent and Gender Respondent

Table 03.8 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.

Relation Role Preferred Agent and Health Literacy Respondent

No significant effect of the role of health literacy of the respondent on the role of the preferred agent was found, neither for the general and elderly population.

Table 03.7 – 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 Preferred Agent			
	Young, M(SD)	Old, M(SD)	p	Cohen's d
Age Respondent	47.52(21.925)	64.46(13.312)	<.001	0.934

Table 03.8 – 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			
		Female, n(%)	Male, n(%)	p	Phi ϕ
General Population	Female	48(73.8)	17(26.2)	.008	0.238
	Male	31(50.8)	30(49.2)		
Elderly Population	Female	27(79.4)	7(20.6)	.013	0.293
	Male	19(51.4)	18(48.6)		

Table 03.9 – 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		
		Peer, n(%)	Expert, n(%)	p
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 03.10 – 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	

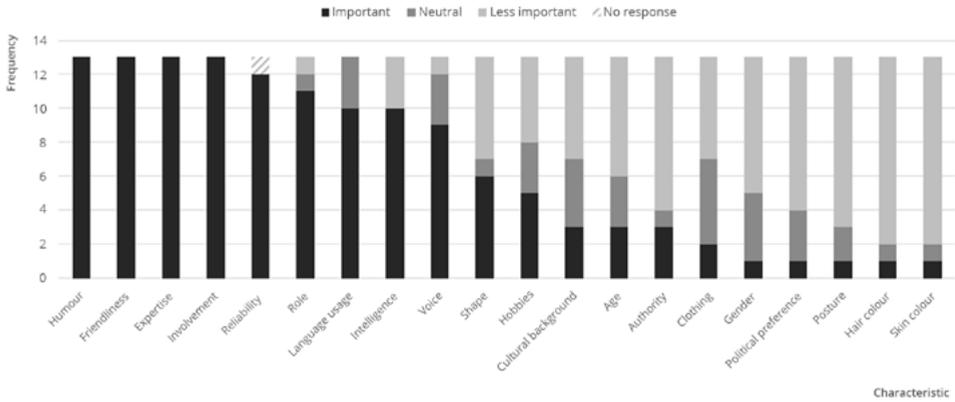


Figure 03.5 – Results of card sorting task in focus groups. Participants rated each characteristic as either important, neutral or less important.

Attitude Elderly Population toward Agent Characteristics

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).

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. 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 03.5** 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.

Attitude toward Design Features

Figure 03.5 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.

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.

Attitude toward Agent Characteristics

Figure 03.5 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 conformation. 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.

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, and 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.** Furthermore, our results show that (3) **females seem to prefer images of agents similar in gender (i.e. female agents), whereas males do not seem to have a clear preference for a particular gender, and that (4) younger people seem to have a preference for images of agents similar in age (i.e. young agents), whereas older people do not seem to have a clear preference for a particular age.**

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.

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 and men are perceived as more assertive, having a higher self-esteem, whereas females are seen as more trustworthy and tenderminded (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) and Nguyen and Masthoff (2007). 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. Research by Nguyen and Masthoff (2007) shows similar findings. They show that people rate images of doctors (wearing a doctor's coat) higher on expertise with respect to the health benefits of exercise compared to sport instructors (wearing fitness clothes), while they rate sport instructors higher on expertise with respect to fitness programs. 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.

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 females are more likely to select an image of a female 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.

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.

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.

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.

Acknowledgements – 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].

04

Effects of ECA Age and Gender on Users' Impressions after Short Interaction

BASED ON:

ter Stal, S., Broekhuis, M., van Velsen, L., Hermens, H., & Tabak, M. (2020). Embodied conversational agent appearance for health assessment of older adults: Explorative study. *JMIR Human Factors*, 7(3). [doi:10.2196/19987](https://doi.org/10.2196/19987)



Abstract

Embodied conversational agents (ECAs) have great potential for health apps but are rarely investigated as part of such apps. To promote the uptake of health apps, we need to understand how the design of ECAs can influence the preferences, motivation, and behavior of users. The study presented in this chapter is one of the first studies that investigates how the appearance of an ECA implemented within a health app affects users' likeliness of following agent advice, their perception of agent characteristics, and their feeling of rapport. In addition, we assessed usability and intention to use. The ECA was implemented within a frailty assessment app in which three health questionnaires were translated into agent dialogues. In a within-subject experiment, questionnaire dialogues were randomly offered by a young female agent or an older male agent. Participants were asked to think aloud during interaction. Afterward, they rated the likeliness of following the agent's advice, agent characteristics, rapport, usability, and intention to use and participated in a semistructured interview. A total of 20 older adults (72.2 ± 3.5 years) participated. The older male agent was perceived as more authoritative than the young female agent ($p=0.03$), but no other differences were found. The app scored high on usability (median 6.1) and intention to use (median 6.0). Participants indicated they did not see an added value of the agent to the health app. Concluding, agent age and gender little influence users' impressions after short interaction but remain important at first glance to lower the threshold to interact with the agent. Thus, it is important to take the design of ECAs into account when implementing them into health apps.



INTRODUCTION

As people get older, they are likely to experience frailty, a decline in functional and cognitive abilities such as walking speed, balance control, and working memory (Malva & Bousquet, 2016; Fried et al., 2001). Through electronic health (eHealth), frailty can be assessed using digital questionnaires. A large population can be targeted, including those who are less mobile and face difficulties in seeing a caregiver to perform frailty assessment. In addition, digital frailty assessments can be performed on a regular basis, be dynamically adapted based on information provided by the user, and provide immediate results. An eHealth app can coach the user in a personalized way toward a healthy lifestyle based on the outcomes of the frailty assessment. Research shows that collecting health data using a digital survey does not affect test reliability with respect to a paper version (Bliven et al., 2001; Kvien et al., 2005; Hess et al., 2008), and several studies showed similar results for a population of older adults (van Velsen et al., 2018; Fanning & McAuley, 2014). In addition, Fanning and McAuley showed that older adults may accept a tablet for health surveys and van Velsen et al. (2018) showed that older adults preferred a tablet survey to a paper survey.

Research shows that the older and more frail adults get, the more they become non-respondents to questionnaires (Hébert, 1996; Hardie, Bakke & Mørkve, 2003), whereas refusal of face-to-face interviewing is less present in this population (Hébert, 1996). To overcome the problem of lack of face-to-face interaction in a digital frailty assessment, an embodied conversational agent (ECA) can provide an alternative. ECAs are more or less autonomous and intelligent software entities with an embodiment used to communicate with the user (Ruttkay et al., 2004). By interacting with the user face to face, ECAs can build trust and rapport—a close and harmonious relationship—leading to companionship and long-term continual use (Vardoulakis et al., 2012).

To establish trust and rapport with the agent, users should have a positive impression of the agent. These impressions can be shaped by static (Cafaro et al., 2016) and dynamic characteristics (Cafaro et al., 2016; Bergmann et al., 2012). Static characteristics mostly relate to an agent's visual appearance, often tested using the so-called zero acquaintance approach, where a person observes the agent without interacting with the agent. Dynamic characteristics include an agent's verbal and non-verbal behaviours and are often tested using a thin-slicing approach, where a person draws inferences about an agent's personality based on short excerpts of social behaviour (Vartanian et al., 2012).

Although ECAs have the potential to be used as eHealth apps such as digital frailty assessments, little is known about how these agents should be designed and how the design affects our impressions of the agents, and no design guidelines exist (chapter 2). In chapter 3, we identified people's first impressions of agents varying in age, gender, and role using a zero acquaintance approach: there was no interaction involved, and participants rated static agent images at first glance. The study shows that characteristics of older and male agents were perceived differently than characteristics of young and female agents, respectively. In addition, older adults seem to prefer a young female over an older male agent. Other research focused on users' perceptions of static agent images at first glance (Forlizzi et al., 2007; Parmar et al., 2018; Zimmerman, 2005), showing that the agent's gender and role affect the user's perception of the agent. However, little research exists on people's impressions after short interactions with agents and how the design of the agents affects these impressions. Therefore, research is needed to investigate how the design of an agent affects users' impressions of the agent during and after actual interaction (using a thin-slicing approach).



The aim of this study is to assess how an agent's appearance, particularly age and gender, affects the users' likeliness of following agent advice and users' perceptions of the agent's characteristics and feeling of rapport after short interaction with the agent. This study builds on previous work (*chapter 3*) by studying users' impressions of agents at first glance (using the zero acquaintance approach) and after a short interaction with the agents (using the thin-slicing approach). As a secondary aim, we investigate the potential of a frailty assessment app with an agent by evaluating its usability and intention to use.

METHODS

Frailty Assessment App

The ECA under study was embedded within a frailty assessment web app developed as part of a larger platform designed to counter frailty by offering older adults training modules in the domains of healthy nutrition and physical and cognitive training to maintain a healthy lifestyle (*Noorman-de Vette, 2019*). Initial and continued use of the platform is stimulated by integrating gamification elements. In this study, we focused on the stand-alone frailty assessment app.

The frailty assessment app consisted of an index page (*Figure 04.1*) and a dialogue page (*Figure 04.2*). On the index page, an agent was displayed next to a blackboard. The blackboard provided a list of available dialogues: introductory small talk, questionnaire assessing aspects of the older adult's health, and small talk explaining the results of the questionnaires*. When a dialogue was finished, the user returned to the index page. Before the questionnaire dialogues were performed, only the introductory small talk was available on the blackboard. In this dialogue, users were introduced to the agent and the goal of the frailty assessment. Afterward, the questionnaire dialogues were unlocked and shown on the blackboard. Three validated questionnaires were implemented to assess the older adult's frailty status covering multiple health domains. The 36-item Short-Form Health Survey (*van der Zee & Sanderman, 1993*) contains 36 multiple-choice questions related to health topics (e.g., physical functioning, social functioning). The Alzheimer Disease Detection (*Galvin et al., 2007*) tests for functional decline in memory using 8 yes or no items. The Mini Nutritional Assessment (*Rubenstein et al., 2001*) tests for malnutrition with 6 multiple-choice questions related to nutrition and weight. We translated the three frailty assessment questionnaires into dialogues between the agent and older adults. When the original questionnaire contained a question with multiple sub questions, we changed these sub questions into multiple questions that were all asked in a separate dialogue step (as a human would have done). After the questionnaires were completed, the result dialogue was unlocked on the blackboard. In this dialogue, users received the outcomes of the assessment.

Only one dialogue was available at a time. Clicking on the start button of a dialogue opened the dialogue page (*Figure 04.2*). A dialogue consisted of multiple dialogue steps. Each dialogue step consist-

* Dialogues available upon request.



Figure 04.1 – Frailty assessment app: opening page introducing agents Sylvia and Egbert.

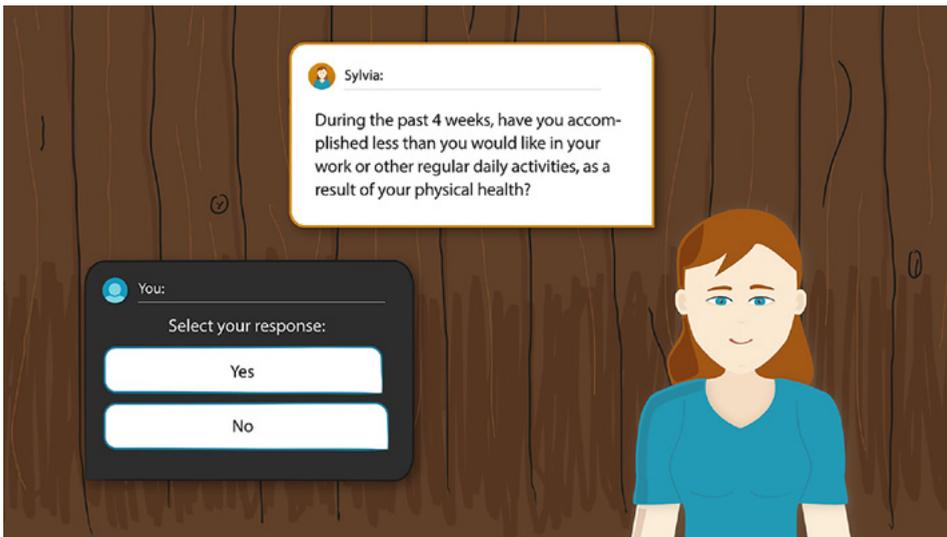


Figure 04.2 – Dialog page with peer agent Sylvia.

ed of a statement by the agent and one or more reply options that could be selected by the user. The statement by the agent was shown in the white box with the orange border and the reply options for the user were listed in the black box. After finishing a dialogue with the agent, the user returned to the index page and available dialogues listed on the blackboard were updated.

The agents used in the frailty assessment app (Figure 04.3) are Sylvia, a young female peer agent, and Egbert, an older male peer agent. By a peer agent, we mean an agent who is not a medical expert. Agent designs were selected based on findings from a previous study, in which the static images of eight agents were evaluated. The agent images differed on three features: age (young or old), gender (male or female), and role (experts had a high level of health expertise, and peers had a low level of health expertise). In an online questionnaire, images of all agents were shown to the participant at once, with participant selecting agent they preferred most (to be their health coach) at first glance. Afterward, participant rated characteristics for each agent. Results showed that a young female agent was preferred most and an older male agent was preferred least in both a general and elderly population (i.e., these designs were extremes in terms of user preference). This study builds on the previous study by evaluating users' impressions of these two agents, both at first

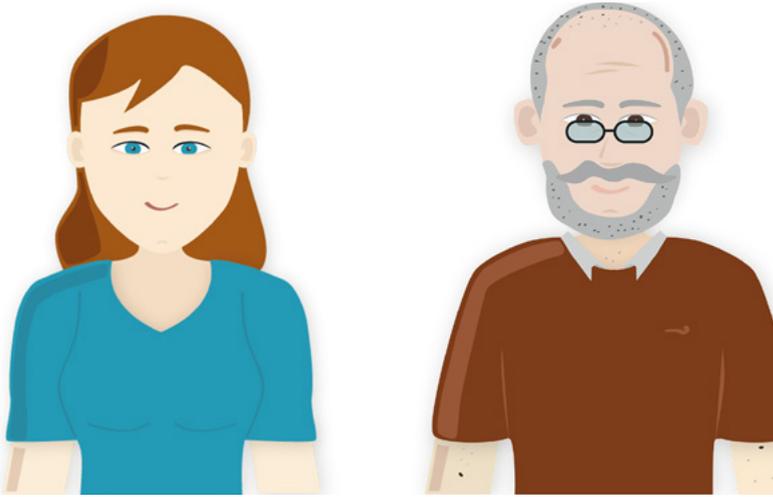


Figure 04.3 – Agents used during the experiment.

glance and after a short interaction with the agents. A blinking eyes animation was implemented for both agents. In addition, when the agent spoke (i.e., when a new dialogue step was loaded), a mouth animation of a fixed duration was played.

Study Design

We applied a within-subject design in which we counterbalanced the order in which agents were presented to participants. Half of the participants started the frailty assessment with the young, female peer agent and finished with the older, male peer agent (**Figure 04.4**, top). The other half of the participants were first presented with the older male peer agent, followed by the young, female peer agent (**Figure 04.4**, bottom). The study was performed in a lab setting, taking place either at a research institute or a local physiotherapy practice. The nature of this general study with healthy volunteers from the general population does not require formal medical ethical approval according to Dutch law. All participants provided their informed consent.

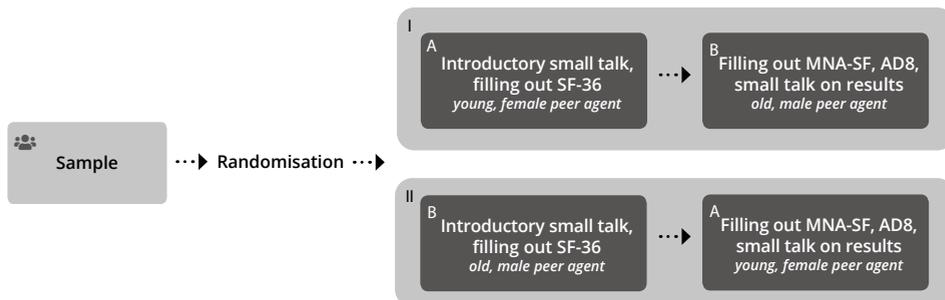


Figure 04.4 – Study design including randomization process.

Participants

Participants should be aged 65 years or above and fluent in the Dutch language in order to be included. In addition, they should be cognitively able to work with an ECA as assessed via the Mini-Mental State Examination, scoring at least 23 out of 30 points (Kok & Verhey, 2002). We recruited the respondents via a Dutch panel of adults that indicated they were interested in participating in research on eHealth. Participants were also recruited via a local physiotherapy practice.

Measurements

Questionnaires

Before interacting with the frailty assessment app, the participant completed the pre-interaction questionnaire gathering the participant's gender, date of birth, education, housing status, technology literacy, health literacy, and state of change for nutrition and physical activity (Prochaska & Velicer, 1997). After interacting with each agent (Figure 04.4), the participant completed the post-interaction questionnaire (Appendix C). To investigate the effect of the agent's appearance, we assessed the following:

- Likelihood of following the agent's advice (on a 7-point Likert scale)
- Agent characteristics ratings (all on 7-point Likert scales): friendliness, authority, involvement, reliability, intelligence
- Agent rapport scale rating (all on 7-point Likert scales) by Acosta and Ward (2011): emotional rapport, cognitive rapport, helpfulness, trustworthiness, likeability, naturalness, enjoyableness, human-likeness, persuasiveness, recommend-ability

Secondarily, we investigated the usability of the frailty assessment app and the intention to use the frailty assessment app on a single 7-point Likert scale.

Thinking Aloud

In order for us to triangulate the quantitative data, participants were asked to think aloud while interacting with the frailty assessment app. Audio was recorded and screen captures were taken. The researcher did not help or support the participant but only reminded the participant to think out loud when necessary.

Interviews

At the end of the session, the participant was interviewed. The interview was semi-structured and guided by asking the user's opinion regarding positive and negative aspects around the effect of the agent's appearance, usability of the frailty assessment app, and intention to use the frailty assessment app.

Data Analyses

SPSS Statistics 25 (IBM Corporation) software was used to perform statistical analyses. Since the underlying data were non-parametric, for all relations testing differences between the two agents, a Wilcoxon signed-rank test was conducted. All tests used a 95% confidence interval. All variables were tested for statistically significant differences between the two agents by means of a model consisting of Wilcoxon signed-rank tests for cross-over designs. Effect size was calculated by $r=Z/\sqrt{N}$, using 0.1, 0.3, and 0.5 as cut-off values for a small, medium, and large effects, respectively. The audio recordings of the thinking aloud sessions and interviews were transcribed and inductive-

ly thematically analysed. In addition, screen captures of the interaction with the frailty assessment app were aligned with the audio recordings. This way, the screen captures were used to verify the thoughts of the participants on the audio recordings. All themes were coded using ATLAS.ti 8 (ATLAS.ti Scientific Software Development GmbH) based on an empirical method proposed by Mays and Pope (1995). One researcher (StS) created a first coding scheme based on the data and then labelled the transcripts. A second researcher (MB) used the coding scheme to code a subset of the data so that a discussion could be held between the first and second coder for improving the coding scheme. The procedure of creating a first coding scheme, labelling the data by two researchers, and discussing the coding scheme was repeated a second time leading to a final coding scheme. The final coding scheme was used by the first coder to code all data for final analyses. The final coding scheme contained the following codes: agent characteristics, appearance agents, interaction with agents, preference agent, content questionnaires, language usage in dialogues, presentation information, interaction with app, design, navigation, general computer interaction, and intention to use.

RESULTS

Participants A total of 21 participants began the study (**Table 04.1**). One participant was not able to complete the protocol due to a lack of computer experience and was excluded. The average age of participants was 72.2 (SD 3.5) years, and 13 males and 7 females participated. Ten participants started with the young, female agent, and ten participants started with the older, male agent.

Agent Appearance

Ratings Questionnaire

Table 04.2 shows the questionnaire results regarding (1) the likeliness of following the agent's advice, (2) users' perceptions of the agent characteristics (e.g., friendliness, expertise), and (3) users' feeling of rapport (e.g., emotional rapport, helpfulness) for both agents. Corresponding box plots can be seen in **Figure 04.5**, **Figure 04.6** and **Figure 04.7**. For the ratings of the likeliness of following the agent's advice, no significant difference between Egbert and Sylvia was found. However, Egbert was rated significantly more authoritative than Sylvia ($p=0.03$), resulting in a medium effect size ($r=0.344$). No significant differences were found between the agents for all other agent characteristics and the rapport scale items. Analysis of the thinking aloud sessions and interviews resulted in the following themes on the effects of agent appearance: agent characteristics, agent appearances, interaction with the agents, and agent preferences.

Users' Perceptions of Agent Characteristics

A few participants indicated they had trouble getting an impression of the agents' personalities or found it difficult to connect personality to ECAs in general. A few others perceived the agents as natural and not artificial. On the other hand, the majority did not perceive the agents as human: they perceived the agents as cartoons, static dolls, computers, or machines.

Table 04.1 – Participant demographics (n=20).

Demographic		Value, n (%)
<i>Education</i>	Elementary school	1 (5)
	High school	1 (5)
	Vocational education	8 (40)
	College	6 (30)
	University	4 (20)
<i>Living situation</i>	Living alone	1 (5)
	Living with a partner	19 (95)
<i>Stage-of-change nutrition</i>	Maintenance	18 (90)
	Precontemplation	2 (10)
<i>Stage-of-change physical activity</i>	Maintenance	13 (65)
	Action	3 (15)
	Contemplation	1 (5)
	Precontemplation	2 (20)
	Unknown	1 (5)
Technology literacy level	Moderate or high	20 (100)
<i>Health literacy level</i>	Moderate or high	19 (95)
	Low	1 (5)
<i>Physical limitations</i>	No risk of facing physical limitations	9 (45)
	Risk of facing physical limitations	10 (50)
	Already faced physical limitations	1 (5)
<i>Cognitive limitations (Mini-Mental State Examination)</i>	No risk of facing cognitive limitations (score ≥ 23)	19 (95)
	Risk of facing cognitive limitation (score < 23)	1 (5)

Table 04.2 – Results of the Wilcoxon signed-rank tests (n=19 or 20) comparing the mean ranks of the ratings of likeliness of following the agent’s advice, agent characteristics, and rapport scale items.

Characteristic	Median Egbert (Q1-Q3)	Median Sylvia (Q1-Q3)	z score	p
Likeliness of following advice	5.0 (3.3-6.0)	6.0 (4.0-6.0)	-1.613	0.11
<i>Agent characteristics</i>				
Friendliness	6.0 (5.0-6.0)	6.0 (5.0-6.0)	-0.264	0.79
Expertise	5.0 (4.0-6.0)	5.0 (4.0-6.0)	-0.966	0.33
Reliability	5.0 (4.0-6.0)	5.0 (4.0-6.0)	-0.276	0.78
Authority	2.0 (2.0-4.0)	2.0 (1.0-4.0)	-2.121	*0.03
Involvement	4.5 (4.0-6.0)	5.0 (4.0-6.0)	-0.158	0.88
<i>Rapport scale</i>				
Emotional rapport	4.0 (2.0-5.0)	4.0 (3.0-5.0)	-1.310	0.19
Cognitive rapport	4.0 (4.0-5.0)	5.0 (3.3-5.8)	-0.829	0.41
Helpfulness	5.0 (4.0-6.0)	5.0 (4.0-6.0)	-0.877	0.38
Trustworthiness	5.0 (4.0-6.0)	5.0 (4.0-6.0)	0	>0.99
Likeability	6.0 (4.0-6.0)	6.0 (4.3-6.0)	-0.604	0.55
Naturalness	5.0 (4.0-6.0)	5.0 (4.0-6.0)	-0.491	0.62
Enjoyability	5.0 (3.0-6.0)	4.0 (4.0-6.0)	-0.182	0.86
Human-likeness	4.0 (3.3-6.0)	4.5 (3.3-5.0)	-0.486	0.63
Persuasiveness	5.0 (4.0-6.0)	5.0 (4.0-6.0)	-0.942	0.35
Recommendability	5.0 (4.0-6.0)	5.0 (4.0-6.0)	-0.368	0.71

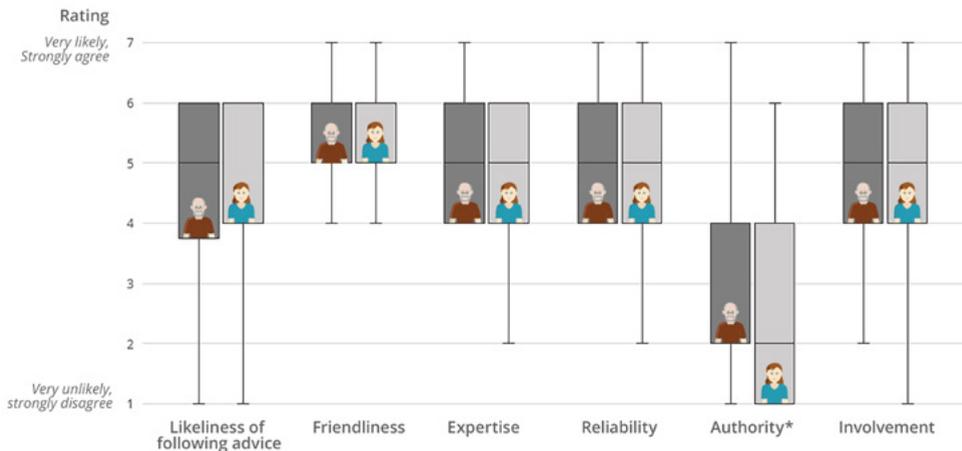


Figure 04.5 – Ratings of the likeliness of following advice and characteristics of the two agents (p < 0.05).

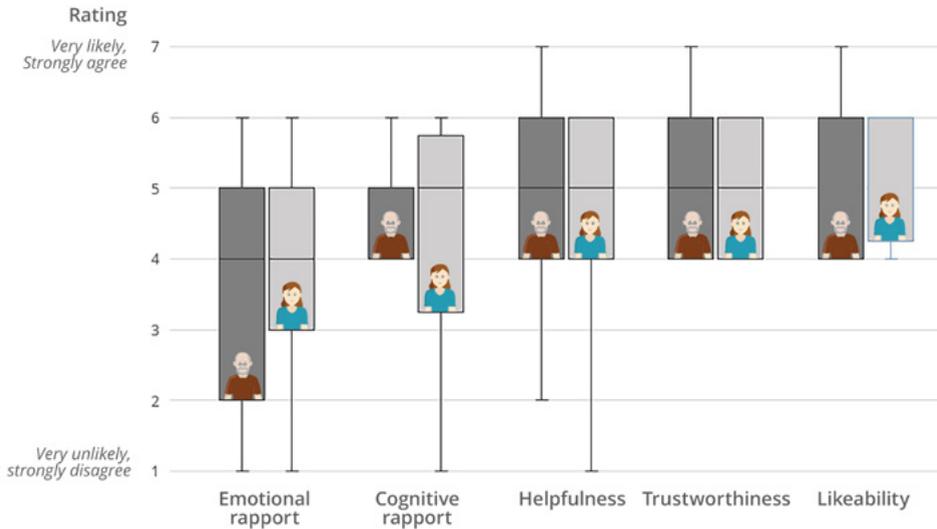


Figure 04.6 – Ratings of the rapport scale items emotional rapport, cognitive rapport, helpfulness, trustworthiness and likeability of the two agents.

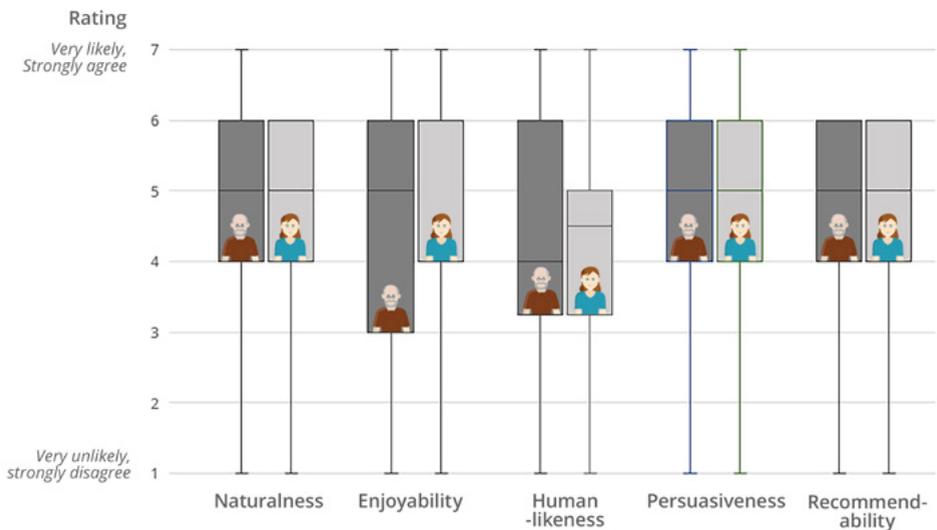


Figure 04.7 – Ratings of the rapport scale items naturalness, enjoyability, human-likeness, persuasiveness and recommendability of the two agents.



It is a computer, it is still interaction from a distance, it does not become personal, it does not have any personality, I do not feel a connection.

— Male, 68 years —



The agents remain computers, you cannot call them friendly or unfriendly, they are computers and I do not connect any human characteristics to them.

— Male, 78 years —

In the interviews, some participants indicated they did not perceive the agents differently with respect to their personality. A few participants explained that both agents used friendly language, whereas others argued the agents were friendly, since they responded in a way that fit the situation and provided compliments. In addition, a participant explained that both agents were not too young or too old and seemed to be modern people due to responses such as “Gosh, how nice.” Also, this participant said he liked that the agents were not too young, since a young agent would not have much experience. One participant particularly indicated that the female agent was friendlier than the male agent, whereas another participant believed that the male agent was more highly educated and more intellectual than the female agent.

Users’ Perceptions of Agent Appearances

A participant indicated that the agents looked like cartoons or drawings, whereas she preferred the agents to look like real humans. This participant also indicated that the blinking eyes and mouth animation were distracting.

The rest of the comments related to the appearance of either one of the agents. One person particularly mentioned the female agent having a friendly face, whereas all other comments related to the male agent. The appearance of the older male agent evoked several associations, such as the agent looking old, and, therefore, unhealthy. Others associated the older male agent with a scientific staff member, a nerd or a male of the type of wearing sandals with socks, because of his glasses and trendy beard. Participants preferred an energetic, spontaneous person and one that is more neutral and clean-shaven. One participant did not like the male agent, because he associated the agent with his or her uncle, having a similar name: a spoiled man with whom you would not be able to connect. Another participant found the male agent more distracting than the female agent, because of his glasses.

Users’ Perceptions of Interaction With the Agents

Several participants explicitly indicated that they expected or would like the agent to speak. One participant expected the agent to speak due to its mouth animation, whereas another had this expectation, since humans interact via speech in real conversations. Another participant pointed out that, due to the absence of agent speech, the user has to multitask: the user simultaneously has to read and answer the questions and pay attention to the agent. Therefore, she would like the agent to speak.

“

Well, I have to read what you say to me, but instead open your mouth yourself!

– Female, 73 years –

Other opinions on the interaction with the agent focused on the naturalness of the interaction.

“

It felt as if there was a real human in front of me.

– Female, 71 years –

Another participant described the interaction as actually talking to someone, and yet another participant described the interaction as having a phone call, in which someone is checking how you are doing. Some participants were less positive. A few participants specifically said that the interaction with the agents was impersonal.

“

Actually, I do not have the feeling I am really communicating with someone.

– Female, 65 years –

Another participant said that she did not take part in a conversation but was simply reading and answering questions. This participant did not establish a connection with the agents.

“

I barely know her.

– Female, 65 years –

“

Understanding each other? Then one would expect interaction.

– Female, 65 years –

Last, some comments related to the implemented small talk. On the one hand, some participants seemed to like the small talk, reflected by them laughing. On the other hand, a participant was irritated by the implemented small talk, she felt being treated like a child.

Agent Preference

The majority of the participants indicated they did not prefer one agent over the other. Most of them indicated they did not have a preference, since they perceived the agents to be similar. Some did not even remember they interacted with two different agents. However, some participants did show a preference. Most participants preferred the female agent, either because they believed she was friendlier or discussed a more interesting topic. Only one participant preferred the male agent but could not say why.

Usability and Intention to Use Frailty Assessment App

Questionnaire results show that the usability of and intention to use the frailty assessment app were high: the 20 usability ratings displayed a median of 6.1 (interquartile range [IQR] 6.1-7.0) and the 20 intention-to-use ratings displayed a median of 6.0 (IQR 4.0-6.0) on a 7-point Likert scale.

During the thinking aloud session and interviews, participants pointed out usability issues of the frailty assessment app or provided suggestions for improvements to the app. The following themes were identified: content questionnaires (mentioned 107 times), language usage in dialogues (mentioned 41 times), presentation information (mentioned 21 times), interaction with app (mentioned 14 times), design (mentioned 7 times), navigation (mentioned 7 times), and general computer interaction (mentioned 6 times).

Most comments or suggested improvements related to the content of the questionnaires and the language in the app. The majority of the participants reported that the questionnaires did not fit their personal situation and contained a lot of repetition or ambiguity. Participants suggested adapting the questionnaires according to previous answers given. In addition, participants commented on the language used: words being ambiguous, too trendy or too old fashioned, unnecessary, patronizing, or not being known by people with a lower education or older adults. Furthermore, participants commented on the length and structure of the sentences and pointed out spelling mistakes. A participant suggested adapting the language in the app to the education of the user. Considerably fewer comments related to the presentation of information, interaction with the app, design or navigation of the app, and general computer interaction. As an example, with respect to navigation, some participants indicated they would like to be able to go back to a previous dialogue step.

With respect to the intention to use, the thinking aloud sessions and interviews showed that a minority of the participants would like to use the app. A participant indicated he would not use the app but would recommend the app to others who might benefit from it. In addition, some participants clearly indicated they would not use the app. The majority of the participants indicated that the agents did not add any benefit to the app, arguing that the app was not personal since answer options were limited and the opportunity to explain them was missing. A participant stated that for the app to be beneficial, it should also provide advice on what actions the user should perform to become more healthy. Another participant explicitly stated that he would use the app when the text was replaced by speech.

DISCUSSION

Principal Findings

Our results show that the appearance of an agent, in particular age and gender, affects users' perceptions of agent authority but does not affect users' perceptions of other agent characteristics, users' feelings of rapport, or users' likeliness of following agent advice. Compared with a young female agent, an older male agent is only seen as more authoritative. These results are not in line with our expectation that agents are perceived differently after a short interaction with a user. To the best of our knowledge, there is no existing research comparing users' impressions of agents at first glance with those after short interactions. But research shows that in human-human interaction, first impressions, formed within milliseconds (*Bar et al., 2006*), are difficult to lose. Therefore, we assumed that the differences in perceptions of characteristics of a static image of a young female agent and an older male agent, as found in a previous study (*chapter 3*), would still be present after a short interaction with these agents. An explanation for this inconsistency could be that impressions in human-agent interaction differ from impressions in human-human interaction. Users' judgments of agents may modify with ongoing interaction, as research shows that agents do have a second chance to make a first impression (*Bergmann et al., 2012; Komatsu et al., 2011*). Therefore, differences in perceptions of both agents may have been present at first glance but disappeared after interaction. Another explanation for the inconsistency between the results of this study and those of our previous study is that users' perceptions of the ECAs might have been influenced by the task of the ECAs: in our previous study, the ECAs were positioned as health coaches, whereas in this study the ECAs did not perform coaching (they only asked health-related questions, needed to perform personalised coaching). Further research is needed to confirm this finding. Future research could study users' perceptions of agent characteristics with a larger study population. Eventually agents will be used in a long-term setting; therefore, it is interesting to research not only users' perceptions at first glance and after short-term interaction, but also after long-term interaction.

How do we explain the difference in perceptions of agent authority after a short interaction? Although research on short-term interaction with an agent indicates that an agent's appearance, including clothing (*Parmar et al., 2018*), racial concordance with the user (*Zhou et al., 2014; 2017*), and similarity with the user (*Zhou et al., 2014; Wissen et al., 2016*), could affect users' perceptions of the agent, to the best of our knowledge there is no research on agent authority after short interaction in particular. From a previous study (*chapter 3*), we see that at first glance, static images of male and older agents are indeed seen as more authoritative than female and young agents, respectively. In addition, the study shows that the differences found in authority are often higher compared with differences found for other characteristics tested, which could explain why the difference in authority level is still present after short interaction. However, since we did not control the age and gender of the agents in this study independently, it is difficult to say whether the difference in perception of agent authority is caused by agent age or gender in particular or solely by the combination. Future research could study which factors actually control the difference, researching users' perceptions of agent authority by independently controlling the age and gender of the agents. In addition, future research could study how an agent's authority is perceived after long-term interaction.

We expect that the effect of the first impression established by agent age and gender on the impression after short interaction is small compared with the effect of other design features, such as the content and language of the messages, (absence of) agent speech, and the amount of embodiment. Our study shows that the majority of participants perceived the agents not as humans but as machines or cartoons and found interaction with the agents impersonal or artificial. They did not have the feeling of being in a conversation. These perceptions may indicate users had a negative adaptation gap (Komatsu et al., 2012), which occurs when a user overestimates the competency of an agent, creating a negative gap between expected and actual competency of the agent and resulting in the user being disappointed. This negative adaptation gap may have been caused by the content and language of agent messages, agents lacking speech, or agents having little embodiment, as supported by remarks made by participants during the thinking aloud sessions and interviews. Therefore, we believe it is important to manage users' expectations of agent characteristics and functionality up front, ensuring users' expectations match actual agent capabilities by explaining what the users can expect from the agent. Future research could study how an agent's content, language, speech, and embodiment affect users' perceptions of the conversation with the agent (e.g., how these factors could make the conversation with an agent more human-like).

Although our study shows agent age and gender have little effect on users' impressions of the agent after short interaction, we believe that adapting these features to the user is important because they affect users' impressions of the agent at first glance (Forlizzi et al., 2007; Nguyen & Masthoff, 2007; Zimmerman et al., 2005), and research shows that people with favourable impressions of someone tend to interact more with that person than they do others who gave unfavourable impressions (Kelley, 1950). Selecting an agent with the right age and gender could thus lower the threshold to interact with the agent and use the app.

Second, our results show that usability of the developed frailty assessment app was judged positively overall; issues identified by participants related to the content or language of the questionnaires. We suggest tailoring the content and language toward the personal characteristics of the user, as confirmed by existing research (Beukema et al., 2017), and adapting the content to previous answers given by the user.

Third, not all participants show an intention to use the app. Research indicates that older adults put effort into learning new digital technologies as long as they are believed to be worthy of time and dedication (e.g., when technology can be used to keep in touch with others to foster relationships (Lindley et al., 2009). Similarly, research shows that the elderly value apps that address a social problem (Waycott et al., 2016). The app used in our study did not address a social problem, which could have resulted in some participants not seeing the added value of the app and not showing an intention to use the app. In addition, intention to use digital technologies in elderly persons is, next to the quality of the technology itself, affected by their personal context (e.g., their ability to concentrate) and social context (e.g., whether family is around to provide technical support) (Waycott et al., 2016). Both factors might have affected participant intentions to use the frailty assessment app in our study.

More specifically, the majority of participants do not believe the agent adds value to the frailty assessment app. Therefore, we suggest updating the design of the agent. We believe that the agent should convey additional information to its message in text via its embodiment. Existing research provides evidence for implementation of animations of the agent's embodiment, showing that animations positively affect users' impressions of the agent (Baylor & Ryu, 2003; Kang et al., 2015; Cowell & Stanney,

2003) and interaction time (Bergmann et al., 2012; Kang et al., 2015). In addition, the use of speech is recommended because it could increase the sense of personality of an agent (Nass & Lee, 2000) and could be used to describe feelings (Veletsianos et al., 2009). Low-literate users could benefit from multiple output modalities (Thies, 2015). Furthermore, participants indicated they would like the app to provide advice on what actions they should perform in order to become more healthy. We see an opportunity for using the agent to provide this advice. As an example, the agent could show videos of exercises to improve physical strength.

Strengths and Limitations

This is the first study that specifically evaluates effects of agent appearance after short interaction with the agent. In addition, this study uses actual health content, which is scarce in research on agent design. Our study also has some limitations. First, the negative adaptation gap between user expectations of agent capabilities and actual agent capabilities suggests the app used might not have been mature enough. The agent conveyed the majority of the information via text. Participants might have been focused on reading the text and therefore paid little attention to the agent, resulting in participants having difficulties in creating impressions of agent characteristics and establishing rapport. Second, interaction time with the agents might have been too short to create impressions of agent characteristics and establish rapport. Third, although we found a difference in users' perceptions of authority of the young female and the older male ECA, it is difficult to identify whether this was caused by the ECA's gender or age (or a combination), since these factors were not independently controlled in the study. Also, other visual features that were different between the two ECAs (e.g. glasses or no glasses) might have had influence.

Toward Digital Frailty Assessment With Embodied Conversational Agents: Recommendations for Future Research

Agent Design Implications

First, convey empathy or emotion using the agent's embodiment. This way, agent design can positively affect users' impressions of the agent and interaction time. Second, reduce the user's cognitive load by providing the agent messages in speech. This way, agent design can positively affect users' impressions of the agent. Third, select an agent appearance that fits the age and gender of the user. This way, agent design can lower the threshold to start using the app.

Prerequisites Frailty Assessment

First, take into account the user's personal situation, such as disabilities and living situation, and adapt the content. Adapt the questionnaire so users do not see questions that do not apply to their situation. Second, save the answers given by the user, and adapt the questionnaire accordingly. This way, users do not have to answer questions that are not applicable to them. Third, adapt the agent's language based on the educational level of user so the language is neither too simple nor too complex.

Conclusions

Our study shows that an agent's appearance, in particular age and gender, only affects users' perceptions of agent authority after short-term interaction. We conclude that adapting agent age and gender to users' preferences is important to lower the threshold to interact, whereas the content and language of the agent's messages and agent speech and embodiment are important factors for users' impressions of the agent after short interaction.

We believe that ECAs have potential to be used in digital frailty assessment, but future research is needed. Future research could study users' perceptions of agents after long-term interaction, whether users' perceptions of agent authority are related to agent age or gender in particular, and how an agent's content, language, speech, and embodiment affect users' perceptions of the conversation with the agent.

Acknowledgements – This work was supported by Interventions on Frailty and Ageing Risks for Elderly People Based on Information and Communication Technology Tools, funded by the Eurostars-2 Programme (no.10824) and the SPRINTT project (Sarcopenia & Physical fRailty IN older people: multi-component Treatment strategies); IMI1 - Call 9, project no. 115621.

05

Effects of ECA Emotion on Users' Perceptions of Rapport after Short Interaction

BASED ON:

ter Stal, S., Jongbloed, G., & Tabak, M. Embodied conversational agents in eHealth: How facial and textual expressions of positive and neutral emotions influence perceptions of mutual understanding (*submitted for publication*)



Abstract

Embodied Conversational Agents (ECAs) could engage users in eHealth by building mutual understanding (i.e. rapport) via emotional expressions, but little research exists on how these expressions affect rapport in this context. In this chapter, we compared an ECA's emotions expressed in text with an ECA's emotions in facial expressions on users' perceptions of rapport. We combined a happy or neutral facial expression with a happy or neutral textual expression, leading to four conditions. Sixty-three participants (mean 48 ± 22 years) had a dialogue with an ECA on healthy living and rated multiple rapport items. Results show no significant difference in overall rapport between the conditions ($p = 0.062$), but a happy textual expression for an ECA with a neutral facial expression shows higher ratings of the individual rapport items helpfulness ($p = 0.019$) and enjoyableness ($p = 0.028$). Future research should investigate users' rapport towards an ECA with different emotions in long-term interaction and how a user's age and personality and an ECA's animations affect rapport building. Optimising rapport building between a user and an ECA could contribute to achieving long-term interaction with eHealth.



INTRODUCTION

eHealth applications seem promising to support users in living healthy. However, most of the applications are used without human support, which makes it difficult to maintain users' engagement. To engage users, the use of Embodied Conversational Agents (ECAs) as health coaches is investigated. ECAs are more or less autonomous and intelligent software entities with an embodiment used to communicate with the user (Ruttkay et al., 2004). Evidence has been found that actively building mutual understanding, or 'rapport', between the ECA and the user can lead to higher user engagement (Gratch, Wang, Gerten, et al., 2007; Gratch, Wang, Okhmatovskaia, et al., 2007).

In research, various models of rapport exist. For example, *Tickle-deggen and Rosenthal (1990)* define rapport in terms of three dimensions: **attentiveness** (i.e. the focus of the interactants is directed towards each other and the interactants experience a sense of mutual interest), **positivity** (i.e. the interactants feel mutual friendliness and caring), and **coordination** (i.e. the interactants experience balance and harmony; they are "in sync"). *Gratch, Wang, Gerten, et al. (2007)* designed a model of "virtual rapport", being a model of rapport specifically for human-ECA interaction. The model splits rapport into three dimensions: **emotional rapport** – an inherently rewarding experience; a feeling of harmony, **cognitive rapport** – sharing of understanding between interactants; a convergence of beliefs or views, a bridging of ideas or perspectives, and **behavioural rapport** – convergence of movements between interactants; increased synchrony, fluidity and coordination in interactants' movements. *Novick and Gris (2014)* created a paralinguistic rapport model for human-ECA interaction, describing similar dimensions as the model by *Gratch, Wang, Gerten, et al. (2007)*, namely: **sense of emotional connection**, **sense of mutual understanding** and **sense of physical connection**.

According to *Tickle-deggen and Rosenthal (1990)*, one of the variables that strongly influences this so called 'rapport', is non-verbal behaviour. More specifically, when there are expressions of positive emotions, rapport can be built (*Tickle-deggen & Rosenthal, 1990*). Thus, we could hypothesise that expressions of positive emotions by the ECA could increase rapport between the ECA and the user and increase the user's engagement. An ECA can express these positive emotions in several ways, such as via speech, gaze behaviour and hand and body gestures. In addition, an ECA can communicate these emotions via **(1) textual expressions and (2) facial expressions**. First, humans are experts in recognising emotions based on facial expressions (*Ekman & Friesen, 1978*). Second, with respect to emotion in text, initially, it was suggested that within computer mediated textual communication, the lack of non-verbal cues might make it more difficult for individuals to recognise expressed emotions (*Short et al., 1976, as cited by Hancock et al. (2007)*). However, people might be more able to adapt to emotional expressions through computer mediated textual communication than previously suggested (*Walther, 1992; Hancock et al., 2007*).

Several studies have researched an ECA's emotional expressions. For example, studies investigated how to model an ECA's emotions in its facial expressions (*Bevacqua et al., 2007; Pelachaud, 2009; Llorach et al., 2019*). These studies are mainly on how to model and implement an ECA's emotions technically and are not primarily user studies. Other studies do focus on users' perceptions, researching effects of positive, negative and neutral emotional expressions of an ECA on students' social judgements, interest and self efficacy (*Kim et al., 2007*) and effects of an ECA's emotional facial expressions and tone of voice on students' emotional states (*Moridis & Economides, 2012*). Although these studies

show positive effects of an ECA's emotional expressions, they do not investigate effects of an ECA's emotional expressions on the relation of the ECA with its users.

Some studies do specifically investigate effects of an ECA's emotional expressions on this relation (such as rapport) of the ECA with its users. *Beale and Creed (2009)* present a structured overview of research into emotional simulation in ECAs. They show that an ECA's emotion (compared to no emotion) positively affects relationship aspects, such as the ECA's perceived caring, trust and likeability. *Loveys et al. (2020)* performed a literature review on an ECA's design features, including an ECA's emotion, and their effects on the quality of the relationship between the ECA and its users. They show that an ECA's emotion does affect the relationship with its users, amongst which rapport. Also, *Cerekovic et al. (2017)* researched two different emotional ECAs and user's perceptions of rapport towards these ECAs, showing that rapport building towards ECAs having different emotions is affected by a user's personality. These studies show that an ECA's emotional expressions might affect users' perceptions of rapport towards an ECA, however, they were not performed in the context of eHealth.

Only a few studies researched effects of an ECA's emotional expressions on the relation with the user in the context of eHealth specifically. *Creed et al. (2015)* describe that an emotionally expressive ECA, in the role of a nutritional coach, is perceived more likeable and caring than an unemotional ECA. In addition, *Amini et al. (2014)* report that an empathic ECA, assessing a user's alcohol dependence, scores higher on perceived sociability, social presence, trust and social influence than a non-empathic version and text-only interface. To the best of our knowledge, only one study (*Ranjbartabar et al., 2019*) explicitly researched effects of an ECA's emotional expressions in eHealth on a user's perception of rapport towards the ECA. *Ranjbartabar et al. (2019)* compared an empathic therapist with a neutral therapist. They show that users interacting with an empathic ECA will not necessarily build more rapport than those interacting with a neutral ECA, but that rapport building towards the ECA depends upon the emotional feelings the user is initially expressing towards the problem addressed by the ECA. Results of these two studies indicate that an ECA's emotional expressions might affect a user's perception of rapport towards an ECA in the context of eHealth, but they show too little evidence to conclude that this is indeed the case.

Thus, although some research has been performed on effects of an ECA's emotional expressions on a users' perception of the relation with the ECA, amongst which rapport, little is known about the effect of an ECA's emotional expression specifically on users' perceptions of rapport in the eHealth context.



Research Objectives

Therefore, this research aims to compare an ECA's emotions expressed in text with an ECA's emotions in facial expressions on users' perceptions of rapport in the context of eHealth.

The expression of positive emotions (e.g. happiness) is compared to the absence of emotions (e.g. 'neutral' emotions). The hypotheses to be tested in this research can be seen in **Table 05.1**.

Table 05.1 – Hypotheses to be tested in this research.

Hypothesis	Description
H1	Users' perceptions of rapport towards an ECA with happy textual expressions are higher than towards an ECA with neutral textual expressions.
H2	Users' perceptions of rapport towards an ECA with happy facial expressions are higher than towards an ECA with neutral facial expressions.
H3	Users' perceptions of rapport towards an ECA with both happy textual and happy facial expressions are higher than towards an ECA with only happy facial or happy textual expressions.

METHOD

Study Design

We researched effects of two independent variables 'type of facial expression' and 'type of textual expression' on perceptions of rapport towards an ECA (dependent variable). For both the facial expression and the textual expression, two types of emotions were examined: (1) neutral, and (2) happiness. As shown in **Table 05.2**, this leads to four conditions to be examined. Participants were randomly assigned to one of the four conditions, instead of being exposed to all four conditions in a row, to prevent their perceptions from being biased by seeing either four times the same dialogue, or four different dialogues. We researched the perceptions of rapport using an online questionnaire in which users interacted in a dialogue with one of the four ECAs.

Table 05.2 – Participants are randomly assigned to one of the four conditions.

		Textual expression	
		Neutral text	Happy text
Facial expression	Neutral face	NFNT	NFHT
	Happy face	HFNT	HFHT

The Dialogues

The topic of the dialogue with the ECA, Laura, was gaining a healthy lifestyle. More specifically, the ECA introduced herself, explained the role of an ECA and the topic of the dialogue, to finally work towards concrete tips on physical activity and healthy nutrition. The dialogues were presented to the participants in Qualtrics, an online survey tool. Every dialogue consisted of multiple dialogue steps. As seen in **Figure 05.1**, each dialogue step consisted of the ECA with its message (presented as one static image in a question in the survey tool) and a set of response options for the user (presented as responses in the survey tool).

The ECA's Facial Expressions

For the ECA's facial expressions, the Facial Action Coding System (FACS) developed by *Ekman and Friesen (1978)* was used. The FACS classifies multiple action units that, when combined, lead to facial expressions of emotions. Examples of these action units are "Frontalis, pars medialis" (raising inner corner of eyebrow), and "pterygoids and digastric" (mouth stretched open). The ECA's happy and neutral facial expressions were developed using the action units as classified by the FACS. For a neutral expression, no muscle contractions take place that lead to an expression of emotion, therefore no action units were active for this expression.

For the happy facial expression, action unit 6 ("Orbicularis oculi, pars orbitalis" – Cheeks raised, eyes narrowed) was combined with action unit 12 ("Zygomaticus major" - Lip corners pulled up and laterally), a common combination to express happiness (*Kanade et al., 2000*). **Table 05.3** shows an image of both an ECA with a happy and neutral facial expression. The ECA's facial expression was static during the complete dialogue (i.e. the ECA was displayed as an image).

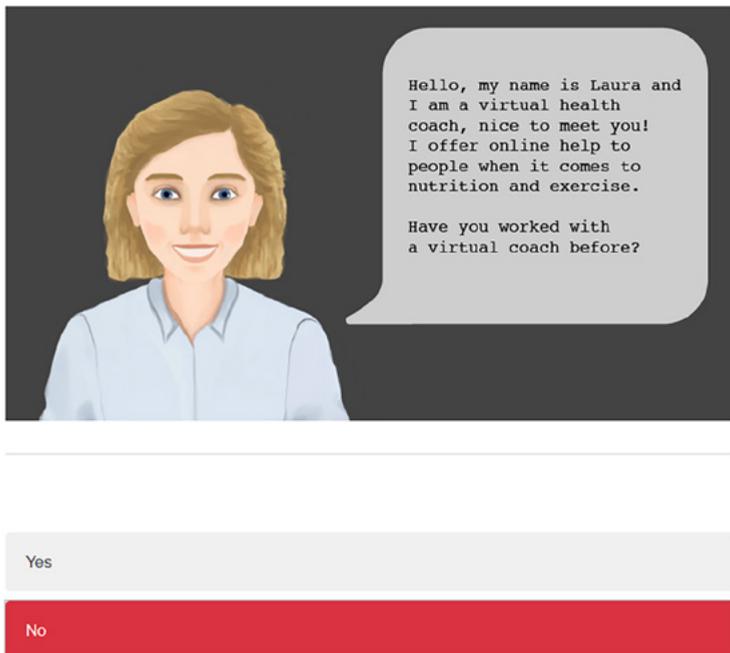


Figure 05.1 – Example step of the dialogue between the ECA and the participant. Every dialogue step consisted of an image of the message presented by the ECA (as a question in the survey tool) and a set of response buttons for the participant (as responses in the survey tool).

The ECA's Textual Expressions

As suggested by *Hancock et al. (2007)*, positive emotions as opposed to negative emotions can be added to text by implementing: (1) more words, (2) fewer negative affect terms, (3), fewer negotiations, (4) fewer assents, (5) less agreement, and (6) more punctuation and exclamation marks. Therefore, compared to the neutral textual expressions, the happy textual expressions contained more words, less negative affect terms, more positive affect terms and more punctuation and exclamation marks.

Table 05.3 shows an example of both a neutral and happy textual expression of the ECA.

Table 05.3 – The differed facial and textual expressions used in this study.

	Neutral expression	Happy expression
Facial expression		
Textual expression (example text)	<p>💬 "Today we are going to discuss your health. If you have a question about your health, you can ask me."</p>	<p>💬 "Today we are going to discuss your health. I can imagine that you want to have a healthy lifestyle. That is why I am here to answer all your questions about healthy nutrition and exercise. I am happy to help!"</p> <p>Difference with respect to neutral: More words, less negative affect terms, more positive affect terms, more punctuation and exclamation marks</p>

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 5 weeks, in October and November 2018. Participants gave an informed consent prior to the questionnaire.

Measurements

The following data were collected via the online questionnaire (see also [Appendix D](#)):

- › **Respondents perceptions of rapport with the ECA** – Measured using an adaptation (*Acosta & Ward, 2011*) of the 'Rapport scale' originally created by *Gratch, Wang, Okhmatovskaia, et al. (2007)*. The scale covers ten different dimension of feelings of rapport: emotional rapport, cognitive rapport, helpfulness, trustworthiness, likeability, naturalness, enjoyableness, human-likeness, persuasiveness, recommend-ability (all rate on 7-point Likert scales ranging from strongly disagree to strongly agree)
- › **Characteristics of respondents** (age, gender and education and technology use)

Procedure

In the introduction to the questionnaire, the context of the research was explained and informed consent was obtained. Users were asked to imagine being supported by a virtual coach on healthy living. Then, the respondent was randomly assigned to one of the four conditions. The respondent engaged in a short dialogue with a ECA, after which he or she rated the ECA's rapport by statements on 7-point Likert scales. Lastly, participants filled out some questions regarding their demographics. The content of the questionnaire can be seen in [Appendix D](#). In total, the questionnaire took around 15 minutes.

Data Analyses

SPSS 25 was used to perform statistical analyses. The respondent's overall rapport towards the ECA was calculated by taking the mean of all ten rapport items. To test for a difference in mean of overall rapport between the four conditions, an ANOVA between subjects test was conducted. Differences were further analysed by performing Bonferroni post-hoc tests. MANOVA between subjects test was conducted to test for a difference in mean value for every individual rapport item across the four conditions. Again, the differences were further analysed by performing Bonferroni post-hoc tests. All tests used a 95% Confidence Interval.

RESULTS

Participants

A total of 63 respondents filled out the complete questionnaire. All questionnaires were filled out in Dutch. Of these respondents, 27 were male and 36 were female. The age of the respondents ranged from 20 to 88 years (mean 48.22 years). The respondents were distributed over the conditions as follows: NFNT: 15 respondents (mean age 54 years), HFNT: 16 respondents (mean age 42 years), NFHT: 19 respondents (mean age 44 years), HFHT: 13 respondents (mean age 57 years). An overview of all characteristics of the respondents can be seen in [Table 05.4](#).

Table 05.4 – Characteristics of the respondents to the questionnaire. Numbers represent number of respondents per condition.

Characteristic		Neutral face, neutral text (n)	Happy face, neutral text (n)	Neutral face, happy text (n)	Happy face, happy text (n)
<i>Gender</i>	Male	7	8	7	5
	Female	8	8	12	8
<i>Education school</i>	Elementary school	-	1	-	1
	High school	3	1	5	2
	Vocational education	1	3	3	1
	College	4	6	2	2
	University	6	5	9	7
	Unknown	1	-	-	-
	<i>Technology usage</i>	Smartphone	15	15	18
	Smartwatch	2	3	1	2
	Tablet	11	7	10	7
	PC/laptop	14	15	17	11
	Game console	2	2	-	1
	Other	2	2	-	1

Table 05.5 – Results of the ANOVA comparing the mean values for overall rapport of the four conditions.

	Neutral face, neutral text, mean (SD)	Happy face, neutral text, mean (SD)	Neutral face, happy text, mean (SD)	Happy face, happy text, mean (SD)	F-value	p
<i>Overall rapport</i>	3.687 (1.087)	4.394 (0.733)	4.574 (0.843)	4.092 (1.245)	2.577	0.062

Overall Rapport

Table 05.5 shows the results of the ANOVA comparing the mean values for overall rapport of the four conditions. Results show no significant difference ($p = 0.062$). To validate that there is not a significant difference in means between any of the conditions, a Bonferroni post-hoc test was performed; again it appears that there is no significant difference in means.

Individual Rapport Items

Table 05.6 shows the results of the MANOVA comparing the mean values for the individual rapport items of the four conditions. A significant difference in mean value for the ECA's helpfulness ($p = 0.024$), enjoyableness ($p = 0.032$) and persuasiveness ($p = 0.021$) was found. These differences are further analysed by performing Bonferroni post-hoc tests, as can be seen in **Figure 05.2**. These tests show that:

- › There is a significant difference between the perception of the ECA's helpfulness of NFNT and NFHT ($p = 0.019$). The presence of a happy textual expressions for an ECA having neutral facial expressions shows a higher score for the participants' perception of the ECA's helpfulness.
- › There is a significant difference between the perception of the ECA's enjoyableness of NFNT and NFHT ($p = 0.028$). The presence of a happy textual expression for an ECA having neutral facial expression shows a higher score for the participants' perception of the ECA's enjoyableness.
- › There is no significant difference between the perception of the ECA's persuasiveness for the four conditions. Emotion in the textual or facial expression of the ECA, thus, does not seem to affect the perception of the ECA's persuasiveness.

Table 05.6 – Results of the MANOVA comparing the mean values of the individual rapport items for the four conditions.

	Neutral face, neutral text, mean (SD)	Happy face, neutral text, mean (SD)	Neutral face, happy text, mean (SD)	Happy face, happy text, mean (SD)	F-value	p
<i>Emotional rapport</i>	3.13 (1.552)	3.69 (1.195)	3.74 (1.147)	3.69 (1.548)	.695	.559
<i>Cognitive rapport</i>	3.53 (1.598)	4.00 (1.211)	4.11 (1.243)	4.23 (1.641)	.685	.565
<i>Helpfulness</i>	4.53 (1.302)	5.06 (.998)	5.79 (.918)	4.92 (1.553)	3382	.024*
<i>Trustworthiness</i>	4.27 (1.223)	4.75 (1.065)	5.42 (1.216)	4.62 (1.446)	2649	.057
<i>Likability</i>	4.07 (1.624)	5.12 (1.204)	5.11 (.994)	5.08 (1.320)	2487	.069
<i>Naturalness</i>	3.00 (1.604)	3.75 (1.438)	3.63 (1.257)	3.23 (1.481)	.921	.436
<i>Enjoyableness</i>	3.53 (1.302)	4.50 (1.033)	4.74 (1.284)	4.15 (1.068)	3148	.032*
<i>Human-likeness</i>	3.67 (1.175)	4.31 (1.014)	4.21 (1.475)	3.85 (1.625)	.804	.497
<i>Persuasiveness</i>	3.60 (1.352)	4.63 (.957)	4.74 (1.368)	3.69 (1.377)	3507	.021*
<i>Recommendability</i>	3.53 (1.407)	4.13 (1.408)	4.26 (1.240)	3.46 (1.613)	1322	.276

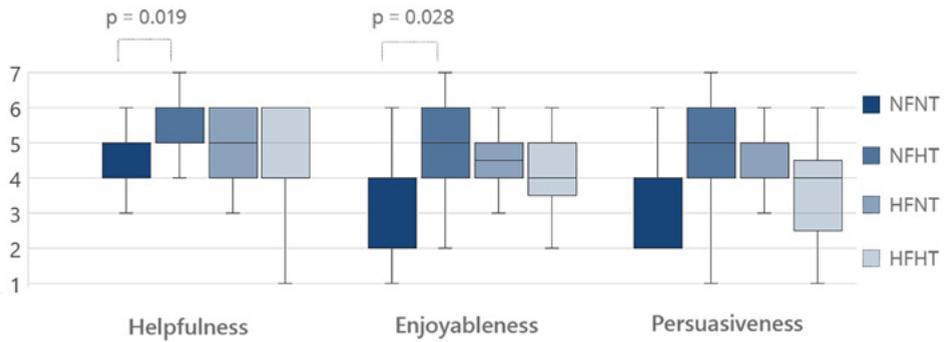


Figure 05.2 – Results of the Bonferroni post-hoc test comparing the mean values for the ECA’s helpfulness, enjoyableness and persuasiveness for the four conditions.

Table 05.7 shows the findings for the hypotheses of this research. This research only partly confirms hypothesis H1 and does not support hypotheses H2 and H3.

Table 05.7 – Findings for the hypotheses tested in this research.

Hypothesis	Explanation	Finding
H1	Users’ perceptions of rapport towards an ECA with happy textual expressions are higher than towards an ECA with neutral textual expressions.	Partially supported (only for the rapport items helpfulness and enjoyableness)
H2	Users’ perceptions of rapport towards an ECA with happy facial expressions are higher than towards an ECA with neutral facial expressions.	Not supported
H3	Users’ perceptions of rapport towards an ECA with both happy textual and happy facial expressions are higher than towards an ECA with only happy facial or happy textual expressions.	Not supported

DISCUSSION

This research compared an ECA's emotions expressed in text with an ECA's emotions in facial expressions on users' perceptions of rapport in the context of eHealth. We compared expressions of positive emotions (e.g. happiness) with absence of emotions (e.g. 'neutral' emotions). Our results show that **users perceive an ECA with a neutral face and happy text as more helpful and enjoyable than an ECA with a neutral face and neutral text**. This partly confirms our first hypothesis stating that users' perceptions of rapport towards an ECA with happy textual expressions are higher than towards an ECA with neutral textual expressions. Our second hypothesis, stating that users' perceptions of rapport towards an ECA with happy facial expressions are higher than towards an ECA with neutral facial expressions, and third hypothesis, stating that users' perceptions of rapport towards an ECA with both happy textual and happy facial expressions are higher than towards an ECA with only happy facial or happy textual expressions, were both rejected. In the rest of this section we discuss our findings in more detail.

First, **overall rapport values found were limited**. As explained by *Tickle-deggen and Rosenthal (1990)*, regardless of the time at which rapport is assessed, it consists of three components – **mutual attention, positivity and coordination** –, but the relative importance of these components changes over time. At initial encounters, development of rapport is strongly connected to the presence of positivity (warmth and friendliness) and attentiveness. In these first interactions humans try to be attentive and pleasant towards each other, which does not require previous experience with another. Only in later stages of interaction, rapport is judged more by the coordination component, which develops when the interaction feels less awkward and involves fewer communication misjudgements. In our study, we assessed rapport after just a short interaction, meaning that participants' rapport building during the study was more likely to be connected to the presence of positivity and attentiveness, and less likely to be connected to coordination. This could explain why many rapport items, such as cognitive rapport and naturalness, were rated low: they might have been stronger related to the coordination component of rapport and were, thus, not yet developed. Therefore, overall rapport values for all conditions might have been relatively low. Eventually, ECAs are developed for long-term engagement with eHealth technologies. Therefore, future work could focus on researching how an ECA's emotional expressions affect users' perceptions of rapport with the ECA after long-term interaction. As the relative importance of different components of rapport changes over time, overall rapport values might be different after long-term interaction. In addition, future research could investigate objective rapport. As suggested by *Cerekovic et al. (2017)*, rapport can also be judged by observers, taking into account aspects, such as a user's facial expressions. A user's subjective rapport towards an ECA might be different from an observers' objective rapport of the user towards an ECA.

Second, **we did not find a difference between the overall rapport values of the four different conditions**. A lack of differences between the conditions could have been explained by the study population: half of the participants in our study were fifty five years or older. As shown by *Beer et al. (2015)*, older adults show a decrease in emotion recognition accuracy for an ECA's emotions of anger, fear and happiness. The older adults that participated in our study might have had difficulties in recognizing the ECA's happy emotions, and therefore, they might have experienced little differences between the conditions. Future work could differentiate between perceptions of younger and older adults specifically. Next to the participant's age, the personality of the participants might have affected our re-

sults. As shown by *Cerekovic et al. (2014;2017)*, a user's personality could affect his or her perceptions of rapport towards an ECA. Future work could investigate how a user's personality affects building of rapport towards an ECA in the eHealth context specifically. By investigating possible differences in rapport building due to a user's age or personality, the interaction with an ECA can be optimised for different users.

Third, we found that **an ECA with a neutral face and happy text was perceived as more helpful and enjoyable than an ECA with a neutral face and neutral text**. This finding confirms that, indeed, humans are able to adapt to emotional expressions through computer mediated textual communication (*Walther, 1992; Hancock et al., 2007*). As the finding for overall rapport, this finding could also relate to the relative importance of the different rapport components over time. Helpfulness and enjoyableness seem to be related to the positivity and attentiveness component which are developed at initial encounters. Our results show that the ECA's helpfulness and enjoyableness indeed score higher than many other rapport items. Since these rapport items might have been developed more, this also explains why there are differences found between the conditions for these items. Based on these results, we suggest to carefully design an ECA's texts in eHealth applications, since it seems that an ECA's emotion in its textual expressions can affect user's perceptions of aspects of rapport in an early stage.

Fourth, **individual rapport items for ECAs having different facial expressions were not rated differently**. Though, we would have expected a difference, since humans are experts in recognising emotions based on facial expressions (*Ekman & Friesen, 1978*). In addition, we expected that positive facial expressions, such as smiling, would lead to a higher perception of warmth (*Tickle-deggen & Rosenthal, 1990*). The lack of perceived difference between the happy and neutral facial expression could be, as explained before, because older adults in our study population might have had difficulties with recognising the ECA's happy emotions in its facial expressions (*Beer et al., 2015*). Moreover, it could be explained by the scenario we used; the agent providing help in the health context. *Tickle-deggen and Rosenthal (1990)* show that when interlocutors smile in a non-helping context, a positive interaction is created, whereas this effect was not found for a helping context. A possible explanation for this difference between the contexts provided by the authors is that, in a helping context, people feel more anxiety and are, therefore, less observant for non-verbal cues. Furthermore, the absence of a difference for the facial expressions could be caused by the use of static 2D ECAs, instead of animated, 3D ECAs. Assessing non-verbal correlates of rapport is not easy (*Tickle-deggen & Rosenthal, 1990*). The impression of facial warmth is based on movements of the mouth, eyes, brow and head in concert with each other. Movements were absent in the ECAs used in this study, which might have resulted in participants having difficulties with assessing the non-verbal correlates of rapport in the ECA's facial expressions. Future work could test the effect of emotion in facial expressions in animated 2D ECAs instead of in static 2D ECAs. Since movements in the face help users to assess facial expressions, the contrast between the happy and neutral facial expressions most likely increases, which might affect the levels of rapport.

Strengths and Limitations

The strength of this study is that we tested the effect of an ECA's emotional expressions in a realistic scenario in the context of eHealth. In addition, this study particularly researched the effect of emotional expressions in 2D ECAs, whereas much research focuses on emotional expressions in 3D modelled ECAs having dynamic animations. 2D ECAs require less technological development, and, therefore, the threshold to implement these ECAs in eHealth is lower. However, testing the effect of emotions via facial expressions in static 2D ECAs was also a limitation of this study, since the contrast between the happy and neutral facial expression might have been less than it could have been in 3D ECAs. In addition, changing an ECA's textual expression only in emotion is complex. Ideally, one would like the content of the neutral and happy textual expressions to be equal, such that differences in users' perceptions of these expressions can be dedicated to the difference in emotion. However, by changing the words of the neutral expression, such that the expression conveys the happy emotion naturally, we might have introduced another, unwanted factor that was different in both expressions. Such a factor might have affected our results. Furthermore, participants' emotional states might have affected our results. For example, a participant that was already feeling sad before interaction with an ECA, might not perceive high rapport towards this ECA, independent of the ECA's expression. Lastly, participants only interacted with the ECA shortly. The interaction time might have been too short to actually build rapport within all components. For future studies we would recommend to create animated 2D ECAs and to evaluate them in long-term interaction.

CONCLUSION

We did not find any difference between an ECA's emotions expressed in text and an ECA's emotions in facial expressions on users' overall perceptions of rapport, i.e. mutual understanding, in the context of eHealth. However, our ECA was perceived as more helpful and enjoyable when happy textual expressions were used for an ECA having a neutral facial expression. Therefore, we suggest to carefully design an ECA's textual expressions, since their framing could influence users' perceptions' of the ECA's helpfulness and enjoyableness in an early stage. Future research should investigate users' rapport towards ECAs with different emotions in long-term interaction, both subjectively and objectively, and how a user's age and personality and an ECA's animations affect rapport building. Optimising users' interaction with ECAs could contribute to achieving long-term interaction with eHealth.

06

ECA Appearance in a Long-term, Daily Life Setting for Self-management

BASED ON:

ter Stal, S., Sloots, J., Ramlal, A., op den Akker, H., Lenferink, A., & Tabak, M. An embodied conversational agent in an eHealth self-management intervention for chronic obstructive pulmonary disease and chronic heart failure: Exploratory study in a real-life setting (*submitted for publication*)



Abstract

Embodied Conversational Agents (ECAs) have the potential to stimulate actual use of eHealth applications. An ECA's design influences the user's perception during short interactions, but daily life evaluations of ECAs in healthcare are scarce. This chapter presents an exploratory, long-term study on the design of ECAs for eHealth. The study investigates how patients perceive the design of the ECA over time with regard to the ECA's characteristics – friendliness, trustworthiness, involvement, expertise and authority –, small talk interaction, and likeliness of following the agent's advice. We developed an ECA within an eHealth self-management intervention for patients with both Chronic Obstructive Pulmonary Disease (COPD) and Chronic Heart Failure (CHF), which we offered for four months. Patients rated five agent characteristics and likeliness of following the agent's advice before use and after three and nine weeks of use. The amount of patients' small talk interaction was assessed by log data. Lastly, individual semi-structured interviews were used to triangulate results. Eleven patients (7 male and 4 female) with COPD and CHF participated (median age 70 years). Patients' perceptions of the agent characteristics did not change over time ($P > 0.05$ for all characteristics) and only one participant finished all small talk dialogues. After three weeks of use, the patients were less likely to follow the agent's advice ($P = .01$). The agent's messages were perceived as non-personalized and the feedback as inappropriate, affecting the agent's perceived reliability. This exploratory study provides first insights into ECA design for eHealth. The first impression of an ECA's design seems to remain during long-term use. To investigate future added value of ECAs in eHealth, perceived reliability should be improved by managing users' expectations of the ECA's capabilities and creating ECA designs fitting individual needs.



INTRODUCTION

The number of people having a chronic disease, such as diabetes, cancer or Chronic Obstructive Pulmonary Disease (COPD) is increasing (*World Health Organization, 2015*). COPD is a chronic lung disease that is progressive, and often accompanied by comorbidities, such as Chronic Heart Failure (CHF) that further increases the risk of COPD exacerbations, hospitalizations, mortality and costs (*Global Initiative for Chronic Obstructive Lung Disease, 2020; Mannino et al., 2008*). Research shows that paper versions of exacerbation action plans tailored for COPD and comorbidities, embedded in a multi-faceted self-management intervention, reduce the duration of COPD exacerbations and the risk of respiratory-related hospitalizations (*Lenferink et al., 2019*).

To further facilitate this chronic disease self-management in daily life, eHealth applications can be used. eHealth applications can provide patients insight in their behaviour and disease by symptom monitoring, and patient-tailored and accessible support in their home setting, supervised by their healthcare professional at a distance (*Tabak et al., 2014*). Although such applications seem promising, many eHealth applications face the problem of its actual use decreasing after several weeks by a lack of user engagement (*Nijland, 2011; Barello et al., 2016; Buimer et al., 2017*). Research indicates that a patient's use of eHealth applications is influenced by extrinsic motivation cues, such as stimulation by care professionals and fellow patients (*Tabak et al., 2014; Buimer et al., 2017; Kelders et al., 2012*). The majority of existing eHealth applications provide such support in the form of plain text or via a text-based question-answer module, whereas face-to-face interaction remains one of the best ways to communicate health information (*Clark et al., 1991; Bickmore et al., 2009*).

A different way of providing (motivational) support includes use of Embodied Conversational Agents (ECAs). ECAs are defined as more or less autonomous and intelligent software entities with an embodiment used to communicate with the user (*Ruttkey et al., 2004*). By face-to-face interaction with the user, ECAs can build trust and rapport – i.e. agreement or sympathy between people or groups (*Cambridge University Press, 2019*). By building trust and rapport, they could create a companionship with the user, leading to long-term and continuous use (*Vardoulakis et al., 2012*) and, thereby, stimulate the actual use of the underlying eHealth application. Just as a human's appearance affects how we evaluate a human, an ECA's appearance affects how we evaluate an ECA. When we interact with another human, or ECA, for the first time, we immediately form initial ideas about the other (*Kelley, 1950; Bergmann et al., 2012*). Furthermore, when we have a positive impression of another human, we tend to interact more with that human. This likely applies to human-agent interaction as well, such that we interact more with ECAs of which we have a positive first impression (*Kelley, 1950; Bergmann et al., 2012*).

ECAs have thus potential to promote engagement with eHealth applications. However, a recent review on the design of ECAs for eHealth (**chapter 2**) shows no clear consensus on the design of ECAs for eHealth. More specifically, the review states that emotion and empathic behaviour seem to positively affect the user's perception of the agent's characteristics, but that these design features do not necessarily lead to users' behaviour change. The review also shows that studies mainly focus on the effect of the ECA design at first glance or after short interaction. But, to gain insight in the possible added value of ECAs in eHealth, it is important to evaluate how the ECAs should be designed for the intended context of long-term use in daily life. Only one study reports on the design of an ECA for eHealth in such a long-term, daily life setting (*Zhou et al., 2014*). In this study, a virtual hospital dis-

charge nurse discussed the patient's diagnosis and post-discharge self-care with the patient once a day at his or her hospital bed. In addition, the agent instructed the patient about medication, follow-up appointments and self-care procedures just before hospital-leave. Questionnaires filled out after the hospital-leave interaction showed that the patient's perceived similarity to this agent was significantly associated with the patients liking the agent and patients' trust in and desire to continue with the agent. In addition, perceived similarity was associated with the patient's working alliance with the agent – which the authors define as 'trust and belief in working with the agent to achieve a therapeutic outcome'.

To develop ECAs to support users in self-management of chronic diseases, such as of COPD and CHF, more research is necessary on how ECA design affects users' perceptions of an ECA in the intended context of use: a long-term, daily life setting. Research should already start in early stages of development of such ECAs, as small-scale eHealth evaluation studies focusing on usability, feasibility and end-user experience allow researchers to gain detailed information that can be used for further improvement of an eHealth application (*Jansen-Kosterink et al., 2016*). The importance of applying user-centred design – designing with end-users instead of for end-users by involving them in all stages of the development process – is increasingly recognized to be valuable in health care (*Moody, 2015; Horne et al., 2013*). By involving users to participate in early stages of development, technical flaws can be understood and overcome (*Nijland, 2011*) and the technology can be developed in such a way to reach clinical value in follow-up larger scale studies.

This study is a first exploratory study on ECA design for eHealth in a long-term, daily life setting. In this study, an ECA is implemented into an eHealth self-management intervention for patients with COPD and CHF, offered for approximately four months.



Research Objectives

The objective of our study is to investigate how users perceive the design of the ECA over time. In particular, how they perceive the agent's characteristics (friendliness, trustworthiness, involvement, expertise and authority) and the agent's small talk, and how likely they are to follow the agent's advice.

METHODS

This study was performed as part of the MATCH study. The aim of the MATCH study was to investigate the feasibility of an eHealth self-management intervention for patients with COPD and CHF over a 4-month period. The ECA was implemented into this eHealth self-management intervention. The MATCH study was approved by the Twente Medical Ethical Committee and registered in the Netherlands Trial register (NL6480).

Participants

People were included for participation to the MATCH study if they: (1) had a clinical diagnosis of both COPD and CHF, (2) had at least two COPD and/or CHF exacerbations and/or at least one hospitalization for COPD and/or CHF in the two years preceding study entry, (3) were at least one week after prednisolone/antibiotics/furosemide course and hospitalization and at least four weeks post-rehabilitation, (4) were at least 40 years of age, (5) were able to understand and read the Dutch language, (6) were able to use a smart phone, tablet or PC, and (7) provided written informed consent prior to participation. People were excluded from participation if they: (1) had terminal cancer or were at the end stage of another serious disease, (2) had another serious lung disease, (3) expected cardiovascular intervention within three months, (4) were enrolled in randomized controlled trials or a trial with study medication, (5) were waiting for a heart or lung transplantation, and (6) received renal dialysis.

The eHealth-supported Self-management Intervention

The self-management intervention was offered through an application on a tablet (eHealth platform, Roessingh Research and Development, Enschede, the Netherlands (*op den Akker et al., 2012*)) and consisted of the following modules, as can be seen in **Figure 06.1**:

› Self-management module:

- › *Daily symptom diary*: registration of symptoms related to COPD (e.g. dyspnea, cough), CHF (e.g. weight, edema) and common comorbidities (depression, anxiety) and classification of symptoms in case of symptom deterioration determined by the patient by comparing the symptoms experienced in the last 24 hours with his or her 'usual' symptoms on his or her 'what are my "usual" symptoms' card. In case of any symptom deterioration, patients were asked to classify each symptom as 'normal', 'slightly increased' or 'significantly increased'. The daily symptom diary was connected to a decision-support system that automatically launched self-management advice in case of worsening of the patient's clinical condition (according to symptoms and weight). The automated decision support system was translated from an evidence-based self-management intervention including paper versions of multi-morbid exacerbation action plans for patients with COPD and comorbidities (*Lenferink et al., 2013; 2019*).
- › *Action list*: a list of actions containing (1) self-management advice determined by the automated decision-support system of the daily symptom diary (e.g. initiate self-treatment, perform relaxation exercises from the exercise module, call the case-manager). In addition, the list contained (2) reminders to measure weight by a smart scale and (3) reminders to complete questionnaires.
- › *Phone numbers*: to contact healthcare providers for support.
- › *Health status*: an overview of patients' health status during the last week, including an indication of no, slightly increased or significantly increased symptoms.

- › **Monitoring module**: a detailed overview of health status, self-reported symptoms, weight and received advice.
- › **Inhaler module**: monitoring of and feedback on inhalation medication adherence and technique (add-on sensor for Ellipta, Amiko Respiro®).
- › **Information module**: information about self-management including patients' diseases and healthy behaviour (*Lenferink et al., 2013; 2019*) and instruction videos to perform breathing exercises.

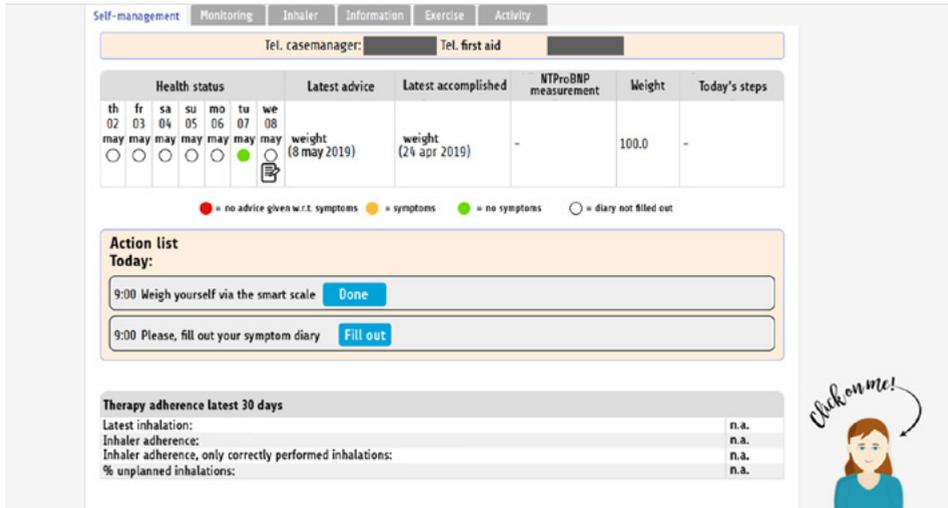


Figure 06.1 – Home page of the MATCH self-management application, showing the patient’s health status and action list, and ECA. The ECA, Sylvia, was always present in the right bottom of the application. The text “click on me” and an arrow pointing to the agent were shown only before the first interaction with the agent started.

- **Exercise module:** a standardized set of breathing, relaxation and physical exercises, accompanied by videos and explanation in text.
- **Activity module:** display of daily physical activity (amount of steps measured by the Fitbit Zip).

Patients were advised to use the self-management module daily by completing the daily symptom diary, monitoring their weight via the smart scale and performing the actions on the action list. Furthermore, they were advised the use the inhaler daily. The use of the monitoring, information, exercise and activity module was voluntarily.

Interaction with Caregivers and Fellow Patients during Self-management

Patients first attended three self-management training sessions (two group sessions and one individual session with the case manager), that amongst other things included information regarding their diseases, training to recognize symptoms and to practice with using the self-management application. Patients started using the application after the first (group) session, so that questions regarding self-management and the technology could be answered during the next two sessions.

For safety during the period of application use, patients were advised to call the case manager (or general practitioner outside office hours) when symptoms did not improve after two days of self-treatment and when they experienced dizziness. In addition, the case manager checked health status of the patients (in the application) once per week, and called the patient when they found this was necessary. During the self-management training, patients were instructed that they could call the case manager in case of any questions or doubts. Further, regular healthcare (e.g. visits to their pulmonary physicians and cardiologists) continued as normal during the study.

The Embodied Conversational Agent

The following agent characteristics found in literature were taken into account when designing the current agent in a creative process with the developers having a description of a persona as outcome:

- ▶ **Gender:** female – Research indicates that people prefer ECAs that fit their task conform stereotypes. For health-related tasks (e.g. providing medical advice) female agents are preferred (Forlizzi et al., 2007; Zimmerman et al., 2005), since these tasks are traditionally being undertaken by women. Furthermore, **chapter 3** shows a general preference for a young, female ECA.
- ▶ **Age:** young adult – Research indicates that people prefer young agents over older agents in the context of health, specifically in self-management for chronically ill elderly (van Wissen et al., 2016). As the authors explain, a younger agent might be found more attractive. Furthermore, **chapter 3** shows a general preference for a young, female ECA.
- ▶ **Cultural background:** grown up in the region Twente, the Netherlands, living in a terraced house with garden – Research indicates people prefer agents having the same cultural background as themselves (Alsharbi & Richards, 2017; Yin et al., 2010; Zhou et al., 2017). The cultural background of the agent is, for example, expressed in the agent’s small talk: the agent talks about activities and events related to her place of living.

In addition, to establish a full persona, additional characteristics of the persona were created. Two examples of reasoning behind the characteristics of the persona are shown below. The persona used as a guideline to write the dialogues can be seen in **Figure 06.2**.

- ▶ **Role:** semi-expert – Since the self-management intervention was supported by a healthcare professional (nurse practitioner COPD and nurse practitioner CHF), we decided not to create a second medical expert agent. In addition, the goal of the interactive dialogues of the agent was to support patients. Therefore, we gave the agent the role of a “semi-expert”: an agent with some experience in chronic diseases (reflected in her career), but that does not act as a doctor or nurse practitioner.
- ▶ **Energy consumers:** Asthma – To trigger users to identify with the agent, we decided that the agent has a chronic lung disease as well. However, to ensure the stories of the agent would not become too negative, focusing on limitations related to the disease, we decided the agent has a mild form of asthma.

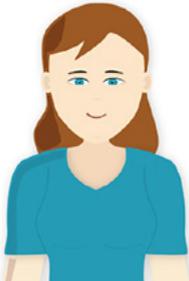
 <p style="font-size: 24px; font-weight: bold; margin-top: 0;">Sylvia</p>	Age	32	Previous job	Pulmonary function analyst at the hospital in Enschede.
	Role	Semi-expert	Current job	Communication officer at the lung agency Enschede
	Place of residence	Enschede, the Netherlands	Hobbies	Hiking, cooking, reading, watching movies and series
	Place of birth	Haaksbergen, the Netherlands	Personal characteristics	Helpful, caring, driven
	Housing status	With husband, son and dog.	Energy consumers	Asthma
	Education	Nursing school	Energy boosters	Family and friends, helping others

Figure 06.2 – The persona of the agent Sylvia used as a guideline for writing the dialogues.

The ECA, Sylvia, was implemented into the application via the use of a Scalable Vector Graphics (SVG) object, including HyperText Markup Language (HTML) animations. The agent blinked her eyes every ten seconds, and moved her mouth a fixed period of time after a new agent sentence appears on the screen (as if she was talking to the user). The ECA was always present in the right bottom on the pages of the self-management application. Before the user interacted with the agent for the first time, the text “click on me” and an arrow pointing to the agent were shown *Figure 06.1*. After the first interaction, this message disappeared. When hovering over the agent, the agent increased in size.

The content and trigger times of the dialogues were created in collaboration with experts on COPD and CHF to ensure that it was in line with patients’ daily practice. Four types of dialogues could be triggered*:

- 1. Action reminders:** dialogues in which the agent reminded the patient of performing actions of the action list of the self-management application, e.g. completing the daily symptom diary, weekly questionnaire or monthly motivation questionnaire, weighing themselves, initiating medication for self-treatment of worsening symptoms, and calling case-manager for support. The agent provided the patient with a general message stating that there were uncompleted actions on the action list, but did not provide the patient with the actual content of these actions.
- 2. Inhaler feedback:** dialogues in which the agent informed the patient about 1) the synchronization of the smart inhaler and 2) the inhalation adherence and technique. More specifically, the first type of dialogue informed the user when the smart inhaler had not synced for either 24 or 72 hours. The second type of dialogue informed the user when the inhalation had been skipped for over 2 days, an extra dose had been taken during the last 7 days, the inhalation time of the last inhalation deviated too much from the average duration of the inspiratory flow and when the position of the device was not optimal.
- 3. Health-related tips:** dialogues in which the agent provided the patient with several health-related tips, such as accessing information sources or small actions to perform in daily life. Some of the tips referred to information provided at pages in the self-management application.
- 4. Small talk:** chit chat dialogues to increase the patient’s engagement (*Kang & Gratch, 2011*), stimulating the use of the underlying application. The small talk dialogues were designed as a daily soap series to trigger the patient’s curiosity about the continuation of the story. The small talk was split up into seven “episodes”, all containing multiple dialogue steps around a certain theme (the introduction and Sylvia’s housing status, husband, child, neighbour, hobbies and dog). When the patient finished an episode of the small talk, the next episode was unlocked the next day. In the meantime, when the patient clicked on the agent, the agent informed the patient that she does not have time to talk until tomorrow (i.e. showed a “wait till tomorrow” message). When the patient finished all seven episodes, the agent told the patient that she had nothing more to say

Small talk dialogues could be triggered by the user by clicking on the agent on the home page of the self-management application. The other dialogues were triggered by the system at predefined trigger times:

- › **Action reminders:** 1, 2 and 3 hours after an action was added to the action list and not yet performed;

* Dialogues available upon request.

- **Inhaler feedback:** when incorrect inhaler use was measured;
- **Health related tips:** each day at 15h;
- **Small talk:** each day at 14h (only when the patient did not yet initiate a small talk dialogue that day by him or herself and the small talk was not yet finished).

Each dialogue consisted of one or multiple dialogue steps, containing one or multiple answer possibilities for the user. The agent message was displayed in text and not communicated via speech. An example of the interface of the dialogue step can be seen in [Figure 06.3](#). Examples of the content of the dialogue steps for every dialogue type can be seen in [Table 06.1](#).

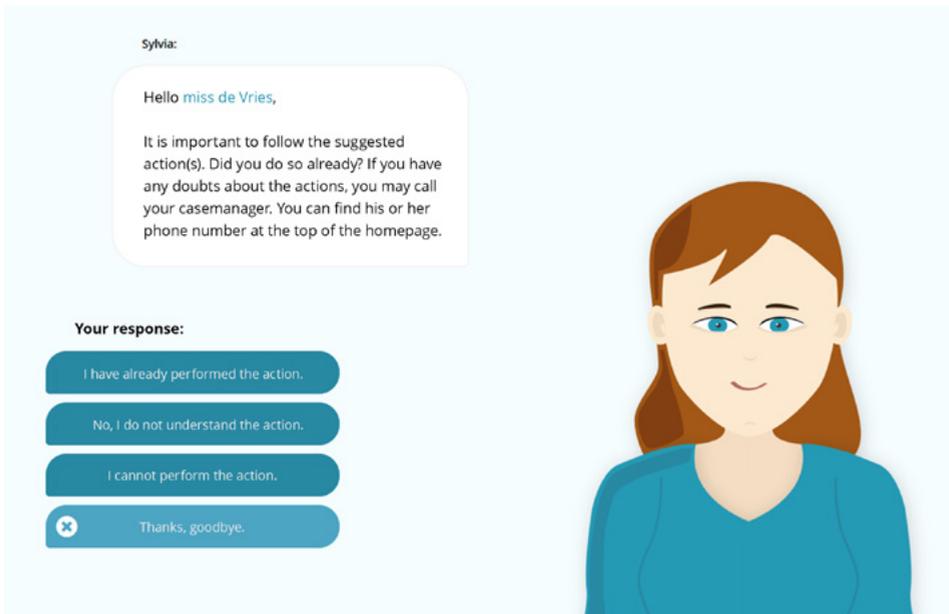


Figure 06.3 – Example of a dialogue in which the agent Sylvia reminds the user to perform an action.

Table 06.1 – An example dialogue step for every dialogue type that could be triggered.

	Action reminders	Inhaler feedback	Health related tips	Small talk
 Message agent	<p>It is important to follow the suggested action(s). Did you do so already? If you have any doubts about the actions, you may call your casemanager. You can find his or her phone number at the top of the homepage.</p>	<p>You did not take your inhalation medication for two mornings or evenings. It is important not to forget to take your inhalation medication.</p>	<p>On the information page you can find lots of information about various aspects of COPD and heart failure. Have a look if you want to know more about your disease.</p>	<p>Nice to meet you! My name is Sylvia de Ridder. I am 32 years old and I will help you using the application. I will remind you of performing actions, but when you feel like a chat, you are welcome to come by as well!</p>
 Reply options patient	<ul style="list-style-type: none"> > I have already performed the action. > No, I do not understand the action. > I cannot perform the action. > Thanks, goodbye. 	<ul style="list-style-type: none"> > I will try to improve my habits. > I do not understand how to perform the inhalation. > I do not have any inhalation medication. > Thanks, goodbye. 	<ul style="list-style-type: none"> > Show me the information page. > Thanks, goodbye. 	<ul style="list-style-type: none"> > What did you study? > That sounds great. Are you interested in coaching? > Nice to meet you too. Is your task related to your job? > I agree. Talk to you later.

Procedure

Figure 06.4 provides an overview of the study procedure. Written informed consent from the participants was obtained prior to study participation. Then, the participants filled out the baseline questionnaire at home (t0). At this point, the participants had not yet seen the application and were not aware of the existence of an agent in the application. The participants were introduced to the agent for the first time in the baseline questionnaire, as a picture of the agent was attached to the questions regarding the agent. The agent was introduced as a hypothetical coach. During the first group session (S1) participants received a tablet, step counter (Fitbit Zip) and smart scale to be used with the application. After this meeting, the participants could already start using the application and sensors provided. In a second, individual meeting, patients practiced with using the eHealth application according to their individual symptoms (S2). In the second group session (S3), some last questions with respect to self-management and the technology were answered. After the second group session, participants received the add-on sensor for the inhaler and afterwards all patients used the application and sensors. Two weeks after the last group session (S3), participants received the intermediate questionnaire (t1). After 9 weeks of use, users received the follow-up questionnaire (t2). Technology usage was logged during the complete period of use. After the end of the use period, the participants were interviewed by an independent interviewer (AR) (t3).

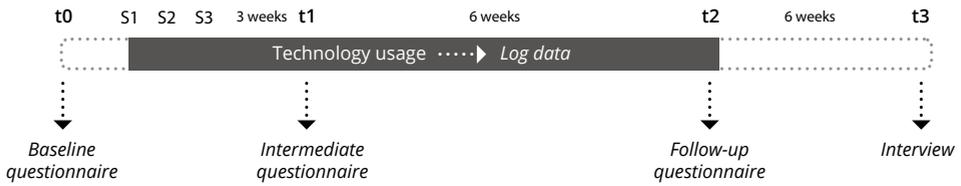


Figure 06.4 – The procedure of the study. S1 = first group session, S2 = individual session, S3 = second group session, t0 = start, t1 = 3 weeks from the start, t2 = 12 weeks from the start, t3 = 18 weeks from the start.

Design and Measurements

We used a mixed-method design, combining both quantitative and qualitative research methods: questionnaires, log data and semi-structured interviews (see [Figure 06.5](#)).

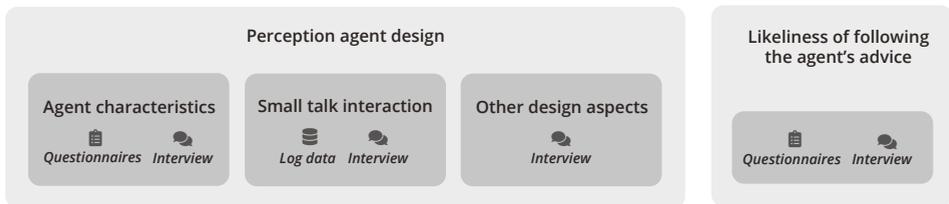


Figure 06.5 – Study measurements to evaluate the user's perception of the agent design and the user's likeliness of following the agent's advice.

The patient's perception of the characteristics of the agent and likeliness of following the agent's advice were measured via I) a baseline questionnaire at t0, II) an intermediate questionnaire at t1 and III) a follow-up questionnaire at t2 (as shown in [Appendix E](#)). These paper self-reported questionnaires assessed the patient's perception of:

- Five characteristics of Sylvia (the agent in the MATCH self-management application): friendliness, trustworthiness, involvement, expertise and authority
- The importance of these five characteristics of an ECA for self-management in general
- The likeliness of following Sylvia's agent's advice

The questions on users' perceptions of these characteristics and likeliness of following an ECA's advice were similar to those of two other studies ([chapter 3](#) and [chapter 4](#)). All items were assessed by ratings on a 7-point Likert scale. In addition, the baseline questionnaire contained questions related to the patient's characteristics.

Furthermore, small talk interaction was analysed using IV) log data. The dialogue history of the small talk of the ECA with the patient was logged on the server. Per patient, the date and time of dialogues triggered by either the system or the user were logged. Furthermore, the patient's selected responses were logged per dialogue step of a triggered dialogue, including a date and time.

Lastly, the patient's impression of the agent's characteristics, the likeliness of following the agent's advice, the small talk and other design aspects were gathered in V) semi-structured interviews.

Data Analyses

Statistical analyses were performed in SPSS 25. 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 – ordinal – variables. In the questionnaires, the five agent characteristics were classified as low (a score from 1 until 3), neutral (a score from 3 until 6) or high (a score from 6 until and including 7) on applicability to Sylvia and importance for an ECA for self-management in general. The same classification was used for the user's likeliness of following advice.

For all relations, a related-samples Friedman's two-way analysis of variance by rank was conducted as appropriate. The Holm–Bonferroni method was used to correct for multiple comparisons: the comparisons of the ratings for the characteristics of Sylvia and ECA for self-management in general and the likeliness of following Sylvia's advice at t0 (before use), t1 (after three weeks of use) and t2 (after nine weeks of use).

The interviews were transcribed by the interviewer (AR) using automatic transcription in Amberscript and a manual check afterwards. Another researcher extracted the interview data focusing on the MATCH-agent or ECAs in general (StS). Then, the remaining interview data was thematically analysed by two researchers independently (StS, MT). All themes were grouped either under 1) the patients' perceptions of agent characteristics, 2) small talk interaction or 3) other design aspects. The themes were coded retrospectively using ATLAS.ti 8, based on the steps proposed in (*Pope & Mays, 2006*): one researcher (StS) created a first coding scheme and labelled all the data accordingly. A second researcher (MT) used the coding scheme to code a subset of the data. Disagreements between 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.

RESULTS

Baseline Demographics

Eleven patients (7 male, 4 female) completed the study procedure until t2, of which nine agreed to participate in the interview at t3. The age of the participants (n = 11) ranged from 49 to 83 years (median = 70). The highest educational degree for the majority of the participants was high school or vocational education, one participant had a university degree. Three participants lived alone, the others with their partner. Four participants indicated that their partner is their informal caregiver, whereas the others said they do not have an informal caregiver. Self-reported tablet skills were high for four participants, three did not have any experience with a tablet yet and the rest had some experience.

Patients' Perceptions of Agent Characteristics

Table 06.2 shows the patients' perceptions of the characteristics of Sylvia and on the importance of the characteristics of an ECA for self-management in general over time. Some participants did not fill out all questions of the questionnaires. Hence, per characteristic only 8 or 9 participants rated the characteristic at both t0, t1 and t2. At t0, t1 and t2 Sylvia was rated high on friendliness, on t1 Sylvia was rated high on reliability and low on authority. For all other characteristics, the median rating of the agent was neutral at t0, t1 and t2. In addition, at each point in time, the agent characteristic authority was rated neutral on importance for an ECA for self-management. Expertise, reliability and involvement were rated high on importance for an ECA for self-management in general. Friendliness was rated high on importance at t0 and t2, and neutral on importance at t1.

Table 06.2 – Comparison of the patients' ratings of the characteristics Sylvia and the patients' ratings of the importance of characteristics for an Embodied Conversational Agent (ECA) for self-management in general at t0 (before use), t1 (after three weeks of use) and t2 (after nine weeks of use) using a Friedman's two-way analysis of variance by rank. IQR = Inter Quartile Range (Q1 – Q3).

	n	Median (IQR) t0	Median (IQR) t	Median (IQR) t2	P
Ratings characteristics Sylvia					
<i>Friendliness</i>	9	6(4.0 – 7.0)	6.0 (4.5 – 6.0)	6.0 (4.0 – 7.0)	.45
<i>Expertise</i>	8	5.0 (4.0 – 7.0)	5.0 (4.0 – 6.0)	5.0 (4.0 – 7.0)	.47
<i>Reliability</i>	9	4.0 (4.0 – 7.0)	6.0 (3.5 – 6.0)	4.0 (3.8 – 7.0)	.77
<i>Authority</i>	8	4.0 (2.3 – 4.8)	2.0 (2.0 – 5.5)	4.0 (2.3 – 5.5)	.64
<i>Involvement</i>	8	5.5 (4.0 – 7.0)	5.0 (4.0 – 6.0)	4.5 (3.3 – 7.0)	.68
Ratings importance characteristics ECA for self-management in general					
<i>Friendliness</i>	9	6.0 (4.0 – 7.0)	5.0 (4.0 – 6.0)	6.0 (4.0 – 7.0)	.43
<i>Expertise</i>	9	7.0 (7.0 – 7.0)	7.0 (6.0 – 7.0)	7.0 (6.5 – 7.0)	.25
<i>Reliability</i>	9	7.0 (6.5 – 7.0)	7.0 (6.0 – 7.0)	7.0 (6.5 – 7.0)	.84
<i>Authority</i>	9	4.0 (2.0 – 6.0)	4.0 (1.5 – 5.5)	4.0 (3.0 – 4.5)	.65
<i>Involvement</i>	9	7.0 (6.0 – 7.0)	6.0 (6.0 – 7.0)	6.0 (4.0 – 7.0)	.78

In the interviews, the patients' commented on some of the above measured characteristics. We identified the themes **friendliness, reliability, expertise, authority**. One participant found the agent – Sylvia – friendly, whereas another did not indicate whether Sylvia was friendly, but stressed that an agent for self-management support should be friendly. Furthermore, Sylvia was not always seen as reliable, supported by a participant indicating that the messages of Sylvia were based on data from a non-reliable Fitbit. Although Sylvia did not provide advice based on the Fitbit data, this participant might have thought this was the case. Another participant indicated that Sylvia sometimes gave tips that did not fit the participant's individual situation – e.g. suggesting to perform physical activity when having filled out symptoms in the diary –, affecting the agent's reliability. One participant especially commented on the agent's expertise, calling Sylvia "a stupid woman". In addition, one participant commented on authority, saying:



She could be your girl next door... If I have medical complaints, I prefer an authority to explain what to do or not to do.

Small Talk Interaction

Figure 06.6 shows how many participants unlocked particular episodes, based on the log data. Seven out of the eleven participants did not finish the first episode. Two participants finished the first episode and, therefore, unlocked episode 2. In addition, one participant unlocked the episodes until episode 4. Finally, one participant finished all seven small talk episodes. Two participants were shown a “wait till tomorrow” message once, since they already finished a small talk episode that day, one participant saw the message three times and one participant five times. Finally, the participant that finished all dialogues was shown the message that the small talk was finished for 45 times, meaning this participant clicked on the agent to receive a new small talk dialogue for 45 times, whereas the dialogues were finished.

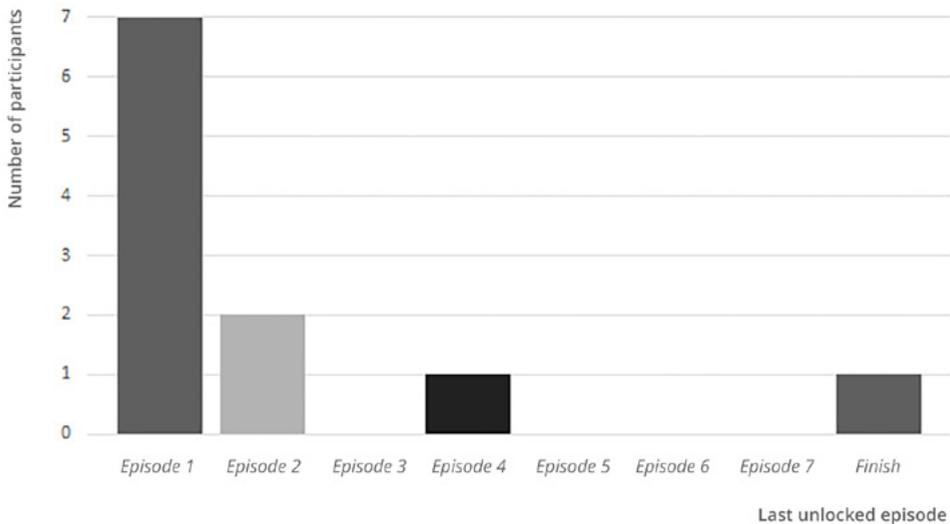


Figure 06.6 – The number of study participants that unlocked a particular small talk episode.

In the interviews, participants had a few comments on the agent’s small talk. One person did not notice that Sylvia talked about her own life. Five participants said not to be interested in small talk. However, two of them thought that people that feel lonely might be interested. One participant (not the participant that finished all the small talk dialogues) showed a more positive attitude towards the agent’s small talk:



Sylvia could talk nicely, she told me many things, for example, that she was lonely.

Other Design Aspects

In addition to the agent characteristics and small talk, the analysis of the interviews resulted in the following themes with respect to the agent design: **the agent's appearance, frequencies of the messages, timing of the messages and the interface design.**

First, the agent's appearance was evaluated. One participant preferred to interact with a photo-realistic nurse, instead of with a computer animated figure, since a photo-realistic nurse would make the interaction more personal. The participant also stated:

“

I am not impressed by a cartoon figure.

Furthermore, the participant described the agent as:

“

A male or female such as on the doors of bathrooms.

Another participant also preferred the agent to look like a nurse, this participant particularly commented on the agent's clothing:

“

Put a white coat and a stethoscope on her.

The participant described that an agent having a white coat and stethoscope would look more authoritative than the current agent in a t-shirt. Lastly, one participant liked that the agent was a woman, since this participant indicated to hate to listen to men.

Second, two participants particularly commented on the frequency of the messages from the agent: in their view, they received too many messages. One of them indicated that, therefore, he or she closed the dialogue before reading it.

Third, with respect to the timing of the messages, one participant would like to receive conformation messages when performing actions (real-time feedback on actions), whereas another suggested that the agent should come back to topics discussed before, as illustrated by:

“

But, then, ask the next day: 'Did you read that? Did you do this?'

In addition, two participants indicated that Sylvia provided unwanted and unsolicited information. One argued that she started to talk about a topic, regardless of whether the user was interested in that topic at a particular moment. This participant, instead, would like to receive the information when asking for it. The other participant argued that Sylvia provided advice when the participant felt well, whereas this participant only wanted to receive advice when not feeling well. Also, two participants said that they did not always have the time to follow the suggested actions or tips when receiving them from the agent. One was really annoyed when receiving messages, like “think about your exercises”, straight in the morning, the other explained:

“*When I have to go to work, I do not have time to watch at a 15-minute video.*”

Finally, two participants found it annoying that the agent already started giving reminders, when opening the application, not having the chance to even perform the action, illustrated by:

“*Look, what really bothered me was that, in the morning, I turned on the device and it [the agent] started with saying: ‘Did you follow the instruction?’ Well, I did not see any instruction yet.*”

Likelihood of Following the Agent's Advice

Table 06.3 shows the results of the related-samples Friedman's two-way analysis of variance by rank, comparing the participant's likelihood of following the agent's advice over time. On t0, t1 and t2 Sylvia scores neutral. A significant difference ($P = .01$) in the distribution of the values over time was found.

As a second step, pairwise comparisons of t0 and t1, t0 and t2 and t1 and t2 show no significant difference between t0 and t2 ($P = .07$) and t1 and t2 ($P = .48$), but do show a significant difference for the pair t0-t1 ($P = .01$). The participant's indicated likelihood of following the agent's advice statistically dropped at t1 compared to t0.

Table 06.3 – Comparison of the ratings of participants' likelihood of following Sylvia's advice at t0 (before use), t1 (after three weeks of use) and t2 (after nine weeks of use) using a Friedman's two-way analysis of variance by rank. IQR = Inter Quartile Range (Q1 – Q3).

	n	Median (IQR) t0	Median (IQR) t1	Median (IQR) t2	P
<i>Likelihood of following advice</i>	9	6.0 (4.5 – 7.0)	3.0 (2.0 – 4.0)	4.0 (2.5 – 5.0)	.01*

In the interviews, the majority of the participants indicated they would not follow the agent's advice. Two of them questioned the agent's reliability. In line with this, another participant indicated that he would first go to a doctor to verify the agent's advice of taking prednisolone. Although it should be noted that the agent did not provide the patient with advice on taking prednisolone directly; the agent only mentioned that there was an uncompleted action on the action list, which might have been taking prednisolone. However, the actual advice was determined by the automatic decision support system. Another participant argued that the agent did not respond to user input and, therefore, did not find the agent's advice valuable. One participant said not to listen to a cartoon figure, and another stated:



I do not listen to a device, I do listen to people.

Furthermore, a participant indicated that adults have their own responsibility, and therefore, this participant did not feel the need for an agent to suggest what to do. One participant argued not having the time to follow the advice, and therefore, not seeing the benefits of the agent's advice. In contrast, three participants said they sometimes did follow the agent's advice. One of these participants sometimes performed the physical exercises advised, since this participant valued the exercises. Another indicated to follow the advice of calling the case manager or reading information pages, but would not follow an advice to start prednisolone. The participant said that being wrongly advised to take prednisolone could have negative health consequences, indicating that he or she believes that the technology's advice is not always correct. Furthermore, a participant indicated to call the case manager if advised, since she would normally also have done so. As explained before, it should be noted that the actions of calling the case manager and taking prednisolone were part of the action list of the self-management application, but were not presented by the agent itself.

General Attitude towards ECA Design

The last theme we identified was **general attitude towards the ECA design**. The theme does not correspond to our main objective, but we present the findings to provide insight into the context of the results described above. The interviews show that the majority of the participants (seven) did not think that Sylvia had any value, illustrated by comments, such as:



I do not have any connection with Sylvia.



Sylvia is not it.

Arguments supporting this opinion were the agent's statements being too obvious, general or simplistic: a participant described that it was clear that the dialogues were not personalized, but a result of a general set of if-else statements.

Also, Sylvia led to lots of frustration and annoyance, as supported by statements, such as:



I found this female extremely annoying.” and “Sylvia was a very irritating woman.

Frustrations were caused by Sylvia providing incorrect feedback on the inhalations and suggesting actions not fitting the user’s health status, as a participant illustrates:



I thought: ‘Gosh, what are you talking about? I’m not complaining about respiratory problems.’

One participant particularly indicated he would like to switch off the agent. On the contrary, the interviews also showed some positive attitudes towards Sylvia. One participant said that the agent triggered laughing, since Sylvia would adapt the conversation to the answers given. This participant explained:



Occasionally, if I felt bad, I could laugh again.

This participant said that Sylvia made the application more personal, for example, by addressing the user by his or her first name:



It [Sylvia] creates a slightly more informal atmosphere, which I always like, I feel a bit more free.

In addition, this participant believes people should get used to interacting with agents:



When you are at the station, you have this as well [...]. You enter the station and then you face a digital agent. This is something we should get used to, I think.

Lastly, participants suggested improvements for the interaction with the agent. First, two participants explained they would like to be able to type a question in an input field and receive a personalized answer. One of them sketched a scenario in which a patient, who is not feeling well, types in a question into an input field, for example “I am feeling stuffy, but have taken prednisolone: what should I do?” and the application would respond with an answer 24-7. It should be noted that one participant did not understand that Sylvia was a digital agent. He thought that Sylvia represented one of the real people involved in the self-management meetings.

Discussion

This exploratory study aimed to investigate how an ECA’s design is perceived by its users when implemented in a long-term, daily life setting. Although the results of this study should be interpreted carefully, since this study is a small-scale study, they provide first insights into ECA design for self-management and guides for follow-up work in terms of both development and evaluation. Our study shows that the patient’s perception of the agent’s characteristics friendliness, expertise, reliability, involvement and authority did not change over time. The majority of the users were not interested in the agent’s small talk and the likeliness of following advice decreased after three weeks of use.

First, our study shows that the perception of the agent’s characteristics at first glance was similar to that after two weeks and three months of use, suggesting that the user’s first impression does not change over time. To the best of our knowledge, there are no studies on how these perceptions change over time. But, in [chapter 4](#) we showed that an agent’s design affects the user’s perception of this agent after short-term interaction, while *Zhou et al. (2014)* showed that this also applies to long-term interaction.

How do you design an agent for self-management that creates positive impressions that persist? Our results suggest that an agent for self-management should be friendly, reliable and involved and should have expertise, since patients rated these characteristics as important. *Cafaro et al. (2012)* found that an agent’s friendliness was related to the user’s number of agent approaches and likeliness of future encounters with the agent. In addition, the characteristics expertise, reliability and involvement are found to be important aspects of persuasive systems (*Oinas-Kukkonen & Harjuma, 2009*), and eHealth applications in particular (*Mohr et al., 2011; van Velsen et al., 2016; Fogg, 2003*). However, taking this together does not provide much evidence on what agent characteristics are especially important. In our study, patients gave higher scores for Sylvia’s reliability and involvement than for Sylvia’s authority. However, patients also indicated that an agent’s authority is less important than expertise and reliability. This emphasizes the importance for future ECA design studies to ask for both the perception of the characteristics of the agent designed (i.e. the scoring) as well as the perceived importance of these characteristics for an agent in the specific context. With respect to the agent’s authority, our study was indecisive. Different from the quantitative data, qualitative data indicated that patients do prefer an agent portraying authority. These contradicting results might be caused by the patients actually meaning that the agent should have expertise, as they indicate in the interviews that ‘the agent should have authority regarding the topic’. Nevertheless, research confirms that people tend to prefer agents designed to fit their task (*Forlizzi et al., 2007; Zimmerman et al., 2005*). In the context of a self-management intervention for COPD and CHF, we could increase the agent’s expertise by having the agent wear a doctor’s coat. Whether this actually results in a better perception of the agent should be further investigated.

In addition, our study showed that a photorealistic agent could result in users being more likely to follow the agent's advice, compared to a static cartoon. *Van Wissen et al. (2016)* indicate that a more realistic agent appearance increases users' likeliness of following the agent's advice and leads to increased learning of students supported by a pedagogical agent (*Baylor & Kim, 2004*). This increased learning might possibly also apply to a patients' learning on chronic disease self-management. In addition, a realistic agent appearance leads to higher user engagement (*Zimmerman et al., 2005; van Wissen et al., 2016; van Vugt et al., 2007*) and a positive perception of the agent's characteristics, such as its trustworthiness and competence (*van Wissen et al., 2016*). On the other hand, we should avoid the agent being too human-like, since then a mismatch between the users' expectations of the agent and the agent's actual capabilities – a so called negative adaptation gap – could be created, resulting in the users being disappointed (*Komatsu et al., 2012*). Future work should investigate the sweet spot between facilitating expertise (through more realism) and managing expectations of intelligence (through reduced realism).

Furthermore, our study showed that the majority of the users was not interested in the agent's small talk. Although we expected that the small talk would increase users' engagement through companionship building with the agent (*Kang & Gratch, 2011*), this seemed not the case. A possible explanation might be that the amount of small talk might have exceeded the amount of health-related content, and, therefore, distracted the patients from the actual goal of the application: self-management. We expect that it is better to adapt the amount of small talk to the user, for example, by tracking the user's interaction in the small talk dialogues and adapt the amount of small talk in the future accordingly – i.e. users that interact in small talk more often receive small talk more frequently. In addition, the content of the small talk could be adapted to the user. Research shows that tailoring health messages towards personal characteristic pays off (*Beukema et al., 2017*), suggesting that a user's demographics might affect the type of small talk the user is most engaged by. Future work could focus on how small talk can be personalized in order to fit the users' personal values and interests.

Lastly, our results show that patients were less likely to follow the agent's advice over time. We expect that the participants had a negative adaptation gap, meaning that their expectations of the agent's capabilities did not match the agent's actual capabilities (*Komatsu et al., 2012*). After a few weeks of use, the users might have realized that the agent's messages did not always fit their situation, resulting in a decrease of their likeliness of following the agent's advice. In addition, the agent design led to frustrations, mainly caused by non-personalized content and inappropriate feedback, affecting the agent's reliability. Such a mismatch of the content of the agent's message with the user's personal situation was also found in **chapter 4**, in which we evaluated ECAs for health assessment of older adults. Personalizing the agent by providing more specific feedback on user input and health-related data (e.g. sensor inhaler data) might improve the likeliness of individuals to follow the agent's advice. However, the technology readiness level (TRL) of the ECA fits the exploratory character of the study, as explained in the staged approach of telemedicine evaluation (*Jansen-Kosterink et al., 2016*). In the first stages of an telemedicine evaluation, i.e. evaluation of feasibility and user experience, exploratory studies are used to investigate and increase quality of technology, while in later stages, clinical value can be researched with more mature technology in large-scale studies (*Nijland, 2011; Jansen-Kosterink et al., 2016*). As a consequence, participants' expectations of the technology, especially that of the agent, might have exceeded the functionalities and quality of the technology used. Therefore, we emphasize the importance of managing the participants' expectations of the technology used in a study; they should match the actual TRL of the technology. For an agent in particular, it needs to be explained what the user can expect from the agent, which allows one to focus on the objective of the study.

Strengths & Limitations

The strength of our research is that we evaluated the perception of an agent's design at an early development stage with the end-users. Furthermore, participants interacted with an agent in a daily life setting and during a longer period of use: a setting which is rare in agent research, mostly consisting of research on short-term interaction with agents in lab settings. However, this long-term, daily life setting put quite some load on participants. Because of the exploratory character of the study, a limited number of patients participated, which should be taken into account when interpreting the results. However, the results can provide guidance for follow-up agent development and evaluation. Furthermore, participants used a Fitbit, smart scale and smart sensorized inhaler in combination with the self-management application. Many participants complained about the sensors not working properly. This might have affected the participant's perception of the agent, since some of the messages of the agent were based on incorrect sensor information. Lastly, the interviews focused on all elements of the self-management intervention, not specifically on the design of the agent. Not all participants provided information related to the research question of this study, and, therefore, we should be careful with interpreting the results of the interviews.

CONCLUSION

This exploratory study provided first insights on ECA design for long-term, daily use. An agent's design is important for patients to establish a good first impression of the agent, which remains during long-term usage. Based on our findings we expect that ECAs do have potential to be used for self-management, but several design aspects should be investigated in order for ECAs to become successful for increasing engagement in eHealth. When designing ECAs for self-management, we recommend designing an agent that is friendly, reliable, involved and that has expertise, implement and evaluate personalized content and small talk with sufficient variation, and find a good balance between small talk and health-related content. Careful consideration should be given to the apparent realism of the agent to find the sweet spot between facilitating expertise (through more realism) and managing expectations of intelligence (through reduced realism). In combination with managing the user's expectations of the agent capabilities, a personalized ECA design fitting individual needs could increase the agent's reliability and, therefore, the user's likeliness of following the agent's advice. This way, the ECA design could be upgraded to a higher TRL of which the added value and clinical benefits can be evaluated in future research.

Acknowledgements – We thank the patients and care professionals for their participation and feedback. Financial support was provided by the Pioneers in Healthcare innovation fund (2016 MATCH project).

07

Effects of ECA Age and Gender in a Multi-Agent Application in a Long-term, Daily Life Setting

BASED ON:

ter Stal, S., Hurmuz, M., Jansen-Kosterink, S., Beinema, T., op den Akker, H., Hermens, H., & Tabak, M. Older adults' satisfaction with embodied conversational agents of similar age and gender in a multi-agent eHealth application (*submitted for publication*)



Abstract

An eHealth application consisting of multiple Embodied Conversational Agents (ECAs) could support and motivate older adults in healthy aging by covering multiple health domains. This chapter explores whether, in such an application, older adults are more satisfied with ECAs that are similar to themselves in age and gender, both at first glance and after long-term use. We performed a workshop including an ECA card-sorting task (n=20, median 72 years) and a 4-week daily life evaluation (n=45, median 64 years) consisting of a baseline and follow-up ECA satisfaction questionnaire, evaluating seven ECAs differing in age and gender. Results do not show that older adults are more satisfied with ECAs that are similar to themselves in age and gender, neither at first glance nor after long-term interaction. Furthermore, preferences for particular ECAs change over time. Evaluating ECAs' appearances in a multi-agent setting is challenging, since it is difficult to differentiate the ECAs only in appearance. Topic and content of the dialogues have to carefully be selected and adapted to older adults.



INTRODUCTION

As people get older, they get more health complaints, mostly as a result of chronic diseases, such as heart disease, cancer and diabetes (*World Health Organization, 2015*). Since our society is aging, the increasing population of older adults increases the demand for care, while available health care professionals become scarce. By supporting older adults in healthy aging via development of a healthy lifestyle and management of their chronic conditions during daily life, health complaints can be prevented or delayed (*World Health Organization, 2015*). This often requires that older adults change their health-related behaviour. However, changing health-related behaviour can be difficult, since it requires effort and motivation. Therefore, professional health coaches can be essential to support and motivate older adults in this process. Yet, because of our aging society, professional health coaches become scarce. Thus, alternative ways to provide coaching are necessary.

One solution investigated is the use of Embodied Conversational Agents (ECAs). ECAs are defined as more or less autonomous and intelligent software entities with an embodiment used to communicate with the user (*Ruttkay et al., 2004*). By simulating face-to-face interaction, they can build trust and rapport (mutual understanding) (*Vardoulakis et al., 2012*), which can be used to support and motivate older adults in healthy aging. Several studies investigated the use of ECAs to provide this support. In the majority of these studies, ECAs provided support on only one domain, such as physical activity (*Albaina et al., 2009; Bickmore et al., 2013, 2005*). However, today's perspective on health does not only cover a person's physical health, but multiple health domains. For example, *Huber et al. (Huber et al., 2016)(p.1)* suggest a new definition of health: "Health as the ability to adapt and to self-manage, in the face of social, physical and emotional challenges". This definition has been operationalised in six dimensions: (1) bodily functions, (2) mental functions perception, (3) spiritual/existential dimension, (4) quality of life, (5) social societal participation, and (6) daily functioning. Therefore, the studies mentioned above lacked holistic coaching – coaching on multiple health domains.

Other studies did research ECAs for holistic coaching, but in these studies support on multiple health domains was provided by just a single ECA (*Justo et al., 2020; Esposito et al., 2018; Tsiourti et al., 2014*). However, research shows that a multi-agent application with ECAs having their own role has advantages above an application with one ECA having multiple roles. Some of the advantages of such an application are: (1) contrary opinions can be debated (*André and Rist, 2001*), (2) it can help users to classify information, (*André and Rist, 2001*), (3) the same message can be conveyed in multiple ways, (4) ideas can be generated based on different viewpoints (*Hayashi, 2012*), and (5) multiple ECAs can be more persuasive than a single ECA (*Kantheraju et al., 2018*). Thus, a multi-agent application consisting of ECAs with expertise in their own health domain allows for holistic coaching and could be a means to engage older adults in healthy aging.

But, how do we design ECAs in a multi-agent eHealth application for older adults? Research on the design of eHealth applications for older adults in general teaches us that personalization of the applications is an important factor in older adults' acceptance of the applications (*Beukema et al., 2017; Panico et al., 2020*). *Heerink et al.* shows that this also applies to ECAs: perceived adaptiveness affects older adults' acceptance of assistive social agent technology. Thus, we have to tailor the ECAs to the older adults' preferences. Just as a human's appearance in human-human interaction, an ECA's appearance affects a user's experience in human-agent interaction (*Bergmann et al., 2012*). But how

should the appearance of ECAs in a multi-agent application be designed such that it fits preferences of older adults? Little research describes how to do so. *André and Rist* suggest that we should not simply reuse the design of an ECA for multiple ECAs (*André and Rist, 2001; André et al., 2018*). As they explain: “characters have to be realized as distinguishable individuals with their own areas of expertise, interest profiles, personalities, and audiovisual appearance, taking into account their specific task in a given context”. However, this suggestion does not provide us with concrete guidelines on how to design the ECAs’ appearances.

Previous research on ECAs for older adults in single agent eHealth applications shows positive effects of similarity in age (*Azevedo et al., 2018*) and gender (for females, as shown in [chapter 3](#)). On the contrary, some research shows that, independent of their own gender and age, older adults prefer female (*Justo et al., 2020; Esposito et al., 2018, 2019*) and younger ECAs (*Justo et al., 2020*). How does similarity in age and gender affect older adults’ interaction in a multi-agent application? To the best of our knowledge, no research has been performed on similarity in age and gender for ECAs in multi-agent applications for older adults.



Research Objectives

Therefore, this research aims to study the effects of similarity in age and gender for older adults at first glance and after long-term use of a multi-agent application to support healthy aging,

We focus on older adults’ preferences for and satisfaction with the ECAs. We hypothesize that, in line with previous work on single agent applications, older adults are more satisfied with ECAs that are similar to themselves in age and gender.

METHODS

The Multi-Agent eHealth Application

The ECA designs subjected to testing were part of a multi-agent eHealth application, called Council of Coaches (*op den Akker et al., 2018*). The multi-agent eHealth application was designed to support older adults with age-related impairments, chronic pain or diabetes type 2, and living healthy in general. The application featured six coaches supporting the user in the following domains: (1) physical activity, (2) nutrition, (3) social activity, (4) cognition, (5) chronic pain, and (6) diabetes. In addition, one ECA provided peer support. The application connected with a Fitbit activity tracker, such that the user could be coached based on his or her activity behaviour.

Users picked their own coaches (based on portrait pictures as shown in [Figure 07.2](#) and small persona descriptions), but the physical activity and nutrition coach were mandatory. In addition, the chronic

pain and diabetes type 2 coaches were only available to those having either chronic pain or diabetes type 2. Users interacted with the ECAs in the main interface (Figure 07.1) through text-based dialogues consisting of a message of an ECA and one or multiple replies for the user. Users chose which ECA to interact with, but other ECAs could join the conversation.

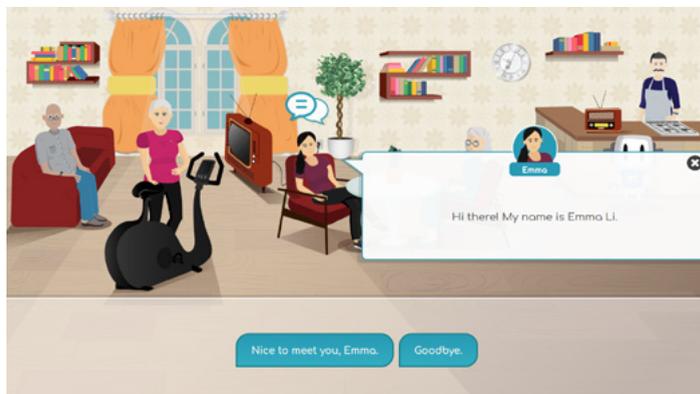


Figure 07.1 – Main interface of the multi-agent application. In this example, the user clicked on coach Emma, opening the dialogue box with a message from the coach and two response options for the user. By clicking on one of the response options the dialogue continues. A next step in the dialogue could either consist of a message from Emma, or a ‘guest’ message from one of the other coaches.

The ECA Designs

We developed a set of seven distinguishable ECA designs varying in age and gender. Since the multi-agent application was developed for older adults, we included both a male and female ECA of above 55 years. In addition, we adapted the clothing of the ECAs to their domain. For example, the physical activity coach was wearing sport clothes and the nutrition coach had an apron. An overview of the personas, including their appearance can be seen in Figure 07.2.

Olivia Simons	François Dubois	Emma Li	Helen Jones	Carlos Silva	Rasmus Johansen	Katarzyna Kowalska
Physical activity	Nutrition	Social	Cognition	Peer	Chronic pain	Diabetes
Amsterdam, Netherlands	Pau, France	Pennsylvania, USA	London, England	Sines, Portugal	Roskilde, Denmark	Lowicz, Poland
52 years ('middle aged')	45 years ('middle aged')	28 years ('young')	64 years ('old')	67 years ('old')	33 years ('young')	45 years ('middle aged')

Figure 07.2 – Personas of the seven ECAs.

Experimental Procedure

We used a mixed-methods design, consisting of a workshop performed with a first set of participants and a 4-week daily life evaluation of the ECA designs as part of the multi-agent eHealth application with a second set of participants. The daily life evaluation was part of larger evaluation of the application as a whole (Hurmuz *et al.*, 2020). **Satisfaction at first glance** was determined by an ECA card-sorting task in the workshop and a baseline satisfaction questionnaire. **Satisfaction after long-term use** was determined after the four-week daily life evaluation with a follow-up satisfaction questionnaire and semi-structured interview. An overview of the methods used can be seen in [Figure 07.3](#).

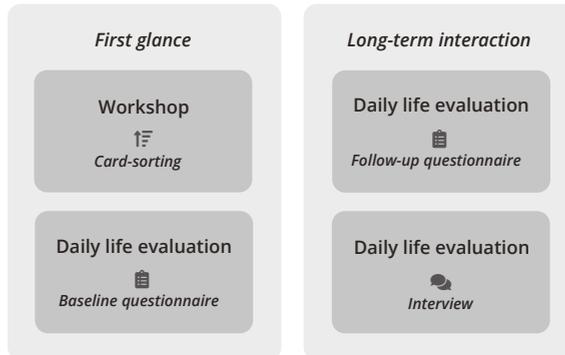


Figure 07.3 – Overview of the methods used to investigate users' satisfaction with the ECAs at first glance and after long-term interaction.

Workshop

The workshop with adults of 55 years or older was performed by the third author in the Netherlands in December 2019. It consisted of a card-sorting task in which the participants sorted the portrait images of the seven ECAs according to preference, and a plenary discussion afterwards. First, participants were introduced to the concept of the multi-agent application and the seven ECAs. Second, participants ranked the portrait images of the ECAs. Participants were asked to rank the ECAs when acting as coaches to develop and maintain a healthy lifestyle. They ranked the ECAs individually on a paper sheet stating on the left "ECA I prefer the least to be my coach" and on the right "ECA I prefer the most to be my coach".

4-week Daily Life Evaluation

The daily life evaluation was performed in Scotland and the Netherlands from February 2020 till May 2020. Participants were 55 years of age or older, or 55 years of age and older and either suffering chronic pain, diagnosed by a general practitioner, or diabetes type 2, diagnosed and controlled by a general practitioner.

At baseline, a researcher installed the application and a Fitbit with each participant individually. Participants chose their coaches either independently, or with help from the application (i.e. coaches were suggested based on results of a baseline health questionnaire). Then, participants filled out the baseline satisfaction questionnaire, rating their satisfaction with each ECA individually (seven items, graded on a scale from 1 - 10, as shown in [Appendix F](#)). Afterwards, participants were asked to wear the Fitbit without using the application for one week to gather baseline data on their physical activity.

After one week, participants could use the application for four weeks*. After four weeks, they filled out the follow-up ECA satisfaction questionnaire (as shown in [Appendix F](#)) and were interviewed. In these individual semi-structured interviews, participants were asked about the coaches they preferred least and most and why (e.g. because of the age, gender, domain, content or personality) and how they perceived these ECAs' personalities (e.g. friendly, reliable). In addition, they could comment on the ECAs' appearances in general.

Data Analyses

All data were analysed using the SPSS 25 statistics program. All tests were conducted using a 95% confidence interval.

Workshop

The ECA preferences were labelled with a score from 1 to 7. A non-parametric Friedman test of differences for repeated measures was conducted to test for any difference in the mean scores of the ECA designs. To test for a difference in mean scores for the three ECA age groups (young, middle, old), a mixed model analysis was conducted. A paired samples t-test was used to assess differences in a preference for ECA gender, for both the female and male participants.

4-week Daily Life Evaluation

Participants rated their satisfaction with each ECA on a 1-10 scale. To test for differences in mean satisfaction score for all ECAs and ECA age groups, mixed model analyses with pair-wise comparisons as post-hoc tests were conducted. Paired samples t-tests and Wilcoxon signed-rank tests were used to test for differences in satisfaction scores for female and male ECAs, both for female and male participants. The statistical tests were performed at baseline and follow-up.

The interviews were transcribed by four researchers and were thematically analysed by the first and second author independently. The themes were coded retrospectively using ATLAS.ti 8. The first author created a first coding scheme and labelled all the data accordingly. The second author used the coding scheme to label the data as well. Disagreements between the first and second author were discussed and overcome.

RESULTS

Workshop

Participants

Twenty people participated in the workshop of which six were female and fourteen were male. The age of the participants ranged from 64 to 78 years (median = 72 years).

* Dialogues available upon request.

Overall Rating ECAs

Figure 07.4 shows that on average, Olivia, the physical activity coach, and Emma, the social coach, scored highest in the card-sorting task ($M = 5.0$ and $M = 4.1$ respectively). Katarzyna, the diabetes coach, and Rasmus, the chronic pain coach, scored lowest. Results show no significant difference between the mean ranks of the scores of the different ECAs ($p = 0.10$, $\chi^2 = 10.799$).

The majority of the participants selected Olivia ($n=8$), as preferred ECA, followed by Helen, the cognition coach ($n=2$). Rasmus, Katarzyna and Carlos, the peer coach, were all mentioned once as the preferred ECA, and François, the nutrition coach, and Emma were never selected as preferred ECA.

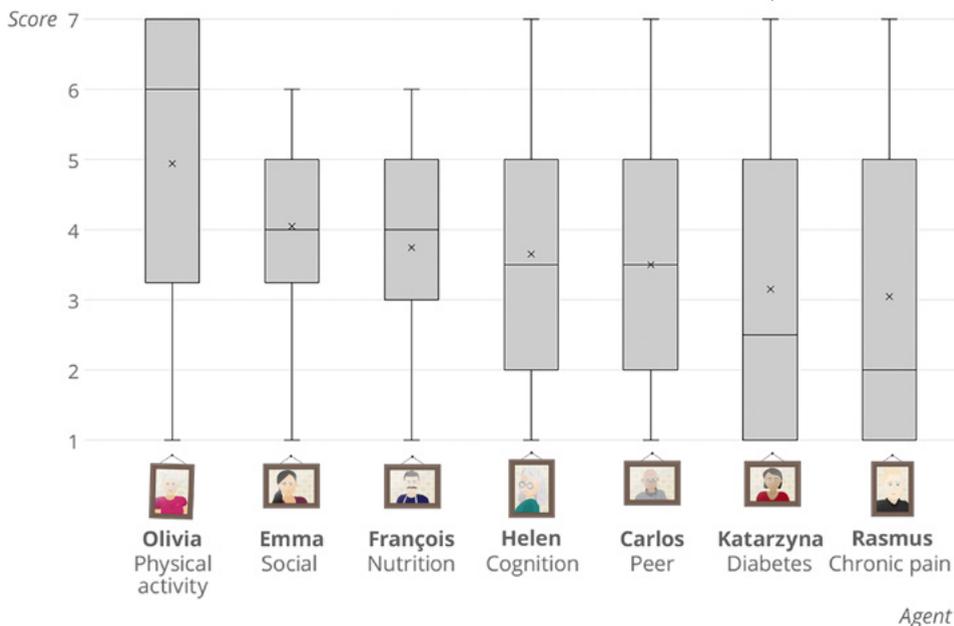


Figure 07.4 – ECA scores of the card-sorting (scale from 1 to 7).

Satisfaction ECAs Similar in Age

Table 07.1 shows no significant difference in the mean scores of the ECA age categories ($p = 0.61$, $F = 0.611$).

Table 07.1 – Results of the mixed model analysis comparing the mean scores of the ECA age categories of the card-sorting.

Score (M±SD) Young ECAs	Score (M±SD) Middle aged ECAs	Score (M±SD) Old ECAs	p	F
3.6±1.0	4.0±0.9	3.6±1.3	0.61	0.611

Satisfaction ECAs Similar in Gender

Table 07.2 shows that for both the female and male participants, neither female nor male ECAs scored significantly higher in the card-sorting task.

Table 07.2 – Results of the paired sample t-tests comparing the mean scores of the female and male ECAs from the card-sorting for both the female and male participants.

	Score (M±SD) Female ECAs	Score (M±SD) Male ECAs	p	t
<i>Females</i>	3.8±0.7	3.1±1.0	0.25	1.311
<i>Males</i>	4.0±0.7	3.6±1.0	0.32	1.030

4-week Daily Life Evaluation

Participants

Forty-five older adults participated. The youngest participant was 55 years and the eldest 87 years (median = 64 years). **Table 07.3** provides the rest of the participants' demographics. Thirty-eight participants finished the study. We analysed the ECA satisfaction scores of forty-five participants at baseline and thirty-eight participants at follow-up.

Table 07.3 – Demographics of the participants in the 4-week daily life evaluation.

Demographic	Value	n
<i>Country</i>	The Netherlands	26
	Scotland	19
<i>Gender</i>	Male	15
	Female	30
<i>Education</i>	Higher vocational education /University	25
	Higher general secondary education / Pre-university	13
	Preparatory secondary vocational education	7
<i>Trying new technologies</i>	Like	29
	Neutral	8
	Dislike	8

Satisfaction All ECAs

Table 07.4 shows the mean satisfaction scores of all ECAs at baseline and follow-up (after four weeks of use). At baseline, Emma scored highest, and Olivia scored lowest. Furthermore, a significant difference in ECA baseline satisfaction scores of all ECAs was found ($p < 0.001$, $F = 300.265$). Post-hoc pairwise comparisons between all ECAs, show that only Emma was rated significantly higher than Carlos ($p = 0.01$), Olivia ($p = 0.01$), Rasmus ($p = 0.01$) and François ($p = 0.02$).

At follow-up, Olivia scored highest, and Rasmus scored lowest. Again, a significant difference in satisfaction scores of all ECAs was found ($p < 0.001$, $F = 13.163$). Post-hoc pairwise comparisons show that Olivia was rated significantly higher than Carlos ($p < 0.001$), Helen ($p = 0.03$) and Emma ($p = 0.04$) and that François was rated significantly higher than Carlos ($p = 0.01$) and Helen ($p = 0.05$). Furthermore, Emma was rated significantly higher than Carlos ($p = 0.02$) and Helen ($p = 0.04$). Helen was rated significantly higher than Carlos ($p = 0.05$). Lastly, results show that the satisfaction scores dropped from baseline to follow-up (significant for all, except Rasmus and Katarzyna).

Table 07.4 – Average satisfaction score per ECA at baseline and follow-up (after four weeks of use). Horizontal: Results of the mixed model analysis comparing the mean satisfaction scores of the different ECAs at baseline and follow up. Vertical: comparison of the mean scores at baseline and follow-up per ECA. The ECAs with a * were not used in the analysis, due to a low n.

	Olivia (M±SD)	François (M±SD)	Emma (M±SD)	Helen (M±SD)	Carlos (M±SD)	Rasmus (M±SD)	Katarzyna (M±SD)	p
Baseline	6.6±2.2	6.9±1.7	7.5±1.1	7.0±1.9	6.8±1.8	7.0±1.4	7.1±1.2	<0.001
Follow-up	5.2±2.5	5.1±2.6	4.4±2.4	4.1±2.5	4.0±2.5	*2.0±1.4	*4.0±4.2	<0.001
p	0.005	0.005	<0.001	<0.001	<0.001	-	-	

Satisfaction ECAs Similar in Age

Table 07.5 shows that, only at follow-up, a significant difference in the mean scores of the ECA age categories was found ($p < 0.001$). Post-hoc pairwise comparisons show that after four weeks of use middle aged ECAs were rated significantly higher than old ECAs ($p < 0.001$) and young ECAs ($p = 0.01$).

Table 07.5 – Results of the mixed model analysis comparing the mean satisfaction scores of the ECA age categories at baseline and follow-up (after four weeks of use).

	Score (M±SD) Young ECAs	Score (M±SD) Middle aged ECAs	Score (M±SD) Old ECAs	p	F
Baseline	7.4±1.1	6.8±1.6	6.9±1.7	0.08	2.371
Follow-up	4.3±2.4	5.1±2.3	4.0±2.4	<0.001	72.397

In the interviews, we coded the following themes for the participants' perceptions on the age of the ECAs: ECA preferred least because of its age, ECA preferred most because of its age, general statement on age of ECAs. Participants did not indicate the ECA's age to be the reason for preferring a coach the least. One participant indicated he most preferred François, since he looked young. In general, six participants would like their coaches to be young, as one indicated: "That greyish is not that appealing" (P5). Another participant said: "Is this a room in a nursing home where patients sit together? When it appeared these patients were the coaches, I thought: 'Well, okay, they are all really old' " (P6). Two participants particularly commented on Carlos looking old, stating "He looked like an old, bald, fat man and I thought this was quite off-putting actually, as that's not how I see myself. You don't want to think of yourself like that." (P43). One participant said that some of the coaches were really old and dusty. Lastly, another participant said that the coach's age does not matter.

Satisfaction ECAs Similar in Gender

Table 07.6 shows that both at baseline and follow-up, there was no significant difference in satisfaction with a particular ECA gender among both female and male participants.

Table 07.6 – Results of the paired sample t-tests (baseline) and Wilcoxon Signed-Rank Tests (follow-up) comparing the mean scores of the female and male ECAs for the females and males.

		Score (M±SD) Female ECAs	Score (M±SD) Male ECAs	p	t
Females	Baseline	6.6±1.6	6.7±1.4	0.65	0.467
	Follow-up	4.2±2.2	4.2±2.0	0.68	-
Males	Baseline	7.0±1.1	7.4±1.6	0.53	0.650
	Follow-up	4.9±2.3	4.8±2.6	0.77	-

In the interviews, we coded the following themes for the participants' perceptions on the gender of the ECAs: **ECA preferred least because of its gender**, **ECA preferred most because of its gender**, **general statement on gender of ECAs**. Participants did not indicate the ECA's gender to be the reason for preferring a coach the most or the least. However, two participants did indicate that, **in general, they do not mind whether their coaches are male or female**.

Older Adults' Preferences for Particular ECAs

In addition to the ECA's age and gender motivating participants' preferences for particular ECAs, we coded the following themes in the interviews motivating the participants' preferences for particular ECAs: **personality, background story, domain, content and small talk**.

The majority of the participants did not prefer one ECA the most. Some found it difficult to make a judgment, because they interacted little. One participant indicated not having a connection with any of the coaches, whereas another said: "I like them all, they are all useful" (P10). Olivia was preferred most, mostly because of her **content**: concrete advice, realistic goals and valuable feedback. A few participants were especially interested in the **domain** of physical activity and, therefore, preferred Olivia. Furthermore, some commented on her **personality**: she was found to be friendly and positive. In addition, François was preferred often. Reasons given related to his **content**: providing concrete advice and interesting recipes. One participant particularly liked François' **personality**: a friendly coach by the way he addressed the user and provided ideas. Another participant found François pleasant and funny. One participant explained: "You can visualize him, he comes across like a character from a book. He was very jovial" (P45). A few participants preferred François, since they were interested in the **domain** of nutrition. One participant preferred both Olivia and Helen, liking their **background stories**. Also, another preferred Helen, but did not know whether this was caused by either her **content** or appearance. Emma was preferred twice, due to her content, providing valuable advice, and her **personality**: she was found to be friendly. One participant preferred Emma, since she talked about the **domain** he or she needed most help with. Carlos was preferred just once, by a participant that enjoyed his **small talk**.

The majority of the participants did not prefer one coach the least, since they believed they all had limited **content**. Two found all coaches equally sweet. From the rest of the participants, the majority

least preferred Carlos, because of a lack of **content**, focusing on **small talk** instead of concrete goals, or because of an **absence of personality**. Also, many participants least preferred François. Reasons given were: not liking his **background story** and his **content** being little concrete, focusing on small talk. One participant was not interested in the **domain** nutrition and a few found François' **personality** annoying, conveyed via stereotype and popular French words. Also, some found him offensive. Only one participant least preferred Emma, since she had little **content**: "I don't have to like her. I just want results. She has to say something to me that's good for me" (P9). Also, just one participant preferred Olivia the least, not liking her **background story**.

Older Adults' Perceptions of Other Design Features

In the interviews, we lastly coded the participants' perceptions of **other ECA design features**. Overall participants said the ECAs' appearances were fine or that they do not care about the appearances. Some participants described the ECAs as simplistic or childish cartoons, digital figures and even robots, and indicated they did not have any connection with them. As one participant indicated: "I found it pretty off-putting and a little condescending" (P43). On the contrary, some thought that it was good that the ECAs were simplistic and cartoonish.

DISCUSSION

This research shows that older adults do not have a preference for ECAs with similar age or gender in a multi-agent eHealth application targeting the following health domains: (1) physical activity, (2) nutrition, (3) social activity, (4) cognition, (5) chronic pain, and (6) diabetes. This was reflected in ECA preference and satisfaction scores at first glance and after four weeks of use.

First, **older adults were not more satisfied with older ECAs**. In fact, it seems that older adults are more satisfied with younger and middle aged ECAs than older ECAs. To the best of our knowledge, there is no existing work in multi-agent applications focusing on the ECAs' age. However, our results might be explained by work on single agent applications. Similar to our results, results of a study by *Justo et al.* show that a majority of older adults prefers an ECA for healthy living to look between 29 and 48 years old. How can we explain this preference? Research indicates that people prefer ECA designs that fit the ECA's domain conform stereotypes (*Forlizzi et al., 2007; Nguyen and Masthoff, 2007; Parmar et al., 2018; Ring et al., 2014; Veletsianos, 2010*). In our study, the ECAs' dialogues were on healthy living. Older ECAs might have been negatively associated with being healthy, conforming stereotypes, resulting in a lower user satisfaction. In addition, a setting consisting of multiple older ECAs might make users more aware of the application being designed for older adults than a single agent application with one older ECA, since it could be perceived as a living room in a nursing home for elderly. But, as our interviews indicated, older adults might not think of themselves as elderly (*Chasteen et al., 2002*). They might not identify themselves with older ECAs, and, thus, prefer younger ECAs.

Second, **older adults were not more satisfied with ECAs of the same gender as themselves**. To the best of our knowledge, there is no existing work in multi-agent applications focusing on the ECAs' gender. However, as explained before, from single agent research we know that people prefer ECA designs that fit the ECA's domain conform stereotypes (*Forlizzi et al., 2007; Nguyen and Masthoff, 2007; Parmar et al., 2018; Ring et al., 2014; Veletsianos, 2010*). In our study, the ECA designs were adapted to

their health domains. By doing so, we might have aligned the gender of the ECAs in line with participants' known gender stereotypes. Since we designed the ECAs in line with both known female and male stereotypes, this might explain the absence of a preference for either of the sexes. Furthermore, we believe that humans perceive a multi-agent setting consisting of both male and female coaches as more natural than a setting of only male or female coaches: a real-life setting with multiple human coaches would most probably also consist of coaches from both sexes.

Overall, we found that **older adults' satisfaction with ECAs changes over time**. In our study, older adults' satisfaction was higher at first glance than after daily life interaction and preferences for particular ECAs changed over time. We expect that an ECA's appearance, such as its gender and age, is more important at first glance than after daily use. At first glance, the ECA's appearance helps users forming impressions of the ECA's personality, as in human-human interaction (*Bar et al., 2006*). Establishing a positive first impression lowers the threshold to interact with the ECAs (*Kelley, 1950*). Then, users' judgments of the ECAs modify with ongoing interaction, as research shows that ECAs do have a second chance to make a first impression (*Bergmann et al., 2012; Komatsu et al., 2012*). After daily interaction with an ECA, we expect that topic (in this case health domain) and content of dialogues becomes more important. This is reflected in our results showing a significant decrease in satisfaction with the ECAs after four weeks of use, and participants' opinions in the interviews. This decrease could be explained by the older adults having a negative adaptation gap, meaning that their expectations of the ECAs' capabilities exceeded the ECAs' actual capabilities (*Komatsu et al., 2012*). After daily use, the older adults might have realized that the content of ECAs' dialogues was not as extensive, smart and personal as they expected. Little personalization and inferior or stigmatized content seen as a threat to their autonomy, might have affected the older adults' satisfaction with the ECAs (*Beukema et al., 2017; Panico et al., 2020; Troncone et al., 2020*). Since the amount and the content of dialogues differed across the ECAs, this adaptation gap might have been larger for some ECAs than for others, which might have affected our results.

Strengths and Limitations

Our study was part of a small-scale evaluation focusing on usability and end-user experience of a multi-agent application for healthy aging having a limited technology readiness level (TRL). According to *Jansen Kosterink et al.*, small-scale eHealth evaluation studies, such as ours, allow researchers to gain detailed information that can be used for further improvement of a eHealth service (*Jansen Kosterink et al., 2016*). These improvements lead to technologies of a higher TRL of which clinical value can be evaluated in extensive large-scale evaluation studies. Thus, our results provide some first insights into older adults' perceptions of ECAs of different ages and gender in a multi-agent application, which are important for further improvement and implementation of the application.

However, our results should be interpreted with caution. First, since our application consisted of seven ECAs, all ECA age groups and gender were present, but not equally balanced (the application consisted of two young, two old and three middle-aged ECAs, and four female and three male ECAs respectively). Second, as a drawback of the evaluation stage of our study, the content of the dialogues was limited (fitting a limited TRL), which might have affected users' perceptions of the different ECA designs. Third, as each ECA coached the user on a different health domain, preferences for particular ECAs might have been a result of users' preferences for particular topics and content of the corresponding dialogues. Though, these study limitations have taught us important aspects for future work.

Guides for Future Work

Future research is needed to study whether older adults indeed prefer a multi-agent eHealth application consisting 1) of younger or middle-aged ECAs over one with only older ECAs and 2) of a mixture of female and male ECAs over one consisting of only male or female ECAs. This could be realized by studying older adults' satisfaction with different ECA age groups in between-subject experiments in which participants are assigned to either a multi-agent eHealth application consisting of younger, middle-aged or older ECAs. Similarly, between-subject experiments in which participants are assigned to either a multi-agent eHealth application consisting of male ECAs, female ECAs or a mixture of both female and male ECAs could be performed to study older adults' satisfaction with ECAs of different gender.

Furthermore, to eliminate influence of preferences for particular health domains and content of the ECAs' dialogues on ECA satisfaction scores, future research could include between-subject experiments in which participants are assigned to multi-agent applications in which the health domains and dialogues are rotated between the different ECAs.

CONCLUSION

Evaluating ECAs' appearances in a multi-agent setting is challenging, since it is difficult to differentiate the ECAs only in appearance. Topic and content of the dialogues have to carefully be selected and adapted to older adults. First, we advise to balance the number of ECAs per age group and gender. Second, future research could study older adults' satisfaction with ECAs when rotating the expertise (health domains) among ECAs of different age and gender, such that satisfaction does not depend on the topic and content of the dialogues.

Acknowledgements – This work was supported by the European Commission's Horizon 2020 Research and Innovation Programme project Council of Coaches [Grant Agreement Number 769553]. We would like to thank Katrien Fischer, Ellis Oude Kempers and Dominic De Franco for their contribution to the interviews and the transcription of them.

08

State-of-the-Art and Design Strategies for ECA Appearance

BASED ON:

ter Stal, S., Kramer, L. L., Bulthuis, R., Hermens, H., & Tabak, M. Embodied conversational agents for eHealth: State-of-the-art and design Strategies for appearance (*submitted for publication*)



Abstract

This chapter aims to create design strategies for an ECA's appearance in eHealth for user engagement. First, we updated our initial state-of-the-art review (*chapter 2*) on design features for ECAs, and their effects. We also included the studies performed in this thesis. The 46 included articles in the updated literature review were of low, fair or moderate quality. For the majority of the design features no or preliminary evidence was found. Next, we synthesised these findings into design strategies, being: 1) adapt the ECA's demographics to the user's preferences, 2) adapt the human-likeness of the ECA to its functionalities, 3) convey an ECA's expertise using its appearance, 4) implement emotion in the ECA's expressions, and 5) implement an ECA's relational behaviour. Lastly, we apply these strategies in a use case: an eHealth application for older workers.



INTRODUCTION

Many eHealth applications encourage users to adapt or sustain a healthy lifestyle, often via behaviour change interventions. The success of such an eHealth application (e.g. to reach efficacy) is affected by users' adherence to the intervention offered. Adherence to an intervention is positively associated with users' engagement: those who are more engaged are significantly less likely to stop using the eHealth application (Yardley et al., 2016; Kohl et al., 2013; Scherer et al., 2017; Crutzen et al., 2011). However, in literature, no clear definition of engagement exists and the definition varies among research areas. For example, the definition in gaming is different from the definition in marketing or education (Yardley et al., 2016; O'Brien & Toms, 2008). As an example, in gaming, engagement refers to the motivation of the player to play and keep playing (Crutzen et al., 2016). O'Brien & Toms de-constructed the term engagement as it applies to people's experiences with technology in general. They define engagement as a quality of user experience characterised by attributes of challenge, positive affect, endurance, aesthetic and sensory appeal, attention, feedback, variety/novelty, interactivity, and perceived user control.

In literature, engagement is often interchanged with many other terms, such as enjoyment – the action or state of deriving gratification from an object – and immersion – a state of high motivation to interact with a technology, while retaining some awareness of one's surroundings (Crutzen et al., 2016). Enjoyment and immersion refer to cognitive and affective processes that take place while the user is interacting with the technology. On the other hand, engagement refers to the motivation to keep using; meaning that highly engaged users not only think about the technology while interacting, but also during daily life (Crutzen et al., 2016). To distinguish between all these terms, Crutzen et al. conceptualise engagement as a motivational construct that goes beyond the specific time that is spent interacting with the technology. This construct consists of several determinants of a user's subjective experience, also including enjoyment and immersion. These determinants of engagement may positively affect a user's adherence to an application and we could thus say that they are important for the success of eHealth applications.

To increase user engagement in eHealth, Embodied Conversational Agents (ECAs) could be implemented. ECAs are more or less autonomous and intelligent software entities with an embodiment used to communicate with the user (Ruttkay et al., 2004). ECAs could contribute to engagement with eHealth applications in several ways. First, ECAs could contribute to engagement by creating user experiences that are more fun, absorbing and intrinsically enjoyable, since ECAs are a form of interactive multi-media (Lefebvre et al., 2010; Yardley et al., 2016). Second, ECAs could contribute to engagement by providing users with social support, which is one of most important persuasive drivers in eHealth (Kelders et al., 2012). ECAs can provide this social support by building trust and rapport (e.g. mutual understanding) with the user, leading to companionship. Especially in long-term care it is important to create such a bonding relationship to engage the user (Kelders et al., 2012).

As such, incorporating ECAs into eHealth applications could possibly promote user engagement. However, it is important that a user's first impression of an ECA, triggered by the ECA's appearance, is positive. As in human-human interaction, people that have a positive first impression of someone tend to interact more with that person (Bergmann et al., 2012; Kelley, 1950). Long-term interaction is essential to sustain a user's engagement. An ECA's appearance is thus important for user engagement, but how should the appearance of ECAs be designed to positively affect the determinants that could elicit user engagement? (Table 08.5)



Table 08.5 – How an ECA’s appearance may contribute to eHealth efficacy. Conceptualisation base on the work of Crutzen et al.

Some research has been performed on developing a set of ECA design features. For example, *Ruttkay et al.* developed a taxonomy for relevant design and evaluation features of ECAs (*Ruttkay et al., 2004*). They distinguish among the ECA’s embodiment – its physical appearances (its looks, speech and/or textual output, hand and body gestures and facial and gaze expressions), mental capacities (its social role, personality, user model, natural language generator and dialogue manager) and the application interface (including background knowledge processing). In addition, *Straßmann and Krämer (2017)* identify design features related to the ECA’s appearance. They categorise the following features: embodiment vs. no embodiment, species, realism, 2D vs. 3D and feature specification (socio-demographic and styling). Although these taxonomies support establishing a common ground for developing an ECA’s appearance, they do not show how to actually design these features, let alone for an eHealth context. As shown by *ter Stal et al. (2020)*, consensus on effects of design features for ECAs in eHealth is far from established. To provide designers, developers and researchers of ECAs in eHealth with practical guidelines for developing an ECA’s appearance, a set of design strategies could be valuable. Some ECA design strategies exist, such as the strategies for pedagogical agents by *Veletsianos et al. (2009)*, but these strategies focus little on the ECA’s appearance.

Research Objectives

Therefore, in this chapter, we will answer the following question: how should the appearance of ECAs be designed to positively affect user engagement?

To answer this question, we take the conceptualisation of engagement (visualised in Figure 1) based on work by *Crutzen et al.* as a guideline. The aim of this research is to create a set of design strategies for an ECA’s appearance in eHealth to contribute to the determinants of user engagement. First, a state-of-the-art review is performed on the design features for ECAs in eHealth, the outcome variables used to measure the effect of these design features and the found effects of each feature, was performed (section 2). Next, we suggest a set of design strategies based on a synthesis of the findings of the state-of-the-art. Lastly, we show how these strategies can be used with an example of an ECA as part of a mobile eHealth application for older workers.

REVIEW STATE-OF-THE-ART

In this section, we describe the state-of-the-art of the design features and effects on (determinants of) user engagement for ECAs in eHealth.

Method

First, a taxonomy of design features for an ECA's appearance was created by combining the categories identified by *Ruttkey et al.* and *Straßmann and Krämer*, as seen in **Table 08.1**. Previously, we conducted a literature review (*ter Stal et al., 2020*) in which we used this taxonomy for article selection. In this original review, we identified the researched design features for ECAs in eHealth, the outcome variables that were used to measure the effect of these design features and what the found effects for each variable were (based on articles published till December 2018). Since the research area of ECAs is growing, this review was repeated for this chapter using the same methodology, but now including articles from December 2018 to July 2020.

Searches were performed in 5 databases and searches were restricted to queries containing terms related to (1) embodied conversational agent and (2) eHealth. As part of this PhD thesis, we complemented the state-of-the-art by our own experimental research, as shown in the **chapters 4, 5 and 7**. Lastly, four researchers in the field of ECAs were asked to take part in a digital co-design session to provide additional articles for the-state-of-the-art, if missing.

Table 08.1 – Taxonomy of design features and categories used in the literature reviews and throughout this chapter.

Main category	Sub category	Design features (examples)
<i>Speech and/or textual output</i> (<i>Ruttkey et al., 2004</i>)	-	Emotion, medium (voice or text)
<i>Facial and gaze expressions</i> (<i>Ruttkey et al., 2004</i>)	-	Emotion, animation
<i>Hand and body gestures</i> (<i>Ruttkey et al., 2004</i>)	-	Animation
Looks (<i>Straßmann & Krämer, 2017</i>)	Species	Agency (human, animal, robot)
	Realism	Rendering style (cartoon, photo-realistic)
	Socio-demographic	Gender, age, role
	Styling	Clothing

Quality Classification Individual Articles

All findings were grouped into a table consisting of the design features and outcome measures of the research. In addition, every article was labelled with a study quality label. Labels were assigned based on the evaluation stage of the article (according to the renewed framework for the evaluation of telemedicine by *Jansen Kosterink et al.*) and the number of study participants in the following way:

1. **Low** – Stage I (technical efficacy): focus on the feasibility and usability of the technology, and less than 50 participants.
2. **Fair** – Stage I (technical efficacy): focus on the feasibility and usability of the technology, and more than 50 participants.
3. **Moderate** – Stage II (specific system objectives): gaining an initial idea about the potential added value for clinical practice and possible working mechanism.
4. **Good** – Stage III (system analysis): technology evaluated in the way they will be implemented in daily clinical practice.
5. **Excellent** – Stage IV (external validity): elaboration of the adoption as addressed in stage III.

Evidence Classification per Design Feature

All articles that were included in the original and updated review of the state-of-the-art and experimental research were grouped according to the design features that they evaluated. Then, the total evidence for every design feature was classified. For the evidence classification, the three domains for determining the strength of evidence (*Lohr, 2004*) were taken into account: 1) the quantity of the articles, 2) the quality of the articles, and 3) the consistency of findings between the articles. Our classification based on these domains was as follows:

- **Strong evidence (★★★★★)** – At least ten studies report effects of the design feature on engagement or engagement determinants and at least ten of them have at least good evidence and do not show inconsistent findings.
- **Solid evidence (★★★)** – At least ten studies report effects of the design feature on engagement or engagement determinants and at least seven of them have at least moderate evidence and do not show inconsistent findings.
- **Moderate evidence (★★★)** – At least five studies report effects of the design feature on engagement or engagement determinants and at least three of them have at least moderate evidence and do not show inconsistent findings.
- **Preliminary evidence (★★)** – At least three studies report effects of the design feature on engagement or engagement determinants and at least two of them have fair evidence and do not show inconsistent findings.
- **No evidence (★)** – The remainder of the articles.

Results

Description Findings Articles Updated Review and Own Experimental Research

For the updated review, we identified 943 articles published from December 2018 to July 2020 through the database searches, of which 10 articles were included. **Figure 08.1** shows the flow diagram of the database searches and article screenings. In addition, three articles of our own experimental research were added (**chapter 4, 5 and 7**). **Table A.3 (Appendix G)** lists general information about the articles found and the design features researched, corresponding outcome variables and results for each article of the updated review and our own experimental research. All articles were categorised in design feature categories. Eight articles investigated multiple design features. The findings of the individual articles of the original review with papers published before December 2018 can be read in *ter Stal et al., 2020*. In the remainder of this section we describe the design features found in the articles

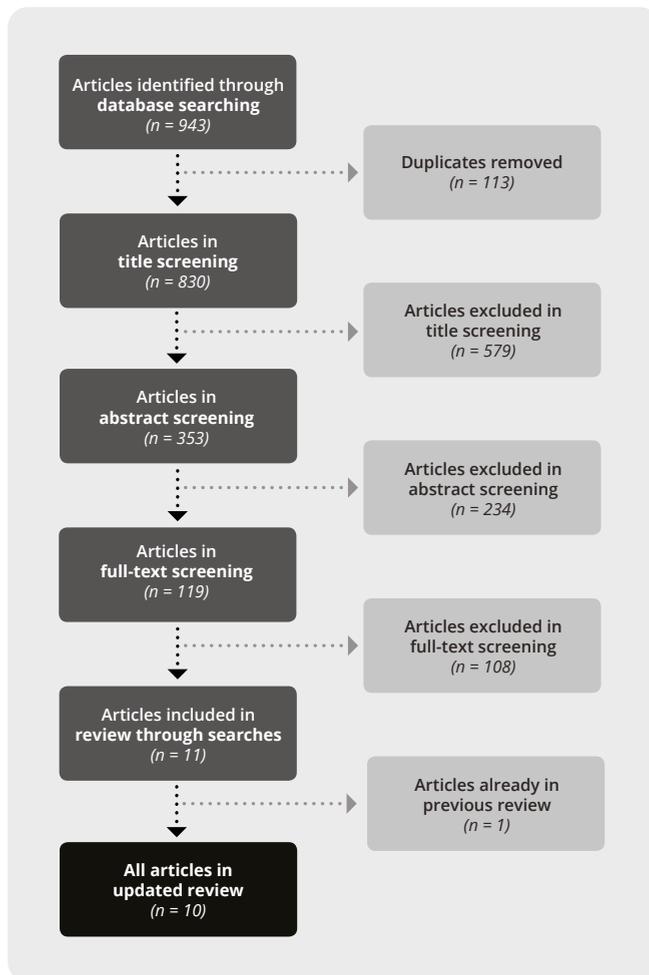


Figure 08.1 – Flow diagram of the updated review of articles published from December 2018 till July 2020.

of the updated review (published from December 2018 to July 2020) and our experimental research, categorised in the main categories described in [Table 08.1](#).

Speech and/or Textual Output

Four articles researched an ECA's emotion in its speech and/or textual output (Azevedo et al., 2018; Kang & Wei, 2018; Lucas et al., 2018; [chapter 5](#)). The articles show positive effects of ECA's with emotion. First, Kang and Wei show a positive effect on the user's (intention towards) behaviour change, usability and user experience (i.e. credibility of the content, perceived interactivity and interaction quality) and the user's perception of the ECA's characteristics (e.g. user's satisfaction with the ECA's support) compared to an ECA without emotion. However, this effect was only found for users that have a need for emotional support, not for users that have a need for instructional support. In [chapter 5](#), we show that presence of a happy textual expression for an ECA with a neutral facial expression results in a higher score for the perception of the rapport items helpfulness and enjoyableness compared to ECA with neutral textual and facial expressions. Furthermore, the articles identified show that effects on the user's (intention towards) behaviour change are larger if an ECA displays positive emotions, compared to negative emotions. For example, memory recall on information about medication (Azevedo et al., 2018) and motivation to perform a physical training (Lucas et al., 2018) improve when an ECA displays positive emotions. In addition, perception of the ECA's characteristics, such as perceived effectiveness, expressiveness and similarity (Azevedo et al., 2018), and perception of the quality of the ECA as fitness trainer (Lucas et al., 2018) are better if the ECA displays positive emotions. Also, the relation with an ECA (Azevedo et al., 2018) and rapport with an ECA improve as a consequence of positive emotions (Lucas et al., 2018). On the contrary, Lucas et al. show some negative effects of an ECA's positive emotions compared to an ECA's negative emotions: users perceive such an ECA as less tough and disagreeable and report lower self-reported values of effort during training. Our original review also shows positive effects of an ECA's emotion on the user's perception of the ECA's characteristics and usability and user experience. However, in contrast to the findings presented in this review, our original review does not show any effect of an ECA's emotion on user's (intention towards) behaviour change.

Only one article researched the medium of an ECA's message. Regarding the medium of an ECA's message, they show positive effects of a verbal ECA message compared to a textual ECA message on the user's (intention towards) behaviour change. For example, the users' intention to learn about colorectal cancer and users' intention to discuss screening options for colorectal cancer with a doctor is larger with verbal messages. This finding is in contrast with the finding of the article researching an ECA's medium as identified in our original review, showing that an ECA's message presented in text is better than a verbal message when it comes to a user's (intention towards) behaviour change (Tielman et al., 2017).

Facial and Gaze Expressions

Three of the articles that researched an ECA's emotion in its speech and/or textual output (Azevedo et al., 2018; Lucas et al., 2018; [chapter 5](#)), researched the ECA's emotion in its facial and gaze expressions in parallel, such that the results described in the previous section also apply to this section.

One article researched an ECA's animation (Zalake et al., 2019). Zalake et al. show positive effects of an ECA with animation compared to a static ECA on the user's (intention towards) behaviour change. Our original review does not identify articles that researched the ECA's animation.

Hand and Body Gestures

The research on animation by *Zalake et al.* described above also applies to an ECA's hand and body gestures, such that the results described in the previous section also apply to this section.

Looks (Species)

Three articles researched the design feature agency (*Gambino et al., 2019; Sin & Munteanu, 2019; Straßmann et al., 2020*). The articles show ambiguous results. *Straßmann et al.* show positive effects of a robot/machine-like ECA over a human-like ECA on the user's perception of the ECA's characteristics (i.e. its likeability and sociability) when filling out a health diary with the ECA. On the other hand, *Gambino et al.* found no effect of the agency on the user's (intention towards) behaviour change, usability and user experience and perception of the ECA's characteristics in the context of an ECA for an annual physical health intake at the doctor. In addition, *Sin and Munteanu* show that users prefer a different species depending on the Information Search Process (ISP) stage they are in when discussing their health with a doctor. For the selection and exploration stage they prefer a voice-only robot, whereas for the collection stage they prefer a video-based human. Findings of the previous review showed mixed results on an ECA's agency as well.

Looks (Realism)

Three articles researched the design feature rendering style (*Azevedo et al., 2018; Mostajeran et al., 2019; Straßmann et al., 2020*). The articles show no effect of the ECA's rendering style (realistic or cartoony) on user's intention to use (*Straßmann et al., 2020*), (intention towards) behaviour change (*Azevedo et al., 2018*), usability and user experience (*Straßmann et al., 2020*), the user's perception of the ECA's characteristics (*Azevedo et al., 2018*), perceived realism, likeability, trustworthiness, competence, attractiveness and sociability (*Straßmann et al., 2020*) and the user's relation with the ECA (*Straßmann et al., 2020*). However, a study by *Mostajeran et al.* shows an effect of the ECA's rendering style. They state that users tend to prefer a realistic ECA over a cartoony ECA in the context of an ECA for balance training. Findings of the previous review also showed ambiguous results on an ECA's rendering style.

Looks (Socio-demographic)

Five articles researched the design feature gender (*Azevedo et al., 2018; Mostajeran et al., 2019; ter Stal et al., 2019; chapter 4, 7*) and show ambiguous results. *Azevedo et al.* found that users prefer female ECAs over male ECAs in the context of an ECA providing information about medication. On the other hand, *Mostajeran et al.* found that users prefer males over females in the context of an ECA for balance training. *ter Stal et al.* show that females tend to prefer ECAs of their own gender (i.e. females). On the contrary, in **chapter 7** we did not find that users are more satisfied with ECAs of their own gender, although this was found in the context of a multi-agent application instead of a single-agent application. In addition, *ter Stal et al.* show that females are rated higher on the characteristics friendliness and involvement, whereas males are rated higher on authority and expertise. Similarly, in **chapter 4** we found that an older male ECA is perceived as more authoritative than a young female ECA. The articles found do not show differences between female and male ECAs on the user's relation with the ECA (e.g. rapport towards the ECA and likeliness of following the ECA's advice) (*ter Stal et al., 2019; chapter 4*). Results on an ECA's gender are similar to that of our original review, which also show ambiguous results.

Also, four articles researched the ECA's age (*Azevedo et al., 2018; ter Stal et al., 2019; chapter 4, 7*). The articles show ambiguous results. *Azevedo et al.* show no effect of an ECA's age on user's (intention

towards) behaviour change (i.e. memory recall on information about medication), regardless of the user's age. However, they also found that compared to younger adults, older adults rate older ECAs higher on perceived effectiveness. *ter Stal et al. (2019)* show that younger people tend to prefer ECAs of their own age (i.e. younger ECAs), and that in general users prefer a young female ECA over an old male ECA. On the contrary, in [chapter 7](#) we do not show that users are more satisfied with ECAs similar in age, although this was researched in the context of a multi-agent application instead of a single-agent application. In addition, *ter Stal et al. (2019)* show that young ECAs are rated higher on the characteristics friendliness, involvement, expertise and reliability, and the relation with the user (i.e. user's likeliness of following the ECA's advice) and that old ECAs are rated higher on the characteristic authority. Also, in [chapter 4](#) we found that an older male ECA is perceived as more authoritative than an young female ECA, although they did not find any effect of the ECA's age on user's perception of ECA's friendliness, involvement, reliability, intelligence and relation with ECA (i.e. rapport towards the ECA and likeliness of following the ECA's advice). Results on an ECA's age are similar to that of our original review, which also show ambiguous results.

Lastly, two articles researched the ECA's role (*Gambino et al., 2019; ter Stal et al., 2019*). Findings by *Gambino et al.* do not show an effect of the ECA's role on usability and user experience and (intention towards) behaviour change. However, an expert ECA has a more positive effect on user's relation with the ECA (i.e. their likeliness of following the ECA's advice) than a peer ECA (*ter Stal et al., 2019*). Furthermore, articles found show mixed results on the user's perception of the ECA's characteristics. *ter Stal et al.* show that an expert ECA is rated higher on expertise, reliability and authority compared to a peer ECA, whereas *Gambino et al.* show that source role (receptionist, doctor or nurse) has no effect on their attitude towards the ECA. Similarly, articles of our original review researching the ECA's role show ambiguous results.

Looks (Styling)

Two articles researched the design feature clothing (*ter Stal et al., 2019; Zalake et al., 2018*). *ter Stal et al.* show that users rate an ECA in professional clothing higher on expertise, reliability and authority and their likeliness of following the ECA's advice compared to an ECA in casual clothing. On the contrary, *Zalake et al.* did not find any effect of an ECA's clothing on user's relation with the ECA (i.e. their trust in the ECA). Our original review also shows positive effects of an ECA's professional clothing on a user's perception of an ECA's characteristics, relation with an ECA and intention to use.

Quality Classification Individual Articles

Table 08.6 shows the findings per outcome measure for the design features researched in all individual articles of the state-of-the-art (from the original, the updated review and our own experimental research). In addition, every article was assigned a quality label. Of the articles, 13 were of low quality, 20 were of fair quality and 13 were of moderate quality.

Evidence Classification per Design Feature

The findings from the individual articles of the state-of-the-art were clustered per design feature for every main design feature category and assigned with an evidence label (**Table 08.7**). Currently, for the majority of the design features there is no evidence, for a limited number of design features evidence is preliminary (emotion, agency, rendering style, clothing, body shape, gender, age, cultural tailoring). Only for the design feature relational, empathetic behaviour moderate evidence was found.

DESIGN STRATEGIES FOR AN ECA'S APPEARANCE IN EHEALTH

Synthesis Findings State-of-the-art into Design Strategies

We selected the findings of the design features of the state-of-the-art (**Table 08.7**) that had preliminary or moderate evidence and grouped them for our design strategies. We discussed our approach with four researchers in the field of ECAs in a digital session. The design strategies are:

- **Strategy 1: Adapt the ECA's Demographics to the User's Preferences**
Based on the design features: gender, age and cultural tailoring
- **Strategy 2: Adapt the Human-likeness of the ECA to its Functionalities**
Based on the design features: species and rendering style
- **Strategy 3: Convey an ECA's Expertise Using its Appearance**
Based on the design features: clothing and body shape
- **Strategy 4: Implement Emotion in the ECA's Expressions**
Based on the design feature: emotion
- **Strategy 5: Implement an ECA's Relational Behaviour**
Based on the design feature: relational behaviour

In the remainder of this section, we describe every strategy in more detail.

Strategy 1: Adapt the ECA's Demographics to the User's Preferences

The state-of-the-art shows that findings on the ECA's demographics often depend on the context. First, with respect to the ECA's gender (preliminary evidence), some research indicates that a female ECA is preferred and more positively evaluated on its characteristics, whereas other research indicates a preference for a male ECA. Some research did not find any difference between a female and male ECA in the user's intention to use and perception of the relation with the ECA. The majority of the research indicates that the preferred gender depends on the context of use and shows positive effects for adapting the ECA's gender to that of the user.

Regarding the ECA's age (preliminary evidence), some research shows that a young ECA is preferred and that a young ECA positively affects the user's intention to use, whereas other research does not show any difference between young and old ECAs with respect to the user's perception of the relation with the ECA. But, again, the majority of the research indicates that the preferred age of the ECA depends upon the context of use and shows positive effects for adapting the ECA's age to that of the user.

Furthermore, with respect to an ECA's cultural background (preliminary evidence), the majority of the research shows positive effects of cultural tailoring (e.g. adapting the skin colour, word choice and directness to that of the user), on the user's perception of the ECA's characteristics, preference and relation with ECA. Only with respect to behaviour change, cultural tailoring is questionable.

Table 08.6 – Articles of the state-of-the-art. Every article is provided with a quality label (1 = low quality, 2 = fair quality, 3 = moderate quality). Findings either show a positive effect (+), negative effect (-), no effect (0) or an effect that depends on the context (~).

Source	Qual. label	Researched design feature	Usage ECA app	Intention to (continue) using ECA app
<i>Alsharbi & Richards (2017)</i>	1	Gender: female (vs. male)		
		Age: young (vs. old)		
		Cultural tailoring (vs. no)		
		Role: friend (vs. professional)		
<i>Amini et al. (2013)</i>	2	Relational, empathic behaviour (vs. no)		+
<i>Amini et al. (2014)</i>	2	Relational, empathic behaviour (vs. no)		+
<i>Azevedo et al. (2018)</i>	2	Emotion: positive (vs. negative)		
		Rendering style: realistic (vs. cartoon)		
		Gender: female (vs. male)		
		Age: young (vs. old)		
<i>Bickmore et al. (2005)</i>	3	Relational, empathic behaviour (vs. no)		+
<i>Bickmore et al. (2009)</i>	3	Providing personal information (vs. no)	+	
<i>Bickmore et al. (2010)</i>	3	Variable behaviour (vs. no)	+	+
<i>Bickmore & Picard (2004)</i>	3	Relational, empathic behaviour (vs. no)		+
<i>Bickmore & Picard (2005)</i>	3	Relational, empathic behaviour (vs. no)		0+
<i>Bickmore & Ring (2010)</i>	1	User control prosody & facial expressions (vs. no)		
<i>Bickmore & Schulman (2007)</i>	1	Relational, empathic behaviour (vs. no)		
<i>Creed & Beale (2012)</i>	3	Emotion (vs. no)		+
<i>Creed et al. (2015)</i>	3	Emotion (vs. no)	0	- 0

(Intention towards) behaviour change	Usability and user experience ECA app	User's perception of ECA's characteristics	User's perception of relation with ECA	User's perception of own characteristics	Preference ECA over other ECAs
					0
					~
					+
					+
	+	+	+		
	+	+	+		
+		+			
0		0			
0					+
0		~			
0		+	+		
	0+	0			
-					
		+	+		
0		+	+		
	0+	0		0	
+	+				
0		0+			
0			0	0	

Source	Qual. label	Researched design feature	Usage ECA app	Intention to (continue) using ECA app
<i>Forlizzi et al. (2007)</i>	1	Rendering style: realistic (vs. cartoon) Gender: female (vs. male)		
<i>Frost et al. (2012)</i>	2	Interactivity (vs. no)		
<i>Gambino et al. (2019)</i>	2	Agency: Human (vs. robot/machine) Role: receptionist (vs. professional)		
<i>Grillon & Thalmann (2008)</i>	1	Variable behaviour (vs. no)		
<i>Kang & Gratch (2011)</i>	2	Providing personal information (vs. no)		
<i>Kang & Wei (2018)</i>	2	Emotion (vs. no)		
<i>Lisetti et al. (2013)</i>	2	User control prosody & facial expressions (vs. no)		+
<i>Lucas et al. (2018)</i>	2	Emotion: positive (vs. negative)		
<i>Malhotra et al. (2016)</i>	1	Gender: female (vs. male)		
<i>Mostajeran et al. (2019)</i>	1	Rendering style: realistic (vs. cartoon)		
<i>Nguyen & Masthoff (2007)</i>	2	Clothing: professional (vs. other)		
<i>Nguyen & Masthoff (2009)</i>	3	Relational, empathic behaviour (vs. no)		
<i>Olafsson et al. (2017)</i>	2	Rap (vs. no)	0	0
<i>Parmar et al. (2018)</i>	2	Clothing: professional (vs. other)		+
<i>Robertson et al. (2015)</i>	1	Rendering style: realistic (vs. cartoon)		
<i>Ring et al. (2014)</i>	2	Rendering style: realistic (vs. cartoon)		
<i>Schmeil & Suggs (2014)</i>	3	Body shape: slim (vs. fat)		0
<i>Silverman et al. (2001)</i>	1	Emotion (vs. no)		
<i>Sin & Munteanu (2019)</i>	1	Agency: Human (vs. robot/machine)		

(Intention towards) behaviour change	Usability and user experience ECA app	User's perception of ECA's characteristics	User's perception of relation with ECA	User's perception of own characteristics	Preference ECA over other ECAs
		+			
		~			
				0 +	
0	0	0			
0	0	0			
		~			
+			+		
~	~	~			
	+	+	+		
- + ~		- ~	+		
		+			
		+			+
		0			~
0 +	0	+			
	0	-	0 +		
+		+	+		+
		+			
		- +			
+					
+	+	+			
					- +

Source	Qual. label	Researched design feature	Usage ECA app	Intention to (continue) using ECA app
<i>Skalski et al. (2007)</i>	1	Body shape: slim (vs. fat)		
		Clothing: professional (vs. other)		
<i>ter Stal et al. (2019)</i>	2	Gender: female (vs. male)		
		Age: young (vs. old)		
		Role: friend (vs. professional)		
<i>chapter 4, ter Stal et al. (2020)</i>	1	Gender: female (vs. male)		
		Age: young (vs. old)		
<i>chapter 7</i>	2	Gender: female (vs. male)		
		Age: young (vs. old)		
<i>chapter 5</i>	2	Emotion (vs. no)		
<i>Straßmann et al. (2020)</i>	2	Agency: Human (vs. robot/machine)		0
		Rendering style: realistic (vs. cartoon)		0
<i>Tielman et al. (2017)</i>	3	Medium message: in text (vs. verbally)		
<i>van Vugt et al. (2006)</i>	2	Body shape: slim (vs. fat)		0
<i>van Wissen et al. (2016)</i>	2	Rendering style: realistic (vs. cartoon)		+
		Age: young (vs. old)		0
<i>Yin et al. (2010)</i>	3	Linguistic tailoring (vs. no)		
		Cultural tailoring (vs. no)		
<i>Zalake et al. (2018)</i>	1	Clothing: professional (vs. other)		
<i>Zalake et al. (2019)</i>	2	Medium message: in text (vs. verbally)		
		Animation (vs. no)		
<i>Zhou et al. (2014)</i>	3	Cultural tailoring (vs. no)		0
<i>Zhou et al. (2017)</i>	3	Cultural tailoring (vs. no)		

(Intention towards) behaviour change	Usability and user experience ECA app	User's perception of ECA's characteristics	User's perception of relation with ECA	User's perception of own characteristics	Preference ECA over other ECAs
+					
		+	+		0
		- +	0		~
		- +	+		~
		-	-		0
		- 0	0		
		- 0	0		
					0
					0 ~
			0 +		
	0	- 0	0		
	0	0	0		
0 +		0			
		-	0		
	0	+			
	0	0			
0					
-					
				0	
-					
+					
		0	-		
- 0	0	0 +			

Table 08.7 – Synthesis of the findings of the state-of-the-art. Findings are grouped per design feature and provided with an evidence level (★ = no evidence, ★★ = preliminary evidence, ★★★ = moderate evidence).

	Design feature	Evid. level	Sources
Speech and/or Textual Output	Emotion (vs. no)	★★	Creed & Beale (2012), Creed et al. (2015), Kang & Wei (2018), Silverman et al. (2001), ter Stal et al. (submitted)
	Emotion: Positive (vs. negative)	★	Azevedo et al. (2018), Lucas et al. (2018)
	Relational, empathic behaviour (vs. no)	★★★	Amini et al. (2013), Amini et al. (2014), Bickmore et al. (2005), Bickmore & Picard (2004), Bickmore & Picard (2005), Bickmore & Schulman (2007), Nguyen & Masthoff (2009)
	Providing personal information (vs. no)	★	Bickmore et al. (2009), Kang & Gratch (2010)
	Variable behaviour (vs. no)	★	Bickmore et al. (2010)
	User control prosody & facial expressions (vs. no)	★	Bickmore & Ring (2010)
	Interactivity (vs. no)	★	Frost et al. (2012)
	Rap (vs. no)	★	Olafsson et al. (2017)
	Medium message: text (vs. verbally)	★	Tielman et al. (2017), Zalake et al. (2019)
	Linguistic tailoring (vs. no)	★	Yin et al. (2010)
Facial and Gaze Expressions	Emotion (vs. no)	★★	Creed & Beale (2012), Creed et al. (2015), Silverman et al. (2001), ter Stal et al. (submitted)
	Emotion: Positive (vs. negative)	★	Azevedo et al. (2018), Lucas et al. (2018)
	Relational, empathic behaviour (vs. no)	★★★	Amini et al. (2013), Amini et al. (2014), Bickmore et al. (2005), Bickmore & Picard (2004), Bickmore & Picard (2005), Bickmore & Schulman (2007), Lisetti et al. (2013)
	Variable behaviour (vs. no)	★	Grillon & Thalmann (2008)
	User control prosody & facial expressions (vs. no)	★	Bickmore & Ring (2010)
	Interactivity (vs. no)	★	Frost et al. (2012)
	Animation (vs. no)	★	Zalake et al. (2019)
	Position to camera: far (vs. near)	★	Zalake et al. (2019)

	Design feature	Evid. level	Sources
<i>Hand & Body Gest.</i>	Relational, empathic behaviour (vs. no)	★★★	Amini et al. (2013), Amini et al. (2014), Bickmore et al. (2005), Bickmore & Picard (2004), Bickmore & Picard (2005)
	Variable behaviour (vs. no)	★	Grillon & Thalmann (2008)
	Interactivity (vs. no)	★	Frost et al. (2012)
	Animation (vs. no)	★	Zalake et al. (2019)
<i>Looks</i>	Agency: Human (vs. robot/machine)	★★	Gambino et al. (2019), Sin & Munteanu (2019), Straßmann et al. (2020)
	Rendering style: realistic (vs. cartoon)	★★	Azevedo et al. (2018), Forlizzi et al. (2007), Mostajeran et al. (2019), Straßmann et al. (2020), Ring et al. (2014), Robertson et al. (2015), Wissen et al. (2016)
	Clothing: professional (vs. other)	★★	Nguyen & Masthoff (2007), Parmar et al. (2018), ter Stal et al. (2020), Zalake et al. (2018)
	Body shape: slim (vs. fat)	★★	Schmeil & Suggs (2014), Skalski et al. (2007), van Vugt et al. (2006)
	Gender: female (vs. male)	★★	Alsharbi & Richards (2017), Azevedo et al. (2018), Forlizzi et al. (2007), Malhotra et al. (2016), Mostajeran et al. (2019), ter Stal et al. (2020), ter Stal et al. (2020), ter Stal et al. (submitted)
	Age: young (vs. old)	★★	Alsharbi & Richards (2017), Azevedo et al. (2018), ter Stal et al. (2020), ter Stal et al. (2020), ter Stal et al. (submitted), Wissen et al. (2016)
	Cultural tailoring (vs. no)	★★	Alsharbi & Richards (2017), Yin et al. (2010), Zhou et al. (2014), Zhou et al. (2017)
	Role: friend (vs. professional)	★	Alsharbi & Richards (2017), ter Stal et al. (2020)
	Role: receptionist (vs. professional)	★	Gambino et al. (2019)

Concluding, these findings suggest that we should not develop a one-size-fits-all ECA appearance. When developing an ECA for eHealth we have to take into account our target users. As with many other aspects of eHealth, personalization matters (*van Velsen et al., 2019; Tielman et al., 2019; Lehto & Oinas-Kukkonen, 2011; Kaptein et al., 2012*). We suggest to adapt the demographics of an ECA to the user's preferences. In order to do so, one should research the preferences of one's target group for the particular context of use. Identify the user's preferences on an ECA's demographics and design a few ECAs with different demographics to fit the user's preferences. For example, when research shows that an ECA's gender is an important factor for a user's impressions of this ECA, provide the user with the option to work with a female, male and gender neutral ECA.

Strategy 2: Adapt the Human-likeness of the ECA to its Functionalities

The state-of-the-art shows that findings on the ECA's agency (preliminary evidence) often depend on the context. On the one hand, research does not show differences between a human-like and robot/machine-like ECA in the effect on a user's (intention towards) behaviour change, usability and user experience and a user's perception of an ECA's characteristics. On the other hand, research shows that users rate a robot ECA more positively on its characteristics. Lastly, research indicates that whether a human-like or robot/machine-like ECA is preferred depends upon the task of the ECA.

In addition, with respect to an ECA's rendering style (preliminary evidence), research found that a realistic ECA is preferred over a cartoon ECA. Furthermore, based on the studied literature it cannot be concluded that a user's perception of the characteristics of a realistic ECA is more positive. However, the majority of the research did not find any difference between a realistic and a cartoon ECA (on user's intention to use, a user's intention towards behaviour change, usability and user experience and a user's perception of an ECA's characteristics and relation with the ECA).

These ambiguous results on an ECA's agency and rendering style might be explained by users having a negative adaptation gap: a mismatch between the users' expectations of the ECA and the ECA's actual capabilities (*Komatsu et al., 2012*). To allow these expectations to match, we suggest to adapt the human-likeness of the ECA to its functionalities, such that the ECA's intelligence expressed by the ECA's appearance matches its actual intelligence. In order to do so, future research is needed. First, one should think of the context of use: when will users interact with the ECA? What is the content of the dialogues? Then, one should study the users' impressions of varying ECA agencies and rendering styles in relation to this context. This way, one may avoid that the human-likeness of the ECA negatively affects users' perceptions of determinants of engagement.

Strategy 3: Convey an ECA's Expertise Using its Appearance

The state-of-the-art shows that findings on the ECA's clothing often depend on the context (preliminary evidence). On the one hand, professional clothing shows positive effects on a user's intention to use, perception of an ECA's characteristics and relation with the ECA. On the other hand, an ECA in casual clothing was preferred and rated more positively on its characteristics and relation with the user. Other research did not find any difference between clothing styles (on user's relation with the ECA).

In addition, the state-of-the-art shows that findings for an ECA's body shape often depend on the context (preliminary evidence). Some research shows that a slim body shape positively affects a user's (intention towards) behaviour change, whereas one article showed quite the opposite: an ECA with a heavier body shape positively affected user's intention to use, intention towards behaviour change and perception of the ECA's characteristics and relation with ECA.

Therefore, we suggest to adapt the ECA's appearance to its task, such that the expertise expressed in the ECA's appearance matches the ECA's expertise in its task. Future research could investigate how users respond to different ECA clothing and body shapes, given the context in which the ECA is used. This way, users' expectations raised by the ECA's appearance could be matched to the actual expertise of the ECA, positively affecting users' impressions of the ECA and their engagement with the eHealth application.

Strategy 4: Implement Emotion in the ECA's Expressions

The state-of-the-art (**Table 08.6**) shows that an ECA's emotion (preliminary evidence) mainly positively affects user's intention to use the ECA, his or her (intention towards) behaviour change, the usability and user experience and user's perception of the ECA's characteristics and relation with the ECA compared to not having emotions.

Therefore, we suggest to implement emotion in an ECA's verbal (text or speech) and non-verbal (facial and gaze expressions and hand and body gestures) behaviour. As an example, an ECA could show a happy expression when the user reached his or her daily step goal, and show a sad expression when the user did not. This way, positive impressions of the ECA can be created, leading to better companionship between the user and the ECA, which could positively affect user's engagement with the underlying eHealth application.

Strategy 5: Implement an ECA's Relational Behaviour

The state-of-the-art shows that implementing an ECA's relational – empathic – behaviour (moderate evidence) mainly positively affects user's intention to use the ECA, his or her (intention towards) behaviour change, the usability and user experience and user's perception of the ECA's characteristics and relation with the ECA. Therefore, we suggest to implement relational behaviour – empathic behaviour – into an ECA's verbal (text or speech) and non-verbal (facial and gaze expressions and hand and body gestures).

For example, when the user reached his or her daily step goal, the ECA could celebrate this achievement with the user by saying "Good job", showing a happy facial expression and raising its thumb. When the user did not manage to reach his or her daily step goal, the ECA could say "I understand that it is not always easy to reach your goal on a workday", and showing understanding in its facial expressions and body language. This way, companionship between the user and ECA is supported, which could positively affect user's engagement with the underlying eHealth application.



Figure 08.2 – Interface of the mobile eHealth application for older workers.

IMPLEMENTATION OF THE DESIGN STRATEGIES: A USE CASE

In this section, we illustrate how the design strategies can be used to create an ECA for eHealth. In this use case, we designed an ECA as part of a mobile eHealth application for older workers.

The Mobile eHealth Application

The mobile eHealth application is a smartphone application running on Android. The main functionality of the application is to help older workers (55+) manage their work-related complaints by providing insight into their physical activity and supporting them in being physically active during leisure time. The interface of the application can be seen in **Figure 08.2**. The application contains the following functionalities:

- **Dashboard (Figure 08.2A)**: a summary of today's physical activity measured by an activity tracker at work and at home (e.g. steps, fatigue level) and exercises performed, including progress towards personal goals. On top, notifications of the ECA are shown. Clicking on the "go" button opens the dialogue interface (**Figure 08.3**).

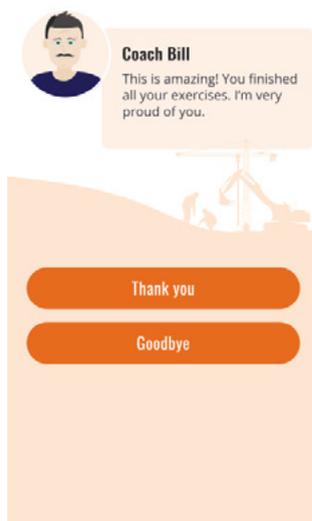


Figure 08.3 – The dialogue interface of the application showing an example dialogue step consisting of an ECA message and two user reply options.



Figure 08.4 – Interface of the ECA selection screen.

- **Visualisation of physical activity (Figure 08.2BC):** detailed information on the user's physical activity per day, week or month, visualised in graphs, including accomplishment of personal goals.
- **6-week activity program (Figure 08.2D):** a 6-week activity program to prevent work-related muscular disorders. Per week, the program offers two body-weight workouts (consisting of five exercises, taking roughly 15 minutes) and one endurance workout (walking/running or biking/ sports cycling activity) a week. The user selects one of the three available programs, targeting either 1) low back pain, 2) fatigue or 3) neck and shoulder pain through stress. Furthermore, the user selects the difficulty level of the program: basic, intermediate or advanced.
- **Exercise explanation with video support (Figure 08.2E):** the user receives an explanation for every exercise in the workouts. The user can watch a video that shows how to perform the exercise. When having performed an exercise, users select the "exercise done" button communicating to the application that they have performed the exercise and then receive the next exercises. User can quit exercising at every point in time. When finishing the exercises, or quitting the training, users will return to the dashboard (Figure 08.2A).
- **Gamification (Figure 08.2F):** the user can earn badges by fulfilling physical activity challenges (e.g. meet your daily step goal 3 days in a row). Every badge is worth a number of experience points, these points are used to determine the user's position on a leader-board.
- **Dialogues with ECA (Figure 08.3):** dialogues between the ECA and the user for motivation and support.

The Embodied Conversational Agent

The aim of the ECA is to engage the user in using the application by providing support and motivating messages. In the dialogues with the user, the ECA:

- Informs about the 6-week activity program
- Provides tips on how to be motivated to perform the exercises
- Provides feedback and compliments on progress with respect to the exercises
- Reminds to do the workouts
- Asks about the user's experience of the workouts
- Informs on benefits for healthy living

The dialogue interface with an example dialogue step between the ECA and the user can be seen in **Figure 08.3**.

Design Strategies Implemented in Use Case

An overview of the implementation of the design strategies for the use case can be seen in **Table 08.2**. In the rest of this section we will explain the approach of each strategy in more detail.

Table 08.2 – Overview of the approach to the design strategies in the application of the use case.

Design Strategy	Approach in Use Case
<i>1. Adapt the ECA's Demographics to the User's Preferences</i>	User can select an ECA from a set of ECA's differing in age and gender
<i>2. Adapt the Human-likeness of the ECA to its Functionalities</i>	ECA's realism is human-like, but not photo-realistic, such that the ECA's appearance matches with the ECA's functionalities and intelligence
<i>3. Convey an ECA's Expertise Using its Appearance</i>	ECA in casual clothing, fitting the role of a coach (not health care professional)
<i>4. Implement Emotion in the ECA's Expressions</i>	Emotion (neutralness, happiness and sadness) implemented in the ECA's facial and textual expressions
<i>5. Implement an ECA's Relational Behaviour</i>	Sympathy and empathy implemented in the ECA's textual expressions

Strategy 1: Adapt the ECA's Demographics to the User's Preferences

To be able to adapt the demographics of the ECA to a set of diverse users, we created six ECAs differing in age and gender (see **Figure 08.4**). Also, the ECAs had their own living situation, interests and physical activity level, reflected in their background stories. At first login, users could select one of these six ECAs (**Figure 08.4A**). This way, the decision to interact with an ECA similar in age and gender was left to the user. Before selecting an ECA, users could read a small description about the ECA in a profile interface available via the info icon below the ECA image (**Figure 08.4B**).

Throughout the rest of this document, we show the implementation of the design strategies for one of the ECAs, namely for the middle-aged male ECA. The strategies were implemented for the other ECAs in a similar way.

Strategy 2: Adapt the Human-likeness of the ECA to its Functionalities

Since the ECA in our application has human-like characteristics, such as showing emotion, sympathy and empathy, we designed a human-like ECA. The messages of the ECA contain some penalisation, based on answers provided by the user in previous dialogues. However, the user cannot interact with the ECA on every topic (different from an interaction with a real human), since the dialogues are pre-scripted. Therefore, the ECA's intelligence is limited. To avoid a mismatch between the user's expectations of the ECA's functionalities and intelligence, we decided not to create a photo-realistic human-like ECA.

Strategy 3: Convey an ECA's Expertise Using its Appearance

The role of the ECA in the application is that of a coach. The coach supports and motivates the user, but is not a medical health care professional. Therefore, we decided to dress the ECA in casual clothing instead of professional clothing (i.e. a doctor's coat). This way, we attempt to match the user's expectations of the expertise of the ECA to the actual expertise of the ECA. Furthermore, the body shape of the ECA was set to normal (not too fat, nor too slim).

Strategy 4: Implement Emotion in the ECA's Expressions

We created three basic emotions: 1) neutral, 2) happy and 3) sad. The set of facial expressions could be expanded by implementing the other basic emotions fear, anger, surprise and disgust. The emotions were connected to the following triggers:

- › **Neutral:** this is the ECA's basic emotion.
- › **Happy:** this emotion is triggered when the ECA compliments the user (e.g. when the user finished a workout).
- › **Sad:** this emotion is triggered when the ECA is disappointed by the user (e.g. when the user did not do a workout).

We implemented the ECA's emotion in both the ECA's facial expressions as well as the ECA's textual expressions. The emotions in the ECA's facial expressions were designed based on the facial action coding system of *Ekman and Friesen*. **Table 08.3** shows how this system was used to create the ECA's facial expressions.

In addition, we implemented emotion into the ECA's textual expressions. For every ECA design we defined a set of words for the neutral, happy and sad emotion. The words differed across the ECAs, such that the words also fit the characteristics of the ECA. A few examples of words expressing emotion for the middle-aged female ECA can be seen in **Table 08.4**. The ECA's emotion in its textual expressions was aligned with the ECA's emotion in its facial expressions.

Table 08.3 – Implementation of emotions into the ECA's facial expressions.

	Neutral	Happy	Sad
Visuals			
Action units	n.a.	06 – Cheek raiser 12 – Lip corner puller	01 – Inner brow raiser 04 – Brow lowerer 15 – Lip corner depressor
Implementation	<ul style="list-style-type: none"> › Lip corners: straight (mouth closed) › Eyes: wide › Eye brows: curved 	<ul style="list-style-type: none"> › Lip corners: pulled (mouth open) › Eyes: a bit narrow › Eye brows: curved 	<ul style="list-style-type: none"> › Lip corners: down (mouth closed) › Eyes: wide › Eye brows: lower and inner side up

Table 08.4 – Implementation of emotions into the ECA’s textual expressions.

	Neutral	Happy	Sad
<i>Textual expressions</i>	Okay	Nice	Hmm
	Allright	Great	Unfortunately
	Agreed	Excellent	That’s too bad
	Good	Well done	Sadly

Strategy 5: Implement an ECA’s Relational Behaviour

The ECA showed sympathy and empathy in the ECA’s messages towards the user in the following way:

- Personal addressing (e.g. with or without salutation, with or without last name, informal or formal addressing)
- Personal greetings (e.g. “Good day mister Janssen”, “Nice to see you again, Anna”)
- Personal goodbyes (e.g. “Have a nice day mister Janssen”, “See you next time, Anna”)
- Understanding (e.g. “I understand”, “I see”)
- Rhetorical questions (e.g. “Isn’t it?”, “Don’t you think?”)
- Shared successes (e.g. “Great to hear!”, “What a marvellous achievement”).
- Shared disappointments (e.g. “That is unfortunate”, “I am sorry to hear”)
- Encouragement on actions (e.g. “I know you can do it”, “Keep on trying”)

The messages differed slightly across the ECAs in order to align them with the characteristics of each ECA. In addition, the ECA’s relational behaviour was partly reflected by aligning its emotions to the emotional state of the user.

DISCUSSION

In this chapter, we developed a set of five design strategies for an ECA’s appearance in eHealth grounded by the state-of-the art, to contribute to determinants of user engagement. The design strategies focus on an ECA’s 1) demographics, 2) human-likeness, 3) conveyed expertise, 4) emotion and 5) relational behaviour.

First, our state-of-the-art analysis showed that the research area of an ECA’s appearance for eHealth is still in its infancy, but that the area is growing: our original review identified 33 articles (from 2001 till December 2018), and our updated review identified ten articles published from December 2018 till July 2020. The state-of-the-art showed a limited body of evidence for how an ECA’s appearance affects user engagement. Not only were the identified studies of low quality, they also focused on a wide range of different ECA design features. For the majority of design features identified in the state-of-the-art, no evidence for effects on user engagement was found. Only for an ECA’s relational,

empathetic behaviour research showed moderate evidence. For an ECA's emotion, agency, rendering style, clothing, body shape, gender, age and cultural tailoring preliminary evidence was found. To the best of our knowledge, there is no state-of-the-art review in other contexts or on other aspects than user engagement. However, we do see that these design features are indeed regularly researched in general or, for example, in the field of pedagogical ECAs (relational behaviour (*Lee et al., 2007; von der Pütten et al., 2009*), emotion (*Kim et al., 2007; Pelachaud, 2009*), agency (*Straßmann & Krämer, 2017; Zimmerman et al., 2005*), rendering style (*Baylor & Kim, 2004; McDonnell et al., 2012; Rosenberg-Kima et al., 2008; Zell et al., 2015; Zibrek et al., 2018*), clothing (*Veletsianos, 2010*), body shape (*Khan & De Angeli, 2009; Khan & Sutcliffe, 2014*), gender (*Cowell & Stanney, 2003; Kim et al., 2007*), age (*Cowell & Stanney, 2003; Lee et al., 2018*), cultural tailoring (*Cowell & Stanney, 2003; Canidate & Hart, 2017*)). Effects of these design features are shown on similar outcome variables as identified in our review, possibly strengthening their relevance for user engagement. One could learn from studies on an ECA's appearance in these other contexts and repeat the studies in the eHealth context. Studies in stage I of the renewed framework for evaluation of telemedicine should be performed with a sufficient number of participants, such that results can be generalised. Eventually, we should work towards studies in stage II, III and IV, such that effects of an ECA's appearance on user engagement can be evaluated in the actual setting of use: daily life.

Second, we created a first set of design strategies for an ECA's appearance in eHealth to guide designers, developers and researchers of ECAs in their future work. We not only identified design features that are investigated in the state-of-the-art research, but also made a first attempt to show how these design features can be designed. Therefore, our strategies go beyond frameworks that provide a solid basis for ECA design by identifying and categorizing design features, amongst which design features for an ECA's appearance (*Ruttkey et al., 2004; Straßmann & Krämer, 2017*). Our strategies are at a similar level of the work of *Veletsianos et al.*, who designed a set of 15 research-based guidelines for designing pedagogical agents for user engagement with learning tasks. *Veletsianos et al.* focus on more aspects than solely the ECA's appearance, and thus provide a more extensive set of strategies, but content-wise, the strategies are similar (example strategy by *Veletsianos et al.*: 'Agents should establish credibility and trustworthiness'). Like our strategies, the strategies by *Veletsianos et al.* provide guiding, but remain open to one's own interpretation: the motivations for the strategies are clearly explain, but guides for practical implementation are little concrete. Different from *Veletsianos et al.*, we present our strategies in combination with a single use case in which we show how one could translate the theoretical design strategies into concrete solutions. However, the limited body of evidence of the state-of-the-art makes that our set of design strategies are not set in stone; they should be re-evaluated by future work. Currently, the number of design strategies is limited and the strategies are very theoretical and little concrete. As explained, more (high quality) research is needed to gain better understanding of what and how particular ECA design features affect user engagement and how an ECA's appearance can best be designed accordingly. Therefore, we encourage other researchers to validate, update and expand the strategies based on future research. In addition, our design strategies only reflect if particular design features affect user engagement, not the extent to which they do so. Yet, adding this to the strategies is complex. First, no clear definition of the term 'user engagement' exists yet when it comes to eHealth, let alone a standardised method to measure the degree of engagement (*Yardley et al., 2016; O'Brien & Toms, 2008*). Second, there is no unified set of outcome variables used as determinants for an ECA's effect on user engagement. Third, studies that do investigate effects of ECAs show a large heterogeneity in study design (*Provoost et al., 2017; Kramer et al., 2020; Scholten et al., 2017*). Thus, to work towards design strategies with 'engagement effect labels', the research area of ECAs in eHealth needs to grow, creating more homogeneity in definitions and methods used.

CONCLUSION

The state-of-the-art shows limited body of evidence for how to design an ECA's appearance in eHealth to promote user engagement. Our design strategies for an ECA's appearance based on this state-of-the-art suggest to take into account an ECA's demographics, human-likeness, conveyed expertise, emotion and relational behaviour. This initial set of design strategies should be elaborated upon by performing future research of high quality, to eventually work towards a more extensive and concrete set of design strategies.

Acknowledgements – This work has been partially supported by the BIONIC project, Horizon 2020 Framework Programme of the European Union for Research and Innovation, SC1-DTH-03-2018, grant agreement 826304.

09

General Discussion

OVERVIEW

The objective of this thesis was to research how to design the appearance of an ECA in eHealth to promote user engagement. First, we reviewed the state-of-the-art of design features for ECAs in eHealth (**chapter 2**). Second, we used this state-of-the-art to scope and perform our experimental research (**chapter 3 – chapter 7**). Lastly, we updated our state-of-the-art review (including the work presented in this thesis) and designed a set of design strategies for an ECA's appearance in eHealth accordingly (**chapter 8**). We conclude that 1) adapting an ECA's demographics to users' preferences, 2) adapting its human-likeness to its functionalities, 3) conveying its expertise and 4) implementing emotion and relational behaviour in its appearance, may have positive effects with respect to user engagement.

In this chapter, we place our results in a broader perspective. First, we discuss what we have learned on the design of an ECA's appearance for eHealth. We focus on an ECA's appearance for positive impressions 1) at first glance, 2) after short interaction, and 3) after long-term interaction (our sub objectives). Second, we describe our vision on the contribution of an ECA's appearance to user engagement. Third, we outline our perspective on future research and development of an ECA's appearance for eHealth and how this future work could lead to design strategies for ECAs. Lastly, we envision how engaging ECAs will shape our future health care.

HOW TO DESIGN AN ECA'S APPEARANCE IN EHEALTH?

In **chapter 2**, we showed that the ECA's speech and/or textual output and its facial and gaze expressions are the most researched design feature categories, mostly studied at first glance and after short-interaction. Within these categories, an ECA's emotion and relational behaviour seem to positively affect the perception of the ECA's characteristics. Relational behaviour also seems to positively affect the relation with the ECA, usability of the ECA application and intention to use the ECA application (**chapter 2**). However, our review additionally shows that these design features do not necessarily lead to users changing their health-related behaviour. *Loveys et al. (2020)* show similar findings in a recent literature review on design features for ECAs in general. Their review shows solid evidence for effects of an ECA's verbal behaviours, as well as non-verbal behaviours (facial expressions, gaze) on the quality of the relationship with its users, users' social perceptions of the ECA and their behaviours towards the ECA. But, a lack of research on the ability of ECAs to change behaviour was also reported by them (*Loveys et al., 2020*). Studies on ECAs in other contexts, such as e-learning, also show positive effects of an ECA's relational behaviour and emotion. For example, users perceive an empathic ECA as a better learning facilitator than a non-empathic ECA (*Kim et al., 2007*), show more rapport towards an emotional ECA than an unemotional ECA (*Acosta & Ward, 2011*) and perceive a caring ECA as more trustworthy than an uncaring ECA (*J. E. R. Lee et al., 2007*). These studies also show positive effects on users' (intention towards) behaviour change, such as an ECA's emotion increasing a user's interest in learning (*Kim et al., 2007*) and cognitive performance (*Berry et al., 2005*) and an ECA's relational behaviour increasing a user's interest in learning (*Kim et al., 2007*) and a user's learning itself (*J. E. R. Lee et al., 2007*).

Thus, implementing an ECA's relational behaviour and emotion may have positive effects with respect to user engagement, but little is known on their contribution to actual behaviour change. Therefore, future research should 1) investigate effects of an ECA's emotion and relational behaviour on users' behaviour towards healthy living, by providing them with ECAs either with or without emotion or relational behaviour in a long-term setting and 2) to investigate effects of an ECA's looks on any possible determiner of user engagement.

In this thesis, we mainly investigated the second: effects of an ECA's looks. In the remainder of this section, we describe what we have learned about the design of an ECA for positive impressions at first glance, after short interaction and after long-term interaction.

ECA Appearance at First Glance

From **chapters 2** and **8**, as well as from a recent review by *Loveys et al. (2020)*, we learned that few studies research effects of an ECA's age and gender, in the context of eHealth or another. Therefore, in **chapter 3**, we studied the effect of an ECA's age (young or old), gender (male or female) and role (peer or expert) **at first glance**, to investigate sub objective 1. Our results showed that users rate ECAs differently at first glance depending on the ECA's age, gender and role, and that users have preferences for particular ECAs at first glance. The general population preferred the young female peer ECA. But, both the general and elderly population preferred an ECA that was similar in a) age and b) gender. Similar to our results people prefer ECAs that are similar to themselves at first glance (*van Wissen et al., 2016*), such as in age and gender (*Alsharbi & Richards, 2017*). On the other hand, other studies show that, in general, a female ECA is preferred over a male ECA at first glance (*Cowell & Stanney, 2003; Justo et al., 2020*). It should be noted that this preference could also depend upon the context: people tend to prefer a particular gender for a particular task conform gender stereotypes (*Forlizzi et al., 2007; Zimmerman et al., 2005*).

Concluding, adapting an ECA's age and gender to the user may contribute to positive impressions of the ECA at first glance, as people tend to prefer an ECA that is similar to themselves. But, the context and the task of the ECA might affect users' preferences for ECA age and gender. Therefore, it seems to be important that for future research and development of an ECA, one should study the preferences for an ECA's age and gender of the target group in the particular context of use.

ECA Appearance for Short Interaction

Next, we investigated whether we would find similar findings **after short interaction** with an ECA (sub objective 2). In **chapter 4**, we translated three validated health assessment questionnaires into dialogues between a user and an ECA. Users interacted with a young female peer ECA in half of the dialogues and with an older male peer ECA in the other half of the dialogues. Whereas users showed a preference for particular ECAs at first glance (**chapter 3**), **chapter 4** showed that users did not prefer one of the two ECAs after short interaction. Users perceived the older male ECA as more authoritative than the young female ECA. They did not perceive the ECAs differently in other characteristics, rapport or likeliness to follow advice. However, other studies show that users disclose more information to an ECA similar in age and perceive such an ECA as more trustworthy (*Y. H. Lee et al., 2018*). In addition, users prefer ECAs of the same gender (*Guadagno et al., 2007*), and are more influenced by ECA of similar gender (*Baylor, 2009*).

Besides the ECA's looks, we investigated the ECA's emotion after short interaction. In [chapter 5](#), we studied effects of an ECA's emotions expressed in text and facial expressions on users' perception of rapport. We combined a happy or neutral text with a happy or neutral facial expression, leading to four conditions. The ECAs were positioned as a health coach on physical activity and healthy nutrition. We did not find a significant difference in overall rapport between the conditions. But, an ECA with a happy text and a neutral facial expression was perceived as more helpful and enjoyable (individual rapport items). As an example, research by *Creed et al. (2015)* also found that an ECA in the role of a nutritional coach having emotion (in its voice and facial expressions) is perceived as more likeable and caring than an unemotional ECA. They found this effect not after short interaction, but after long-term interaction: users interacted with the ECA for 49 days. Other research also shows that an ECA's emotion might have positive effects on how users perceive an ECA. In 2009, *Beale and Creed* presented a structured overview of research into emotional simulation in ECAs. They describe that results of studies on ECA emotion are often inconclusive and contradictory, but that several studies have highlighted the potential of ECA emotion to enhance interaction. Furthermore, they highlight that there is no strong evidence to suggest that emotional ECAs hinder an interaction. They conclude that synthetic emotions expressed by ECAs have the potential to influence user attitudes and behaviour. More recent work by *Loveys et al. (2020)* suggests that the expression of positive emotion improves social perceptions and behaviours towards ECAs.

Thus, 1) an ECA's age and gender may have less contribution to positive impressions of the ECA after short interaction than at first glance and 2) emotion in an ECA's textual expressions may contribute to positive impressions of the ECA after short-term interaction. Therefore, future research should investigate whether these hypotheses can be confirmed and whether they also apply after interactions with ECAs in the long-term.

ECA Appearance for Long-term Interaction

[Chapter 6](#) and [chapter 7](#) both describe a study on long-term interaction with an ECA in daily life, researching an ECA's appearance in the actual context of use (sub objective 3). In [chapter 7](#), we continued our work on an ECA's age and gender. We studied an ECA's appearance in a multi-agent eHealth application that provided users with holistic coaching on healthy ageing. [Chapter 7](#) showed that, in a multi-agent application, older adults were not more satisfied with ECAs that are similar to themselves in age and gender, neither at first glance nor after long-term interaction. With respect to age, in fact, it seems that older adults are more satisfied with younger and middle aged ECAs than older ECAs. This is different from findings in [chapter 3](#). Although not a long-term daily life study, a study by *Justo et al. (2020)* shows similar results; a majority of older adults prefers an ECA for healthy living to look between 29 and 48 years old. We expect that this preference for younger looking ECAs can be explained by people preferring ECA designs that fit the ECA's domain conform stereotypes (*Forlizzi et al., 2007; Zimmerman et al., 2005; Nguyen & Masthoff, 2007*). When interacting with older ECAs in dialogues on healthy living, these older ECAs might be negatively associated with being healthy, conforming stereotypes, resulting in a lower user satisfaction.

In [chapter 6](#), we developed an ECA within an eHealth self-management intervention for patients with both Chronic Obstructive Pulmonary Disease (COPD) and Chronic Heart Failure (CHF), offered for four months in daily life. Taking into account the findings of [chapter 3](#) and [chapter 4](#), we developed a young, female ECA. [Chapter 6](#) showed that the patients' perceptions of the ECA's characteristics did not change over time, but that patients' likeliness of following the ECA's advice dropped significantly from before interaction to after three weeks of use. From then on, no difference in patients' likeliness

of following the ECA's advice was found. A few studies researched ECAs in a long-term, daily life setting: an ECA for hospital discharge (*Zhou et al., 2014*) and promotion of physical activity (*Bickmore et al., 2005, 2009, 2010, 2013*), healthy nutrition (*Creed et al., 2015*) and medication adherence (*Bickmore et al., 2010*). Although these studies indicate that an ECA has potential to promote a healthy lifestyle, some of them show that users' interaction with the ECA decreased over time (*Bickmore et al., 2005, 2009, 2013*). Only one of these studies particularly showed how users' perceptions of the ECA changed over time, showing a decrease in enjoyment of an ECA's stories and level of engagement with the ECA (*Bickmore et al., 2009*). This decrease in user perception might be related to the decrease in likeliness of following the ECA's advice in our study. As discussed in [chapter 6](#), we expect that the decrease in likeliness of following advice could be explained by participants having a negative adaptation gap, meaning that their expectations of the ECAs' capabilities exceeded the ECAs' actual capabilities (*Komatsu et al., 2012*). Such a negative adaptation gap might be related to the technology readiness level (TRL). In our study, the TRL was about 5 – i.e. the technology was a large scale prototype tested in the intended environment, and not yet a fully operational system at (pre)commercial scale (*European Commission, 2014*). This TRL fitted the small-scale, exploratory character of the study aimed to investigate the feasibility of eHealth for self-management. As such, the content of the ECA's messages might not have been as personalised as users expected, creating a negative adaptation gap. As a result, users became frustrated. This frustration might have negatively influenced their perception of an ECA's characteristics, as in our case, affecting the ECA's reliability and users' likeliness of following the ECA's advice. Such a negative adaptation gap might also explain part of the results of [chapter 7](#), stating that preferences for particular ECA designs changed over time and satisfaction with all ECAs decreased over time.

Concluding, we expect that in long-term interaction of older adults with an ECA, 1) a young ECA may contribute to positive impressions and 2) that the content of an ECA's dialogues has a larger effect on users' impressions of the ECA than its appearance. The next step would be to investigate whether these hypotheses can be confirmed.

ECAS FOR USER ENGAGEMENT IN EHEALTH

Throughout this thesis, we have learned about the design of an ECA's appearance for user engagement. In this section, we discuss our thoughts on effects of an ECA's appearance on user engagement in relation to effects of other ECA aspects.

We have learned that the TRL of the technology used, in this case the ECA, is important. For exploratory studies one does not need to develop a fully functioning technology. But, when investigating effects of particular ECA design features in daily life, it is essential that the TRL of the ECA is sufficient. This does not only apply to the ECA's appearance, but to the ECA offered as a whole: its appearance, content and behaviour. We have learned that, when it comes to design for engagement, one cannot separate an ECA's appearance from other ECA design aspects, such as its content. As shown in interviews presented in [chapter 4, 6 and 7](#), users may have difficulties in separating their perceptions of an ECA's appearance from their perceptions of other ECA aspects, such as its messages. Users judge ECAs as

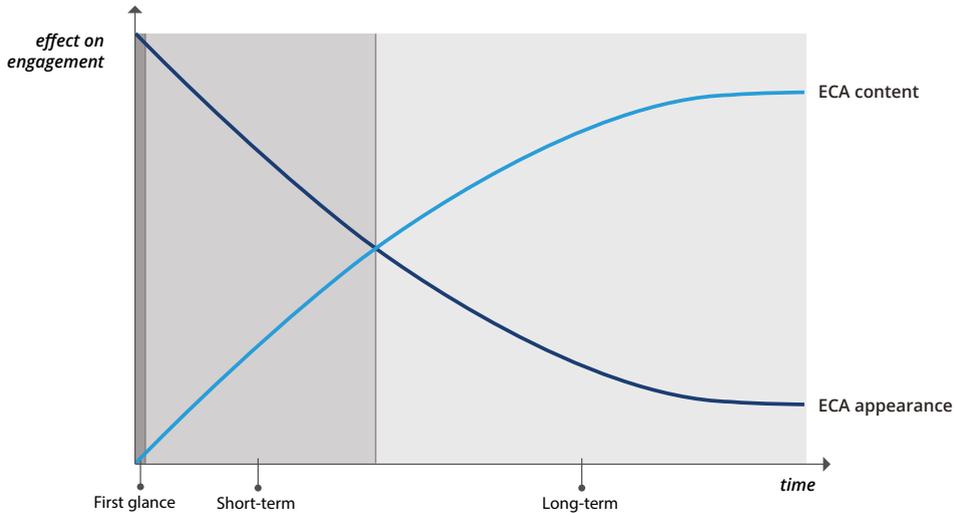


Figure 09.1 – Our illustration of the relation between an ECA’s appearance and its content on engagement in eHealth over time.

a whole. An ECA’s appearance might have zero effect on user engagement, as long as the content of the ECA’s messages is not optimised for it. This also works the other way around: perfectly designed ECA content might not affect user engagement, when users are annoyed by an ECA’s appearance. In, [Figure 09.1](#) we illustrate our thoughts on the effects of different ECA aspects on engagement over time. We distinct three phases:

- **Phase 1: First glance** – At first glance, the user is not yet exposed to the ECA’s content, such that only the ECA’s appearance influences a user’s engagement with the ECA. As in human-human interaction, we immediately form impressions of an ECA’s characteristics at first encounter. Therefore, an ECA’s appearance can be designed to lower the threshold to interact with the ECA and to manage a user’s expectations of the ECA’s capabilities by adapting its appearance to its actual capabilities. As described in this thesis, especially an ECA’s static characteristics (its looks), such as its demographics, human-likeness and conveyed expertise, can contribute to lowering this threshold.
- **Phase 2: Short interaction** – During short interactions with an ECA, both the ECA’s appearance and its content affect user engagement. With respect to an ECA’s appearance, dynamic characteristics, such as emotion and relational behaviour, become important to build a relation with the user in order to keep the user engaged.
- **Phase 3: Long-term interaction** – To establish long-term interaction, an ECA’s content becomes more important than an ECA’s appearance. Content with sufficient variation and personalisation are necessary to keep a user’s engagement in the long-run.

Furthermore, we suggest to manage the users’ expectations of the ECA offered. Expectation management allows one to focus on the objective of the study instead of indirectly testing the TRL of a

technology. Therefore, users' expectations should match the actual TRL of the ECA. One can do so, by explaining the ECA's role and functionalities in the application in an early stage.

TOWARDS DESIGN STRATEGIES FOR ECAS IN EHEALTH

Our review of the state-of-the-art of design features for an ECA's appearance in eHealth ([chapter 2](#) and update in [chapter 8](#)) showed that the research area is still in its infancy, but growing. [Chapter 8](#) showed that current studies on an ECA's appearance for eHealth are of low, fair or moderate quality. For the majority of the design features researched no or preliminary evidence exists for effects on user engagement. In general, the use of ECAs for eHealth is growing, as indicated by recently published reviews on the applicability of ECAs in eHealth (*Provoost et al., 2017; Kramer et al., 2020; Scholten et al., 2017*). These reviews all show potential of ECAs for eHealth. But, they also indicate that future research is needed, since evaluating effects of ECAs is difficult due to a large heterogeneity in study design and outcome measures. Consequently, **consensus on how to design an ECA's appearance for eHealth is far from established.**

So, what about future research on an ECA's appearance to stimulate user engagement in eHealth? Future studies on ECA design features are needed. As described in [chapter 8](#), we should work towards studies in stage II, III and IV of the renewed framework for evaluation of telemedicine (*Jansen-Kosterink et al., 2016*). These type of studies evaluate an ECA's appearance in the actual setting of use: daily life. These future studies might be reproductions of studies on ECAs in general or in other contexts. Looking back at the findings of this thesis, we observe that findings of studies on ECAs in eHealth do not differ a lot from similar studies for ECAs in general or in other contexts. Therefore, we expect that these findings can possibly also be applied to the eHealth context and the other way around. However, some studies indicate that users' preferences for and perceptions of a particular ECA depend on its task. We seem to associate particular tasks with particular appearances, according to human stereotypes (*Forlizzi et al., 2007; Zimmerman et al., 2005; Nguyen & Masthoff, 2007*). As an ECA's task in eHealth (e.g. promoting physical activity) can be very different from tasks of ECAs in other contexts (e.g. providing maths instructions or support in booking a flight), the ECA appearance that contributes most to a user's engagement might differ across contexts. Therefore, we recommend to reproduce studies for ECAs in general or in other contexts, to investigate whether results can indeed be applied to the eHealth context as well. But, before we do so, we have take into account several aspects. Below, we discuss aspects that we came across throughout this thesis and that should be considered while performing future research on the design of ECAs.

First, the large variety in what and how particular ECAs are researched makes it difficult to compare findings of multiple studies and create a grounded understanding. Therefore, we believe that it is important to strive for more homogeneity among studies on an ECA's appearance. Researchers in related contexts support this note, such as *de Filippis et al. (2020)*: 'dedicated tools and methods for assessing the quality of interaction with chatbots should be developed and used to enable the generation of comparable evidence'. *Fitriani et al. (2019)* show that, currently, the majority of questionnaires used in ECA research are only reported in one paper. They argue that a unified ECA measurement

instrument should be created. We agree. Besides, we suggest to introduce more unified reporting. As an example, in gaming research it is quite common to provide a summary of the most important characteristics of the game researched, often in the form of a table. For example, *Tabak et al. (2020)* and *Rose and Unni (2020)* provide tables in which they list aspects of their games, such as the target group, setting, story and game mechanics. We believe that a similar description of characteristics would be valuable for research on an ECA's appearance. In our literature review ([chapter 2](#) and [chapter 8](#)), we made a first attempt to create such tables. Per study, we described the ECA's topic and target group, what ECA design features were researched and what the outcome measures and study outcomes were. When all studies on an ECA's appearance would incorporate such a summary of study and design characteristics, comparison of findings of studies investigating the same design feature will become easier and with it the making of general statements on the effects of such features.

Second, we need to work towards a shared understanding of what engagement in eHealth actually is. We need to have a uniform definition throughout the research area. Currently, this is not the case: many different conceptualisations of engagement are used (*O'Brien & Toms, 2008; Cole-Lewis et al., 2019; Yardley et al., 2016; Kelders & Kip, 2019*). Several researchers conceptualised engagement for eHealth in general (*Kelders & Kip, 2019*), or for Digital Behaviour Change Interventions (DBCIs) in particular (*Cole-Lewis et al., 2019; Yardley et al., 2016; Perski et al., 2017*). However, despite these conceptualisations, there is not one definition that is used throughout the research area.

Third, when having a shared understanding of what engagement is, how do and should we measure it? We need to work towards a unified measurement of engagement for eHealth, such that we can compare results of different studies. Currently, this is not the case: methods used to measure engagement in eHealth vary a lot (*Short et al., 2018*). Also, they often actually measure attributes that predict engagement and not engagement itself (*Short et al., 2018*). Some unified measurements for engagement in eHealth are being developed, such as the Twente Engagement with eHealth Technologies Scale (TWEETS) by *Kelders and Kip (2019)* and the DBCI engagement scale by *Perski et al. (2020)*. Although these are two great examples to work towards more uniform ways of measuring engagement, a 'golden standard' when it comes to measuring engagement in eHealth does not yet exist.

As a next step, results of future work which take the aspects described above into account, can be used to work towards design strategies for ECAs in eHealth. We believe that is important to work towards an elaborated set of design strategies, such that designers could easily develop an ECA's appearance for a particular target group and context. In [chapter 8](#), we created an initial set of five design strategies for an ECA's appearance in eHealth based on state-of-the-art research including our own work. To the best of our knowledge, no other design strategies for an ECA's appearance in eHealth exist. Our strategies go beyond frameworks that provide a solid basis for ECA design by identifying and categorising design features, including features for an ECA's appearance (*Ruttkay et al., 2004; Straßmann & Krämer, 2017*). Our work is similar to the work of *Veletsianos et al. (2009)* who designed 15 research-based guidelines for designing engaging pedagogical agents. Although *Veletsianos et al. (2009)* provide a more extensive set of strategies, they do not solely focus on an ECA's appearance, the content of our strategies is similar. Like the strategies by *Veletsianos et al. (2009)* our strategies remain open to one's own interpretation: the motivations for the strategies are clearly explained, but guides for practical implementation are little concrete. Therefore, our initial set of design strategies should be validated, updated and extended by future work.

As a possible result of extended work, we believe that it is important to add, what we call, ‘engagement effect labels’ to the design strategies. One could imagine that particular outcome variables have a larger effect on user engagement than others, reflected in the ‘engagement effect labels’ of the design strategies (e.g. labelled as either low, medium or high). The extent to which different ECA design features affect engagement can be compared, the engagement effect relatively to that of other design features can be classified, and design strategies can be provided with an ‘engagement effect label’.

An elaborated set of design strategies for an ECA’s appearance could contribute to the design of successful ECAs for eHealth. Yet, an ECA’s appearance is not the only aspect we have to think about when developing ECAs. As describe in section 3, users judge ECAs as a whole. Therefore, we suggest to take a more holistic perspective on designing ECAs for eHealth and work towards a broader set of design strategies. Besides strategies for an ECA’s appearance, this set could contain, for example, strategies for an ECA’s content or interaction. Such a broader set of design strategies could raise awareness among researchers, developers and designers about the importance of particular ECA aspects when it comes to user engagement. In that sense, the strategies provide direction for what aspects should be researched and evaluated with the target group of the ECA being developed. As described before, the relative impact of these aspects on user engagement most likely depends upon the context in which the ECA is used. Therefore, depending on the context of use, some design strategies might have more value to look into than others.

To speed up creating the body of knowledge, we suggest to publish such a set of ECA design strategies for eHealth and resources openly online, such that researchers can build upon the work of others. This suggestion is in line with open science, capturing the following shift in perspective on science: “from the standard practices of publishing research results in scientific publications towards sharing and using all available knowledge at an earlier stage in the research process” (*European Commission, 2016*) (p.33). Through open science, knowledge on how to design ECAs for eHealth will become widely accessible to experts in different fields. This is important, since designing ECAs for eHealth is complex, involving expertise from many different fields, ranging from health science and behavioural science to computer science. In addition, open science allows people and businesses outside the scientific community that are interested in developing or using ECAs to contribute to the research. By sharing knowledge between many different parties with expertise in different fields, the development of ECAs for eHealth could receive a boost.

TOWARDS FUTURE HEALTH CARE WITH ECAS

How will engaging characters, such as Sylvia (**chapter 1**), support us in our daily lives in the future? How will they look like? Will they be 3D web-based characters (*Zalake et al., 2020*), or projected using Augment Reality (*Reinhardt et al., 2020*)? How will we interact with them and what will we talk about?

We expect that ECAs will gradually play a larger role in our lives. First, we will interact with them on all devices that we use throughout the day, from our smartphone and tablet to our smart refrigerators. Eventually, we will take our coach with us everywhere we go via our smart glasses or maybe even

via a chip implanted into our body. This way, our coach will be available to support us throughout the day. Our coach will be a holographic projection and interactions with our coach will be similar to interactions with other humans: we will interact via speech and show emotion and empathy towards each other. The underlying eHealth application will not just show information on a screen, but via augmented projections into our real-life surroundings. Our coach will present information about our lifestyle via graphs, images and videos while 'standing' next to it. With the rise of Artificial Intelligence (AI), ECAs will become more intelligent and advanced. Consequently, our coach will know all about our interests and struggles and will respond to it, such that our conversations with our coach will become more personalised and engaging.

We are certain that future technology will make this scenario possible. But the introduction of such personal health coaches into our future lives also faces challenges. First, do we really want such a health coach embedded in our daily life 24/7? Are we up to such a radical change? Some people are tech-minded and adopt new technologies early, but in general, people have difficulties with it. We are connected to our routines, fear the unknown (*Lazanyi, 2018*) and fear that technology will manipulate us (*Vollmer, 2018*). Especially a lack of trust in new technology is one of main reasons holding us back from it (*Lazanyi, 2018*). Also, we fear that we lack the ability to think as computer algorithms (*Vollmer, 2018*). When we allow a computer to make decisions for us, we want to understand its rationale, especially in most crucial decisions (i.e. life or death situations) (*Kolasinska et al., 2019*), but we are not always able to understand. This as a reason why Explainable Artificial Intelligence (XAI) – AI that generates solutions that can be understood by humans – is getting more and more attention.

Second, with respect to intelligent assistants in particular, main concerns for adoption are fear for loss of control or to be taken over by machines (*Vollmer, 2018*). Are we ready to transmit a particular level of control to our personal health coach, such that not we, but our coach determines how our life looks like: what we eat, when we walk, when we go to bed? Other concerns against intelligent assistants are privacy and security issues (*Feng et al., 2019; Fruchter & Liccardi, 2018; Burbach et al., 2019*). A health coach will always be 'on', listening to triggers in our surroundings. Where will data gathered by the coach be stored? What will it be used for? Who will be able to access the data and algorithms? What if people with malicious intents can access the data and algorithms? As a consequence, we might ask ourselves the question whether our conversations remain private. By mining and analysing lots of personal data, our coach can generate our 'life portrait' and extract valuable information from it. This, way, we might unconsciously lose control over our privacy (*Feng et al., 2019*).

Lastly, will personal health coaches result in us escaping reality and living in a 'virtual world'? How will our relationship with our personal health coach affect relationships we have with others? With the rise of new technologies, communication has never been this easy, but as a drawback, we communicate without engagement and trust our peers only to a limited extent (*Lazanyi, 2018*). This results in people having a lack of social support and eventually getting lonely (*Lazanyi, 2018*). Thus, we already tend to loose connection with others, will a personal health coach decrease our connection even more?

The questions portrayed above illustrate some of the challenges that lay ahead of us when it comes to the adoption of ECAs as personal health coaches. We believe that engaging ECAs have potential to support us in developing and maintaining a healthy lifestyle, but that user-centred design is the key for acceptance and adoption. End-users should be incorporated early in the design process, such that ECAs can be designed according to their wishes and needs. Furthermore, we believe that, as scientists or engineers we should act according to our moral responsibility when developing new technologies, such as personal health coaches. It is our task, together with the government, to research ethics

and safety risks of new technologies, and to regulate them via laws, safety standards and industry supervision (*Feng et al., 2019*). This allows us to create ECAs that really serve us as health companions. We will be engaged in our relationship and listen more carefully to what our coach has to say. We will become more aware of what a healthy lifestyle is, why it is important and how to improve our own lifestyle. Consequently, we will take an extra walk, change our eating habits and learn about mindfulness. Together with our health coaches, we will live longer and healthier and happier ever after.



Appendices
References
Summary
Samenvatting
Dankwoord
About the Author
Progress Range

APPENDIX A – DETAILS LITERATURE REVIEW

Table A.1 – General information of the articles included in the review.

Source	Goal application	N	Gender	Age	Education	Cultural background	Other	Eval. stage	Study design
Alsharbi & Richards (2017)	[M] Medical treatment advice for anxiety and post-trauma stress disorders	11	9 females, 2 males	[C] 11-15		[C] 9 Northern-Western European, 2 Oceania	Playing computer games; all > 3h a week	I	S
Amini et al. (2013)	[A] Changing unhealthy alcohol consumption behaviours	51	37% females, 63% males		[S] University students	[C/H/AA/A] 55% White, 27% Hispanic, 16% African American, 2% Asia	Drinking at least one drink a month	I	E
Amini et al. (2014)	[A] Changing unhealthy alcohol consumption behaviours	81						I	E
Bickmore et al. (2005)	[P] Promoting physical exercise	91						II	E
Bickmore et al. (2009)	[P] Promoting walking behaviour	26	21 females, 5 males	[A] 54-67	[C] 92% had some college education	[C/AA] 80% Caucasian, 20% African American		II	E
Bickmore et al. (2010)	[P] Promoting walking behaviour	E1: 24 E2: 26	E1: 17 females, 7 males; E2: 21 females, 5 males	E1: [A] 55-67 E2: [A] 54-67	E2: [C] 92% had some college education	E2: [C/AA] 80% Caucasian, 20% African American		E1: II E2: II	E
Bickmore & Picard (2004)	[P] Promote physical exercise	60	60% female		[S] 69% students			II	E
Bickmore & Picard (2005)	[P] Motivating to performing physical exercises	101						II	E
Bickmore & Ring (2010)	[O] Promoting pre-conception care	9	All female	[A] 20-24	[C] 1: high school level, & at least some college	[AA] All African American		I	E
Bickmore & Schulman (2007)	[M] Calming users using negative mood induction	16	9 females, 7 males	[A] 18-30	[S] 15 were students			I	E

Source	Goal application	N	Gender	Age	Education	Cultural background	Other	Eval. stage	Study design
Creed & Beale (2012)	[N] Changing nutrition habits	50	28 females, 20 males					II	E
Creed et al. (2015)	[N] Improving fruit and vegetable consumption	45	17 females, 26 males, 1 not specified	[A] M = 30.53 (SD = 9.33)	[S / U] Graduates and undergraduates school of computer science	[C] UK and three other countries	No dietary plan, no history of eating disorders, in contemplation, preparation or action stage	II	E
Forlizzi et al. (2007)	[T / P] Being a medical advisor, athletic trainer	24	15 females, 9 males	[A] 18-36, M = 25.7				I	S
Frost et al. (2012)	[P] Raise intrinsic motivation in children exercising	61	39 girls, 22 boys	[C] M = 9.2 (SD = 0.94)		[C] Dutch		I	E
Grillon & Thalmann (2008)	[M] Reducing social phobia (in public speaking)	12		[A] 25-35				I	E
Kang & Gratch (2010)	[M] Psychotherapy	57	53% females, 47% males	[A] M = 30.68 (SD = 10.08)				I	E
Lisetti et al. (2013)	[A] Changing unhealthy alcohol consumption behaviours	81			[S] University students			I	E
Malhotra et al. (2016)	[O] Assisting people with cognitive disabilities in daily life	22	14 females, 8 males	[E] Elderly			All have Alzheimer's disease	I	E
Nguyen & Masthoiff (2007)	[P] Promoting physical exercise	E1: 51 E2: 59			E2: [S/U] University staff and graduate students			I	E1: Q, S E2: E
Nguyen & Masthoiff (2009)	[M] Performing behaviour modification via the mood manipulation technique	84	47 females, 37 males	[A] M = 25.4 (SD = 7.9)	[S/U] University students and staff			II	E

Source	Goal application	N	Gender	Age	Education	Cultural background	Other	Eval. stage	Study design
Olafsson et al. (2017)	[P / N] Motivating to exercise and good nutrition	84	60% male	[A]M = 22.8 (SD = 2.4)	[S] Students	[A/C/AA] 79% Asian, 18% White, 2% Hispanic, and 1% Black	Like rap	I	E
Parmar et al. (2018)	[T] Discussing the importance of having a healthcare proxy and providing information about receiving life-support via a medical ventilator	E1: 43 E2: 308	E1: 20 females, 23 males, E2: 175 females, 133 males	E1: [A] 16-62, E2: [A] 18-76, M = 38 (SD = 12.32)				I	E1: S E2: E
Ring et al. (2014)	[T] Performing social interactions on user's favourite books and medical counseling (cancer related)	E1: 67, E2: 47	E1: 31 females, 36 males, E2: 16 females, 31 males					I	E1: E E2: E
Robertson et al. (2015)	[T] Helping patients (with lower health literacy and socio-economic status) to understand the benefits and drawbacks of different treatment choices in response to prostate cancer		Location I: Men with prostate cancer and a few wives of patients, location II (target group): All male		Location I: [C] Higher education than location II, location II: [L] Lower education than location I	Location I: [C] Caucasian, location II: [AA] African American	Location I: Higher health literacy and socio-economic status than location II	I	Q
Schmeil & Suggs (2014)	[N / P] Change eating and or physical activity behaviours	512	219 females, 293 males	[A]M = 33				II	E
Silverman et al. (2001)	[T] Helping with dealing with a variety of heart attack scenarios	18			[C / U] Cardiologists and general practitioners			I	E
Skalski et al. (2007)	[T] Providing information about risks and treatment of problems associated with high blood pressure	125	83 females, 42 males	M = 21.56				II	E

Source	Goal application	N	Gender	Age	Education	Cultural background	Other	Eval. stage	Study design
Tielman et al. (2017)	[M] Providing psychoeducation: expressive writing for Post-Traumatic Stress Disorder therapy	46	12 females, 38 males	M = 22.9 (SD = 3.8)	[S / U] University staff and students	[C] All had Dutch as first language		II	E
van Vugt et al. (2006)	[O] General eHealth context	80	56 females, 24 males	M = 23 (SD = 7.8)	[S] University students of Free University of Amsterdam		Body mass index: M = 22 kg/m ² (SD = 3.8)	I	E
Wissen et al. (2016)	[P / T] Providing information about importance physical activity, reminders for medication intake and appointments with health care professionals	64	34 females, 30 males	[E] < 65: 6.3 %, 65-70: 62.5%, > 70: 31.3%	[C / U] Highly educated: 59.4% => associate degree	[C] 95.3% Caucasian	<ul style="list-style-type: none"> 54.7% living with spouse Chronic illnesses with medication intake 	I	E
Yin et al. (2010)	[P] Promoting physical exercise	43	44.2% female	18-65		[H/C] 65.0% Latinos: Speak, read and understand English and Spanish: A) Anglo-American born in U.S., in Spanish-speaking Latin American countries for at least two months, English as first language, or B) Latino/Latina born in Spanish-speaking Latin American country, in U.S. for at least two months, Spanish as first language.		II	E
Zhou et al. (2014)	[T] Performing hospital discharge	764	50.8% female	18-90 (M = 49)		[A/A/C/H] 52.6% African American, 29.1% Caucasian, 11.4% Hispanic		III	E
Zhou et al. (2017)	[P] Promoting physical activity to young Chinese adults living in the U.S.	49	46.9% male	[C/A] 9-34, M = 26.1 (SD = 3.6)	[C/U] 85.7% college graduate or higher	[A] Speak and read English, born in China, moved to U.S. at least by age of 16, lived in U.S. for at least six months (time in U.S.: M = 2.8 years, SD = 1.8)		II	E

Table A.2 – Overview of the articles included in the review. For every article, the feature category, design feature(s), outcome variable(s), methods and outcomes are listed.

Source	Feature category	Design feature(s)	Outcome variable(s)	Method	Outcomes
Speech and/or textual output, facial and gaze expressions and hand and body gestures					
Creed & Beale (2012)	Speech and/or textual output, facial and gaze expressions	Presence or absence of emotion	<ul style="list-style-type: none"> [AC] Caring [AC] Likeability [AC] Trustworthiness [AC] Intelligence [AC] Feeling of positivity [AC] Feeling of support [AC] Level of emotion agent [U] Time viewing educational resources [U] Number of articles and recipies viewed [IT] Desire to work with agent again [O] General opinion via open questions 	<p>Two conditions:</p> <ul style="list-style-type: none"> Emotional (E): emotional facial expressions, appropriate emotion in audio, response manipulated to be appropriate to responses user Non-emotional (NE): neutral facial expressions, neutral audio, neutral responses 	<ul style="list-style-type: none"> Caring, likeability: E > NE Trustworthiness, intelligence, feeling of positivity, feeling of support: no difference Time viewing educational resources, number of articles viewed: no difference Desire to work with agent again: <ul style="list-style-type: none"> Not interacting again? NE: 12 people, E: 2 people Many people preferred to use the website only Open questions: <ul style="list-style-type: none"> Both conditions: visual appearance was liked NE: perceived less positively, neutral voice disliked NE: comments on visuals E: emotional voice liked E: positive comments related to the emotional expressions (liked that agent was friendly, warm, and cheerful)
Creed et al. (2015)	Speech and/or textual output, facial and gaze expressions	Presence or absence of emotion	<ul style="list-style-type: none"> [AR] Working Alliance Inventory (WAI) [B] Fruit and vegetable intake baseline and after intervention [B] Average fruit and vegetable intake [B] Days per week eating >=5 portions [U] Dropout [U] Number of days logged in [U] Number of times logged in [U] Amount of articles and recipies viewed [U] Number of times viewed different resources [UC] Self-efficacy [UC] Decisional balance 	<p>Two conditions:</p> <ul style="list-style-type: none"> Emotional (E): facial expressions and recorded speech Non-emotional (NE): neutral expression and voice 	<ul style="list-style-type: none"> WAI: no difference Fruit and vegetable intake compared to baseline: NE > E Average fruit and vegetable intake: no difference Days per week eating >=5 portions: no difference Dropout: no difference Logins, articles viewed and resources accessed: no difference Self-efficacy: no difference Decisional balance: no difference

Emotion

Source	Feature category	Design feature(s)	Outcome variable(s)	Method	Outcomes
Silverman et al. (2001)	Speech and/or textual output, facial and gaze expressions	Amount of emotional behaviour	<ul style="list-style-type: none"> [B] Intention shift (pre and post intervention) for: <ul style="list-style-type: none"> Self care during attack Waiting to call Calling/seeing a doctor Calling 9-1-1 [B] Scores on knowledge quiz (tests: knowledge on heart attack symptoms and delay issues) [UX] Usability [AC] Logic (and believable) [AC] Pedagogy 	<p>Four prototypes (focus groups):</p> <ul style="list-style-type: none"> Multimedia versions (M): <ul style="list-style-type: none"> Full Emotive Basic Agent Didactic versions (D): <ul style="list-style-type: none"> Pop-up Text Didactic/Text 	<ul style="list-style-type: none"> Intention shift: <ul style="list-style-type: none"> Pop-up Text (140.6%) > Full Emotive (125%) > Didactic/Text and Basic Agent (37.5%) Shift for Full Emotive, Basic Agent, Didactic/Text Shift in call 9-1-1: no difference D and M Shift in self care: M (140%) > D (56.3%) Intention to call or see a doctor: M > D Knowledge quiz: all: change in knowledge, after-the-fact learning: D best, post-intervention: M > D Usability: <ul style="list-style-type: none"> Overall: positive Highest for Full Emotive Drops for each media-crippled version Didactic/Text: lowest ratings (despite being the shortest module with the least complexity) Pedagogy: full Emotive, Basic Agent, Pop-up Text result in: (1) vividness that helps better remember results, (2) learning about symptoms overall, where symptoms occur in the heart, and about delay issues, (3) subjects learning more how time slips by during a heart attack event Logical (and believable); all conditions: logical and believable, least: Basic Agent, Full Emotive > Basic Agent
Emotion					
Amiri et al. (2013)	Facial and gaze expressions	Presence or absence of empathy	<ul style="list-style-type: none"> [UX] Attitude (ATT) [IT] Intention to Use (ITU) [UX] Perceived Enjoyment (PENJ) [UX] Perceived Ease of Use (PEOU) [AC] Perceived Sociability (PS) [UX] Perceived Usefulness (PU) [AR] Social Presence (SP) [AR] Trust (TRUST) [AR] Anxiety (ANX) [AR] Social Influence (SI) [AC] Anthropomorphism (ANT) [AC] Likeability (LIKE) [AC] Animacy (ANIM) [AC] Perceived Intelligence (PI) [AR] Perceived Safety (PSA) 	<p>Two conditions:</p> <ul style="list-style-type: none"> Non-empathic (NE): neutral facial expression, does not empathize with user Empathic (E): facial expressions, head gestures, smile, head posture mimicry, eye brow movement, mutual gaze, lip, synchronized verbal expressions 	<ul style="list-style-type: none"> ATT, ITT, PENJ, PEOU, PS, PU, SP, TRUST, SI, ANT, LIKE, ANIM, PI, PSA: E > NE ANX: No difference E and NE
Relational, empathic behaviour					



Relational, empathic behaviour

Source	Feature category	Design feature(s)	Outcome variable(s)	Method	Outcomes
Amini et al. (2014)	Facial and gaze expressions	Presence or absence of empathy	<ul style="list-style-type: none"> [UX] Attitude (ATT) [IT] Intention to Use (ITU) [UX] Perceived Enjoyment (PENJ) [UX] Perceived Ease of Use (PEOU) [AC] Perceived Sociability (PS) [UX] Perceived Usefulness (PU) [AR] Social Presence (SP) [AR] Trust (TRUST) [AR] Anxiety (ANX) [AR] Social Influence (SI) 	<p>Three conditions:</p> <ul style="list-style-type: none"> Non-empathic (NE): neutral facial expression, does not empathize with user Empathic (E): facial expressions, head gestures, smile, head posture mimicry, eye brow movement, mutual gaze, lip, synchronized verbal expressions Text-only (T) 	<ul style="list-style-type: none"> ATT, ITU, PENJ, PS, PU, SP, TRUST, E > NE, E > T ATT, ITU, PEOU, SP: NE > T PENJ, PS, PU, TRUST, SI: T > NE
Bickmore et al. (2005)	Speech and/or textual output, facial and gaze expressions, hand and body gestures	Presence or absence of relational behaviour	<ul style="list-style-type: none"> [AR] Working Alliance Inventory (on day 7 and 27) [AC] Likability agent [AC] Closeness relationship agent [UX] Usefulness discussions agent [IT] Desire to continue working with agent [U] Nr. of pages with educational content accessed [B] Nr. of days a week physical activity performed 	<p>Three conditions:</p> <ul style="list-style-type: none"> Control (web forms only, no agent) Non-relational (NR): talks about exercise behavior, negotiates and follows up on exercise goals, suggestions to overcome obstacles to exercise, cognitive-behavioral techniques for health behavior change Relational (R): social dialogue, empathetic feedback, meta-relational communication, humor, close forms of address, nonverbal immediacy 	<ul style="list-style-type: none"> Bond-subscale of WAI: R > NR (on day 7 and 27) Overall WAI: no differences Likability agent, desire to continue: R > NR Educational pages viewed: R and NR > C Physical activity: R and NR > C
Bickmore & Picard (2004)	Speech and/or textual output, facial and gaze expressions, hand and body gestures	Presence or absence of caring	<ul style="list-style-type: none"> [AR] Working Alliance Inventory (WAI) [IT] Willingness to continue [AR] Selection type of farewell messages 	<p>Two conditions:</p> <ul style="list-style-type: none"> Relational, caring (R): empathetic feedback, concerned facial expressions Non-relational, non-caring (NR): no feedback, no concerned facial expressions 	<ul style="list-style-type: none"> Perception caring: R > NR WAI items "Agent cares about me", "Agent is genuinely concerned", "Agent likes me" and "Agent and I trust in one another": R > NR Willingness to continue working: R > NR Selection of sentimental farewell: R > NR

Source	Feature category	Design feature(s)	Outcome variable(s)	Method	Outcomes
Bickmore & Picard (2005)	Speech and/or textual output, facial and gaze expressions, hand and body gestures	Presence or absence of relational abilities	<ul style="list-style-type: none"> [AR] Working Alliance Inventory (WAI) [AC] Likability agent [AR] Closeness relationship [IT] Desire to continue [IT] Desire to continue using the overall FITTrack system [AR] Sentimentality farewell [O] Opinion on whom helped best to perform exercises [B] Minutes moderate activity a day [B] Days a week > 30 minutes moderate activity [B] Steps per day [B] Days per week > 10,000 steps [U] Average nr. of educational pages accessed per session 	<p>Three conditions:</p> <ul style="list-style-type: none"> Relational (R): daily social dialogue, frequent meta-relational dialogues, friend role, small social distance, daily empathy exchanges, humour, daily continuity behaviours, high non-verbal immediacy Non-relational (NR): no social dialogue, minimal meta-relation dialogues, stranger role, large social distance, no empathy exchanges, no humour, no continuity behaviours, low non-verbal immediacy Control (C): no agent 	<ul style="list-style-type: none"> Bond subscale WAI: R > NR Likability: R > NR Closeness relationship: R > NR Desire to continue: R > NR Desire to continue system: no difference Sentimentality farewell: R > NR Who helped best: no difference Who helped best: agents > friends, family and workout buddy Physical activity: for both groups increase after intervention Physical activity: for both groups decrease during follow up (decrease R > NR) and NR > C Days a week > 30 minutes moderate activity: R and NR > C Pages accessed: R and NR > C
Bickmore & Schulman (2007)	Speech and/or textual output, facial and gaze expressions	Expressivity user, empathic accuracy	<ul style="list-style-type: none"> [B] Change in positive and negative affect [UX] Satisfaction [IT] Willingness to continue working [AC] Likelihood [AC] Perception of caring 	<p>Two conditions:</p> <ul style="list-style-type: none"> EXPRESSIVITY: high expressivity dialogue, free speech input EMPATHY: high empathic accuracy and empathic feedback, including happy facial display and close-up facial display for other feeling states, input limited to predefined answers 	<p>Empathic accuracy more important than user expressivity:</p> <ul style="list-style-type: none"> Change in positive affect from pre-induction to post-intervention: EMPATHY > EXPRESSIVITY Satisfaction: EMPATHY > EXPRESSIVITY Females showed greater desire to continue with EXPRESSIVITY agent than males
Lisetti et al. (2013)	Facial and gaze expressions	Presence or absence of empathic abilities	<ul style="list-style-type: none"> [UX] Attitude (ATT) [IT] Intention to Use (ITU) [UX] Perceived Enjoyment (PENJ) [UX] Perceived Ease of Use (PEOU) [AC] Perceived Sociability (PS) [UX] Perceived Usefulness (PU) [AR] Social Presence (SP) [AR] Trust (TRUST) [AR] Anxiety (ANX) [AR] Social Influence (SI) [AC] Anthropomorphism (ANT) [AC] Likability (LIKE) [AC] Animacy (ANIM) [AC] Perceived Intelligence (PI) [AR] Perceived Safety (PSA) 	<p>Three conditions:</p> <ul style="list-style-type: none"> Text-only interface (T) Non-empathic ECA (NE): neutral facial expressions, does not empathize with user, no reflection on user's answers, ignores user's emotional state ECA with some empathic abilities (E): emotional facial expressions (happy, sad, concerned, surprised, and neutral); head gestures (nodding); big/subtle smile; head posture mimicry (pitch, yaw, roll); eyebrow movement and lip synchronized verbal reflections 	<ul style="list-style-type: none"> ATT: E > NE, E > T. All means positive ITU: E > NE, E > T, NE and E positive, T negative mean PENJ: E > NE, E > T. All means positive PEOU: No differences. All means positive PS: E > NE, E > T, T > NE PU: E > NE, E > T. All means positive SP: E > NE, E > T, E positive NE and T negative mean Trust: E > NE, E > T, E and T positive, NE negative mean ANX: No differences. All means positive SI: No differences. All means positive ANT: E > NE LIKE: E > NE, E > T. All means positive ANIM: E > NE, E positive, NE negative mean PI: E > NE, T > NE. All means positive PSA: E > NE, E > T. All means positive

Relational, empathic behaviour

Source	Feature category	Design feature(s)	Outcome variable(s)	Method	Outcomes
Nguyen & Masthoff (2009)	Speech and/or textual output, facial expressions, hand and body gestures	Modality (animated vs no visuals), level of empathy (non-empathy, empathy, empathy and expression)	<ul style="list-style-type: none"> [B] PANAS-S (short-form positive and negative affect schedule questionnaire) before mood induction, after mood induction and after web credibility judgment task [B] Performance on second and third task (time reading and judging, total number of thoughts recalled) [LUX] Judgement of web credibility (biased, comprehensive, caring, trustworthy, expert) [AC] Attitude toward agent: enjoyability, caring, likeability, trustworthiness, empathy [AC] Perception of agent design (language, appearance, gestures and voice) 	<p>2 x 3 experiment:</p> <ul style="list-style-type: none"> Modality: <ul style="list-style-type: none"> Animated (A): animated, synthesized voice No visual (NV): no visual representation, no voice Intervention: <ul style="list-style-type: none"> Empathy (E): small talk, polite and friendly, acknowledging mood change, offering means to correct agent's judgments, conveying sympathy and empathy, acknowledging validation emotional state Empathy and expressivity (EE): same as EMPATHY, encourages subjects to freely express feelings Non-empathy (NE): ignores feelings, asks distracting questions <p>Three steps:</p> <ul style="list-style-type: none"> First: mood manipulation Second: participant looks for info about alcoholism and consequences on a low and high credible website Third: measuring effect of mood on subject's judgements of memory 	<ul style="list-style-type: none"> Change in negative mood: rise for all conditions, no differences between conditions Change in positive mood: <ul style="list-style-type: none"> Decline for all conditions E: no differences between modalities NE: NV better recovery positive mood > A Time reading website: E and EE > NE Total number of thoughts recalled: <ul style="list-style-type: none"> Recall positive thoughts > recall negative thoughts Recall more positive thoughts for website with high credibility and more negative thoughts for website with low credibility E: NV > A Credibility websites: no effect modality or intervention Attitude towards agent: <ul style="list-style-type: none"> The more positive? the more empathic the agent was perceived, and the more positive mood change occurred before mood induction and after intervention Enjoyability, caring, trustworthiness, likeability: E and EE > NE EE: Empathy, trustworthiness, enjoyability, caring, likeability: A > NV E and NE: no differences: modalities A: Enjoyability, caring, likeability, trustworthiness: E and EE > NE NV: no differences: intervention Perceived empathy? EE < E and NE Perception agent design: no differences
Bickmore et al. (2009)	Speech and/or textual output	Person form stories (first person or third person)	<ul style="list-style-type: none"> For each completed conversation: <ul style="list-style-type: none"> [UX] Enjoyability stories [AC] Engagement [AC] Perceived dishonesty counselor [U] System usage 	<p>Two conditions:</p> <ul style="list-style-type: none"> 1st person (narratives as being part of the agent's own life story) 3rd person (narratives as being part from the life of a human friend of the agent with a similar role and occupation) 	<ul style="list-style-type: none"> Enjoyability stories: 1st > 3rd, decreased over time in both conditions Engagement: no difference, decreased over time in both conditions Perceived dishonesty: no difference, both agents not dishonest System usage: 1st > 3rd, both: decrease over time in both conditions
Kang & Gratch (2010)	Speech and/or textual output	Amount of self-disclosure (via providing backstories)	<ul style="list-style-type: none"> [AR] Co-presence [AR] Social attraction [B] User's self disclosure (verbally) 	<p>Three conditions:</p> <ul style="list-style-type: none"> High disclosure Low disclosure Non-disclosure 	<ul style="list-style-type: none"> High disclosure: more co-presence and social attraction than low-disclosure, non-disclosure High disclosure: people revealed a medium level of personal information more often than in low-disclosure, non-disclosure

Relational, empathic behaviour

Providing personal information

Source	Feature category	Design feature(s)	Outcome variable(s)	Method	Outcomes
Bickmore et al. (2010)	Speech and/or textual output	Variability content, person form stories (first person or third person)	<p>Experiment 1:</p> <ul style="list-style-type: none"> [UX] Perceived repetitiveness of the interaction [IT] Desire to continue [U] System usage [B] Number of steps <p>Experiment 2:</p> <ul style="list-style-type: none"> [AC] Self-report engagement [AC] Enjoyability stories [IT] Looking forward to talk [AC] Perceived dishonesty [U] System usage [B] Number of steps 	<p>Experiment 1, two conditions:</p> <ul style="list-style-type: none"> Non-variable (NV): same dialogue structure and language, appearance and setting never change Variable (V): one of five different dialogue structures randomly selected, topic randomly selected, one of five different background images randomly selected <p>Experiment 2, two conditions:</p> <ul style="list-style-type: none"> 1st person 3rd person 	<p>Experiment 1:</p> <ul style="list-style-type: none"> Perceived repetitiveness: NV > V Desire to continue: V > NV System usage: V > NV Number of steps: NV > V <p>Experiment 2:</p> <ul style="list-style-type: none"> Self-report engagement: no difference, decrease over time in both conditions Enjoyability stories: 1st > 3rd Looking forward: no difference Perceived dishonesty: no difference System usage: 1st > 3rd Number of steps: 3rd > 1st
Grillon & Thalmann (2008)	Facial and gaze expressions, hand and body gestures	Attitude (bored, attentive, random or changed based on eye contact user)	<ul style="list-style-type: none"> [AC] Interested/indifferent [AC] Engaged/distracted [AC] Behave normal /abnormal 	<p>Four attitude conditions:</p> <ul style="list-style-type: none"> Always bored (looks away) Always attentive (positive attitude) Randomly changed attitude (attentive and bored) Changed attitude on eye contact behaviour (user looks = attentive behaviour, user looks away = boring behaviour) 	<p>Character changing behavior to human eye contact behavior seems more normal than a character that randomly changes behavior:</p> <ul style="list-style-type: none"> Mean interested/indifference: bored: 5 / attentive: 1.67 / random: 3.50 / tracked: 3.50 / tracked (only people that looked): 2.20 / tracked (without people that did not look): 2.71 Mean engaged/distracted: bored: 4.92 / attentive: 1.92 / random: 3.58 / tracked: 3.67 / tracked (only people that looked): 2.80 / tracked (without people that did not look): 3.29 Mean normal/abnormal: bored: 3.83 / attentive: 2.50 / random: 3.33 / tracked: 3.17 / tracked (only people that looked): 2.80 / tracked (without people that did not look): 3.00 Interviews: tendency from to find the tracked version more realistic than either the bored version or the random version

Variable behaviour



Source	Feature category	Design feature(s)	Outcome variable(s)	Method	Outcomes
Bickmore & Ring (2010)	Speech and/or textual output, facial and gaze expressions	Degrees of control over the agent's verbal and nonverbal behavior	<ul style="list-style-type: none"> [AC] Helpfulness in creating a story [AC] Confidence that the agent would properly re-tell the story [O] Possibility to express oneself [UX] Ease of use [UX] Overall satisfaction 	<p>Five conditions:</p> <ul style="list-style-type: none"> TEXT (user enters story in a text box, story displayed as text) AGENT (user enters story in a text box, story is spoken by an ECA with non-verbal behaviour and prosody and neutral face) FACE (same as AGENT, but, in addition, the user can specify the agent's facial display (happy, sad or neutral)) VOICE (same as AGENT, but, in addition, the user can vary the agent's prosody: speed, volume, baseline pitch, emphasis and pauses) ALL (Combination of VOICE and FACE) 	<p>Users more satisfied when stories are delivered by a virtual agent compared to plain text, more satisfied with tools to control agent's prosody compared to facial display of emotion and most satisfied when they can control all aspects of the agent's delivery:</p> <ul style="list-style-type: none"> For all measures: TEXT rated lowest, ALL highest, and VOICE second highest Helpfulness: difference between TEXT and other conditions and AGENT and other conditions, but not between FACE, VOICE and ALL Satisfaction: difference between TEXT and FACE, TEXT and VOICE, and TEXT and ALL Confidence, possibility expression and ease of use: no differences between conditions
Frost et al. (2012)	Speech and/or textual output, facial and gaze expressions, hand and body gestures	Movement agent, interactivity	<ul style="list-style-type: none"> [LUC] Intrinsic motivation (IM) [LUC] Co-variates: parasocial interaction (PSI) and Self-efficacy (SE) 	<p>Three conditions:</p> <ul style="list-style-type: none"> Non-interactive virtual coach (NIVC); no compliments, looks away from viewer, directive instructions in impersonal manner and remains stationary Non-interactive moving virtual coach (NIMVC): same audio track as in NIVC one, moves along with viewer to model behavior Interactive virtual coach (IVC): compliments, looks directly at camera, personal and non-directive feedback cues, moves, models behavior 	<p>Parasocial interaction fully explains variation in intrinsic motivation, self-efficacy does not:</p> <ul style="list-style-type: none"> IM, PSI: higher in IVC than other conditions SE: higher in IVC and NIMVC compared to NIVC SE: NIMVC higher than NIVC Model IM: higher IVC compared to NIMC Model IM & SE: IM different for IVC and NIVC Model IM & SE & PSI: no difference IM for NIMVC and NIVC

Source	Feature category	Design feature(s)	Outcome variable(s)	Method	Outcomes
Olafsson et al. (2017)	Speech and/or textual output	Presence of rap music in the dialogue	<p>Self-reports:</p> <ul style="list-style-type: none"> [UX] Satisfaction [IT] Desire to continue working with agent [AC] Trust [AR] Closeness relationship [AC] Naturalness [AC] Knowledgeability [AC/UC] Perceived similarity [AC] Liking of the agent <p>Behavioural measures:</p> <ul style="list-style-type: none"> [O] Selection agent to have third conversation with [U] Number of turns of talk with agent in third conversation <p>Interview:</p> <ul style="list-style-type: none"> [OI] Concepts [AC] Opinion most engaging agent 	<p>Four conditions:</p> <ul style="list-style-type: none"> RN (rap nutrition) RE (rap exercise) NoRE (no rap exercise) NoRN (no rap nutrition) 	<p>Self-reports:</p> <ul style="list-style-type: none"> Satisfaction: no difference Desire to continue working: no difference Trust? NoR > R Closeness relationship: R > NoR Naturalness, knowledgeability, perceived similarity, liking of the agent: no difference <p>Behavioural measures:</p> <ul style="list-style-type: none"> Selection agent third conversation: R > NoR (62% vs 38%) Nr. of turns third conversation: no difference <p>Interview:</p> <ul style="list-style-type: none"> Concepts reliability and engagement: R > NoR, information and knowledge: NoR > R Most engaging agent: R > NoR (65% vs 35%)
Tielman et al. (2017)	Speech and/or textual output	Presentation mode psycho-education (ECA providing info verbally or by showing text)	<p>Primary measures:</p> <ul style="list-style-type: none"> [O] Suggestion adherence (stressfulness chosen memory) [O] Task adherence (quality of memory description) [O] Subjective adherence (subjective experience recollecting memory) <p>Mediating measures:</p> <ul style="list-style-type: none"> [B] Recollection of information [AC] Attitude to agent <p>Exploratory measures:</p> <ul style="list-style-type: none"> [UC] Psychoeducation attitude [UX] Function usefulness 	<p>Psychoeducation by:</p> <ul style="list-style-type: none"> A) a virtual agent that asks to read psychoeducation in text on the screen B) a virtual agent that provides the psychoeducation verbally <p>Then, participants perform an expressive writing task (describing one of the worst memories in a digital diary)</p>	<ul style="list-style-type: none"> Agent attitude, psychoeducation attitude, usefulness diary, usefulness psychoeducation: significantly higher than the neutral Subjective adherence, usefulness agent, usefulness of the virtual agent speaking: no significant difference from neutral One group considered agent and voice a useful addition, another group did not No effect presentation mode on agent attitude and recollection Suggestion adherence: positively correlated with subjective and task adherence Subjective and task adherence: not correlated Agent attitude and psychoeducation attitude: positively correlated Correlations recollection and task adherence, and agent attitude and task adherence No effect recollection and agent attitude on suggestion adherence and subjective adherence Task adherence: text > VA

Agent message

Rap



Source	Feature category	Design feature(s)	Outcome variable(s)	Method	Outcomes
Yin et al. (2010)	Speech and/or textual output, socio-demographic	Presence or absence of linguistic and cultural tailoring	<ul style="list-style-type: none"> Independent variables: <ul style="list-style-type: none"> [AC/UC] Linguistic congruity [AC/UC] Cultural congruity [UC] Need for cognition Dependent variables: <ul style="list-style-type: none"> [B] Persuasion; differences in pre- and post ranking task consisting of a list of statements about exercise agent [AC] Perceived trustworthiness 	<ul style="list-style-type: none"> 2 x 2 experiment: <ul style="list-style-type: none"> Spanish (Latina) English (Anglo-American) Low cultural congruity High cultural congruity Agent designs: adapted to culture in appearance, behaviour and values and background Anglo-American: focus on user's well-being, Latina: interest in family and friends 	<ul style="list-style-type: none"> Persuasion: no effects of linguistic congruity Persuasion, trustworthiness: interaction effect of cultural congruity and need-for-cognition Need for cognition: <ul style="list-style-type: none"> People predisposed to be influenced by peripheral cues (low need for cognition)? Trust and persuasion: own > different culture People predisposed to process information deeply (high need for cognition)? Persuasion: different > own culture
Looks					
Forlizzi et al. (2007)	Soci-demographics, species	Gender, species (human or abstract)	<ul style="list-style-type: none"> [AC] Likability agent per task (athletic trainer, medical advisor) 	Rating multiple static images differing in, amongst others, gender and species	<ul style="list-style-type: none"> Athletic trainer: male > female Medical advisor: female > male Both: humans (females and males) > abstract
Ring et al. (2014)	Realism	Shading styles (realistic, cartoon shaded), character proportions (of a human, cartoon)	<ul style="list-style-type: none"> [AC] Realism [AC] Appeal [AC] Familiarity [AC] Feriness [AC] Friendliness [AC] Trustworthiness [UX] Easiness to interact with [IT] Desire to continue working [AC] Likability [AC] Caring [AC] Appropriateness [AC] Quality of motion 	Experiment 1, four conditions: <ul style="list-style-type: none"> Shaded agent - social task (user's favorite books and movies) Realistic agent - social task Shaded agent - medical task (cancer related topics) Realistic agent - medical task Experiment 2, four conditions: <ul style="list-style-type: none"> Proportions human - social task Proportions cartoon - social task Proportions human - medical task Proportions cartoon - medical task 	Experiment 1: <ul style="list-style-type: none"> Familiarity, trustworthiness, appropriateness: medical > social Friendliness: social > medical, shaded > realistic Likability: shaded > realistic (only social task) Caring: trend toward significance: medical task: realistic > shaded, social task: shaded > realistic Experiment 2: <ul style="list-style-type: none"> Familiarity, trustworthiness, appropriateness, caring: medical > social Friendliness: cartoon > human Appropriateness: medical task: human > cartoon Appropriateness social task: cartoon > human

Source	Feature category	Design feature(s)	Outcome variable(s)	Method	Outcomes
Robertson et al. (2015)	Realism	Amount of realism (photographic, highly stylized renderings to abstract)	<p>Early focus groups:</p> <ul style="list-style-type: none"> [AC] Opinion on four rendering styles of static agent images (photograph, vectorized photo, realistic illustration and cartoon) <p>Final focus groups:</p> <ul style="list-style-type: none"> [AC] Satisfaction, feasibility, acceptability, agent as patient decision aid and general opinion on agent prototype (cartoon style) with interaction 	<p>Early focus groups:</p> <ul style="list-style-type: none"> Four rendering styles presented <p>Final focus groups:</p> <ul style="list-style-type: none"> Agent prototype in cartoon style 	<p>Early focus group Emory: negative reaction towards cartoon (unappealing)</p> <p>Early focus group Grady Memorial Hospital:</p> <ul style="list-style-type: none"> Cartoon: dislike, but acceptable Vectorized photo: looked "fake" Photographic agent: "good" and "looked cool" Preference: cartoon > photographic <p>Final focus groups:</p> <ul style="list-style-type: none"> Grady: cartoon is "OK" Emory: cartoon disliked, prefer "real person" Satisfaction: Grady > Emory Feasibility and acceptability: Grady < Emory Agent as a patient decision aid: low rating
Wissen et al. (2016)	Socio-demographic, realism	Similarity (age: young or old), Familiarity (familiar: a representation of an actual female health coach/nurse, rendering (photo-realistic, stylized)	<ul style="list-style-type: none"> User acceptance: [AC] Friendliness [AC] Competence [AC] Attractiveness [AC] Likeability [AC] Trustworthiness [IT] Following advice [IT] Continue interaction [AC] Satisfaction Attitude towards use: [UX] Easy to use [UX] Liking working with management [UX] Useful to support health goal [UX] Enable health goal achievement [UX] Easier to manage health [UX] Useful [UX] Make life more interesting [UX] Help when necessary 	<p>3 x 2 design:</p> <ul style="list-style-type: none"> Appearance: 3 between-subject conditions: older (unfamiliar) vs. younger (unfamiliar) vs. familiar Rendering style: 2 within-subject conditions: stylized vs. photo-realistic 	<p>User acceptance:</p> <ul style="list-style-type: none"> All outcome variables: photo-realistic > stylized Friendliness: older > familiar All other outcome variables: no difference between older, younger and familiar <p>Attitudes towards use of VC:</p> <ul style="list-style-type: none"> No effects of rendering style Four of seven variables: effect of appearance Easy to use, like working with and enable health goal achievement: younger > familiar Easy to use, like working with and enable health goal achievement: no differences between younger and older, and older and familiar Easier to manage health and useful: older > familiar

Rendering style

Source	Feature category	Design feature(s)	Outcome variable(s)	Method	Outcomes
Parmar et al. (2018)	Styling	Agent's attire, design virtual environment	<p>Experiment 1:</p> <ul style="list-style-type: none"> [AC] Professionalism <p>Experiment 2:</p> <ul style="list-style-type: none"> [AC] Professionalism [AC] Credibility information [AC] Trustworthiness [AC] Reassurance [AC] Caring [AC] Friendliness [AR] Comfortability to talk to [AR] Perception face-to-face interaction [IT] Willingness to continue [B] Commitment to choose a healthcare proxy 	<p>Experiment 1:</p> <ul style="list-style-type: none"> Pre-ranking 12 agents with different outfits and offices on professionalism Agents differing in 1) Attire: A) dressed to fit health care (professional) / B) not role-appropriate (casual) and 2) Environment: A) professional setting (medical examination room) and B) no setting (empty room) <p>Experiment 2, four conditions (attire and environment):</p> <ul style="list-style-type: none"> CASUAL EMPTY CASUAL FULL PROF-EMPTY PROF-FULL 	<p>Experiment 1:</p> <ul style="list-style-type: none"> Agent's attire: lab coat and stethoscope most professional, casual clothing without a stethoscope least professional Environment: office setting most professional, empty room least professional <p>Experiment 2:</p> <ul style="list-style-type: none"> Professionalism: white coat > casual and full office setting > empty room Credibility, trustworthiness, reassurance, caring, friendliness, comfortability to talk to, perception face-to-face conversation, willingness to continue: white coat > casual Commitment to choosing a healthcare proxy: white coat > casual Reassurance: full office > empty room Comfortability to talk to: white coat in empty room > casual in empty room
Clothing					

Source	Feature category	Design feature(s)	Outcome variable(s)	Method	Outcomes
Nguyen & Masthoff (2007)	Socio-demographic, styling	Profession	<p>Experiment 1:</p> <ul style="list-style-type: none"> [AC] Perceived age, gender, profession, attractiveness, trustworthiness, expertise [AC] Preferred agent for learning about fitness programs and health benefits of exercise <p>Experiment 2:</p> <ul style="list-style-type: none"> [O] Credibility messages 	<p>Experiment 1:</p> <ul style="list-style-type: none"> Opinion on 16 images varying in age, gender, profession (doctors and sport instructors) and estimated attractiveness, trustworthiness and expertise <p>Experiment 2 (3 x 2 experiment):</p> <ul style="list-style-type: none"> Credibility agent (based on experiment 1): <ul style="list-style-type: none"> High credibility image Low credibility image No image Content message: <ul style="list-style-type: none"> Fitness programs Health benefits of exercise 	<p>Experiment 1:</p> <ul style="list-style-type: none"> Expertise health benefits of exercise: doctors > sport instructors Expertise fitness programs: sport instructors > doctors Both doctors and sport instructors perceived as experts in fitness programs and health benefits Preference fitness program: sport instructor (40) > doctor (9) Preference health benefits of exercise: sport instructor (24) > doctor (23) Correlation expertise and attractiveness, and preferred agent for fitness programs, but not for health benefits (participants like to learn about fitness programs from an attractive looking expert) Per topic, image most often selected? highest credibility (combination of attractiveness, trustworthiness and expertise) Reasons choosing agent? healthy (but not overly fit) appearance, friendliness and expertise No correlation gender and age participant and gender and age preferred agent <p>Experiment 2:</p> <ul style="list-style-type: none"> Credibility messages: highly credible image > lowly credible image Credibility and trust/expertise: no image very similar to highly credible image
Schmeil & Suggs (2014)	Styling	Representation health status: body shape, body animations and poses and skin texture	<ul style="list-style-type: none"> [B] Comparison pre-test and post-test: motivation to engage in more physical activity, intention to engage in more physical activity, motivation to engage in healthier diet, intention to engage in a healthier diet [IT] Motivation and intention to access the tool again 	<p>Four conditions:</p> <ul style="list-style-type: none"> (1) Agent tailored to current health status: a non-animated character (2) Two agents: one tailored to current health status; other shows possible future health status (both still poses) (3) Agent tailored to and models current health behavior (animated) (4) Two agents: one tailored to current health status and current behaviour, other shows possible future changes in health status (both animated) 	<ul style="list-style-type: none"> Motivation to engage in more physical activity: positive change for all conditions, largest in (4) Intention to engage in more physical activity: only change for (2) and (4), largest change in (2) Motivation to engage in healthier diet: positive change for (4) Intention to engage in healthier diet: positive change all conditions, largest in (4) Motivation to revisit the tool: no differences between conditions

Body shape

Body shape

Source	Feature category	Design feature(s)	Outcome variable(s)	Method	Outcomes
Skalski et al. (2007)	Styling	Interactivity (interactive or not), source attractiveness (attractive or unattractive)	<p>Manipulation check:</p> <ul style="list-style-type: none"> [AC] Attitude toward agent: source credibility (knowledgeable, intelligence, competence), likeable, warm, pleasing, friendly, attractiveness (appealing, attractive, nice looking) [AR] Perceived interactivity <p>Parameters in causal models:</p> <ul style="list-style-type: none"> Interactivity Perceived interactivity Social presence Source thoughts Message processing Attitude toward blood pressure <p>[B] Behavioural intention: sign up for actual appointment to test blood pressure</p>	<p>Four conditions:</p> <ul style="list-style-type: none"> Interactive (I) (ability to talk to source, control order in which message elements are presented), attractive (A) (pleasant-looking woman) Interactive (I), unattractive (UA) (generated by computer programs) Not interactive (NI) (source discussed categories one by one, no ability to select order), attractive (A) Not interactive (NI), unattractive (UA) 	<ul style="list-style-type: none"> Variable check: <ul style="list-style-type: none"> Attractiveness: A > UA Rest attitude variables: no difference A and UA Perceived interactivity: not high in both conditions, but I > NI Attractive source model: interactivity induction (+)-> perceived interactivity (+)-> social presence (+)-> positive source thoughts / message processing -> (+) attitude (all, but positive source thoughts were significant paths) Unattractive source model: interactivity induction (+) -> perceived interactivity (+) -> social presence -> (-) negative source thoughts / (+) message processing -> (- and +) attitude (all, but message processing were significant paths) Combined model: interactivity induction (+) -> perceived interactivity (+) -> social presence (+) -> positive source thoughts / message processing -> attitude (all, but path from source thoughts to attitude significant) Behavioral intention: 42% signed up for test
van Vugt et al. (2006)	Styling	<p>Body shape:</p> <ul style="list-style-type: none"> Similarity between user and interface character ideal (slim) and non-ideal body shapes (fat) 	<p>Dependent variables:</p> <ul style="list-style-type: none"> [AR] User involvement [AR] User distance [IT] Intention to use the agent <p>Co-variables:</p> <ul style="list-style-type: none"> [AC/UC] Similarity [AC] Valence [AC] Affordance [AC] Aesthetics [AC] Perceived realism [AC] Perceived ethics (trustworthiness and credibility) 	<p>A 2 (similarity: similar versus dissimilar body shapes) x 2 (valence: positive (slim) versus negative (fat)) factorial design</p>	<ul style="list-style-type: none"> Similarity, valence, perceived affordance and aesthetics: no effect on dependents Perceived ethics, perceived realism: effect on dependents Two-way interaction effects of similarity and ethics, and of valence and ethics on dependents Effect perceived realism on user involvement and user distance: the more agent perceived as realistic, the higher user involvement and the lower user distance Effect perceived ethics on user involvement and intention to use: the more agent perceived as ethically good, the higher user involvement and intention to use Effect perceived ethics on user distance and intention to use: similar > dissimilar Effect perceived ethics on user involvement: fat > slim Realism: fat > slim

Source	Feature category	Design feature(s)	Outcome variable(s)	Method	Outcomes
Gender	Alsharbi & Richards (2017)	Age, gender, ethnicity, role	<ul style="list-style-type: none"> [AC] Pre-test: Preference age, gender, ethnicity and role [AC] Ranking task: Preferred agent out of 12 agents [AC/AR] Post-test: rapport preferred agent 	<ul style="list-style-type: none"> Pre-test: opinion on age (younger, older or same age), gender (female, male or no preference), ethnicity (same, other, no preference), role (teacher, doctor, parent or friend), looks and environmental context Ranking task: ranking 12 agents differing in gender (male or female), age (young or old) and ethnicity (Asia, white, Mediterranean) Post-test: rapport preferred agent 	<ul style="list-style-type: none"> Frequencies pre-test: age (0, 4, 7), gender (6, 1, 4), ethnicity (8, 1, 2), role (majority: friend) Ranking task: majority preferred young female Mediterranean, majority preferred the old male Asian the least. Post-test: agent was liked, agent was weird, experience was comfortable, interacting with character does not help to solve a problem
	Forlizzi et al. (2007): See rendering style				
	Malhotra et al. (2016)	Socio-demographic	Gender	<p>Comparison of levels of emotional dimensions of respondents with the known levels in literature. The emotional dimensions:</p> <ul style="list-style-type: none"> [AC] Evaluation (valence) [AC] Potency (power/dominance) [AC] Activity/arousal 	<ul style="list-style-type: none"> Six audio-visual prompts (different steps of handwashing task) with five different emotional deliveries by a male and female character
Age	Alsharbi & Richards (2017): See rendering style				
	Wissen et al. (2016): See rendering style				
Cultural background	Alsharbi & Richards (2017): See rendering style				
	Yin et al. (2010): See speech and/or textual output, facial and gaze expressions and hand and body gestures				
	Zhou et al. (2014)	Socio-demographic	Racial concordance, perceived similarity	<ul style="list-style-type: none"> [AC/UC] Similarity [AR] Working alliance [AC] Satisfaction [UX] Easiness to talk [IT] Desire to continue [AC] Likability [AC] Trust [AC] Closeness relationship [AC] Caring [IT] Likelihood of following advice [O] Pref. talking to doctor or agent 	<p>Four conditions:</p> <ul style="list-style-type: none"> A) Black agent, multiple interactions B) Black agent, single interaction C) White agent, multiple interactions D) White agent single interaction

Source	Feature category	Design feature(s)	Outcome variable(s)	Method	Outcomes
Zhou et al. (2017)	Socio-demographic	Cultural adaptation (Chinese vs American agent)	<p>Before and after interaction:</p> <ul style="list-style-type: none"> [B] Persuasion: attitude towards regular exercise [B] Decisional balance: one's relative weighing of the pros and cons of exercise <p>After interaction:</p> <ul style="list-style-type: none"> [LUX] Satisfaction [AC] Attitudes toward agent 	<ul style="list-style-type: none"> The agent designs: <ul style="list-style-type: none"> English-speaking, animated American (A): US living room, more hand gestures Chinese (C): room with interior design elements traditional Chinese culture Cultural dimensions dialogues: <ul style="list-style-type: none"> (1) power distance, A: recommended, C: use of should/have (2) individualism vs collectivism, A: potential effect for individual, C: potential effect for family (3) uncertainty avoidance, A: step-by-step advice, C: year-long plan (4) long-term vs short-term orientation, A: focus on fun side of exercise, C: importance of persistence and long-term benefits (5) indulgence vs restraint Small talk: A: local sports teams, basketball, C: Chinese Lunar New Year, traditions, trending TV shows Exercises suggested: A: hiking, cycling, weightlifting, simple exercises, C) badminton, jogging, aerobic exercises 	<ul style="list-style-type: none"> Attitudinal change towards exercise: both positive, but American > Chinese Change decisional balance: no differences (both a significant increase) No significant interaction between condition and participants' acculturation level on either persuasion or decisional balance. Both conditions: satisfied with agent, liked agent, trusted agent, agent easy to use Feeling of how interesting agent is: near significant, Chinese > American No other differences
Cultural background					
Role	Alsharbi & Richards (2017): <i>See rendering style</i> Nguyen & Masthoff (2009): <i>See relational, empathic behaviour</i>				

Outcome variables: U = usage / IT = intention to (continue) using / B = (intention towards) behaviour change / UX = usability and user experience / AC = agent characteristics (e.g. demographics, personality, styling) / AR = relation with agent / UC = user characteristics / O = other

APPENDIX B – QUESTIONNAIRE CH3

Beste deelnemer,

Bij Roessingh Research and Development doen we onderzoek naar het inzetten van virtuele karakters voor coaching op het gebied van gezond leven. In het kader van dit onderzoek zijn wij geïnteresseerd in uw mening over het ontwerp van verschillende coaches. U krijgt zo een aantal afbeeldingen van verschillende coaches te zien, waarna u wordt gevraagd om deze te beoordelen. Uw mening zal ons helpen bij het ontwikkelen van gepersonaliseerde coaches. Het invullen van de vragenlijst duurt ongeveer 15 minuten.

Heeft u vragen over deze studie? Dan kunt u contact opnemen met:

[REMOVED]

Door de vragenlijst te starten gaat u akkoord met het gebruik van uw gegevens voor onderzoeksdoeleinden. Uw data worden anoniem opgeslagen. Op onderzoek bij de RRD zijn de privacyregels van de Europese Commissie van toepassing zoals die gelden voor wetenschappelijk onderzoek. Ons privacy reglement kunt u vinden op onze website.

Allereerst willen we u een aantal vragen over uzelf stellen.

1. Wat is uw geslacht?

- Man
- Vrouw
- Anders

2. Wat is uw geboortedatum? (dd mm jjjj)

3. Wat is uw hoogst afgeronde opleiding? (Als uw opleiding er niet bij staat, kruis dan de opleiding aan die het meest op de door u gevolgde opleiding lijkt)

- Basisonderwijs
- Middelbaar onderwijs (MAVO, VMBO, HAVO, VWO)
- Middelbaar beroepsonderwijs (MBO)
- Hoger beroepsonderwijs (HBO)
- Wetenschappelijk onderwijs (WO)

4. Wat is uw woonsituatie?

- Ik woon samen met mijn vriend(in)/partner
- Ik woon samen met een vriend, familielid of iemand anders

- > Ik woon alleen
- > Anders, namelijk:

5. Welke apparaten gebruikt u thuis?

- > Smartphone
- > Smartwatch
- > Tablet
- > PC / laptop
- > Spelcomputer
- > Anders, namelijk:

6. Hoeveel uur spendeert u per dag aan onderstaande activiteiten?

(U mag hierbij werktijd meenemen)

	0u	0-1u	1-2u	2-3u	3-4u	>4u
Zoekopdrachten op internet	0	0	0	0	0	0
Offline gebruik van mijn computer	0	0	0	0	0	0
E-mailen	0	0	0	0	0	0
Online chatten / social media	0	0	0	0	0	0
Telefoneren	0	0	0	0	0	0
Teksberichten sturen / whatsappen	0	0	0	0	0	0
Videospellen spelen	0	0	0	0	0	0
Naar de radio luisteren	0	0	0	0	0	0
Tv kijken	0	0	0	0	0	0
Boeken en tijdschriften lezen	0	0	0	0	0	0
Met iemand spreken (face-to-face)	0	0	0	0	0	0

7. Vindt u dat u op dit moment een gezond eetpatroon heeft?

- > Ja, ik heb al langer dan zes maanden een gezond eetpatroon.
- > Ja, ik heb sinds de laatste zes maanden een gezond eetpatroon.
- > Nee, maar ik ben wel van plan om binnen dertig dagen gezonder te gaan eten.
- > Nee, maar ik ben wel van plan binnen zes maanden gezonder te gaan eten.
- > Nee, en ik ben ook niet van plan om binnen zes maanden gezonder te gaan eten.

In de volgende vraag wordt de term “regelmatig lichamelijk actief” gebruikt, hiermee wordt bedoeld: tenminste 5 dagen in de week minimaal een half uur per dag op een middelmatig of zwaarder niveau lichamelijk actief zijn. U kunt daarbij bijvoorbeeld denken aan activiteiten zoals zwemmen, stofzuigen, dansen, in de tuin werken, fietsen, wandelen, fitness of andere soortgelijke activiteiten. Dat half uur per dag hoeft niet aaneengesloten te zijn, maar kan ook in tijdsblokkjes (van minimaal 5 minuten) over de dag verdeeld zijn.

8. Welke van de onderstaande vijf uitspraken past het best bij u?

(Selecteer de uitspraak die het beste bij u past)

- Op dit moment ben ik niet lichamelijk actief en ik ben niet van plan lichamelijker actiever te worden in de komende 6 maanden.
- Op dit moment ben ik niet lichamelijk actief, maar ik denk er over om lichamelijk actiever te worden in de komende 6 maanden.
- Op dit moment ben ik lichamelijk actief, maar niet regelmatig.
- Op dit moment ben ik regelmatig lichamelijk actief, maar ben ik daar pas in de afgelopen 6 maanden mee begonnen.
- Op dit moment ben ik regelmatig lichamelijk actief en ben dat al langer dan 6 maanden.

9. Hoe vaak ervaart u problemen bij het begrijpen van teksten (denk aan folders) over uw gezondheid of een ziekte?

- Nooit
- Zelden
- Soms
- Vaak
- Altijd

10. Hoe zelfverzekerd voelt u zich wanneer u medische formulieren invult?

- Helemaal niet zelfverzekerd
- Een beetje zelfverzekerd
- Redelijk zelfverzekerd
- Zelfverzekerd
- Heel zelfverzekerd

11. Hoe vaak helpt iemand u bij het lezen van brochures, formulieren of brieven van het ziekenhuis, de apotheek of uw huisarts?

- Nooit
- Zelden
- Soms
- Vaak
- Altijd

U krijgt zo een aantal afbeeldingen van virtuele karakters te zien. U wordt gevraagd naar uw voorkeur voor een van deze karakters die als uw coach u zou kunnen gaan helpen bij het verkrijgen van een gezonde levensstijl. Bekijkt u rustig alle afbeeldingen.

12. Stel, u heeft behoefte aan een coach voor gezond leven, naar welke karakter gaat dan uw voorkeur uit?

[SET OF IMAGES OF ALL AGENT DESIGNS, PRESENTED IN RANDOM ORDER]

13. Waarom gaat uw voorkeur uit naar dit karakter?

U krijgt nu alle karakters een voor een te zien. Bekijkt u nogmaals de afbeelding van het karakter. U wordt vervolgens gevraagd om een aantal eigenschappen van het karakter te beoordelen.

[ALL AGENT DESIGNS PRESENTED INDIVIDUALLY, IN RANDOM ORDER]

14. Hoe waarschijnlijk is het dat u advies van dit karakter zal opvolgen?

Onwaarschijnlijk								Waarschijnlijk
	0	0	0	0	0	0	0	

15. Ik vind het karakter:

	Ze er oneens	Oneens	Enigszins oneens	Neutraal	Enigszins eens	Eens	Ze er eens
Vriendelijk	0	0	0	0	0	0	0
Deskundig	0	0	0	0	0	0	0
Betrouwbaar	0	0	0	0	0	0	0
Autoritair	0	0	0	0	0	0	0
Betrokken	0	0	0	0	0	0	0

Dit is het einde van de vragenlijst. Wanneer u nog opmerkingen heeft kunt u deze hieronder kwijt.

APPENDIX C – POST-CONDITION QUESTIONNAIRE CH4

Beste deelnemer,

Bij Roessingh Research and Development doen we onderzoek naar het invullen van kwetsbaarheidsvragenlijsten in dialoog met een virtuele coach. Zo hopen we het invullen leuker te maken en kwetsbaarheid vroegtijdig te kunnen opsporen. Uiteindelijk is het van belang dat de gebruikers het prettig vinden om vragenlijsten af te nemen met de virtuele coach. Daarom evalueren we de huidige applicatie met toekomstige gebruikers.

Om de applicatie goed te kunnen laten aansluiten bij de toekomstige gebruikers willen we u graag een aantal vragen stellen over de virtuele coach en uw ervaring met het systeem.

Alvast bedankt voor uw medewerking!

Namens de onderzoekers Roessingh Research and Development

U heeft zojuist geïnteracteed met uw virtuele coach, [Egbert/Sylvia] (zie afbeelding hieronder). We vragen u om een aantal eigenschappen van uw virtuele coach te beoordelen.

[IMAGE EGBERT OR SYLVIA]

1. Hoe waarschijnlijk is het dat u advies van uw virtuele coach zal opvolgen?

Selecteer 1 antwoordmogelijkheid.

Onwaarschijnlijk				Waarschijnlijk			
0	0	0	0	0	0	0	0

2. Ik vind mijn virtuele coach:

Selecteer 1 antwoordmogelijkheid per eigenschap.

	Zeer oneens	Oneens	Enigszins oneens	Neutraal	Enigszins eens	Eens	Zeer eens
Vriendelijk	0	0	0	0	0	0	0
Deskundig	0	0	0	0	0	0	0
Betrouwbaar	0	0	0	0	0	0	0
Autoritair	0	0	0	0	0	0	0
Betrokken	0	0	0	0	0	0	0



3. Ik had het gevoel een klik te hebben met de virtuele coach:

Selecteer 1 antwoordmogelijkheid

Zeer oneens	Oneens	Enigszins oneens	Neutraal	Enigszins eens	Eens	Zeer eens
0	0	0	0	0	0	0

4. Ik denk dat de virtuele coach en ik elkaar begrepen:

Selecteer 1 antwoordmogelijkheid

Zeer oneens	Oneens	Enigszins oneens	Neutraal	Enigszins eens	Eens	Zeer eens
0	0	0	0	0	0	0

5. De virtuele coach leek bereid te zijn mij te helpen:

Selecteer 1 antwoordmogelijkheid.

Zeer oneens	Oneens	Enigszins oneens	Neutraal	Enigszins eens	Eens	Zeer eens
0	0	0	0	0	0	0

6. De virtuele coach leek betrouwbaar:

Selecteer 1 antwoordmogelijkheid.

Zeer oneens	Oneens	Enigszins oneens	Neutraal	Enigszins eens	Eens	Zeer eens
0	0	0	0	0	0	0

7. De virtuele coach leek sympathiek:

Selecteer 1 antwoordmogelijkheid.

Zeer oneens	Oneens	Enigszins oneens	Neutraal	Enigszins eens	Eens	Zeer eens
0	0	0	0	0	0	0

8. Mijn gesprek met de virtuele coach voelde natuurlijk aan:

Selecteer 1 antwoordmogelijkheid.

Zeer oneens	Oneens	Enigszins oneens	Neutraal	Enigszins eens	Eens	Zeer eens
0	0	0	0	0	0	0

9. Ik vond de interactie met de virtuele coach gemakkelijk:

Selecteer 1 antwoordmogelijkheid.

Zeer oneens	Oneens	Enigszins oneens	Neutraal	Enigszins eens	Eens	Zeer eens
0	0	0	0	0	0	0

10. De virtuele coach was menselijk:

Selecteer 1 antwoordmogelijkheid.

Zeer oneens	Oneens	Enigszins oneens	Neutraal	Enigszins eens	Eens	Zeer eens
0	0	0	0	0	0	0

11. De virtuele coach was overtuigend:

Selecteer 1 antwoordmogelijkheid.

Zeer oneens	Oneens	Enigszins oneens	Neutraal	Enigszins eens	Eens	Zeer eens
0	0	0	0	0	0	0

12. Ik zou de virtuele coach aanbevelen aan anderen:

Selecteer 1 antwoordmogelijkheid.

Zeer oneens	Oneens	Enigszins oneens	Neutraal	Enigszins eens	Eens	Zeer eens
0	0	0	0	0	0	0

U heeft zojuist gewerkt met een systeem waarin u vragenlijsten met uw virtuele coach heeft afgenomen. We vragen u nu om het systeem te beoordelen.

14. Over het algemeen, hoe moeilijk of makkelijk was het om de vragenlijsten met het systeem te voltooien?

Selecteer 1 antwoordmogelijkheid.

Heel moeilijk			Heel makkelijk			
0	0	0	0	0	0	0

15. Ik zou dit systeem opnieuw willen gebruiken voor het invullen van gezondheidsvragenlijsten:

Selecteer 1 antwoordmogelijkheid.

Zeer oneens	Oneens	Enigszins oneens	Neutraal	Enigszins eens	Eens	Zeer eens
0	0	0	0	0	0	0

Dit is het einde van de vragenlijst.



APPENDIX D – QUESTIONNAIRE CH5

Beste deelnemer,

Bij Roessingh Research and Development doen we onderzoek naar het inzetten van virtuele karakters voor coaching op het gebied van gezond leven. In het kader van dit onderzoek zijn wij geïnteresseerd in hoe u een interactie met een virtuele coach ervaart. U zult straks een simpel dialoog houden, waarna u gevraagd wordt uw mening te geven over de coach. Uw mening zal ons helpen bij het ontwikkelen van gepersonaliseerde coaches. Het invullen van de vragenlijst duurt ongeveer 15 minuten.

Als u de vragenlijst op een mobiel apparaat invult, wordt u geadviseerd om voor het eerste deel van de vragenlijst het apparaat horizontaal te houden. Dit om de dialoog met de virtuele coach goed leesbaar te maken. Heeft u vragen over deze studie? Dan kunt u contact opnemen met: [REMOVED]

Door de vragenlijst te starten gaat u akkoord met het gebruik van uw gegevens voor onderzoeksdoeleinden. Uw data worden anoniem opgeslagen. Op onderzoek bij de RRD zijn de privacyregels van de Europese Commissie van toepassing zoals die gelden voor wetenschappelijk onderzoek. Ons privacy reglement kunt u vinden op onze website.

Questions in Case of Neutral Textual Expression

1. Hallo, mijn naam is Laura en ik ben een virtuele gezondheidscoach. Heb je al eens met een virtuele coach gewerkt? (>2a&2b)

2a. Ja. (>3a)

2b. Nee. (>3b)

3a. Ik hoop dat je door deze voorgaande ervaring een idee hebt van wie en wat ik ben. (>4)

3b. Een virtuele coach is een digitaal persoon die door middel van gesprekken met de gebruiker ondersteuning kan bieden binnen het behalen van doelen. (>4)

4. Waar gaan we het vandaag over hebben? (>5)

5. We gaan het over je gezondheid hebben. Als je vragen hebt over je gezondheid kun je die aan mij stellen. (>6a&6b)

6a. Ik heb een vraag over voeding. Hoe kan ik ervoor zorgen dat ik genoeg vitamines binnenkrijg? (>7a)

6b. Ik heb een vraag over beweging. Hoe weet ik dat ik genoeg beweeg op een dag? (>7b)

7a. Probeer genoeg fruit en groente te eten. (>8a&8b)

7b. Probeer 30 minuten per dag te bewegen. (>8c&8d)

8a. Zijn er ook dingen die ik niet zou moeten eten? (>9a)

8b. Dit was alles wat ik wilde weten. Bedankt voor de tip. (>9b)

8c. Maakt het uit wat voor soort beweging het is? (>9c)

8d. Dit was alles wat ik wilde weten. Bedankt voor de tip. (>9d)

- 9a. Probeer te minderen met vet eten, cafeïne en suiker. (end)
- 9b. Geen probleem. (end)
- 9c. Nee, zo lang het je hartslag maar verhoogt. (end)
- 9d. Geen probleem. (end)

Questions in Case of Happy Textual Expression

1. Hallo, mijn naam is Laura en ik ben een virtuele gezondheidscoach, leuk je te ontmoeten! Ik bied mensen online ondersteuning op het gebied van voeding en beweging. Heb je al eens eerder met een virtuele coach gewerkt? (>2a&2b)

2a. Ja. (>3a)

2b. Nee. (>3b)

3a. Dat is leuk om te horen, ik hoop dat je door deze eerdere ervaring al een beetje een idee hebt van wie en wat ik ben. Zo niet, dan is dat helemaal niet erg! Na dit gesprek zal dat veel duidelijker zijn. (>4)

3b. Een virtuele coach is een digitaal persoon die door middel van gesprekken met de gebruiker ondersteuning kan bieden bij het behalen van doelen. Gefeliciteerd met je eerste online coaching gesprek! Ik hoop dat je nu een beter beeld hebt van wie en wat ik ben. (>5)

4. Waar gaan we het vandaag over hebben? (>5)

5. Vandaag gaan we het over je gezondheid hebben. Ik kan me voorstellen dat je graag een gezonde levensstijl wilt hebben. Daarom ben ik hier om al je vragen rondom gezond eten en bewegen te beantwoorden. Ik help je graag! (>6a&6b)

6a. Ik heb een vraag over voeding. Hoe kan ik ervoor zorgen dat ik genoeg vitamines binnenkrijg? (>7a)

6b. Ik heb een vraag over beweging. Hoe weet ik dat ik genoeg beweeg op een dag? (>7b)

7a. Dat is een goede vraag. Om genoeg vitamines binnen te krijgen is het belangrijk om genoeg groente en fruit te eten, en natuurlijk om gevarieerd te eten. Beantwoordt dit je vraag? (>8a&8b)

7b. Dat is een goede vraag. De richtlijn is om ieder dag 30 minuten te bewegen. Als je op de fiets naar je werk gaat dan haal je dit misschien al! Beantwoordt dit je vraag? (>8c&8d)

8a. Nog niet helemaal. Zijn er ook dingen die ik niet zou moeten eten? (>9a)

8b. Dit was alles wat ik wilde weten. Bedankt voor de tip. (>9b)

8c. Nog niet helemaal. Maakt het uit wat voor soort beweging het is? (>9c)

8d. Dit was alles wat ik wilde weten. Bedankt voor de tip. (>9d)

9a. Het klinkt logisch, maar probeer niet te veel vet eten, cafeïne en suiker binnen te krijgen op een dag. Ik weet dat dit lastig kan zijn maar je zult zien dat ook jij het kunt. Veel succes met het realiseren van je gezondheidsdoelen! (end)

9b. Mocht je in de toekomst nog vragen hebben dan kun je die aan mij stellen. Veel succes met het realiseren van je gezondheidsdoelen! (end)

9c. Nee dat maakt niet uit, zo lang het je hartslag maar verhoogt. Als je er op gaat letten zal je zien dat je zomaar op die 30 minuten zit. Je kunt het! Veel succes met het realiseren van je gezondheidsdoelen. (end)

9d. Mocht je in de toekomst nog vragen hebben dan kun je die aan mij stellen. Veel succes met het realiseren van je gezondheidsdoelen! (end)

-
1. U mag nu reageren op enkele statements die betrekking hebben op de virtuele coach. Geef aan in hoeverre u het met de statements eens bent.

	Zeer oneens	Oneens	Enigszins oneens	Neutraal	Enigszins eens	Eens	Zeer eens
Ik had het gevoel een klik te hebben met de virtuele coach.	0	0	0	0	0	0	0
Ik denk dat de virtuele coach en ik elkaar begrepen.	0	0	0	0	0	0	0
De virtuele coach leek bereid zijn mij te helpen.	0	0	0	0	0	0	0
De virtuele coach leek betrouwbaar.	0	0	0	0	0	0	0
De virtuele coach leek sympathiek.	0	0	0	0	0	0	0

2. Geef aan in hoeverre u het eens bent met deze statements die betrekking hebben op de virtuele coach.

	Zeer oneens	Oneens	Enigszins oneens	Neutraal	Enigszins eens	Eens	Zeer eens
Mijn gesprek met de virtuele coach voelde natuurlijk aan.	0	0	0	0	0	0	0
Ik vond mijn interactie met de virtuele coach gemakkelijk.	0	0	0	0	0	0	0

	Zeer oneens	Oneens	Enigszins oneens	Neutraal	Enigszins eens	Eens	Zeer eens
De virtuele coach leek menselijk.	0	0	0	0	0	0	0
De virtuele coach was overtuigend.	0	0	0	0	0	0	0
Ik zou de virtuele coach aan anderen aanbevelen.	0	0	0	0	0	0	0

Graag willen we nog een aantal vragen over uzelf stellen.

3. Wat is uw geslacht?

- > Man
- > Vrouw
- > Anders

4. Wat is uw geboortedatum? (dd mm jjjj)

5. Wat is uw hoogst afgeronde opleiding? (Als uw opleiding er niet bij staat, kruis dan de opleiding aan die het meest op de door u gevolgde opleiding lijkt)

- > Basisonderwijs
- > Middelbaar onderwijs (MAVO, VMBO, HAVO, VWO)
- > Middelbaar beroepsonderwijs (MBO)
- > Hoger beroepsonderwijs (HBO)
- > Wetenschappelijk onderwijs (WO)

6. Wat is uw woonsituatie?

- > Ik woon samen met mijn vriend(in)/partner
- > Ik woon samen met een vriend, familielid of iemand anders
- > Ik woon alleen
- > Anders, namelijk:

7. Welke apparaten gebruikt u thuis?

- > Smartphone
- > Smartwatch
- > Tablet

- > PC / laptop
- > Spelcomputer
- > Anders, namelijk:

8. Hoeveel uur spendeert u per dag aan onderstaande activiteiten?

(U mag hierbij werktijd meenemen)

	0u	0-1u	1-2u	2-3u	3-4u	>4u
Zoekopdrachten op internet	0	0	0	0	0	0
Offline gebruik van mijn computer	0	0	0	0	0	0
E-mailen	0	0	0	0	0	0
Online chatten / social media	0	0	0	0	0	0
Telefoneren	0	0	0	0	0	0
Teksberichten sturen / whatsappen	0	0	0	0	0	0
Videospellen spelen	0	0	0	0	0	0
Naar de radio luisteren	0	0	0	0	0	0
Tv kijken	0	0	0	0	0	0
Boeken en tijdschriften lezen	0	0	0	0	0	0
Met iemand spreken (face-to-face)	0	0	0	0	0	0

APPENDIX E – QUESTIONNAIRES CH6

Agent Questions Baseline Questionnaire

Binnen het MATCH project doen we onderzoek naar het inzetten van een digitaal persoon voor het geven van adviezen op het gebied van gezond leven. In het kader van dit onderzoek zijn wij geïnteresseerd in uw mening over het ontwerp van een virtuele (digitale) coach. U krijgt zo een afbeelding van een virtuele coach te zien, waarna u wordt gevraagd om deze te beoordelen. U krijgt zo een aantal vragen over deze virtuele coach. Uw mening zal ons helpen bij het ontwikkelen van virtuele/digitale coaches die advies op maat kunnen geven.

U krijgt nu een afbeelding van een virtueel karakter te zien. Bekijkt u rustig de afbeelding. U wordt vervolgens gevraagd om een aantal eigenschappen van het karakter te beoordelen.

[IMAGE AGENT]

1. Hoe waarschijnlijk is het dat u advies van dit karakter zal opvolgen?

Selecteer 1 antwoordmogelijkheid.

Onwaarschijnlijk							Waarschijnlijk
0	0	0	0	0	0	0	

2. Ik vind het karakter:

Selecteer 1 antwoordmogelijkheid per eigenschap.

	Zeer oneens	Oneens	Enigszins oneens	Neutraal	Enigszins eens	Eens	Zeer eens
Vriendelijk	0	0	0	0	0	0	0
Deskundig	0	0	0	0	0	0	0
Betrouwbaar	0	0	0	0	0	0	0
Autoritair (gezaghebbend)	0	0	0	0	0	0	0
Betrokken	0	0	0	0	0	0	0

3. Hoe belangrijk vindt u:

Selecteer 1 antwoordmogelijkheid per eigenschap.

	Zeer oneens	Oneens	Enigszins oneens	Neutraal	Enigszins eens	Eens	Zeer eens
Vriendelijkheid	0	0	0	0	0	0	0
Deskundigheid	0	0	0	0	0	0	0
Betrouwbaarheid	0	0	0	0	0	0	0
Autoriteit (gezag)	0	0	0	0	0	0	0
Betrokkenheid	0	0	0	0	0	0	0

Agent Questions Intermediate and Follow-up Questionnaire

Binnen het MATCH project doen we onderzoek naar het inzetten van een digitaal persoon voor het geven van adviezen op het gebied van gezond leven. In het kader van dit onderzoek zijn wij geïnteresseerd in uw mening over het ontwerp van een virtuele (digitale) coach. U heeft nu een tijdje met een virtuele coach, Sylvia, mogen interacteren op het portaal. U krijgt zo een aantal vragen over deze virtuele coach. Uw mening zal ons helpen bij het ontwikkelen van virtuele/digitale coaches die advies op maat kunnen geven.

Op het MATCH portaal heeft u nu een tijdje kennis kunnen maken met uw virtuele coach, Sylvia (zie afbeelding hieronder). We vragen u om een aantal eigenschappen van uw virtuele coach te beoordelen.

[IMAGE AGENT]

1. Hoe vaak heeft u het advies van uw virtuele coach opgevolgd?

Selecteer 1 antwoordmogelijkheid.

Nooit							Vaak
0	0	0	0	0	0	0	0

2. Ik vind mijn virtuele coach:

Selecteer 1 antwoordmogelijkheid per eigenschap.

	Zeer oneens	Oneens	Enigszins oneens	Neutraal	Enigszins eens	Eens	Zeer eens
Vriendelijk	0	0	0	0	0	0	0
Deskundig	0	0	0	0	0	0	0
Betrouwbaar	0	0	0	0	0	0	0
Autoritair (gezagheb- bend)	0	0	0	0	0	0	0
Betrokken	0	0	0	0	0	0	0

3. Hoe belangrijk vindt u:

Selecteer 1 antwoordmogelijkheid per eigenschap.

	Zeer oneens	Oneens	Enigszins oneens	Neutraal	Enigszins eens	Eens	Zeer eens
Vriendelijkheid	0	0	0	0	0	0	0
Deskundigheid	0	0	0	0	0	0	0
Betrouwbaarheid	0	0	0	0	0	0	0
Autoriteit (gezag)	0	0	0	0	0	0	0
Betrokkenheid	0	0	0	0	0	0	0

APPENDIX F – QUESTIONNAIRES CH7

Agent Satisfaction Questions Baseline Questionnaire (EN/NL)

EN: Please score each coach on a scale of 1 (low) to 10 (high).

NL: Wat voor rapportcijfer geeft u deze coaches?

1. Olivia Simons (EN: physical activity / NL: lichamelijke activiteiten)

[PORTRAIT IMAGE OLIVIA]

1 2 3 4 5 6 7 8 9 10

2. François Dubois (EN: nutrition / NL: voeding)

[PORTRAIT IMAGE FRANCOIS]

1 2 3 4 5 6 7 8 9 10

3. Emma Li (EN: social / NL: sociaal)

[PORTRAIT IMAGE EMMA]

1 2 3 4 5 6 7 8 9 10

4. Helen Jones (EN: cognitive / NL: cognitief)

[PORTRAIT IMAGE HELEN]

1 2 3 4 5 6 7 8 9 10

5. Carlos Silva (EN: peer & support / NL: ondersteuning)

[PORTRAIT IMAGE CARLOS]

1 2 3 4 5 6 7 8 9 10

6. Rasmus Johansen (EN: chronic pain / NL: chronische pijn)

[PORTRAIT IMAGE RASMUS]

1 2 3 4 5 6 7 8 9 10

7. Katarzyna Kowalska (EN: diabetes / NL: diabetes)

[PORTRAIT IMAGE KATARZYNA]

1 2 3 4 5 6 7 8 9 10

Agent Satisfaction Questions Follow-up Questionnaire (EN / NL)

1. [PORTRAIT IMAGE OLIVIA]

EN: Please score the physical activity coach (Olivia Simons) on a scale from 1 (low) to 10 (high)

NL: Wat voor rapportcijfer geeft u aan de lichamelijke activiteiten coach (Olivia Simons)?

1 2 3 4 5 6 7 8 9 10

2. [PORTRAIT IMAGE FRANCOIS]

EN: Please score the nutrition coach (François Dubois) on a scale from 1 (low) to 10 (high)

NL: Wat voor rapportcijfer geeft u aan de voedingscoach (François Dubois)?

1 2 3 4 5 6 7 8 9 10

3. [PORTRAIT IMAGE EMMA]

EN: Please score the social coach (Emma Li) on a scale from 1 (low) to 10 (high)

NL: Wat voor rapportcijfer geeft u aan de sociale coach (Emma Li)?

1 2 3 4 5 6 7 8 9 10

4. [PORTRAIT IMAGE HELEN]

EN: Please score the cognitive coach (Helen Jones) on a scale from 1 (low) to 10 (high)

NL: Wat voor rapportcijfer geeft u aan de cognitieve coach (Helen Jones)?

1 2 3 4 5 6 7 8 9 10

5. [PORTRAIT IMAGE CARLOS]

EN: Please score the peer & support coach (Carlos Silva) on a scale from 1 (low) to 10 (high)

NL: Wat voor rapportcijfer geeft u aan de ondersteuning coach (Carlos Silva)?

1 2 3 4 5 6 7 8 9 10

6. [PORTRAIT IMAGE RASMUS]

EN: Please score the chronic pain coach (Rasmus Johansen) on a scale from 1 (low) to 10 (high)

NL: Wat voor rapportcijfer geeft u aan de chronische pijn coach (Rasmus Johansen)?

1 2 3 4 5 6 7 8 9 10

7. [PORTRAIT IMAGE KATARZYNA]

EN: Please score the diabetes coach (Katarzyna Kowalska) on a scale from 1 (low) to 10 (high)

NL: Wat voor rapportcijfer geeft u aan de diabetes coach (Katarzyna Kowalska)?

1 2 3 4 5 6 7 8 9 10

APPENDIX G – DETAILS UPDATE

LITERATURE REVIEW

Table A.3 – Overview of the articles included in the update of the literature review (articles published from December 2018 till July 2020). For every article, the feature category, design feature(s) researched, application goal, number of participants, evaluation stage, outcome variable(s), methods and outcomes are listed.

Source	Feature category	Design feature(s)	Goal app.	N	Eval. st.	Outcome variable(s)	Method	Outcomes
Speech and/or textual output, facial and gaze expressions and hand and body gestures								
Azevedo et al. (2018)	Speech and/or textual output, facial and gaze expressions Realism, Socio-demographic	Framing message (gain or loss framed) in facial and textual expressions: realism (emoji, photorealistic and cartoon), age (young, old), gender	[T] Medication information	360	I	<ul style="list-style-type: none"> [B] Memory for medication information [AR] Affective responses to the information [AC] Perception of ECA's teaching effectiveness and expressiveness [AC] Perceived level of similarity with each CA 	<p>A.2 (participant age: older=>60 and younger) x.2 (message frame: gain vs loss) x.3 (CA levels of realism: photorealistic vs cartoon vs emoji))</p>	<ul style="list-style-type: none"> Memory recall: older > young adults Memory recall: no difference old or young ECAs Affective responses: older > younger adults Framed gains: positive affect on affective response, framed loss: negative affect on affective response Memory recall: no effect perceived similarity age Perceived effectiveness: perceived similarity age > no perceived similarity age Perceived effectiveness: older adults rated older ECAs higher Perceived effectiveness older and younger ECAs: no difference older and younger users Perceived similarity: no effect age user Memory recall, perceived effectiveness, expressiveness, similarity: gain > loss (no effect of user age) Memory recall: no effect realism Perceived effectiveness, expressiveness, similarity: cartoon + realistic > emoji Female > male (general), no effect user age
Kang & Wei (2018)	Speech and/or textual output	Emotional support vs. informational support	[M] Stress management	132	I	<ul style="list-style-type: none"> [UC] User's general social support preference (emotional or informational) [AR] Perception of agent's support (emotional or informational) [AC] Likability [AC] Trustworthiness [AC] Support satisfaction [UX] Content credibility [UX] Interaction quality [B] Behavioral intention [UX] Perceived interactivity 	<p>2 (an agent's social support: informational vs. emotional) x.2 (a user's general social support needs: informational vs. emotional)</p>	<ul style="list-style-type: none"> Trustworthiness and likability: User with emotional support needs? matched support agent > unmatched support agent User with informational support needs? No difference matched and unmatched support <p>Other variables:</p> <ul style="list-style-type: none"> Matched support for users with emotional support need? Higher rating content credibility, perceived interactivity, interaction quality, support satisfaction, and behavioral intention Informational social support need? Social support type did not influence the outcome variables via trustworthiness and likability

Emotion

	Source	Feature category	Design feature(s)	Goal app.	N	Eval. st.	Outcome variable(s)	Method	Outcomes
Emotion	Lucas et al. (2018)	Speech and/or textual output, facial and gaze expressions	Message tone (positively vs. negatively valenced)	[P] Fitness training	212	I	<ul style="list-style-type: none"> [B] Motivation during workout [AC] Perceived quality of the trainer [AC] Personality trainer (toughness, disagreeableness, lack of character) [AR] Rapport [B] Self-reported effort during workout [B] Heart rate increase during workout 	<ul style="list-style-type: none"> 3 (trainer type) x 2 (message tone): <ul style="list-style-type: none"> Trainer type: virtual agent, an avatar agent or a real person via skype Message tone: positively valenced or negatively valenced 	<ul style="list-style-type: none"> Motivation (real person): neg. val. > pos. val Motivation (agent, avatar): pos. val. > neg. val. Quality trainer (real person): neg. val. > pos. val. Quality trainer (agent, avatar): pos. val. > neg. val. Toughness trainer (all trainer types): neg. val. > pos. val., but more penalty for using negatively valenced messages for agents and avatars than for humans Rapport: avatar and real person > agent Rapport: pos. val > neg. val. Self-reported effort: pos. val. > neg. val. Heart rate: neg. val > pos. val.
	ter Stal et al. (subm.)	Speech and/or textual output, facial and gaze expressions	Happy or neutral textual and facial expressions	[N/P] Tips for healthy physical activity and nutrition	63	I	<ul style="list-style-type: none"> [AR] Rapport scale items 	<ul style="list-style-type: none"> Short dialogue with ECA, 2 (textual expression) x 2 (facial expression): <ul style="list-style-type: none"> Textual: happy or neutral Facial: happy or neutral 	<ul style="list-style-type: none"> No significant difference in overall rapport Happy textual expression for ECA with neutral facial expression? Helpfulness and enjoyableness higher
	Ziaki et al. (2019)	Speech and/or textual output, facial and gaze expressions, hand and body gestures	Voice (vs. text), animation (vs. no animation), position ECA to the 'camera' (far or near)	[T] Colorectal cancer screening	1400	I	<ul style="list-style-type: none"> [B] Intention to learn more about colorectal cancer [B] Intention to discuss screening options with a doctor 	<ul style="list-style-type: none"> Three conditions: <ul style="list-style-type: none"> Animated 3D virtual human with voice vs. static 3D virtual human with text vs. text-only 	<ul style="list-style-type: none"> Intention to learn: animated > text-only Intention to discuss with doctor: animated > text-only and animated > static No effect camera position on intention to learn and to discuss with doctor



	Source	Feature category	Design feature(s)	Goal app.	N	Eval. sc.	Outcome variable(s)	Method	Outcomes
Animation	Zalake et al. (2019). See agent message.								
	Looks								
	Gambino et al. (2019)	Species, socio-demographic	Agency, role	[T] Annual physical health intake	283	I	<ul style="list-style-type: none"> [AC] Believe in machine heuristic [AC] Power usage [AC] Attitude toward the automated agent chat service [UX] Attitude toward the chat service [B] Behavioural intention toward the chat service 	3 (Role: Receptionist, Nurse, Doctor) x 3 (Type: Human, Avatar, Robot) factorial experiment	<ul style="list-style-type: none"> Machine heuristic and power usage significant predictors for attitude towards the automated agent, attitude toward the chat service, and behavioral intention towards the chat service Power usage higher? Effect of a user's belief in the machine heuristic is greater Source type and source role had no effect on either of the variables
	Sin & Munteanu (2019)	Species	Agency (human vs. machine/robot)	[T] Hy-per-tension, osteoporosis, shingles and type II diabetes	5	I	<ul style="list-style-type: none"> [O] Post-task interview: perceptions, confidence, and adoption of the respective interfaces [O] Post-session interview: compared their experiences based on fidelity and agency 	<ul style="list-style-type: none"> 2x2 design: fidelity (voice-only vs. embodied) and agency of doctor (human vs. machine/robot) Post-task's and post-session interviews: study findings related to the Information Search Process (ISP) 	<ul style="list-style-type: none"> ISP stage 2 selection, 3 exploration: voice-only robot preferred ISP stage 5 collection: video-based human preferred
	Agency								

Source	Feature category	Design feature(s)	Goal app.	N	Eval. s.t.	Outcome variable(s)	Method	Outcomes
Sträßmann et al. (2020)	Species, realism	Agency (cartoon and realistic agent vs. machine-like agent), realism (cartoon vs. realistic humanoid agent)	[O] Filling in a health diary and schedule appointments in a calendar	130	I	<ul style="list-style-type: none"> [AC] Perceived realism [AC] Likability [AC] Trustworthiness [AC] Competence [AC] Attractiveness [AC] Liking of the agent [UX] Perceived usefulness [IT] Intention to use [AR] Trust [UX] Enjoyment [UX] Ease of use [AC] Sociability 	<ul style="list-style-type: none"> Four conditions: two humanoid agents (a cartoon and a realistic agent, one machine-like agent and a voice-only) Three dimensions: <ul style="list-style-type: none"> (1) embodiment (humanoid agents and machine-like agent vs. voice-only) (2) species (cartoon and realistic agent vs. machine-like agent) (3) realism (cartoon vs. realistic humanoid agent) 	<ul style="list-style-type: none"> Seniors perceived agents as more realistic and attractive than students did Competence: voice-only > other three (embodied characters) Likability: machine-like > humanoid (cartoon and realistic) Seniors evaluated cartoon agent and voice-only as more realistic than students did In general, seniors liked agents more than students did Liking agent: machine-like > humanoid (cartoon and realistic) Students rated agents as easier to use than seniors did Perceived usefulness: seniors evaluated cartoon higher and showed a higher intention to use cartoon than students did Seniors reported higher bonding for cartoon and voice-only compared to students Seniors showed higher trust in cartoon than students did In general seniors evaluated agents as more sociable than students Sociable: voice only > humanoid and machine-like. Machine-like > humanoid Seniors rated cartoon as more sociable than students did Seniors found interaction more enjoyable than students did

Agency

Source	Feature category	Design feature(s)	Goal app.	N	Eval. st.	Outcome variable(s)	Method	Outcomes
Azevedo et al. (2018). <i>See speech and/or textual output, facial and gaze expressions and hand and body gestures.</i>								
Mostajeran et al. (2019)	Realism, socio-demo-graphic	Rendering style, gender	[P] Balance training	E1: 5 E2: 47	I	<p>Experiment 1:</p> <ul style="list-style-type: none"> [AC] Rating agent (five points Likert good - bad) [O] Coach preference [O] General opinion <p>Experiment2:</p> <ul style="list-style-type: none"> [AC] Rating agent (five points Likert good - bad) [O] Coach preference [O] General opinion [O] What design feature they would like to personalize 	<p>Two focus groups (Experiment 1 and experiment 2) with survey and discussion on five different coaches::</p> <ul style="list-style-type: none"> a) realistic male b) realistic female c) cartoony male d) cartoony female e) realistic polygonal <p>3D reconstruction of a real person</p>	<p>Experiment 1:</p> <ul style="list-style-type: none"> Rating: realistic male scored highest Preference coach: 60% preferred realistic male. General: mood coach should look more optimistic and happy, friendly (not grumpy) General: clothing should be clear and colourful <p>Experiment 2:</p> <ul style="list-style-type: none"> Rating: 3D reconstruction highest, followed by realistic male Preference: 53.3% realistic 3D reconstruction, 28.8% realistic female, 24.4% realistic male, 11.1% cartoon male and 4.4% cartoon female Aspects to personalize: gender: 48%, clothes: 28%, colour clothes: 24%, voice: 61%, age: 41%, size: 26%, nothing: 15% General: speech should be clear and natural Size agent: average human size: 66%, depends on space room: 20%, smaller than average: 8.8%, bigger than average: 4.4% Interaction medium: speech: 5.6%, body gesture: 33%, real physical touchpad: 24%, virtual touchpad: 17%, no interaction at all: 24% Distance to coach: 2-3m: 57.4%, 1-2m: 42.5%
Straßmann et al. (2020). <i>See species.</i>								
Rendering style								

Source	Feature category	Design feature(s)	Goal app.	N	Eval. s.c.	Outcome variable(s)	Method	Outcomes
Clothing	Styling	Agent's clothing/role, age and gender	[O] Lifestyle in general	E1: 155 E2: 13	I	<p>Experiment 1:</p> <ul style="list-style-type: none"> [O] Preferred ECA at first glance [AR] Likelihood of following ECA's advice [AC] Friendliness, trustworthiness, involvement, expertise and authority <p>Experiment 2:</p> <ul style="list-style-type: none"> [AC] Importance set of 20 predefined agent characteristics 	<p>Experiment 1:</p> <ul style="list-style-type: none"> Selection preferred static agent image out of 8 agent images (differing in age, gender and role) Per image: rating likelihoods of following advice and agent characteristics on 7-point Likert scales <p>Experiment 2:</p> <ul style="list-style-type: none"> Sorting of characteristics on importance in card-sorting task Discussion 	<p>Experiment 1:</p> <ul style="list-style-type: none"> General: preferred: young females, least preferred: older males Young > old agents: friendliness, involvement (general and elderly pop.), expertise and reliability (general pop.), likelihood of following advice (general pop., elderly pop.) Old > young agents: authority (general and elderly pop.) Female > male: friendliness, involvement (general pop.) Male > female: authority (general and elderly pop.), expertise (general pop.) Male = female: likelihood of following advice (general pop.) Experts > peers: expertise, reliability, authority, likelihood of following advice (general and elderly pop.), involvement (elderly pop.) Preference same age and gender as oneself (general and elderly pop.) <p>Experiment 2:</p> <ul style="list-style-type: none"> Importance: Dynamic > static features
	Zalaket al. (2018)	Styling	Agent's attire	[M] Coping with chronic stress	13	I	<ul style="list-style-type: none"> [AR] Trust in agent [O] Proportion of time focused on agent's face (compared to setting) 	<p>The VH had either professional attire (doctor's coat) or general attire (shirt)</p>



Source	Feature category	Design feature(s)	Goal app.	N	Eval. st.	Outcome variable(s)	Method	Outcomes
Azevedo et al. (2018). <i>See speech and/or textual output, facial and gaze expressions and hand and body gestures.</i>								
Mostajeran et al. (2019). <i>See rendering style.</i>								
ter Stal et al. (2020). <i>See clothing.</i>								
Gender	Socio-demo-graphic	Age, gender	[O] Filling in health questionnaires	20	I	<ul style="list-style-type: none"> [AC] Friendliness, trustworthiness, involvement, expertise and authority [AR] Likelihood of following ECA's advice [AR] Rapport scale items 	Within-subject experiment: users filled out health questionnaires with both a) a young female agent and b) an old male agent	<ul style="list-style-type: none"> User's perception of agent's friendliness, involvement, reliability, intelligence, and relation with agent (i.e. rapport towards the agent and likelihood of following the agent's advice); no effect age and gender Agent's authority: older male > young female
	Socio-demo-graphic	Age, gender in multi-agent application	[O] Holistic healthy aging	E1: 20 E2: 45	I	<ul style="list-style-type: none"> [O] E1: preference ECA as coach [O] E2: satisfaction with ECAs 	<ul style="list-style-type: none"> E1: Card-sorting task of 7 static ECA images E2: 4-week daily life evaluation of a multi-agent application with baseline and follow-up satisfaction questionnaire rating 7 ECAs differing in age and gender 	<ul style="list-style-type: none"> Older adults not more satisfied with ECAs similar to themselves in age and gender, neither at first glance nor after long-term interaction. E1: no significant difference rating ECA age groups and ECA gender E2: more satisfied with middle-aged ECAs, no significant difference in satisfaction with ECA gender
Azevedo et al. (2018). <i>See speech and/or textual output, facial and gaze expressions and hand and body gestures.</i>								
ter Stal et al. (2020). <i>See clothing.</i>								
ter Stal et al. (2020). <i>See gender.</i>								
ter Stal et al. (2020). <i>See gender.</i>								
Gambino et al. (2019). <i>See species.</i>								
ter Stal et al. (2020). <i>See clothing.</i>								
Age								
Role								

Goal application: A = alcohol consumption / M = mental health / N = nutrition / P = physical activity / T = medical treatment / O = other
Evaluation stage: I = technical efficacy / II = specific system objectives / III = system analysis / IV = external validity
Outcome variables: U = usage / IT = intention to (continue) using / B = (intention towards) behaviour change / UX = usability and user experience / AC = agent characteristics (e.g. demographics, personality, styling) / AR = relation with agent / UC = user characteristics / O = other

REFERENCES

Chapter 1

- 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. *International Conference on Intelligent Virtual Agents* (pp. 126–138). doi:10.1007/978-3-642-33197-8_13
- Bickmore, T. W., Carusob, L., Clough-Gorrb, K., & Heeren, T. (2005). 'It's just like you talk to a friend' relational agents for older adults. *Interacting with Computers*, 17(6), 711–735. doi:10.1016/j.intcom.2005.09.002
- Boumans, R., van Meulen, F., Hindriks, K., Neerincx, M., & Rikkert, M. G. (2019). Robot for health data acquisition among older adults: A pilot randomised controlled cross-over trial. *BMJ Quality and Safety*, 28(10), 793–799. doi:10.1136/bmjqs-2018-008977
- Chen, W., FitzGerald, J. M., Sin, D. D., & Sadatsafavi, M. (2017). Excess economic burden of comorbidities in COPD: A 15-year population-based study. *The European Respiratory Journal*, 650(1), 1–10. doi:10.1183/13993003.00393-2017
- Cole-Lewis, H., Ezeanochie, N., & Turgiss, J. (2019). Understanding health behavior technology-engagement: Pathway to measuring digital behavior change interventions. *JMIR Formative Research*, 3(4), e14052. doi:10.2196/14052
- Craig, J., & Patterson, V. (2005). Introduction to the practice of telemedicine. *Journal of Telemedicine and Telecare*, 11(1), 3–9. doi:10.1177/1357633X0501100102
- Crutzen, R., Cyr, D., & De Vries, N. K. (2011). Bringing loyalty to e-health: Theory validation using three internet-delivered interventions. *Journal of Medical Internet Research*, 13(3). doi:10.2196/jmir.1837
- Eysenbach, G. (2001). What is e-health? *Journal of Medical Internet Research*, 3(2), 1–5. doi:10.2196/jmir.3.2.e20
- Fatehi, F., & Wootton, R. (2012). Telemedicine, telehealth or e-health? A bibliometric analysis of the trends in the use of these terms. *Journal of Telemedicine and Telecare*, 18(8), 460–464. doi:10.1258/jtt.2012.GTH108
- Heinzelmann, P. J., Lugn, N. E., & Kvedar, J. C. (2005). Telemedicine in the future. *Journal of Telemedicine and Telecare*, 11(8), 384–390. doi:10.1258/135763305775013554
- Kelders, S. M., Kok, R. N., Ossebaard, H. C., & Van Gemert-Pijnen, J. E. W. C. (2012). Persuasive system design does matter: A systematic review of adherence to web-based interventions. *Journal of Medical Internet Research*, 14(6), 1–24. doi:10.2196/jmir.2104
- Kelley, H.H. (1950). The warm-cold variable in first impressions of persons. *Journal of Personality*, 18(4), 431–439. doi:10.1111/j.1467-6494.1950.tb01260.x
- Kohl, L. F., Crutzen, R., & De Vries, N. K. (2013). Online prevention aimed at lifestyle behaviors: A systematic review of reviews. *Journal of Medical Internet Research*, 15(7), 1–13. doi:10.2196/jmir.2665

- Kreps, G. L., & Neuhauser, L. (2010). New directions in eHealth communication: Opportunities and challenges. *Patient Education and Counseling*, 78(3), 329–336. doi:10.1016/j.pec.2010.01.013
- Lefebvre, C. R., Tada, Y., Hilfiker, S. W., & Baur, C. (2010). The assessment of user engagement with eHealth content: The eHealth engagement scale. *Journal of Computer-Mediated Communication*, 15(4), 666–681. doi:10.1111/j.1083-6101.2009.01514.x
- Nijland, N. (2011). *Grounding eHealth* [Doctoral dissertation, University of Twente]. Research Information University of Twente. doi:10.3990/1.9789036531337
- O'Brien, H. L., & Toms, E. G. (2008). What is user engagement? A conceptual framework for defining user engagement with technology. *Journal of the American Society for Information Science and Technology*, 59(6), 938–955. doi:10.1002/asi.20801
- Pagliari, C., Sloan, D., Gregor, P., Sullivan, F., Detmer, D., Kahan, J. P., ... MacGillivray, S. (2005). What is eHealth (4): A scoping exercise to map the field. *Journal of Medical Internet Research*, 7(1), 1–20. doi:10.2196/jmir.7.1.e9
- Preller, L., Helmink, J., Schaars, D., Aalbers, M., Wagemakers, A., & van Brussel-Visser, M. (2011). BeweegKuur: Development of a combined lifestyle intervention and implementation in local settings. *Proceedings of HEPA Europe* (pp. 38–38)
- RIVM - National Institute for Public Health and the Environment. (2014). *A healthier Netherlands. Key findings from the Dutch 2014 public health status and foresight report*. Retrieved from: <http://www.rivmjaarverslag2012.nl>
- Ruttkey, 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*, pp. 27–66. doi:10.1007/1-4020-2730-3_2
- Scherer, E. A., Ben-Zeev, D., Li, Z., & Kane, J. M. (2017). Analyzing mHealth engagement: joint models for intensively collected user engagement data. *JMIR mHealth and uHealth*, 5(1), e1. doi:10.2196/mhealth.6474
- Straßmann, C., & Krämer, N. C. (2017). A categorization of virtual agent appearances and a qualitative study on age-related user preferences. *International Conference on Intelligent Virtual Agents* (pp. 413–422). doi:10.1007/978-3-319-67401-8_51
- U.N. Department of Economic and Social Affairs. (2017). *World Population Prospects: 2017 Revision*. Retrieved from: <http://www.un.org/en/development/desa/population/events/other/21/index.shtml>
- van Boven, J. F. (2017). Costly comorbidities of COPD: The ignored side of the coin? *The European Respiratory Journal*, 50(1), 9–11. doi:10.1183/13993003.00917-2017
- Vanfleteren, L. E., Spruit, M. A., Wouters, E. F., & Franssen, F. M. (2016). Management of chronic obstructive pulmonary disease beyond the lungs. *The Lancet Respiratory Medicine*, 4(11), 911–924. doi:10.1016/S2213-2600(16)00097-7
- Veletsianos, G., Miller, C., & Doering, A. (2009). Enali: A research and design framework for virtual characters and pedagogical agents. *Journal of Educational Computing Research*, 41(2), 171–194. doi:10.2190/EC.41.2.c
- World Health Organization. (2010). *Global Recommendations on Physical Activity for Health*
- World Health Organization. (2012). *Telemedicine: Opportunities and Developments in Member States*. Retrieved from: http://www.who.int/goe/publications/goe_telemedicine_2010.pdf
- World Health Organization. (2015). *World Report on Ageing and Health*. Retrieved from: <https://www.who.int/ageing/publications/world-report-2015/en>

World Health Organization. (2020). *Healthy Diet*. Retrieved from: <http://www.who.int/en/news-room/fact-sheets/detail/healthy-diet>

Yardley, L., Spring, B. J., Riper, H., Morrison, L. G., Crane, D. H., Curtis, K., ... Blandford, A. (2016). Understanding and promoting effective engagement with digital behavior change interventions. *American Journal of Preventive Medicine*, 51(5), 833–842. doi:10.1016/j.amepre.2016.06.015

Zorginstituut Nederland. (n.d.). *Gecombineerde Leefstijlinterventie (GLI) (Zvw)*. Retrieved from: <https://www.zorginstituutnederland.nl/Verzekerde+zorg/gecombineerde-leefstijlinterventie-gli-zvw>

Chapter 2

Acosta, J. C., & Ward, N. G. (2011). Achieving rapport with turn-by-turn, user-responsive emotional coloring. *Speech Communication*, 53(9–10), 1137–1148. doi:10.1016/j.specom.2010.11.006

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* (pp. 11–15). doi:10.1145/3057039.3057080

Amini, R., Lisetti, C., & Yasavur, U. (2014). Emotionally responsive virtual counselor for behavior-change health interventions. *International Conference on Design Science Research in Information Systems* (pp. 433–437). doi:10.1007/978-3-319-06701-8_40

Amini, R., Lisetti, C., Yasavur, U., & Rische, N. (2013). On-demand virtual health counselor for delivering behavior-change health interventions. *IEEE International Conference on Healthcare Informatics* (pp. 46–55). doi:10.1109/ICHI.2013.13

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. *International Conference on Intelligent Tutoring Systems* (pp. 592–603). doi:10.1007/978-3-540-30139-4_56

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., Gruber, A., & Picard, R. (2005). Establishing the computer-patient working alliance in automated health behavior change interventions. *Patient Education and Counseling*, 59(1), 21–30. doi:10.1016/j.pec.2004.09.008

Bickmore, T., & Ring, L. (2010). Making it personal: end-user authoring of health narratives delivered by virtual agents. *International Conference on Intelligent Virtual Agents* (pp. 399–405). doi:10.1007/978-3-642-15892-6_43

Bickmore, T., & Schulman, D. (2007). Practical approaches to comforting users with relational agents. *Proceedings of ACM CHI 2007: Conference on Human Factors in Computing Systems* (pp. 2291–2296). doi:10.1145/1240866.1240996

- Bickmore, T., Schulman, D., & Yin, L. (2009). Engagement vs. deceit: Virtual humans with human autobiographies. *International Workshop on Intelligent Virtual Agents* (pp. 6–19). doi:10.1007/978-3-642-04380-2_4
- Bickmore, T., Schulman, D., & Yin, L. (2010). Maintaining engagement in long-term interventions with relational agents. *Applied Artificial Intelligence*, 24(6), 648–666. doi:10.1080/08839514.2010.492259
- Bickmore, T. W., Carusob, L., Clough-Gorrb, K., & Heeren, T. (2005). 'It's just like you talk to a friend' relational agents for older adults. *Interacting with Computers*, 17(6), 711–735. doi:10.1016/j.intcom.2005.09.002
- 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. *International Conference on Human Factors in Computing Systems* (pp. 1265–1274). doi:10.1145/1518701.1518891
- Bickmore, T. W., Picard, R. W. (2004). Towards caring machines. *CHI'04 Extended Abstracts on Human Factors in Computing Systems* (pp. 1489–1492). doi:10.1145/985921.986097
- Bickmore, T. W., Picard, R. W. (2005). Establishing and maintaining long-term human-computer relationships. *ACM Transactions on Computer-Human Interaction*, 12(2), 293–327. doi:10.1145/1067860.1067867
- Clark, H. H., & Brennan, E. S. (1991). Grounding in communication. *Perspectives on Socially Shared Cognition*, 13, pp. 222–233. doi:10.1037/10096-006
- Cowell, A., & Stanney, K. (2003). Embodiment and interaction guidelines for designing credible, trustworthy embodied conversational agents. *International Workshop on Intelligent Virtual Agents* (pp. 301–309). doi:10.1007/978-3-540-39396-2_50
- Creed, C., & Beale, R. (2012). User interactions with an affective nutritional coach. *Interacting with Computers*, 24(5), 339–350. doi:10.1016/j.intcom.2012.05.004
- Creed, C., Beale, R., & Cowan, B. (2015). The impact of an embodied agent's emotional expressions over multiple interactions. *Interacting with Computers*, 27(2), 172–188. doi:10.1093/iwc/iwt064
- Dryer, D. C. (1999). Getting personal with computers: How to design personalities for agents. *Applied Artificial Intelligence*, 13(3), 273–295. doi:10.1080/088395199117423
- Forlizzi, J., Zimmerman, J., Mancuso, V., & Kwak, S. (2007). How interface agents affect interaction between humans and computers. *Proceedings of the 2007 Conference on Designing Pleasurable Products and Interfaces* (pp. 209–221). doi:10.1145/1314161.1314180
- Frost, J., Boukris, N., & Roelofsma, P. (2012). We like to move it move it! Motivation and parasocial interaction. *CHI '12 Extended Abstracts on Human Factors in Computing Systems* (pp. 2465–2470). doi:10.1145/2212776.2223820
- Grillon, H., & Thalmann, D. (2008). Eye contact as trigger for modification of virtual character behavior. *2008 Virtual Rehabilitation* (pp. 205–211). doi:10.1109/ICVR.2008.4625161
- Guadagno, R., Blascovich, J., Bailenson, J., & McCall, C. (2007). Virtual humans and persuasion: The effects of agency and behavioral realism. *Media Psychology*, 10(1), 1–22. doi:10.1109/ICVR.2008.4625161
- Gulz, A., & Haake, M. (2005). Social and visual style in virtual pedagogical agents. *Workshop on Adapting the Interaction Style to Affective Factors associated with the 10th International Conference on User Modeling* (pp. 24–29)

- Jansen-Kosterink, S., Vollenbroek Hutten, M. M., & Hermens, H. J. (2016). A renewed framework for the evaluation of telemedicine. *eTELEMED 2016: The Eighth International Conference on eHealth, Telemedicine, and Social Medicine* (pp. 57–62)
- Kang, S. H., & Gratch, J. (2011). People like virtual counselors that highly-disclose about themselves. *Annual Review of Cybertherapy and Telemedicine*, 167, pp. 143–148
- 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
- Khan, R., & Angeli, A. D. (2009). The attractiveness stereotype in the evaluation of conversational agents. *IFIP Conference on Human-Computer Interaction* (pp. 85–97). doi:10.1007/978-3-642-03655-2_10
- Khan, R. F., & Sutcliffe, A. (2014). Attractive agents are more persuasive. *International Journal of Human-Computer Interaction*, 30(2), 142–150. doi:10.1080/10447318.2013.839904
- Kim, Y., Baylor, A. L., & Reed, G. (2003). The impact of image and voice of pedagogical agents. *E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education* (pp. 2237–2240).
- 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
- Kramer, L. L., ter Stal, S., Mulder, B. C., de Vet, E., & van Velsen, L. (2019). Developing embodied conversational agents for coaching people towards a healthy lifestyle: A scoping review. *Journal of Medical Internet Research*, 22(2), e14058. doi: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., & Coiera, E. (2018). Conversational agents in healthcare: A systematic review. *Journal of the American Medical Informatics Association*, 25(9), 1248–1258. doi:10.1093/jamia/ocy072
- Lee, J. E. R., Nass, C., Brave, S. B., Morishima, Y., Nakajima, H., & Yamada, R. (2007). The case for caring colearners: The effects of a computer-mediated colearner agent on trust and learning. *Journal of Communication*, 57(2), 183–204. doi:10.1111/j.1460-2466.2007.00339.x
- 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
- Lisetti, C., Amini, R., Yasavur, U., & Rische, N. (2013). I can help you change! An empathic virtual agent delivers behavior change health interventions. *ACM Transactions on Management Information Systems*, 4(4), 1–28. doi:10.1145/2544103
- McDonnell, R., Breidt, M., & Bülthoff, H. H. (2012). Render me real? Investigating the effect of render style on the perception of animated virtual humans. *ACM Transactions on Graphics (TOG)*, 31(4), 1–11. doi:10.1145/2185520.2185587
- Malhotra, A., Hoey, J., König, A., & van Vuuren, S. (2016). A study of elderly people's emotional understanding of prompts given by virtual humans. *Proceedings of the 10th EAI International Conference on Pervasive Computing Technologies for Healthcare* (pp. 13–16)
- Nguyen, H., & Masthoff, J. (2007). Is it me or is it what I say? Source image and persuasion. *International Conference on Persuasive Technology* (pp. 231–242). doi:10.1007/978-3-540-77006-0_29

- Nguyen, H., & Masthoff, J. (2009). Designing empathic computers: The effect of multi-modal empathic feedback using animated agent. *Proceedings of the 4th International Conference on Persuasive Technology* (pp. 1–9) doi:10.1145/1541948.1541958
- Nijland, N. (2011). *Grounding eHealth* [Doctoral dissertation, University of Twente]. Research Information University of Twente. doi:10.3990/1.9789036531337
- Olafsson, S., Kimani, E., Asadi, R., Bickmore, T., & Science, I. (2017). That's a rap. *International Conference on Intelligent Virtual Agents* (pp. 325–334). doi:10.1007/978-3-319-67401-8_41
- 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. *International Conference on Intelligent Virtual Agents* (pp. 301–306). doi:10.1145/3267851.3267915
- Pelachaud, C. (2009). Studies on gesture expressivity for a virtual agent. *Speech Communication* 51(7), 630–639. doi:10.1016/j.specom.2008.04.009
- 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). doi:10.2196/jmir.6553
- von der Pütten, A. M., Krämer, N. C., & Gratch, J. (2009). Who's there? Can a virtual agent really elicit social presence? *Proceedings of the PRESENCE 2009 - The 12th Annual International Workshop on Presence*
- Ring, L., Utami, D., & Bickmore, T. (2014). The right agent for the job? The effects of agent visual appearance on task domain. *International Conference on Intelligent Virtual Agents* (pp. 374–384). doi:10.1007/978-3-319-09767-1_49
- Rist, T., André, E., Baldes, S., Gebhard, P., Klesen, M., Kipp, M., Rist, P., & Schmitt, M. (2004). A review of the development of embodied presentation agents and their application fields. *Life-Like Characters* (pp. 377–404). doi:10.1007/978-3-662-08373-4_16
- Robertson, S., Solomon, R., Riedl, M., Gillespie, T. W., Chociemski, T., & Master, V. (2015). The visual design and implementation of an embodied conversational agent in a shared decision-making context (eCoach). *International Conference on Learning and Collaboration Technologies* (pp. 427–437). doi:10.1007/978-3-319-20609-7
- 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
- Ruttikay, 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*, pp. 27–66. doi:10.1007/1-4020-2730-3_2
- Schmeil, A., & Suggs, S. (2014). "How am I doing?" - Personifying health through animated characters. *International Conference of Design, User Experience, and Usability* (pp. 91–102). doi:10.1007/978-3-319-07635-5_10
- Scholten, M. R., Kelders, S. M., van Gemert-Pijnen, J. E. (2017). Self-guided web-based interventions: Scoping review on user needs and the potential of embodied conversational agents to address them. *Journal of Medical Internet Research*, 19(11), 1–19. doi:10.2196/jmir.7351
- Silverman, B. G., Holmes, J., Kimmel, S., Branas, C., & Ivins, D. (2001). Modeling emotion and behavior in animated personas to facilitate human behavior change: the case of the HEART-SENSE game. *Health Care Management Science*, 4(3), 213–228. doi:10.1023/A:1011448916375

- Skalski, P., Tamborini, R., & Skalski, P. (2007). The role of social presence in interactive agent-based persuasion. *Media Psychology*, 10(3), 385–413. doi:10.1080/15213260701533102
- Straßmann, C., & Krämer, N. C. (2017). A categorization of virtual agent appearances and a qualitative study on age-related user preferences. *International Conference on Intelligent Virtual Agents* (pp. 413–422). doi:10.1007/978-3-319-67401-8_51
- Tielman, M. L., Neerincx, M. A., van Meggelen, M., Franken, I. (2017). How should a virtual agent present psychoeducation? Influence of verbal and textual presentation on adherence. *Technology and Health Care*, 25, pp. 1081–1096. doi:10.3233/THC-170899
- Vardoulakis, L. P., Ring, L., Barry, B., Sidner, C. L., & Bickmore, T. (2012). Designing relational agents as long term social companions for older adults. *International Conference on Intelligent Virtual Agents* (pp. 289–302). doi:10.1007/978-3-642-33197-8_30
- 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
- Veletsianos, G., Miller, C., & Doering, A. (2009). Enali: A research and design framework for virtual characters and pedagogical agents. *Journal of Educational Computing Research*, 41(2), 171–194. doi:10.2190/EC.41.2.c
- van Vugt, H. C., Konijn, E. A., Hoorn, J. F., Veldhuis, J. (2006). Why fat interface characters are better e-health advisors. *International Conference on Intelligent Virtual Agents* (pp. 1–13). doi:10.1007/11821830_1
- WHO (1946). *Constitution of the world health organization*, pp. 1-18
- van Wissen, A., Vinkers, C., & van Halteren, A. (2016). Developing a virtual coach for chronic patients: A user study on the impact of similarity, familiarity and realism. *International Conference on Persuasive Technology* (pp. 263–275). doi:10.1007/978-3-319-31510-2_23
- Yin, L., Bickmore, T., & Cortés, D. E. (2010). The impact of linguistic and cultural congruity on persuasion by conversational agents. *International Conference on Intelligent Virtual Agents* (pp. 343–349). doi:10.1007/978-3-642-15892-6_36
- Zell, E., Aliaga, C., Jarabo, A., Zibrek, K., Gutierrez, D., McDonnell, R., & Botsch, M. (2015). To stylize or not to stylize? The effect of shape and material stylization on the perception of computer-generated faces. *ACM Transactions on Graphics (TOG)*, 34(6), 1–12. doi:10.1145/2816795.2818126
- Zibrek, K., Kokkinara, E., & McDonnell, R. (2018). The effect of realistic appearance of virtual characters in immersive environments-does the character's personality play a role?. *IEEE Transactions on Visualization and Computer Graphics*, 24(4), 1681–1690. doi:10.1109/TVCG.2018.2794638
- Zhou, S., Bickmore, T., Paasche-Orlow, M., & Jack, B. (2014). Agent-user concordance and satisfaction with a virtual hospital discharge nurse. *International Conference on Intelligent Virtual Agents* (pp. 528–541). doi:10.1007/978-3-319-09767-1_63
- Zhou, S., Zhang, Z., & Bickmore, T. (2017). Adapting a persuasive conversational agent for the Chinese culture. *International Conference on Culture and Computing* (pp. 89–96). doi:10.1109/Culture.and.Computing.2017.42

Chapter 3

- Acosta, J. C., & Ward, N. G. (2011). Achieving rapport with turn-by-turn, user-responsive emotional coloring. *Speech Communication*, 53(9–10), 1137–1148. doi:10.1016/j.specom.2010.11.006
- 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* (pp. 11–15). 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). 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. *International Conference on Intelligent Tutoring Systems* (pp. 592–603). doi:10.1007/978-3-540-30139-4_56
- Bergmann, K., Eyssele, 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. *International Conference on Intelligent Virtual Agents* (pp. 126–138). 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. *International Conference on Human Factors in Computing Systems* (pp. 1265–1274). 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. *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, pp. 222–233. doi:10.1037/10096-006

- Cowell, A., & Stanney, K. (2003). Embodiment and interaction guidelines for designing credible, trustworthy embodied conversational agents. *International Workshop on Intelligent Virtual Agents* (pp. 301–309). 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(7), 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*, pp. 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. *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), 1–22. doi:10.1109/ICVR.2008.4625161
- 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. *E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education* (pp. 2237–2240).
- 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). 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. (2007). Is it me or is it what I say? Source image and persuasion. *International Conference on Persuasive Technology* (pp. 231–242). doi:10.1007/978-3-540-77006-0_29
- Nguyen, H., & Masthoff, J. (2009). Designing empathic computers: The effect of multi-modal empathic feedback using animated agent. *Proceedings of the 4th International Conference on Persuasive Technology* (pp. 1–9) doi:10.1145/1541948.1541958
- Nijland, N. (2011). *Grounding eHealth* [Doctoral dissertation, University of Twente]. Research Information University of Twente. doi:10.3990/1.9789036531337

- Numamaker, 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., & Harjumaa, 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* (pp. 1292–1303). 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. *International Conference on Intelligent Virtual Agents* (pp. 301–306). 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. *Qualitative research in health care*, 3, 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
- Ruttikay, 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*, pp. 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. *International Conference on Intelligent Virtual Agents* (pp. 153–160). 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. *International Conference on Intelligent Virtual Agents* (pp. 289–302). 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

- van Wissen, A., Vinkers, C., & van Halteren, A. (2016). Developing a virtual coach for chronic patients: A user study on the impact of similarity, familiarity and realism. *International Conference on Persuasive Technology* (pp. 263–275). 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. *Proceedings of Designing Pleasurable Products and Interfaces* (pp. 233–248). doi:10.1184/R1/6470366.v1

Chapter 4

- Acosta, J. C., & Ward, N. G. (2011). Achieving rapport with turn-by-turn, user-responsive emotional coloring. *Speech Communication*, 53(9–10), 1137–1148. doi:10.1016/j.specom.2010.11.006
- 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., & Ryu, J. (2003). The effects of image and animation in enhancing pedagogical agent persona. *Journal of Educational Computing Research*, 28(4), 373–394. doi:10.2190/v0wq-nw-gn-jb54-fat4
- 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. *International Conference on Intelligent Virtual Agents* (pp. 126–138). doi:10.1007/978-3-642-33197-8_13
- Beukema, S., van Velsen, L., Jansen-Kosterink, S., & Karreman, J. (2017). "There Is something we need to tell you...": Communicating health-screening results to older adults via the internet. *Telemedicine and e-health*, 23(9), 741–746. doi:10.1089/tmj.2016.0210
- Bliven, B. D., Kaufman S. E., Spertus J. A. (2001). Electronic collection of health-related quality of life data: validity, time benefits, and patient preference. *Quality of Life Research*, 10(1), 15–21. doi:10.1023/a:1016740312904
- Cafaro, A., Vilhjálmsson, 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
- Cowell, A., & Stanney, K. (2003). Embodiment and interaction guidelines for designing credible, trustworthy embodied conversational agents. *International Workshop on Intelligent Virtual Agents* (pp. 301–309). doi:10.1007/978-3-540-39396-2_50
- Hardie, J. A., Bakke, P. S., & Mørkve, O. (2003). Non-response bias in a postal questionnaire survey on respiratory health in the old and very old. *Scandinavian Journal of Public Health*, 31(6), 411–417. doi: 10.1080/14034940210165163
- Hébert, R., Bravo, G., Korner-Bitensky, N., & Voyer, L. (1996). Refusal and information bias associated with postal questionnaires and face-to-face interviews in very elderly subjects. *Journal of Clinical Epidemiology*, 49(3), 373–381. doi:10.1016/0895-4356(95)00527-7
- Hess, R., Santucci, A., McTigue, K., Fischer, G., & Kapoor, W. (2008). Patient difficulty using tablet computers to screen in primary care. *Journal of General Internal Medicine*, 23(4), 476–480. doi:10.1007/s11606-007-0500-1
- Fanning, J., & McAuley, E. (2014). A comparison of tablet computer and paper-based questionnaires in healthy aging research. *JMIR Research Protocols*, 3(3), e38. doi:10.2196/resprot.3291

- Fiorlizzi, J., Zimmerman, J., Mancuso, V., & Kwak, S. (2007). How interface agents affect interaction between humans and computers. *Proceedings of the 2007 Conference on Designing Pleasurable Products and Interfaces* (pp. 209–221). doi:10.1145/1314161.1314180
- Fried, L. P., Tangen, C. M., Walston, J., Newman, A. B., Hirsch, C., Gottdiener, J., ... & McBurnie, M. A. (2001). Frailty in older adults: evidence for a phenotype. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 56(3), M146–M157. doi:10.1093/gerona/56.3.M146
- Galvin, J. E., Roe, C. M., Coats, M. A., & Morris, J. C. (2007). Patient's rating of cognitive ability: using the AD8, a brief informant interview, as a self-rating tool to detect dementia. *Archives of neurology*, 64(5), 725–730. doi:10.1001/archneur.64.5.725
- Kang, S. H., Feng, A. W., Leuski, A., Casas, D., & Shapiro, A. (2015). The effect of an animated virtual character on mobile chat interactions. *Proceedings of the 3rd International Conference on Human-Agent Interaction* (pp. 105–112).
- Kelley, H.H. (1950). The warm-cold variable in first impressions of persons. *Journal of Personality*, 18(4), 431–439. doi:10.1111/j.1467-6494.1950.tb01260.x
- Kok, R., & Verhey, F. (2002). *Gestandaardiseerde Mini-Mental State Examination*. Retrieved from: <https://meetinstrumentenzorg.nl/wp-content/uploads/instrumenten/MMSE-meetinstr-ge-stand.pdf>
- Komatsu, T., Kurosawa, R., & Yamada, S. (2012). How does the difference between users' expectations and perceptions about a robotic agent affect their behavior?. *International Journal of Social Robotics*, 4(2), 109–116. doi:10.1007/s12369-011-0122-y
- Kvien, T. K., Mowinckel, P., Heiberg, T., Dammann, K. L., Dale, Ø., Aanerud, G. J., ... & Uhlig, T. (2005). Performance of health status measures with a pen based personal digital assistant. *Annals of the Rheumatic Diseases*, 64(10), 1480–1484. doi:10.1136/ard.2004.030437
- Lindley, S. E., Harper, R., & Sellen, A. (2009). Desiring to be in touch in a changing communications landscape: attitudes of older adults. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1693–1702). doi:10.1145/1518701.1518962
- Malva, J. O., & Bousquet, J. (2016). Operational definition of active and healthy ageing: Roadmap from concept to change of management. *Maturitas*, 84, 3–4. doi:10.1016/j.maturitas.2015.11.004
- Mays, N., & Pope, C. (1995). Qualitative research: observational methods in health care settings. *BMJ*, 311(6998), 182–184. doi:10.1136/bmj.311.6998.182
- Nass, C., & Lee, K. M. (2000). Does computer-generated speech manifest personality? An experimental test of similarity-attraction. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (pp. 329–336). doi:10.1145/332040.332452
- Noorman-de Vette, F. (2019). *Designing Game-Based eHealth Applications Strategies for Sustainable Engagement of Older Adults* [Doctoral dissertation, University of Twente]. Research Information University of Twente. doi:10.3990/1.9789036547994
- 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. *International Conference on Intelligent Virtual Agents* (pp. 301–306). doi:10.1145/3267851.3267915
- Prochaska, J. O., & Velicer, W. F. (1997). The transtheoretical model of health behavior change. *American Journal of Health Promotion*, 12(1), 38–48. doi:10.4278/0890-1171-12.1.38

- Rubenstein, L. Z., Harker, J. O., Salvà, A., Guigoz, Y., & Vellas, B. (2001). Screening for undernutrition in geriatric practice: developing the short-form mini-nutritional assessment (MNA-SF). *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 56(6), M366–M372. doi:10.1093/gerona/56.6.m366
- Ruttkey, 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*, pp. 27–66. doi:10.1007/1-4020-2730-3_2
- 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, 102409. doi:10.1016/j.ijhcs.2020.102409
- ter Stal S., Tabak, M., op den Akker, H., Beinema, T., Hermens, H. (2019). Who do you prefer? The effect of age, gender and role on users' first impressions of embodied conversational agents in eHealth. *International Journal of Human-Computer Interaction*, 36(9), 881–892. doi:10.1080/10447318.2019.1699744
- Thies, I. M. (2015). User interface design for low-literate and novice users: Past, present and future. *Foundations and Trends in Human-Computer Interaction*, 8(1), 1–72. doi:10.1561/11000000047
- van der Zee, K., & Sanderman, R. (1993). *Het meten van de algemene gezondheidstoestand met de rand-36*. Retrieved from: <https://www.umcg.nl/EN/Research/InstitutesProgrammes/SHARE/researchtools/tools/Paginas/RAND36.aspx>
- Vardoulakis, L. P., Ring, L., Barry, B., Sidner, C. L., & Bickmore, T. (2012). Designing relational agents as long term social companions for older adults. *International Conference on Intelligent Virtual Agents* (pp. 289–302). 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., Miller, C., & Doering, A. (2009). Enali: A research and design framework for virtual characters and pedagogical agents. *Journal of Educational Computing Research*, 41(2), 171–194. doi:10.2190/EC.41.2.c
- van Velsen, L., Frazer, S., N'dja, A., Ammour, N., Del Signore, S., Zia, G., Hermens, H. (2018). The reliability of using tablet technology for screening the health of older adults. *Studies in Health Technology Informatics*, 247, 651–655. doi:10.3233/978-1-61499-852-5-651
- Waycott, J., Vetere, F., Pedell, S., Morgans, A., Ozanne, E., & Kulik, L. (2016). Not for me: Older adults choosing not to participate in a social isolation intervention. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 745–757). doi:10.1145/2858036.2858458
- van Wissen, A., Vinkers, C., & van Halteren, A. (2016). Developing a virtual coach for chronic patients: A user study on the impact of similarity, familiarity and realism. *International Conference on Persuasive Technology* (pp. 263–275). doi:10.1007/978-3-319-31510-2_23
- Zhou, S., Bickmore, T., Paasche-Orlow, M., & Jack, B. (2014). Agent-user concordance and satisfaction with a virtual hospital discharge nurse. *International Conference on Intelligent Virtual Agents* (pp. 528–541). doi:10.1007/978-3-319-09767-1_63

- Zhou, S., Zhang, Z., & Bickmore, T. (2017). Adapting a persuasive conversational agent for the Chinese culture. *International Conference on Culture and Computing* (pp. 89–96). doi:10.1109/Culture.and.Computing.2017.42
- Zimmerman, J., Ayoob, E., Forlizzi, J., & McQuaid, M. (2005). Putting a face on embodied interface agents. *Proceedings of Designing Pleasurable Products and Interfaces* (pp. 233–248). doi:10.1184/R1/6470366.v1

Chapter 5

- Acosta, J. C., & Ward, N. G. (2011). Achieving rapport with turn-by-turn, user-responsive emotional coloring. *Speech Communication*, 53(9–10), 1137–1148. doi:10.1016/j.specom.2010.11.006
- Amini, R., Lisetti, C., & Yasavur, U. (2014). Emotionally responsive virtual counselor for behavior-change health interventions. *International Conference on Design Science Research in Information Systems* (pp. 433–437). doi:10.1007/978-3-319-06701-8_40
- Beale, R., & Creed, C. (2009). Affective interaction: How emotional agents affect users. *International Journal of Human Computer Studies*, 67(9), 755–776. doi:10.1016/j.ijhcs.2009.05.001
- Beer, J. M., Smarr, C. A., Fisk, A. D., & Rogers, W. A. (2015). Younger and older users' recognition of virtual agent facial expressions. *International Journal of Human Computer Studies*, 75, 1–20. doi:10.1016/j.ijhcs.2014.11.005
- Bevacqua, E., Mancini, M., Niewiadomski, R., & Pelachaud, C. (2007). An expressive ECA showing complex emotions. *Proceedings of the AISB Annual Convention* (pp. 208–216)
- Cerekovic, A., Aran, O., & Gatica-Perez, D. (2014). How do you like your virtual agent?: Human-agent interaction experience through nonverbal features and personality traits. *International Workshop on Human Behavior Understanding* (pp. 1–15). doi:10.1007/978-3-319-11839-0_1
- Cerekovic, A., Aran, O., & Gatica-Perez, D. (2017). Rapport with virtual agents: What do human social cues and personality explain? *IEEE Transactions on Affective Computing*, 8(3), 382–395. doi:10.1109/TAFFC.2016.2545650
- Creed, C., Beale, R., & Cowan, B. (2015). The impact of an embodied agent's emotional expressions over multiple interactions. *Interacting with Computers*, 27(2), 172–188. doi:10.1093/iwc/iwt064
- Ekman, P., & Friesen, W. (1978). *Facial Action Coding System: A technique for the measurement of facial movement*. Palo Alto, California: Consulting Psychologists Press
- Gratch, J., Wang, N., Gerten, J., Fast, E., & Duffy, R. (2007). Creating rapport with virtual agents. *International Conference on Intelligent Virtual Agents* (pp. 125–138). doi:10.1007/978-3-540-74997-4_12
- Gratch, J., Wang, N., Okhmatovskaia, A., Lamothe, F., Morales, M., van der Werf, R. J., & Morency, L. P. (2007). Can virtual humans be more engaging than real ones?. *International Conference on Human-Computer Interaction* (pp. 286–297). doi:10.1007/978-3-540-73110-8_30
- Hancock, J. T., Landrigan, C., & Silver, C. (2007). Expressing emotion in text-based communication. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 929–932). doi:10.1145/1240624.1240764

- Kanade, T., Cohn, J. F., & Tian, Y. (2000). Comprehensive database for facial expression analysis. *Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition* (pp. 484–490). doi:10.1109/AFGR.2000.840611
- 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
- Llorach, G., Agenjo, J., Blat, J., & Sayago, S. (2019). Web-based embodied conversational agents and older people. *Perspectives on Human-Computer Interaction Research with Older People* (pp. 119–135). doi:10.1007/978-3-030-06076-3_8
- Loveys, K., Sebaratnam, G., Sagar, M., & Broadbent, E. (2020). The effect of design features on relationship quality with Embodied Conversational Agents: A systematic review. *International Journal of Social Robotics*, pp. 1–20. doi:10.1007/s12369-020-00680-7
- Moridis, C. N., & Economides, A. A. (2012). Affective learning: Empathetic agents with emotional facial and tone of voice expressions. *IEEE Transactions on Affective Computing*, 3(3), 260–272. doi:10.1109/T-AFFC.2012.6
- Novick, D., & Gris, I. (2014). Building rapport between human and ECA: A pilot study. *International Conference on Human-Computer Interaction* (pp. 472–480). doi:10.1007/978-3-319-07230-2_45
- Pelachaud, C. (2009). Studies on gesture expressivity for a virtual agent. *Speech Communication* 51(7), 630–639. doi:10.1016/j.specom.2008.04.009
- Ranjbartabar, H., Richards, D., Bilgin, A., & Kutay, C. (2019). First impressions count! The role of the human's emotional state on rapport established with an empathic versus neutral virtual therapist. *IEEE Transactions on Affective Computing*. doi:10.1109/TAFFC.2019.2899305
- Ruttkey, 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*, pp. 27–66. doi:10.1007/1-4020-2730-3_2
- Tickle-deggen, L., & Rosenthal, R. (1990). The nature of rapport and its nonverbal correlates. *Psychological Inquiry*, 1(4), 285–293.
- Walther, J. B. (1992). Interpersonal effects in computer-mediated interaction: A relational perspective. *Communication Research*, 19(1), 52–90. doi:10.1177/009365092019001003

Chapter 6

- op den Akker, H., Tabak, M., Perianu, M. M., Jones, V. M., Hofs, D. H. W., Tönis, T., ... & Hermens, H. J. (2012). Development and evaluation of a sensor-based system for remote monitoring and treatment of chronic diseases—the continuous care & coaching platform. *International Symposium on eHealth Services and Technologies* (pp. 19–27). doi:10.5220/0004474700190027
- 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* (pp. 11–15). doi:10.1145/3057039.3057080

- Barello, S., Triberti, S., Graffigna, G., Libreri, C., Serino, S., Hibbard, J., & Riva, G. (2016). eHealth for patient engagement: A systematic review. *Frontiers in Psychology*, 6. doi:10.3389/fpsyg.2015.02013
- Baylor, A. L., & Kim, Y. (2004). Pedagogical agent design: The impact of agent realism, gender, ethnicity, and instructional role. *International Conference on Intelligent Tutoring Systems* (pp. 592–603). doi:10.1007/978-3-540-30139-4_56
- Bergmann, K., Eyssele, 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. *International Conference on Intelligent Virtual Agents* (pp. 126–138). doi:10.1007/978-3-642-33197-8_13
- Beukema, S., van Velsen, L., Jansen-Kosterink, S., & Karreman, J. (2017). "There Is something we need to tell you...": Communicating health-screening results to older adults via the internet. *Telemedicine and e-health*, 23(9), 741–746. doi:10.1089/tmj.2016.0210
- 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. *International Conference on Human Factors in Computing Systems* (pp. 1265–1274). doi:10.1145/1518701.1518891
- Buimer, H. P., Tabak, M., van Velsen, L., van der Geest, T., & Hermens, H. (2017). Exploring determinants of patient adherence to a Portal-Supported Oncology Rehabilitation Program: Interview and data log analyses. *JMIR Rehabilitation and Assistive Technologies*, 4(2), e12. doi:10.2196/rehab.6294
- Cafaro, A., Vilhjálmsson, 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. *International Conference on Intelligent Virtual Agents* (pp. 67–80). doi:10.1007/978-3-642-33197-8_7
- Cambridge University Press. (2019). *Rapport*. Retrieved from: <https://dictionary.cambridge.org/dictionary/english/rapport>
- Clark, H. H., & Brennan, E. S. (1991). Grounding in communication. *Perspectives on Socially Shared Cognition*, 13, pp. 222–233. doi:10.1037/10096-006
- Fogg, B. (2011). Computers as persuasive social actors. *Persuasive Technology*, pp. 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. *Proceedings of the 2007 Conference on Designing Pleasurable Products and Interfaces* (pp. 209–221). doi:10.1145/1314161.1314180
- Global Initiative for Chronic Obstructive Lung Disease (GOLD). (2020). *Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Pulmonary Disease*. Retrieved from: <https://goldcopd.org/gold-reports>
- Horne, M., Khan, H., & Corrigan, P. (2013). *People Powered Health: Health for People, by People and with People*. London: Nesta.
- Jansen-Kosterink, S., Vollenbroek Hutten, M. M., & Hermens, H. J. (2016). A renewed framework for the evaluation of telemedicine. *eTELEMED 2016: The Eighth International Conference on eHealth, Telemedicine, and Social Medicine* (pp. 57–62)
- Kang, S. H., & Gratch, J. (2011). People like virtual counselors that highly-disclose about themselves. *Annual Review of Cybertherapy and Telemedicine*, 167, pp. 143–148

- Kelders, S. M., Kok, R. N., Ossebaard, H. C., & Van Gemert-Pijnen, J. E. W. C. (2012). Persuasive system design does matter: A systematic review of adherence to web-based interventions. *Journal of Medical Internet Research*, *14*(6), 1–24. doi:10.2196/jmir.2104
- Kelley, H.H. (1950). The warm-cold variable in first impressions of persons. *Journal of Personality*, *18*(4), 431–439. doi:10.1111/j.1467-6494.1950.tb01260.x
- Komatsu, T., Kurosawa, R., & Yamada, S. (2012). How does the difference between users' expectations and perceptions about a robotic agent affect their behavior?. *International Journal of Social Robotics*, *4*(2), 109–116. doi:10.1007/s12369-011-0122-y
- Lenferink, A., Frith, P., van der Valk, P., Buckman, J., Sladek, R., Cafarella, P., ... & Effing, T. (2013). A self-management approach using self-initiated action plans for symptoms with ongoing nurse support in patients with Chronic Obstructive Pulmonary Disease (COPD) and comorbidities: The COPE-III study protocol. *Contemporary Clinical Trials*, *36*(1), 81–89. doi:10.1016/j.cct.2013.06.003
- Lenferink, A., van der Palen, J., van der Valk, P. D., Cafarella, P., van Veen, A., Quinn, S., ... & Effing, T. W. (2019). Exacerbation action plans for patients with COPD and comorbidities: a randomised controlled trial. *European Respiratory Journal*, *54*(5). doi:10.1183/13993003.02134-2018
- Mannino, D. M., Thorn, D., Swensen, A., & Holguin, F. (2008). Prevalence and outcomes of diabetes, hypertension and cardiovascular disease in COPD. *European Respiratory Journal*, *32*(4), 962–969. doi:10.1183/09031936.00012408
- 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
- Moody, L. (2015). User-centred health design: reflections on D4D's experiences and challenges. *Journal of Medical Engineering & Technology*, *39*(7), 395–403. doi:10.3109/03091902.2015.1088086
- Nijland, N. (2011). *Grounding eHealth* [Doctoral dissertation, University of Twente]. Research Information University of Twente. doi:10.3990/1.9789036531337
- Oinas-Kukkonen, H., & Harjumaa, M. (2009). Persuasive systems design: Key issues, process model, and system features. *Persuasive Systems Design*, *24*(28), 485–500. doi:10.17705/1cais.02428
- Pope, C., & Mays, N. (2006). Qualitative methods in health research. *Qualitative research in health care*, *3*, pp. 1–11. doi:10.1002/9780470750841.ch1
- Ruttikay, 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*, pp. 27–66. doi:10.1007/1-4020-2730-3_2
- ter Stal, S., Broekhuis, M., van Velsen, L., Hermens, H., & Tabak, M. (2020). Embodied conversational agent appearance for health assessment of older adults: Explorative study. *JMIR Human Factors*, *7*(3). doi:10.2196/19987
- 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*, 102409. doi:10.1016/j.ijhcs.2020.102409

- ter Stal S., Tabak, M., op den Akker, H., Beinema, T., Hermens, H. (2019). Who do you prefer? The effect of age, gender and role on users' first impressions of embodied conversational agents in eHealth. *International Journal of Human-Computer Interaction*, 36(9), 881–892. doi:10.1080/10447318.2019.1699744
- Tabak, M., Brusse-Keizer, M., van der Valk, P., Hermens, H., & Vollenbroek-Hutten, M. (2014). A telehealth program for self-management of COPD exacerbations and promotion of an active lifestyle: A pilot randomized controlled trial. *International Journal of Chronic Obstructive Pulmonary disease*, 9, pp. 935–944. doi:10.2147/COPD.S60179
- Vardoulakis, L. P., Ring, L., Barry, B., Sidner, C. L., & Bickmore, T. (2012). Designing relational agents as long term social companions for older adults. *International Conference on Intelligent Virtual Agents* (pp. 289–302). doi:10.1007/978-3-642-33197-8_30
- 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
- van Vugt, H. C., Konijn, E. A., Hoorn, J. F., Keur, I., & Eliéns, A. (2007). Realism is not all! User engagement with task-related interface characters. *Interacting with Computers*, 19(2), 267–280. doi:10.1016/j.intcom.2006.08.005
- van Wissen, A., Vinkers, C., & van Halteren, A. (2016). Developing a virtual coach for chronic patients: A user study on the impact of similarity, familiarity and realism. *International Conference on Persuasive Technology* (pp. 263–275). doi:10.1007/978-3-319-31510-2_23
- World Health Organization. (2015). World Report on Ageing and Health. Retrieved from: <https://www.who.int/ageing/publications/world-report-2015/en>
- Yardley, L., Spring, B. J., Riper, H., Morrison, L. G., Crane, D. H., Curtis, K., ... Blandford, A. (2016). Understanding and promoting effective engagement with digital behavior change interventions. *American Journal of Preventive Medicine*, 51(5), 833–842. doi:10.1016/j.amepre.2016.06.015
- Yin, L., Bickmore, T., & Cortés, D. E. (2010). The impact of linguistic and cultural congruity on persuasion by conversational agents. *International Conference on Intelligent Virtual Agents* (pp. 343–349). doi:10.1007/978-3-642-15892-6_36
- Zhou, S., Bickmore, T., Paasche-Orlow, M., & Jack, B. (2014). Agent-user concordance and satisfaction with a virtual hospital discharge nurse. *International Conference on Intelligent Virtual Agents* (pp. 528–541). doi:10.1007/978-3-319-09767-1_63
- Zhou, S., Zhang, Z., & Bickmore, T. (2017). Adapting a persuasive conversational agent for the Chinese culture. *International Conference on Culture and Computing* (pp. 89–96). doi:10.1109/Culture.and.Computing.2017.42
- Zimmerman, J., Ayoob, E., Forlizzi, J., & McQuaid, M. (2005). Putting a face on embodied interface agents. *Proceedings of Designing Pleasurable Products and Interfaces* (pp. 233–248). doi:10.1184/R1/6470366.v1

Chapter 7

- Albaina, I. M., Visser, T., Van Der Mast, C. A., & Vastenburger, M. H. (2009). Flowie: A persuasive virtual coach to motivate elderly individuals to walk. *International Conference on Pervasive Computing Technologies for Healthcare* (pp. 1–7). doi:10.4108/ICST.PERVASIVEHEALTH2009.5949
- André, E., & Rist, T. (2001). Presenting through performing: on the use of multiple lifelike characters in knowledge-based presentation systems. *Knowledge-Based Systems*, 14(1–2), 3–13. doi:10.1016/S0950-7051(00)00096-4
- André, E., Rist, T., Van Mulken, S., Klesen, M., & Baldes, S. (2000). The automated design of believable dialogues for animated presentation teams. *Embodied conversational agents*, 220–255. doi:10.7551/mitpress/2697.003.0010
- Azevedo, R. F., Morrow, D., Graumlich, J., Willemsen-Dunlap, A., Hasegawa-Johnson, M., Huang, T. S., ... & Halpin, D. J. (2018). Using conversational agents to explain medication instructions to older adults. *AMIA Annual Symposium Proceedings* (pp. 185–194)
- Bar, M., Neta, M., & Linz, H. (2006). Very first impressions. *Emotion*, 6(2), 269–278. doi:10.1037/1528-3542.6.2.269
- 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. *International Conference on Intelligent Virtual Agents* (pp. 126–138). doi:10.1007/978-3-642-33197-8_13
- Beukema, S., van Velsen, L., Jansen-Kosterink, S., & Karreman, J. (2017). "There Is something we need to tell you...": Communicating health-screening results to older adults via the internet. *Telemedicine and e-health*, 23(9), 741–746. doi:10.1089/tmj.2016.0210
- Bickmore, T. W., Carusob, L., Clough-Gorrb, K., & Heeren, T. (2005). 'It's just like you talk to a friend' relational agents for older adults. *Interacting with Computers*, 17(6), 711–735. doi:10.1016/j.intcom.2005.09.002
- Bickmore, T. W., Silliman, R. A., Nelson, K., Cheng, D. M., Winter, M., Henault, L., & Paasche-Orlow, M. K. (2013). A randomized controlled trial of an automated exercise coach for older adults. *Journal of the American Geriatrics Society*, 61(10), 1676–1683. doi:10.1111/jgs.12449
- Chasteen, A. L., Schwarz, N., & Park, D. C. (2002). The activation of aging stereotypes in younger and older adults. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 57(6), P540–P547. doi:10.1093/geronb/57.6.P540
- Esposito, A., Amorese, T., Cuciniello, M., Esposito, A. M., Troncone, A., Torres, M. I., ... & Cordasco, G. (2018). Seniors' acceptance of virtual humanoid agents. *Italian Forum of Ambient Assisted Living* (pp. 429–443). doi:10.1007/978-3-030-05921-7_35
- Esposito, A., Amorese, T., Cuciniello, M., Riviello, M. T., Esposito, A. M., Troncone, A., ... & Cordasco, G. (2019). Elder user's attitude toward assistive virtual agents: the role of voice and gender. *Journal of Ambient Intelligence and Humanized Computing*. doi:10.1007/s12652-019-01423-x
- Forlizzi, J., Zimmerman, J., Mancuso, V., & Kwak, S. (2007). How interface agents affect interaction between humans and computers. *Proceedings of the 2007 Conference on Designing Pleasurable Products and Interfaces* (pp. 209–221). doi:10.1145/1314161.1314180
- Hayashi, Y., & Ogawa, H. (2012). Facilitating creative interpretations on collaboration with multiple conversational agents. *Proceedings of the Asia Pacific Conference on Computer Human Interaction* (pp. 443–449).

- Heerink, M., Kröse, B., Evers, V., & Wielinga, B. (2010). Assessing acceptance of assistive social agent technology by older adults: The Almere model. *International Journal of Social Robotics*, 2(4), 361–375. doi:10.1007/s12369-010-0068-5
- Huber, M., van Vliet, M., Giezenberg, M., Winkens, B., Heerkens, Y., Dagnelie, P. C., & Knottnerus, J. A. (2016). Towards a 'patient-centred' operationalisation of the new dynamic concept of health: A mixed methods study. *BMJ open*, 6(1). doi:10.1136/bmjopen-2015-010091
- Hurmuz, M. Z., Jansen-Kosterink, S. M., op den Akker, H., & Hermens, H. J. (2020). User experience and potential health effects of a conversational agent-based electronic health intervention: Protocol for an observational cohort study. *JMIR Research Protocols*, 9(4), e16641. doi:10.2196/16641
- Jansen-Kosterink, S., Vollenbroek Hutten, M. M., & Hermens, H. J. (2016). A renewed framework for the evaluation of telemedicine. *eTELEMED 2016: The Eighth International Conference on eHealth, Telemedicine, and Social Medicine* (pp. 57–62)
- Justo, R., Ben Letaifa, L., Palmero, C., Gonzalez-Fraile, E., Torp Johansen, A., Vázquez, A., ... & Escalera, S. (2020). Analysis of the interaction between elderly people and a simulated virtual coach. *Journal of Ambient Intelligence and Humanized Computing*. doi:10.1007/s12652-020-01983-3
- Kanharaju, R. B., De Franco, D., Pease, A., & Pelachaud, C. (2018). Is two better than one? Effects of multiple agents on user persuasion. *International Conference on Intelligent Virtual Agents* (pp. 255–262). doi:10.1145/3267851.3267890
- Kelley, H.H. (1950). The warm-cold variable in first impressions of persons. *Journal of Personality*, 18(4), 431–439. doi:10.1111/j.1467-6494.1950.tb01260.x
- Komatsu, T., Kurosawa, R., & Yamada, S. (2012). How does the difference between users' expectations and perceptions about a robotic agent affect their behavior?. *International Journal of Social Robotics*, 4(2), 109–116. doi:10.1007/s12369-011-0122-y
- Nguyen, H., & Masthoff, J. (2007). Is it me or is it what I say? Source image and persuasion. *International Conference on Persuasive Technology* (pp. 231–242). doi:10.1007/978-3-540-77006-0_29
- op den Akker, H., op den Akker, R., Beinema, T., Banos, O., Heylen, D., Bedsted, B., ... & Hermens, H. (2018). Council of coaches: a novel holistic behavior change coaching approach. *International Conference on Information and Communication Technologies for Ageing Well and e-Health, ICT4AWE* (pp. 219–226). doi:10.5220/0006787702190226
- Panico, F., Cordasco, G., Vogel, C., Trojano, L., & Esposito, A. (2020). Ethical issues in assistive ambient living technologies for ageing well. *Multimedia Tools and Applications*. doi:10.1007/s11042-020-09313-7
- 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. *International Conference on Intelligent Virtual Agents* (pp. 301–306). doi:10.1145/3267851.3267915
- Ring, L., Utami, D., & Bickmore, T. (2014). The right agent for the job? The effects of agent visual appearance on task domain. *International Conference on Intelligent Virtual Agents* (pp. 374–384). doi:10.1007/978-3-319-09767-1_49
- Ruttikay, 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*, pp. 27–66. doi:10.1007/1-4020-2730-3_2

- ter Stal S., Tabak, M., op den Akker, H., Beinema, T., Hermens, H. (2019). Who do you prefer? The effect of age, gender and role on users' first impressions of embodied conversational agents in eHealth. *International Journal of Human-Computer Interaction*, 36(9), 881–892. doi:10.1080/10447318.2019.1699744
- Troncone, A., Amorese, T., Cuciniello, M., Saturno, R., Pugliese, L., Cordasco, G., ... & Esposito, A. (2020). Advanced assistive technologies for elderly people: A psychological perspective on seniors' needs and preferences (part A). *Acta Polytechnica Hungarica*, 17(2), 163–189. doi:10.12700/APH.17.2.2020.2.10
- Tsiourti, C., Joly, E., Wings, C., Moussa, M. B., & Wac, K. (2014). Virtual assistive companions for older adults: Qualitative field study and design implications. *International Conference on Pervasive Computing Technologies for Healthcare* (pp. 57–64). doi:10.4108/icst.pervasive-health.2014.254943
- Vardoulakis, L. P., Ring, L., Barry, B., Sidner, C. L., & Bickmore, T. (2012). Designing relational agents as long term social companions for older adults. *International Conference on Intelligent Virtual Agents* (pp. 289–302). doi:10.1007/978-3-642-33197-8_30
- 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
- World Health Organization. (2015). *World Report on Ageing and Health*. Retrieved from: <https://www.who.int/ageing/publications/world-report-2015/en>

Chapter 8

- 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* (pp. 11–15). doi:10.1145/3057039.3057080
- Amini, R., Lisetti, C., Yasavur, U., & Rishe, N. (2013). On-demand virtual health counselor for delivering behavior-change health interventions. *IEEE International Conference on Healthcare Informatics* (pp. 46–55). doi:10.1109/ICHI.2013.13
- Amini, R., Lisetti, C., & Yasavur, U. (2014). Emotionally responsive virtual counselor for behavior-change health interventions. *International Conference on Design Science Research in Information Systems* (pp. 433–437). doi:10.1007/978-3-319-06701-8_40
- Azevedo, R. F., Morrow, D., Graumlich, J., Willemsen-Dunlap, A., Hasegawa-Johnson, M., Huang, T. S., ... & Halpin, D. J. (2018). Using conversational agents to explain medication instructions to older adults. *AMIA Annual Symposium Proceedings* (pp. 185–194)
- Bickmore, T., Gruber, A., & Picard, R. (2005). Establishing the computer-patient working alliance in automated health behavior change interventions. *Patient Education and Counseling*, 59(1), 21–30. doi:10.1016/j.pec.2004.09.008
- Bickmore, T. W., Picard, R. W. (2004). Towards caring machines. *CHI'04 Extended Abstracts on Human Factors in Computing Systems* (pp. 1489–1492). doi:10.1145/985921.986097

- Bickmore, T. W., Picard, R. W. (2005). Establishing and maintaining long-term human-computer relationships. *ACM Transactions on Computer-Human Interaction*, 12(2), 293–327. doi:10.1145/1067860.1067867
- Bickmore, T., & Ring, L. (2010). Making it personal: end-user authoring of health narratives delivered by virtual agents. *International Conference on Intelligent Virtual Agents* (pp. 399–405). doi:10.1007/978-3-642-15892-6_43
- Bickmore, T., & Schulman, D. (2007). Practical approaches to comforting users with relational agents. *Proceedings of ACM CHI 2007: Conference on Human Factors in Computing Systems* (pp. 2291–2296). doi:10.1145/1240866.1240996
- Bickmore, T., Schulman, D., & Yin, L. (2009). Engagement vs. deceit: Virtual humans with human autobiographies. *International Workshop on Intelligent Virtual Agents* (pp. 6–19). doi:10.1007/978-3-642-04380-2_4
- Bickmore, T., Schulman, D., & Yin, L. (2010). Maintaining engagement in long-term interventions with relational agents. *Applied Artificial Intelligence*, 24(6), 648–666. doi:10.1080/08839514.2010.492259
- Baylor, A. L., & Kim, Y. (2004). Pedagogical agent design: The impact of agent realism, gender, ethnicity, and instructional role. *International Conference on Intelligent Tutoring Systems* (pp. 592–603). 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. *International Conference on Intelligent Virtual Agents* (pp. 126–138). doi:10.1007/978-3-642-33197-8_13
- Canidate, S., & Hart, M. (2017). The use of avatar counseling for HIV/AIDS health education: The examination of self-identity in avatar preferences. *Journal of Medical Internet Research*, 19(12), e365. doi:10.2196/jmir.6740
- Cowell, A., & Stanney, K. (2003). Embodiment and interaction guidelines for designing credible, trustworthy embodied conversational agents. *International Workshop on Intelligent Virtual Agents* (pp. 301–309). doi:10.1007/978-3-540-39396-2_50
- Creed, C., & Beale, R. (2012). User interactions with an affective nutritional coach. *Interacting with Computers*, 24(5), 339–350. doi:10.1016/j.intcom.2012.05.004
- Creed, C., Beale, R., & Cowan, B. (2015). The impact of an embodied agent's emotional expressions over multiple interactions. *Interacting with Computers*, 27(2), 172–188. doi:10.1093/iwci/iwt064
- Crutzen, R., Cyr, D., & De Vries, N. K. (2011). Bringing loyalty to e-health: Theory validation using three internet-delivered interventions. *Journal of Medical Internet Research*, 13(3). doi:10.2196/jmir.1837
- Crutzen, R., van't Riet, J., & Short, C. E. (2016). Enjoyment: A conceptual exploration and overview of experimental evidence in the context of games for health. *Games for Health Journal*, 5(1), 15–20. doi:10.1089/g4h.2015.0059
- Ekman, P., & Friesen, W. (1978). *Facial Action Coding System: A technique for the measurement of facial movement*. Palo Alto, California: Consulting Psychologists Press
- Forlizzi, J., Zimmerman, J., Mancuso, V., & Kwak, S. (2007). How interface agents affect interaction between humans and computers. *Proceedings of the 2007 Conference on Designing Pleasurable Products and Interfaces* (pp. 209–221). doi:10.1145/1314161.1314180

- Frost, J., Boukris, N., & Roelofsma, P. (2012). We like to move it move it! Motivation and parasocial interaction. *CHI '12 Extended Abstracts on Human Factors in Computing Systems* (pp. 2465–2470). doi:10.1145/2212776.2223820
- Gambino, A., Kim, J., & Sundar, S. S. (2019). Digital doctors and robot receptionists: User attributes that predict acceptance of automation in healthcare facilities. *CHI Conference on Human Factors in Computing Systems* (pp. 1–6). doi:10.1145/3290607.3312916
- Grillon, H., & Thalmann, D. (2008). Eye contact as trigger for modification of virtual character behavior. *2008 Virtual Rehabilitation* (pp. 205–211). doi:10.1109/ICVR.2008.4625161
- Jansen-Kosterink, S., Vollenbroek Hutten, M. M., & Hermens, H. J. (2016). A renewed framework for the evaluation of telemedicine. *eTELEMED 2016: The Eighth International Conference on eHealth, Telemedicine, and Social Medicine* (pp. 57–62)
- Kang, S. H., & Gratch, J. (2011). People like virtual counselors that highly-disclose about themselves. *Annual Review of Cybertherapy and Telemedicine*, 167, pp. 143–148
- Kang, J., & Wei, L. (2018). "Give me the support I want!" The effect of matching an embodied conversational agent's social support to users' social support needs in fostering positive user-agent interaction. *International Conference on Human-Agent Interaction* (pp. 106–113). doi:10.1145/3284432.3284462
- 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
- Kelders, S. M., Kok, R. N., Ossebaard, H. C., & Van Gemert-Pijnen, J. E. W. C. (2012). Persuasive system design does matter: A systematic review of adherence to web-based interventions. *Journal of Medical Internet Research*, 14(6), 1–24. doi:10.2196/jmir.2104
- Kelley, H.H. (1950). The warm-cold variable in first impressions of persons. *Journal of Personality*, 18(4), 431–439. doi:10.1111/j.1467-6494.1950.tb01260.x
- Khan, R., & Angeli, A. D. (2009). The attractiveness stereotype in the evaluation of conversational agents. *IFIP Conference on Human-Computer Interaction* (pp. 85–97). doi:10.1007/978-3-642-03655-2_10
- Khan, R. F., & Sutcliffe, A. (2014). Attractive agents are more persuasive. *International Journal of Human-Computer Interaction*, 30(2), 142–150. doi:10.1080/10447318.2013.839904
- 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
- Kohl, L. F., Crutzen, R., & De Vries, N. K. (2013). Online prevention aimed at lifestyle behaviors: A systematic review of reviews. *Journal of Medical Internet Research*, 15(7), 1–13. doi:10.2196/jmir.2665
- Komatsu, T., Kurosawa, R., & Yamada, S. (2012). How does the difference between users' expectations and perceptions about a robotic agent affect their behavior?. *International Journal of Social Robotics*, 4(2), 109–116. doi:10.1007/s12369-011-0122-y
- Kramer, L. L., ter Stal, S., Mulder, B. C., de Vet, E., & van Velsen, L. (2019). Developing embodied conversational agents for coaching people towards a healthy lifestyle: A scoping review. *Journal of Medical Internet Research*, 22(2), e14058. doi:10.2196/14058
- Lee, J. E. R., Nass, C., Brave, S. B., Morishima, Y., Nakajima, H., & Yamada, R. (2007). The case for caring colearners: The effects of a computer-mediated colearner agent on trust and learning. *Journal of Communication*, 57(2), 183–204. doi:10.1111/j.1460-2466.2007.00339.x

- 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
- Lefebvre, C. R., Tada, Y., Hilfiker, S. W., & Baur, C. (2010). The assessment of user engagement with eHealth content: The eHealth engagement scale. *Journal of Computer-Mediated Communication*, 15(4), 666–681. doi:10.1111/j.1083-6101.2009.01514.x
- Lehto, T., & Oinas-Kukkonen, H. (2011). Persuasive features in web-based alcohol and smoking interventions: a systematic review of the literature. *Journal of medical Internet research*, 13(3), e46. doi:10.2196/jmir.1559
- Lisetti, C., Amini, R., Yasavur, U., & Rishe, N. (2013). I can help you change! An empathic virtual agent delivers behavior change health interventions. *ACM Transactions on Management Information Systems*, 4(4), 1–28. doi:10.1145/2544103
- Lohr, K. N. (2004). Rating the strength of scientific evidence: relevance for quality improvement programs. *International Journal for Quality in Health Care*, 16(1), 9–18. doi:10.1093/intqhc/mzh005
- Lucas, G. M., Boberg, J., Traum, D., Artstein, R., Gratch, J., Gainer, A., ... & Nakano, M. (2018). Getting to know each other: The role of social dialogue in recovery from errors in social robots. *International Conference on Human-Robot Interaction* (pp. 344–351). doi:10.1145/3171221.3171258
- Malhotra, A., Hoey, J., König, A., & van Vuuren, S. (2016). A study of elderly people's emotional understanding of prompts given by virtual humans. *Proceedings of the 10th EAI International Conference on Pervasive Computing Technologies for Healthcare* (pp. 13–16)
- McDonnell, R., Breidt, M., & Bühlhoff, H. H. (2012). Render me real? Investigating the effect of render style on the perception of animated virtual humans. *ACM Transactions on Graphics (TOG)*, 31(4), 1–11. doi:10.1145/2185520.2185587
- Mostajeran, F., Katzakis, N., Ariza, O., Freiwald, J. P., & Steinicke, F. (2019). Welcoming a holographic virtual coach for balance training at home: two focus groups with older adults. *IEEE Conference on Virtual Reality and 3D User Interfaces* (pp. 1465–1470). doi:10.1109/VR.2019.8797813
- Nguyen, H., & Masthoff, J. (2007). Is it me or is it what I say? Source image and persuasion. *International Conference on Persuasive Technology* (pp. 231–242). doi:10.1007/978-3-540-77006-0_29
- Nguyen, H., & Masthoff, J. (2009). Designing empathic computers: The effect of multi-modal empathic feedback using animated agent. *Proceedings of the 4th International Conference on Persuasive Technology* (pp. 1–9) doi:10.1145/1541948.1541958
- O'Brien, H. L., & Toms, E. G. (2008). What is user engagement? A conceptual framework for defining user engagement with technology. *Journal of the American Society for Information Science and Technology*, 59(6), 938–955. doi:10.1002/asi.20801
- Olafsson, S., Kimani, E., Asadi, R., Bickmore, T., & Science, I. (2017). That's a rap. *International Conference on Intelligent Virtual Agents* (pp. 325–334). doi:10.1007/978-3-319-67401-8_41
- 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. *International Conference on Intelligent Virtual Agents* (pp. 301–306). doi:10.1145/3267851.3267915
- Pelachaud, C. (2009). Studies on gesture expressivity for a virtual agent. *Speech Communication* 51(7), 630–639. doi:10.1016/j.specom.2008.04.009

- 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). doi:10.2196/jmir.6553
- Robertson, S., Solomon, R., Riedl, M., Gillespie, T. W., Chociemski, T., & Master, V. (2015). The visual design and implementation of an embodied conversational agent in a shared decision-making context (eCoach). *International Conference on Learning and Collaboration Technologies* (pp. 427–437). doi:10.1007/978-3-319-20609-7
- Ring, L., Utami, D., & Bickmore, T. (2014). The right agent for the job? The effects of agent visual appearance on task domain. *International Conference on Intelligent Virtual Agents* (pp. 374–384). doi:10.1007/978-3-319-09767-1_49
- 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
- Ruttkey, 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*, pp. 27–66. doi:10.1007/1-4020-2730-3_2
- Scherer, E. A., Ben-Zeev, D., Li, Z., & Kane, J. M. (2017). Analyzing mHealth engagement: joint models for intensively collected user engagement data. *JMIR mHealth and uHealth*, 5(1), e1. doi:10.2196/mhealth.6474
- Schmeil, A., & Suggs, S. (2014). "How am I doing?" - Personifying health through animated characters. *International Conference of Design, User Experience, and Usability* (pp. 91–102). doi:10.1007/978-3-319-07635-5_10
- Scholten, M. R., Kelders, S. M., van Gemert-Pijnen, J. E. (2017). Self-guided web-based interventions: Scoping review on user needs and the potential of embodied conversational agents to address them. *Journal of Medical Internet Research*, 19(11), 1–19. doi:10.2196/jmir.7351
- Silverman, B. G., Holmes, J., Kimmel, S., Branas, C., & Ivins, D. (2001). Modeling emotion and behavior in animated personas to facilitate human behavior change: the case of the HEART-SENSE game. *Health Care Management Science*, 4(3), 213–228. doi:10.1023/A:1011448916375
- Sin, J., & Munteanu, C. (2019). An information behaviour-based approach to virtual doctor design. *International Conference on Human-Computer Interaction with Mobile Devices and Services* (pp. 1–6). doi:10.1145/3338286.3344391
- Skalski, P., Tamborini, R., & Skalski, P. (2007). The role of social presence in interactive agent-based persuasion. *Media Psychology*, 10(3), 385–413. doi:10.1080/15213260701533102
- ter Stal, S., Broekhuis, M., van Velsen, L., Hermens, H., & Tabak, M. (2020). Embodied conversational agent appearance for health assessment of older adults: Explorative study. *JMIR Human Factors*, 7(3). doi:10.2196/19987
- ter Stal S., Tabak, M., op den Akker, H., Beinema, T., Hermens, H. (2019). Who do you prefer? The effect of age, gender and role on users' first impressions of embodied conversational agents in eHealth. *International Journal of Human-Computer Interaction*, 36(9), 881–892. doi:10.1080/10447318.2019.1699744
- Straßmann, C., & Krämer, N. C. (2017). A categorization of virtual agent appearances and a qualitative study on age-related user preferences. *International Conference on Intelligent Virtual Agents* (pp. 413–422). doi:10.1007/978-3-319-67401-8_51

- Straßmann, C., Krämer, N. C., Buschmeier, H., & Kopp, S. (2020). Age-related differences in the evaluation of a virtual health agent's appearance and embodiment in a health-related interaction: Experimental lab study. *Journal of Medical Internet Research*, 22(4), e13726. doi:10.2196/13726
- Tielman, M. L., Neerincx, M. A., & Brinkman, W. P. (2019). Design and evaluation of personalized motivational messages by a virtual agent that assists in post-traumatic stress disorder therapy. *Journal of Medical Internet Rresearch*, 21(3), e9240. doi:10.2196/jmir.9240
- Tielman, M. L., Neerincx, M. A., van Meggelen, M., Franken, I. (2017). How should a virtual agent present psychoeducation? Influence of verbal and textual presentation on adherence. *Technology and Health Care*, 25, pp. 1081–1096. doi:10.3233/THC-170899
- van Velsen, L., Broekhuis, M., Jansen-Kosterink, S., & op den Akker, H. (2019). Tailoring persuasive electronic health strategies for older adults on the basis of personal motivation: Web-based survey study. *Journal of Medical Internet Research*, 21(9), e11759. doi:10.2196/11759
- van Vugt, H. C., Konijn, E. A., Hoorn, J. F., Veldhuis, J. (2006). Why fat interface characters are better e-health advisors. *International Conference on Intelligent Virtual Agents* (pp. 1–13). doi:10.1007/11821830_1
- van Wissen, A., Vinkers, C., & van Halteren, A. (2016). Developing a virtual coach for chronic patients: A user study on the impact of similarity, familiarity and realism. *International Conference on Persuasive Technology* (pp. 263–275). doi:10.1007/978-3-319-31510-2_23
- 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
- Veletsianos, G., Miller, C., & Doering, A. (2009). Enali: A research and design framework for virtual characters and pedagogical agents. *Journal of Educational Computing Research*, 41(2), 171–194. doi:10.2190/EC.41.2.c
- von der Pütten, A. M., Krämer, N. C., & Gratch, J. (2009). Who's there? Can a virtual agent really elicit social presence? *Proceedings of the PRESENCE 2009 - The 12th Annual International Workshop on Presence*
- Yardley, L., Spring, B. J., Riper, H., Morrison, L. G., Crane, D. H., Curtis, K., ... Blandford, A. (2016). Understanding and promoting effective engagement with digital behavior change interventions. *American Journal of Preventive Medicine*, 51(5), 833–842. doi:10.1016/j.amepre.2016.06.015
- Yin, L., Bickmore, T., & Cortés, D. E. (2010). The impact of linguistic and cultural congruity on persuasion by conversational agents. *International Conference on Intelligent Virtual Agents* (pp. 343–349). doi:10.1007/978-3-642-15892-6_36
- Zalake, M., Woodward, J., Kapoor, A., & Lok, B. (2018). Assessing the impact of virtual human's appearance on users' trust levels. *International Conference on Intelligent Virtual Agents* (pp. 329–330). doi:10.1145/3267851.3267863
- Zalake, M., Tavassoli, F., Griffin, L., Krieger, J., & Lok, B. (2019). Internet-based tailored virtual human health intervention to promote colorectal cancer screening: Design guidelines from two user studies. *International Conference on Intelligent Virtual Agents* (pp. 73–80). doi:10.1145/3308532.3329471

- Zell, E., Aliaga, C., Jarabo, A., Zibrek, K., Gutierrez, D., McDonnell, R., & Botsch, M. (2015). To stylize or not to stylize? The effect of shape and material stylization on the perception of computer-generated faces. *ACM Transactions on Graphics (TOG)*, 34(6), 1–12. doi:10.1145/2816795.2818126
- Zhou, S., Bickmore, T., Paasche-Orlow, M., & Jack, B. (2014). Agent-user concordance and satisfaction with a virtual hospital discharge nurse. *International Conference on Intelligent Virtual Agents* (pp. 528–541). doi:10.1007/978-3-319-09767-1_63
- Zhou, S., Zhang, Z., & Bickmore, T. (2017). Adapting a persuasive conversational agent for the Chinese culture. *International Conference on Culture and Computing* (pp. 89–96). doi:10.1109/Culture.and.Computing.2017.42
- Zibrek, K., Kokkinara, E., & McDonnell, R. (2018). The effect of realistic appearance of virtual characters in immersive environments-does the character's personality play a role?. *IEEE Transactions on Visualization and Computer Graphics*, 24(4), 1681–1690. doi:10.1109/TVCG.2018.2794638
- Zimmerman, J., Ayoob, E., Forlizzi, J., & McQuaid, M. (2005). Putting a face on embodied interface agents. *Proceedings of Designing Pleasurable Products and Interfaces* (pp. 233–248). doi:10.1184/R1/6470366.v1

Chapter 9

- Acosta, J. C., & Ward, N. G. (2011). Achieving rapport with turn-by-turn, user-responsive emotional coloring. *Speech Communication*, 53(9–10), 1137–1148. doi:10.1016/j.specom.2010.11.006
- 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* (pp. 11–15). doi:10.1145/3057039.3057080
- 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
- Beale, R., & Creed, C. (2009). Affective interaction: How emotional agents affect users. *International Journal of Human Computer Studies*, 67(9), 755–776. doi:10.1016/j.ijhcs.2009.05.001
- 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., Schulman, D., & Yin, L. (2009). Engagement vs. deceit: Virtual humans with human autobiographies. *International Workshop on Intelligent Virtual Agents* (pp. 6–19). doi:10.1007/978-3-642-04380-2_4
- Bickmore, T., Schulman, D., & Yin, L. (2010). Maintaining engagement in long-term interventions with relational agents. *Applied Artificial Intelligence*, 24(6), 648–666. doi:10.1080/08839514.2010.492259
- Bickmore, T. W., Carusob, L., Clough-Gorrb, K., & Heeren, T. (2005). 'It's just like you talk to a friend' relational agents for older adults. *Interacting with Computers*, 17(6), 711–735. doi:10.1016/j.intcom.2005.09.002

- Bickmore, T. W., Puskar, K., Schlenk, E. A., Pfeifer, L. M., & Sereika, S. M. (2010). Maintaining reality: relational agents for antipsychotic medication adherence. *Interacting with Computers*, 22(4), 276–288. doi:10.1016/j.intcom.2010.02.001
- Bickmore, T. W., Silliman, R. A., Nelson, K., Cheng, D. M., Winter, M., Henault, L., & Paasche-Orlow, M. K. (2013). A randomized controlled trial of an automated exercise coach for older adults. *Journal of the American Geriatrics Society*, 61(10), 1676–1683. doi:10.1111/jgs.12449
- Burbach, L., Halbach, P., Plettenberg, N., Nakayama, J., Ziefle, M., & Valdez, A. C. (2019). "Hey, Siri", "Ok, Google", "Alexa". Acceptance-relevant factors of virtual voice-assistants. *International Professional Communication Conference* (pp. 101–111). doi:10.1109/ProComm.2019.00025
- Cole-Lewis, H., Ezeanochie, N., & Turgiss, J. (2019). Understanding health behavior technology-engagement: Pathway to measuring digital behavior change interventions. *JMIR Formative Research*, 3(4), e14052. doi:10.2196/14052
- Cowell, A., & Stanney, K. (2003). Embodiment and interaction guidelines for designing credible, trustworthy embodied conversational agents. *International Workshop on Intelligent Virtual Agents* (pp. 301–309). doi:10.1007/978-3-540-39396-2_50
- Creed, C., Beale, R., & Cowan, B. (2015). The impact of an embodied agent's emotional expressions over multiple interactions. *Interacting with Computers*, 27(2), 172–188. doi:10.1093/iwc/iwt064
- de Filippis, M. L., Federici, S., Mele, M. L., Borsci, S., Bracalenti, M., Gaudino, G., ... & Simonetti, E. (2020). Preliminary results of a systematic review: Quality assessment of conversational agents (chatbots) for people with disabilities or special needs. *International Conference on Computers Helping People with Special Needs* (pp. 250–257). doi:10.1007/978-3-030-58796-3_30
- European Commission. (2016). *Open Science, Open to the World – A Vision for Europe*
- European Commission. (2014). *Horizon 2020 – Work Programme 2014-2015 General Annexes*. Retrieved from: http://ec.europa.eu/research/participants/data/ref/h2020/wp/2014_2015/annexes/h2020-wp1415-annex-g-trl_en.pdf
- Feng, C., Yuguo, J., Guiqin, L., & Zhiyuan, G. (2019). Ethical dilemma and countermeasure in artificial intelligence engineering. *Proceedings of the 2019 International Conference on Modern Educational Technology* (pp. 111–114). doi:10.1145/3341042.3341061
- Fitriane, S., Buijnes, M., Richards, D., Abdulrahman, A., & Brinkman, W. P. (2019). What are we measuring anyway? A literature survey of questionnaires used in studies reported in the intelligent virtual agent conferences. *International Conference on Intelligent Virtual Agents* (pp. 159–161). doi:10.1145/3308532.3329421
- Forlizzi, J., Zimmerman, J., Mancuso, V., & Kwak, S. (2007). How interface agents affect interaction between humans and computers. *Proceedings of the 2007 Conference on Designing Pleasurable Products and Interfaces* (pp. 209–221). doi:10.1145/1314161.1314180
- Fruchter, N., & Liccardi, I. (2018). Consumer attitudes towards privacy and security in home assistants. *CHI Conference on Human Factors in Computing Systems* (pp. 1–6). doi:10.1145/3170427.3188448
- Guadagno, R., Blascovich, J., Bailenson, J., & McCall, C. (2007). Virtual humans and persuasion: The effects of agency and behavioral realism. *Media Psychology*, 10(1), 1–22. doi:10.1109/ICVR.2008.4625161

- Jansen-Kosterink, S., Vollenbroek Hutten, M. M., & Hermens, H. J. (2016). A renewed framework for the evaluation of telemedicine. *eTELEMED 2016: The Eighth International Conference on eHealth, Telemedicine, and Social Medicine* (pp. 57–62)
- Justo, R., Ben Letaifa, L., Palmero, C., Gonzalez-Fraile, E., Torp Johansen, A., Vázquez, A., ... & Escalera, S. (2020). Analysis of the interaction between elderly people and a simulated virtual coach. *Journal of Ambient Intelligence and Humanized Computing*. doi:10.1007/s12652-020-01983-3
- Kelders, S. M., & Kip, H. (2019). Development and initial validation of a scale to measure engagement with eHealth technologies. *CHI Conference on Human Factors in Computing Systems* (pp. 1–6). doi:10.1145/3290607.3312917
- 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
- Kolasinska, A., Lauriola, I., & Quadrio, G. (2019). Do people believe in Artificial Intelligence? A cross-topic multicultural study. *International Conference on Smart Objects and Technologies for Social Good* (pp. 31–36). doi:10.1145/3342428.3342667
- Komatsu, T., Kurosawa, R., & Yamada, S. (2012). How does the difference between users' expectations and perceptions about a robotic agent affect their behavior?. *International Journal of Social Robotics*, 4(2), 109–116. doi:10.1007/s12369-011-0122-y
- Kramer, L. L., ter Stal, S., Mulder, B. C., de Vet, E., & van Velsen, L. (2019). Developing embodied conversational agents for coaching people towards a healthy lifestyle: A scoping review. *Journal of Medical Internet Research*, 22(2), e14058. doi: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., & Coiera, E. (2018). Conversational agents in healthcare: A systematic review. *Journal of the American Medical Informatics Association*, 25(9), 1248–1258. doi:10.1093/jamia/ocy072
- Lazanyi, K. (2018). Readiness for Artificial Intelligence. *International Symposium on Intelligent Systems and Informatics* (pp. 235–238). doi:10.1109/SISY.2018.8524740
- Lee, J. E. R., Nass, C., Brave, S. B., Morishima, Y., Nakajima, H., & Yamada, R. (2007). The case for caring colearners: The effects of a computer-mediated colearner agent on trust and learning. *Journal of Communication*, 57(2), 183–204. doi:10.1111/j.1460-2466.2007.00339.x
- 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
- Loveys, K., Sebaratnam, G., Sagar, M., & Broadbent, E. (2020). The effect of design features on relationship quality with Embodied Conversational Agents: A systematic review. *International Journal of Social Robotics*, pp. 1–20. doi:10.1007/s12369-020-00680-7
- Nguyen, H., & Masthoff, J. (2007). Is it me or is it what I say? Source image and persuasion. *International Conference on Persuasive Technology* (pp. 231–242). doi:10.1007/978-3-540-77006-0_29
- O'Brien, H. L., & Toms, E. G. (2008). What is user engagement? A conceptual framework for defining user engagement with technology. *Journal of the American Society for Information Science and Technology*, 59(6), 938–955. doi:10.1002/asi.20801

- Perski, O., Blandford, A., Garnett, C., Crane, D., West, R., & Michie, S. (2020). A self-report measure of engagement with digital behavior change interventions (DBCIs): Development and psychometric evaluation of the "DBCI Engagement Scale". *Translational Behavioral Medicine*, 10(1), 267–277. doi:10.1093/tbm/ibz039
- Perski, O., Blandford, A., West, R., & Michie, S. (2017). Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. *Translational Behavioral Medicine*, 7(2), 254–267. doi:10.1007/s13142-016-0453-1
- 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). doi:10.2196/jmir.6553
- Reinhardt, J., Hillen, L., & Wolf, K. (2020). Embedding conversational agents into AR: Invisible or with a realistic human Body?. *International Conference on Tangible, Embedded, and Embodied Interaction* (pp. 299–310). doi:10.1145/3374920.3374956
- Rose, T. M., & Unni, E. J. (2020). A pilot evaluation of I'M HAPPY: An interactive module to Halt abuse of prescriptions in preteens and youth. *Games for Health Journal*, 9(4), 273–278. doi:10.1089/g4h.2019.0037
- Ruttkey, 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*, pp. 27–66. doi:10.1007/1-4020-2730-3_2
- Scholten, M. R., Kelders, S. M., van Gemert-Pijnen, J. E. (2017). Self-guided web-based interventions: Scoping review on user needs and the potential of embodied conversational agents to address them. *Journal of Medical Internet Research*, 19(11), 1–19. doi:10.2196/jmir.7351
- Short, C. E., DeSmet, A., Woods, C., Williams, S. L., Maher, C., Middelweerd, A., ... & Hingle, M. D. (2018). Measuring engagement in eHealth and mHealth behavior change interventions: viewpoint of methodologies. *Journal of Medical Internet Research*, 20(11), e292. doi:10.2196/jmir.9397
- Straßmann, C., & Krämer, N. C. (2017). A categorization of virtual agent appearances and a qualitative study on age-related user preferences. *International Conference on Intelligent Virtual Agents* (pp. 413–422). doi:10.1007/978-3-319-67401-8_51
- Tabak, M., de Vette, F., van Dijk, H., & Vollenbroek-Hutten, M. (2020). A game-based, physical activity coaching application for older adults: Design approach and user experience in daily life. *Games for Health Journal*, 9(3), 215–226. doi:10.1089/g4h.2018.0163
- Vaidyam, A. N., Wisniewski, H., Halamka, J. D., Kashavan, M. S., & Torous, J. B. (2019). Chatbots and conversational agents in mental health: a review of the psychiatric landscape. *The Canadian Journal of Psychiatry*, 64(7), 456–464. doi:10.1177/0706743719828977
- Veletsianos, G., Miller, C., & Doering, A. (2009). Enali: A research and design framework for virtual characters and pedagogical agents. *Journal of Educational Computing Research*, 41(2), 171–194. doi:10.2190/EC.41.2.c
- Vollmer, A. L. (2018). Fears of intelligent robots. *International Conference on Human-Robot Interaction* (pp. 273–274). doi:10.1145/3173386.3177067
- van Wissen, A., Vinkers, C., & van Halteren, A. (2016). Developing a virtual coach for chronic patients: A user study on the impact of similarity, familiarity and realism. *International Conference on Persuasive Technology* (pp. 263–275). doi:10.1007/978-3-319-31510-2_23
- Yardley, L., Spring, B. J., Riper, H., Morrison, L. G., Crane, D. H., Curtis, K., ... Blandford, A. (2016). Understanding and promoting effective engagement with digital behavior change interventions. *American Journal of Preventive Medicine*, 51(5), 833–842. doi:10.1016/j.amepre.2016.06.015

- Zalake, M., de Siqueira, A. G., Vaddiparti, K., Antonenko, P., Hamza-Lup, F., & Lok, B. (2020). Towards rapid development of conversational virtual humans using Web3D technologies. *International Conference on 3D Web Technology* (pp. 1–2). doi:10.1145/3424616.3424727
- Zhou, S., Bickmore, T., Paasche-Orlow, M., & Jack, B. (2014). Agent-user concordance and satisfaction with a virtual hospital discharge nurse. *International Conference on Intelligent Virtual Agents* (pp. 528–541). doi:10.1007/978-3-319-09767-1_63
- Zimmerman, J., Ayoob, E., Forlizzi, J., & McQuaid, M. (2005). Putting a face on embodied interface agents. *Proceedings of Designing Pleasurable Products and Interfaces* (pp. 233–248). doi:10.1184/R1/6470366.v1

SUMMARY

As we live longer and more people develop chronic diseases, the demand for care increases. The focus of our health care has to shift from curing acute complaints to (secondary) prevention of complaints via long-term care and coaching. A solution widely investigated to offer this long-term care and coaching is eHealth. But, for eHealth applications to be effective, use of the applications is crucial. In that sense, engagement with an eHealth application is necessary; those who are more engaged are significantly less likely to stop using it. To engage users in long-term care and coaching, Embodied Conversational Agents (ECAs) can be used. As humans, we immediately form ideas about such ECAs based on their appearance when we interact with them for the first time, similar to what we do when interacting with another human for the first time. Then, when we have a positive first impression of another human, we tend to interact more with that human. This might also apply to our interaction with ECAs. Therefore, ECAs could stimulate users to start or continue interacting with the underlying eHealth application (at first glance or after short interaction respectively) to eventually reach long-term interaction. By researching how an ECA's appearance should be designed to trigger positive impressions in these different phases of interaction, we can stimulate users to continue to next phase and eventually reach long-term engagement.

The main aim of this thesis was:



Main objective

How to design the appearance of an ECA in eHealth to promote user engagement?

The sub objectives of this thesis were:



Subobjectives

- 1. How to design the appearance of an ECA in eHealth for positive impressions at first glance?**
- 2. How to design the appearance of an ECA in eHealth for positive impressions after short interaction?**
- 3. How to design the appearance of an ECA in eHealth for positive impressions after long-term interaction?**

Chapter 2 investigated the state-of-the-art research on an ECA's appearance in eHealth. The chapter presented a literature review identifying the researched design features for ECAs in eHealth, the outcome variables that were used to measure the effect of these design features and what the found effects for each variable were. Results showed that an ECA's speech and/or textual output and its facial and gaze expressions were the most common design features. Little research was performed on the ECA's looks. The measured effect of these design features was often on the perception of the ECA's and user's characteristics, relation with the ECA, system usage, intention to use, usability and behaviour change. Furthermore, results showed that emotion and relational behaviour seem to positively affect the perception of the ECA's characteristics and that relational behaviour also seems to positively affect the relation with the ECA, usability and intention to use. However, results showed that these design features do not necessarily lead to behaviour change. Overall, chapter 2 showed that consensus on design features of ECAs in eHealth is far from established.

As chapter 2 showed that little research is performed on the ECA's looks, **chapter 3** explored users' perceptions of ECAs with different looks. The chapter described effects of an ECA's age (young or old), gender (male or female) and role (peer or expert) on perceptions of the ECAs at first glance (sub objective 1). In an online survey, participants were shown a set of eight static ECA images, positioned as coaches on healthy living. The chapter investigated which ECAs participants preferred at first glance, how they perceived the characteristics of these ECAs, and how likely they were to follow the advice of these ECAs. Results showed that people do have preferences for particular ECAs at first glance, and do rate ECAs different depending on the ECA's age, gender and role. 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 3) people seem to prefer an agent image that is similar in a) age and b) gender.

Chapter 4 investigated whether similar findings could be found after short interaction with an ECA (sub objective 2). Three validated health assessment questionnaires were translated into dialogues between a user and an ECA. Participants interacted with the ECAs in a lab setting. In half of the dialogues, participants interacted with a young, female peer ECA and in the other half of the dialogues with an older, male peer ECA. Results showed that the older male ECA was perceived as more authoritative than the young female ECA, but that no other differences were found between the ECAs. Besides effects of an ECA's looks, this thesis investigated effects of an ECA's emotion after short interaction. **Chapter 5** presented a study in which an ECA's emotions expressed in text and an ECA's emotions in facial expressions were compared on their effect on users' perceptions of rapport. A happy or neutral facial expression was combined with a happy or neutral textual expression. The ECAs were presented to the participants in an online survey and were positioned as health coaches on healthy living. Participants interacted with one of the four coaches. Results showed no significant difference in overall rapport between the conditions, but that a happy textual expression for an ECA with a neutral facial expression resulted in higher ratings of the individual rapport items helpfulness and enjoyableness.

Chapter 6 and 7 described experimental studies on long-term interaction with an ECA in daily life. In these chapters, an ECA's appearance was researched in the actual context of use (sub objective 3). **Chapter 6** presented a study in which an ECA was developed within an eHealth self-management intervention for patients with both Chronic Obstructive Pulmonary Disease (COPD) and Chronic Heart Failure (CHF), which was offered for four months. Patients interacted with the application at home.

Results showed that patients' perceptions of the ECA's characteristics did not change over time, but that after three weeks of use, the patients were less likely to follow the ECA's advice. The ECA's messages were perceived as non-personalised and the feedback as inappropriate, affecting the ECA's perceived reliability. **Chapter 7** described a study in which an ECA's appearance was researched in a multi-agent eHealth application that provided holistic coaching on healthy ageing. Participants interacted with the application at home. In line with findings of chapter 3, the chapter explored whether, in such an application, older adults are more satisfied with ECAs that are similar to themselves in age and gender, both at first glance and after long-term use. Results did not confirm this hypothesis. In addition, results showed that preferences for particular ECAs change over time.

Lastly, **chapter 8** presented a set of design strategies for an ECA's appearance in eHealth to contribute to user engagement. The initial literature review on the state-of-the-art research on an ECA's appearance for eHealth was updated in this chapter. Results showed that current studies on an ECA's appearance for eHealth are of low, fair or moderate quality. For the majority of the design features researched no or preliminary evidence exists for effects on user engagement. Only for an ECA's relational, empathic behaviour research shows moderate evidence and for an ECA's emotion, agency, rendering style, clothing, body shape, gender, age and cultural tailoring research shows preliminary evidence. The findings of the state-of-the-art with at least preliminary evidence were synthesised into the following design strategies: 1) adapt the ECA's demographics to the user's preferences, 2) adapt the human-likeness of the ECA to its functionalities, 3) convey an ECA's expertise using its appearance, 4) implement emotion in the ECA's expressions, and 5) implement an ECA's relational behaviour. These strategies should be validated, updated and extended in future research. Lastly, the chapter showed how these strategies can be used by an implementation example of an ECA as part of a mobile physical activity application.

SAMENVATTING

De vraag naar zorg toe neemt toe, omdat mensen langer leven en meer chronische ziektes ontwikkelen. De focus van onze gezondheidszorg moet daarom verschuiven van het genezen van acute klachten naar (secundaire) preventie van klachten, door middel van langdurige zorg en begeleiding. Een oplossing die hiervoor veel wordt onderzocht is eHealth (zorg op afstand met behulp van ICT). Om te zorgen dat een eHealth-applicatie hierin een positief effect heeft, is het belangrijk dat deze langdurig gebruikt wordt. In dat opzicht is 'engagement' (betrokkenheid) bij eHealth-applicaties noodzakelijk: mensen die meer betrokken zijn, zullen beduidend minder snel stoppen. Om gebruikers te betrekken bij langdurige zorg en begeleiding kunnen 'Embodied Conversational Agents (ECAs)' (virtuele karakters) worden ingezet. De eerste keer dat er contact gelegd wordt met zulke ECAs, vormen we direct een beeld op basis van hun uitstraling. Dit is vergelijkbaar met hoe we gewoonlijk van mensen een beeld vormen. Als we een positieve eerste indruk hebben van iemand zijn we geneigd meer met deze persoon te interacteren. Wellicht geldt dit ook voor onze interactie met ECAs. ECAs kunnen daarom mensen stimuleren de onderliggende eHealth-applicatie langdurig te blijven gebruiken. Door te onderzoeken hoe de uitstraling van een ECA bijdraagt aan positieve indruk van deze ECA, stimuleren we de betrokkenheid en het gebruik van de ECA en de onderliggende eHealth-applicatie.

Het hoofddoel van dit proefschrift was:



Hoofddoel

Hoe kunnen we de uitstraling van een ECA in een eHealth-applicatie ontwerpen om de betrokkenheid van gebruikers te stimuleren?

De subdoelen van dit proefschrift waren:



Subdoelen

1. Hoe kunnen we de uitstraling van een ECA in een eHealth-applicatie ontwerpen, zodat deze bij gebruikers in eerste instantie een positieve indruk oproept?
2. Hoe kunnen we de uitstraling van een ECA in een eHealth-applicatie ontwerpen, zodat deze na een korte interactie bij gebruikers een positieve indruk oproept?
3. Hoe kunnen we de uitstraling van een ECA in een eHealth-applicatie ontwerpen, zodat deze na langdurige interactie bij gebruikers een positieve indruk oproept?

In **hoofdstuk 2** richtte het onderzoek zich op het in kaart brengen van bestaand onderzoek naar uitsluiting van ECAs in eHealth. Dit literatuuronderzoek bestudeerde de onderzochte ontwerpkenmerken voor ECAs in eHealth, de onderzochte uitkomstmaten gebruikt om het effect van deze kenmerken te meten en de gevonden effecten van deze ontwerpkenmerken. Resultaten lieten zien dat ECA's spraak en/of tekst en gezichtsuitdrukkingen de meest voorkomende ontwerpkenmerken zijn. Er is weinig onderzoek gedaan naar een ECA's uiterlijk. Uitkomstmaten waarop het effect van ontwerpkenmerken vaak werd gemeten waren: de perceptie van de ECA's kenmerken, de relatie met de ECA, systeemgebruik, gebruiksintentie, gebruiksvriendelijkheid en gedragsverandering. Daarnaast lijken emotie en relationeel gedrag een positief effect te hebben op de perceptie van de ECA's kenmerken. Relationeel gedrag lijkt ook een positief effect te hebben op de relatie met de ECA, gebruiksvriendelijkheid en gebruikersintentie. De resultaten laten echter zien dat deze kenmerken niet zo zeer leiden tot gedragsverandering. Samenvattend blijkt uit dit hoofdstuk dat er nog lang geen consensus is over de ontwerpkenmerken van ECA's in eHealth.

Uit hoofdstuk 2 bleek dat er weinig onderzoek is gedaan naar het uiterlijk van ECAs. Als gevolg daarvan is in **hoofdstuk 3** onderzoek gedaan naar hoe gebruikers het uiterlijk van ECAs waarnemen. In dit hoofdstuk werd onderzocht hoe gebruikers in eerste instantie ECAs leeftijd (jong of oud), geslacht (man of vrouw) en expertise (gelijke bekwaamheid of expert) waarnemen (subdoel 1). In een online vragenlijst kregen deelnemers acht afbeeldingen van ECAs te zien. Voorafgaand werden de deelnemers geïnformeerd dat deze ECAs coaches zijn op het gebied van gezond leven. Er werd onderzocht welke ECAs in eerste instantie de voorkeur kregen, hoe de kenmerken van deze ECAs overkwamen en hoe waarschijnlijk ze advies van deze ECAs zouden opvolgen. Resultaten lieten zien dat mensen in eerste instantie wel degelijk voorkeur hebben voor bepaalde ECAs en dat ze ECAs anders beoordelen op grond van leeftijd, geslacht en rol. Voor zowel een algemene als oudere populatie geldt: 1) dat men de voorkeur lijkt te geven aan afbeeldingen van jonge, vrouwelijke ECAs boven oude, mannelijke ECAs; 2) dat men kenmerken van ECAs anders lijkt waar te nemen en dat de waarschijnlijkheid van het opvolgen van het advies anders is op grond van (a) leeftijd, (b) geslacht en (c) rol; en 3) dat men de voorkeur lijkt te geven aan een afbeelding van een ECA die vergelijkbaar is in (a) leeftijd en (b) geslacht.

In **hoofdstuk 4** werd onderzocht of vergelijkbare resultaten konden worden gevonden na korte interactie met een ECA (subdoel 2). Drie gevalideerde vragenlijsten werden vertaald in dialogen tussen een gebruiker en een ECA. Deelnemers hadden interactie met de ECAs in een lab-omgeving. De helft van de dialogen werd gevoerd met een jonge, vrouwelijke ECA met gelijke expertise, de andere helft van de dialogen met een oudere, mannelijke ECA, ook met gelijke expertise. De oudere, mannelijke ECA werd als meer autoritair gezien dan de jonge, vrouwelijke ECA. Verder werden er geen verschillen tussen de ECAs gevonden. Naast effecten van het uiterlijk van ECAs onderzocht dit proefschrift effecten van emotie van ECAs na korte interactie. In **hoofdstuk 5** werd vergeleken hoe een ECA's emotie in tekst en gezichtsuitdrukkingen effect had op hoe de gebruiker de verstandhouding met de ECA ervaarde. Een vrolijke of neutrale gezichtsuitdrukking was gecombineerd met een vrolijke of neutrale tekst, resulterend in vier onderzoekscondities. In een online vragenlijst werden de ECAs als coaches op het gebied van gezond leven voorgelegd aan de deelnemers. Deelnemers hadden contact met een van de vier coaches. Resultaten lieten zien dat er geen significant verschil was tussen de verschillende condities op hoe de gebruiker de verstandhouding met de ECA ervaarde. Een ECA met een vrolijke tekst en een neutrale gezichtsuitdrukking resulteerde wel in een hogere beoordeling van behulpzaamheid en plezier (sub-items van verstandhouding).

Hoofdstuk 6 en 7 beschreven experimentele studies over langdurige interactie met ECAs. In deze hoofdstukken werd het uiterlijk van een ECA onderzocht bij interactie met een ECA in het dagelijks leven (subdoel 3). **Hoofdstuk 6** beschreef een studie waarin een ECA werd ontwikkeld binnen een eHealth zelf-managementinterventie, die voor vier maanden werd aangeboden aan patiënten met zowel chronisch obstructieve longziekte (COPD) als chronisch hartfalen (CHF). Patiënten gebruikten de applicatie thuis. Resultaten toonden aan dat de waarneming van patiënten van de ECA's kenmerken niet veranderde gedurende de vier maanden, maar dat de patiënten na drie weken gebruik minder geneigd waren advies van de ECA op te volgen. Deelnemers zagen de berichten van de ECA als niet-gepersonaliseerd en de ECA's feedback als ongepast, wat de betrouwbaarheid van de ECA aantastte. **Hoofdstuk 7** beschreef een studie waarin het uiterlijk van een ECA werd onderzocht in een multi-agent eHealth-applicatie – een applicatie met meerdere ECAs – die holistische begeleiding bij gezond ouder worden aanbod. De deelnemers gebruikten de applicatie thuis. In lijn met de bevindingen van hoofdstuk 3 werd onderzocht of ouderen in een dergelijke toepassing in eerste instantie en na langdurig gebruik meer tevreden zijn met ECAs die vergelijkbaar zijn met henzelf. De resultaten bevestigden deze hypothese niet. Bovendien toonden de resultaten aan dat de voorkeuren voor bepaalde ECAs veranderden in de loop van de tijd.

Tot slot werd in **hoofdstuk 8** een reeks ontwerpstrategieën voor een ECA's uitstraling in eHealth beschreven om bij te dragen aan de betrokkenheid van de gebruikers. De oorspronkelijke literatuurstudie naar de stand van zaken over uitstraling van ECAs in eHealth werd in dit hoofdstuk geactualiseerd. Resultaten toonden aan dat de huidige onderzoeken naar de uitstraling van ECAs in eHealth van lage, redelijke of gemiddelde kwaliteit zijn. Voor het merendeel van de onderzochte ontwerpkenmerken bestaat er geen of slechts voorlopig bewijs voor effecten op betrokkenheid. Alleen onderzoek naar een ECA's relationeel, empathisch gedrag levert gemiddeld bewijs en onderzoek naar een ECA's emotie, representatie, stijl, kleding, lichaamsvorm, geslacht, leeftijd en culturele aanpassing levert een zekere mate van bewijs. De bevindingen van bestaand onderzoek met ten minste voorlopig bewijs werden gesynthetiseerd in de volgende ontwerpstrategieën: 1) pas de demografische kenmerken van een ECA aan aan de voorkeuren van de gebruiker, 2) pas de menselijke gelijkenis van een ECA aan aan zijn/haar functionaliteiten, 3) gebruik de uitstraling van een ECA om zijn/haar expertise over te brengen, 4) implementeer emotie in ECA's uitdrukkingen en 5) implementeer een ECA met relationeel gedrag. Deze strategieën moeten worden gevalideerd, bijgewerkt en uitgebreid in toekomstig onderzoek. Tot slot werd in het hoofdstuk laten zien hoe deze strategieën kunnen worden gebruikt door middel van een implementatievoorbeeld van een ECA als onderdeel van een mobiele, fysieke activiteitenapplicatie.

DANKWOORD

Het zit erop! Na 3,5 jaar hard werken is mijn proefschrift klaar. Wat een zoektocht is het geweest. Ik schrijf dit dankwoord in tijden van de lockdown naar aanleiding van COVID-19, zittend aan de eettafel, annex mijn bureau. Een geschikte tijd om te reflecteren op mijn promotietraject.

Terugkijkend, kan ik zeggen dat ik drie ontzettend leuke en leerzame jaren heb gehad bij RRD. Ik heb mogen meewerken aan het organiseren van open dagen, focusgroepen en co-creatiesessies. Ik heb mijn kennis mogen verrijken door bezoek aan symposia en het volgen van vele cursussen en meegewerkt in verschillende nationale en Europese projecten. Uiteraard mag ik de buitenlandse tripjes – inclusief een whisky-proeverij op een klassiek schip in Dundee, ontbijt met chocoladetaart in Madrid en een boottocht over de prachtige Seine in Parijs – niet vergeten. Wat een voorrecht! Deze bijzondere ervaringen hebben me zeker verrijkt en richting gegeven aan mijn toekomstige carrière. Niet alleen heb ik nieuwe inspirerende mensen ontmoet, ook werd maar weer eens bevestigd welke mensen in mijn omgeving belangrijk zijn voor mij. Het voltooien van mijn proefschrift had ik niet zonder deze mensen gekund. Een aantal van hen wil ik in het bijzonder bedanken.

Allereerst wil ik graag iedereen bedanken die heeft meegewerkt aan een van mijn onderzoeken, zoals leden van het RRD-panel en lokale fysiotherapeuten. Zonder jullie had ik dit proefschrift niet kunnen schrijven. Wat is immers user-centered design zonder eindgebruikers?

Daarnaast wil ik Monique, Hermie en Harm bedanken voor hun begeleiding. Monique, met enige tegenzin opende ik altijd je feedback. Mijn teksten kwamen vaak rood terug. Dit was in eerste instantie niet zo leuk, maar je altijd kritische houding heeft de kwaliteit van mijn proefschrift absoluut naar een hoger niveau getild. Dank hiervoor! Ook bedankt voor je positieve woorden wanneer ik het even niet meer zo zag zitten. Hermie, bedankt voor de inspirerende en filosofische sessies over wetenschap, hologrammen en natuurlijk virtuele coaches! Aan je enthousiasme voor technologie leek nooit een einde te komen. En Harm, bedankt voor het overnemen van de begeleiding toen Monique met zwangerschapsverlof was, al in een vroeg stadium van mijn promotie. Je altijd nuchtere instelling en enthousiasme voor virtuele coaches hebben me zeker een stapje in goede richting geholpen.

Tessa en Roos, mijn (digitale?) paranimfen, bedankt! Allereerst, Tessa, mijn 'partner in crime'. We zijn op dezelfde dag begonnen, hebben drie jaar lang in dezelfde kamer gewerkt en hebben op dezelfde dag onze spullen ingepakt. Regelmatig heb ik mijn ideeën aan je mogen voorleggen, nooit was je te beroerd om mee te denken. Helaas moet ik nu zonder mijn 'rubber duck' door het leven. Verder gingen er een hele nieuwe werelden voor me open op het gebied van katten, paarden en films. Tessa, bedankt voor de leuke tijd samen!

Roos, ook jou heb ik voor het eerst ontmoet bij RRD. Tijdens onze gezamenlijke reis naar Ierland heb ik je beter leren kennen, en wat ben ik daar blij om! Essentieel onderdeel van mijn promotie waren onze klagkwartiertjes. Soms moet je gewoon even je hart luchten: zeuren over feedback, over papers die weer eens afgewezen zijn, of gewoon over het feit dat je even geen zin hebt. Lekker klagen, om daarna ons positieve petje weer uit de tas te halen. Voor BIONIC hebben we op het einde van mijn tijd bij RRD ook inhoudelijk nog even mogen samenwerken. Je wilde altijd meedenken en verzetten een berg werk, zodat ik me kon focussen op mijn onderzoek. Bedankt daarvoor.

En natuurlijk ook alle andere collega's van RRD, bedankt! Wat zijn jullie fijne mensen. Altijd kon ik bij jullie binnenlopen en waren jullie bereid om mee te denken. Een groep gelijkgestemden, waarvan ik me echt onderdeel voelde. Mensen met hetzelfde gevoel voor humor, en die wel houden van een beetje sarcasme. Heerlijk! Ik heb me erg vermaakt tijdens onze koffiemomentjes, spetterende lunchpauzes, wandelingen en juniorenuities. Zonder jullie had ik mijn promotie zeker niet afgemaakt.

Alice en Michel, bedankt voor jullie ongelimiteerde gastvrijheid. Zonder jullie steun stond ik niet waar ik nu ben. Bedankt voor alle keren dat ik mee heb mogen eten, kopjes thee heb mogen drinken en natuurlijk voor de wandelingen met Sæla!

En dan oma, wat ben ik blij dat ik je ken. Wat er ook gebeurt, je bent altijd positief. Ik hoop dat ik later ook zo in het leven kan staan. Je hebt me geleerd dat je nooit te oud bent om te leren. Met nieuwsgierigheid en interesse red je het wel. Deze interesse heb je ook zeker in je kleinkind(eren). Regelmatig belde je me 's avonds weer eens op om te vertellen dat er een interessante documentaire op was over de zorg, technologie of wetenschap. Daarnaast heb je als iPad-deskundige je generatie vertegenwoordigd als deelnemer in menig RRD-onderzoek. Wie wil er nou niet zo'n oma?

Pap en mam, uiteraard had ik dit proefschrift nooit kunnen schrijven zonder jullie. Bedankt voor jullie steun en vertrouwen. Van jullie heb ik vaardigheden meegekregen waar ik niet zonder had gekund tijdens mijn promotietraject: hard werken, verder kijken dan je neus lang is en een kritische houding (wellicht soms een beetje een té kritische houding, pap...?). Pap, bedankt voor je inhoudelijk interesse en meedenken. Mam, bedankt voor je relativerende noot en je PA-momentjes.

Minke en Joris, jullie ook bedankt voor jullie interesse en steun! Onze gezellige avondjes samen, met af en toe een fikse discussie over de wetenschap, hebben de benodigde ontspanning geboden tijdens mijn promotie. Bedankt hiervoor. Dit geldt natuurlijk ook voor iedereen waarmee ik regelmatig heb gesport, een terrasje heb gepakt of lekker een avondje heb geneflixt. Bedankt!

En dan natuurlijk Frank, bedankt voor alles! Onverwachts kwam je op mijn pad, zo halverwege mijn promotie. Door COVID-19 zitten we inmiddels vrijwel 24/7 op elkaars lip. Regelmatig heb je dan ook mijn gezeur rondom mijn promotie moeten aanhoren, sorry daarvoor! Bedankt voor het vertrouwen dat je altijd in mij hebt getoond en al die keren dat je spontaan met een taartje, bloemetje of vt-wonen thuis kwam om een klein promotie-succesje te vieren. Super lief! En natuurlijk bedankt voor elke keer dat we samen hebben gelachen, gek hebben gedaan en ons suf hebben gepuzzeld. Hopelijk mogen we dat nog heel lang.

Als laatste wil ik graag mijn huidige collega's van Ecare bedanken. Jullie enthousiasme en waardering hebben mij het laatste zetje gegeven om mijn promotie goed af te ronden. Met mijn proefschrift en vergaarde vaardigheden op zak hoop ik in de toekomst samen met jullie nog veel mooie producten te ontwikkelen.

ABOUT THE AUTHOR

Silke ter Stal was born in Enschede, the Netherlands, on May 31st 1992. In July 2015, Silke obtained her bachelor's degree in Creative Technology and in August 2017, she obtained her master's degree in Human Media Interaction, both at the University of Twente (Enschede, the Netherlands). During her studies, Silke's interest in creating products with societal impact by combining design with technology grew. This is reflected in her bachelor thesis on the design of games for interactive gait rehabilitation on a pressure sensitive LED floor and her master thesis on the design of an interactive toy for storytelling for children.

Immediately after obtaining her master's degree, Silke started as a junior researcher and PhD candidate at Roessingh Research and Development (RRD) in Enschede. Silke was involved in different eHealth projects, such as the national projects MATCH (Pioneers in Healthcare) and PACO (ZonMw), and the international projects GOAL (Horizon 2020), Council of Coaches (Horizon 2020), BIONIC (Horizon 2020) and FRAIL (Eurostars-2). Results of these projects have been described in this thesis.

In September 2020, Silke started working as a UI/UX designer at Ecare in Enschede and finished her PhD thesis in parallel. Currently, Silke works full-time at Ecare, where she is responsible for the design of usable and accessible software for both caretakers and caregivers.



I like to work at the intersection of design and technology, such that I can combine my creative and analytic mind. Analysing users' needs and creating products with rich user experiences and societal impact gives me energy.

— Silke ter Stal —

Journal Publications

- Kramer, L. L., **ter Stal, S.**, Mulder, B. C., de Vet, E., & van Velsen, L. (2019). Developing embodied conversational agents for coaching people towards a healthy lifestyle: A scoping review. *Journal of Medical Internet Research*, 22(2), e14058. doi:10.2196/14058
- ter Stal, S.**, Broekhuis, M., van Velsen, L., Hermens, H., & Tabak, M. (2020). Embodied conversational agent appearance for health assessment of older adults: Explorative study. *JMIR Human Factors*, 7(3). doi:10.2196/19987
- 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, 102409. doi:10.1016/j.ijhcs.2020.102409

ter Stal S., Tabak, M., op den Akker, H., Beinema, T., Hermens, H. (2019). Who do you prefer? The effect of age, gender and role on users' first impressions of embodied conversational agents in eHealth. *International Journal of Human-Computer Interaction*, 36(9), 881–892. doi:10.1080/10447318.2019.1699744

Conference Papers

Broekhuis, M., van Velsen, L., **ter Stal, S.**, Weldink, J., & Tabak, M. (2019). Why my grandfather finds difficulty in using eHealth: Differences in usability evaluations between older age groups. In *International Conference on Information and Communication Technologies for Ageing Well and e-Health* (pp. 48–57). doi:10.5220/0007680800480057

Catala, A., Theune, M., Reidsma, D., **ter Stal, S.**, & Heylen, D. (2017). Exploring children's use of a remotely controlled surfacebot character for storytelling. *International Conference on Intelligent Technologies for Interactive Entertainment* (pp. 120–129). doi:10.1007/978-3-319-73062-2_9

Van Delden, R., Janssen, J., **ter Stal, S.**, Deenik, W., Meijer, W., Reidsma, D., & Heylen, D. (2016). Personalization of gait rehabilitation games on a pressure sensitive interactive LED Floor. In *Personalization in Persuasive Technology Workshop* (pp. 60–73).

Jansen-Kosterink, S., Bulthuis, R., **ter Stal, S.**, Pnevmatikakis, A., Kyriazakos, S., Pomazanskyi, A., & op den Akker, H. (2019). GOAL: An eHealth application for rewarding healthy behaviour. The first experiences of older adults. *International Conference on Information and Communication Technologies for Ageing Well and e-Health* (pp. 58–66). doi:10.5220/0007680900580066

Jansen-Kosterink, S., Bulthuis, R., **ter Stal, S.**, van Velsen, L., Pnevmatikakis, A., Kyriazakos, S., ... & op den Akker, H. (2019). The results of an iterative evaluation process of an Mhealth application for rewarding healthy behaviour among older adults. *International Conference on Information and Communication Technologies for Ageing Well and e-Health* (pp. 62–78). doi:10.1007/978-3-030-52677-1_4

ter Stal, S., Catala, A., Theune, M., & Reidsma, D. (2019). Designing a smart toy interactive setting for creating stories. *Interactivity, Game Creation, Design, Learning, and Innovation* (pp. 601–610). doi:10.1007/978-3-030-53294-9_44

Conference Contributions

op den Akker, H., Beinema, T., **ter Stal, S.**, Broekhuis, M. (2018). Designing virtual agents for persuasive health coaching. *Workshop at the 8th Supporting Health By Technology Symposium (HealthByTech)*. 1st of June 2018, Enschede, the Netherlands

ter Stal, S. (2019). Promotieonderzoek: Designing embodied conversational agents for eHealth *Blitz presentation at Symposium Conversational Agents in Gezondheidscommunicatie*. 18th of November 2019, Amsterdam, the Netherlands

Submitted for Publication

- ter Stal, S.**, Hurmuz, M., Jansen-Kosterink, S., Beinema, T., op den Akker, H., Hermens, H., & Tabak, M. Older adults' satisfaction with embodied conversational agents of similar age and gender in a multi-agent eHealth application (*submitted for publication*)
- ter Stal, S.**, Jongbloed, G., & Tabak, M. Embodied conversational agents in eHealth: How facial and textual expressions of positive and neutral emotions influence perceptions of mutual understanding (*submitted for publication*)
- ter Stal, S.**, Kramer, L. L., Bulthuis, R., Hermens, H., & Tabak, M. Embodied conversational agents for eHealth: State-of-the-art and design Strategies for appearance (*submitted for publication*)
- ter Stal, S.**, Sloots, J., Ramlal, A., op den Akker, H., Lenferink, A., & Tabak, M. An embodied conversational agent in an eHealth self-management intervention for chronic obstructive pulmonary disease and chronic heart failure: Exploratory study in a real-life setting (*submitted for publication*)

PROGRESS RANGE

The following publications have been published in the Progress range by Roessingh Research and Development, Enschede, the Netherlands. Copies can be ordered, when available, via info@rrd.nl.

1. Pot, J.W.G.A, Boer H., van Harten, W.H., Hermens, H.J., Seydel, E.R. Comprehensive Need-Assessment. Ontwikkeling van een meetinstrument voor zorgbehoeften en kwaliteitsbeoordeling door patiënten. September 1994, ISBN 90-25452-01-22
2. van Leerdam, N.G.A, Hermens, H.J. Revalidatietechnologie in Euregio. July 1995, ISBN 90-75452-02-0
3. Duda, L., van Noort, L.O., Röseler, S., Greitemann, B.O.L, van Harten, W.H., Klazinga, N.S. Rehabilitation in Germany and the Netherlands, A comparison of two rehabilitation systems. August 1995, ISBN 90-75452-03-9
4. Hermens, H.J., Nene, A.V., Zilvold, G. Electrophysiological Kinesiology, Proceedings of the 11th congress of the International Society of Electrophysiology and Kinesiology in Enschede, the Netherlands 1996. October 1996, ISBN 90-75452-04-7
5. van Harten, W.H. Bouwen aan een kwaliteitssysteem in de revalidatiezorg. Een poging tot constructieve technology assessment van een kwaliteitssysteem in een gezondheidszorginstelling. December 1997, ISBN 90-75452-07-1
6. Baardman, G., IJzerman, M.J. Design and evaluation of a hybrid orthosis for people with paraplegia. November 1997, ISBN 90-75452-8-X
7. Hutten, M.M.R. Lumbar Dynamometry: A useful method for assessment of patients with chronic low back pain? November 1999, ISBN 90-75452-13-6
8. van der Salm, A., van Harten W.H., Maathuis, C.G.B. Ketenkwaliteit Cerebrale Parese Zorg. Een beschrijving van de cerebrale parese zorg en mogelijke verbeteringen hierin. April 2001, ISBN 90-75453-19-5
9. Nederhand, M.J. Muscle activation patterns in post traumatic neck pain. March 2003, ISBN 90-75452-27-6
10. Jannink, M.J.A. Usability of custom-made orthopaedic shoes in patients with degenerative disorders of the foot. September 2004, ISBN 90-75452-28-4
11. Blokhorst, M.G.B.G. State-dependent factors and attention in whiplash associated disorder. January 2005, ISBN 90-365-2111-4
12. Buurke, J.H. Walking after stroke. Co-ordination Patterns & Functional Recovery. February 2005, ISBN 90-365-140-8
13. van der Salm, A. Spasticity reduction using electrical stimulation in the lower limb of spinal cord injury patients. October 2005, ISBN 90-365-2253-6
14. Snoek, G.J. Patient preferences for reconstructive interventions of the upper limb in tetraplegia. December 2005, ISBN 90-365-2255-2

15. de Kroon, J.R. Therapeutic electrical stimulation of the upper extremity in stroke. December 2005, ISBN 90-365-2269-2
16. van Dijk, H. Motor skill learning, Age and Augmented feedback. March 2006, ISBN 90-365-202-9
17. Mes, C.A.J. Improving non-optimal results in chronic pain treatment. January 2007, ISBN 90-365-2435-0
18. Voerman, G.E. Musculoskeletal neck-shoulder pain: a new ambulant myofeedback intervention approach. March 2007, ISBN 90-365-2460-1
19. Kallenberg, L.A.C. Multi-channel array EMG in chronic neck-shoulder pain. March 2007, ISBN 90-365-2459-8
20. Huis in 't Veld, M.H.A. Work-related neck-shoulder pain. The role of cognitive-behavioural factors and remotely supervised treatment. December 2007, ISBN 978-90-365-2584-8
21. Fleuren, J.F.M. Assessment of spasticity. From EMG to patients' perception. October 2009, ISBN 978-90-365-2869-6
22. Reenalda, J. Dynamic sitting to prevent pressure ulcers in spinal cord injured. October 2009, ISBN 978-90-365-2884-9
23. Prange, G.B. Rehabilitation robotics. Stimulating restoration of arm function after stroke. October 2009, ISBN 978-90-365-2901-3
24. Vos-van der Hulst, M. Prognostic factors and underlying mechanisms in chronic low back pain. February 2010, ISBN 978-90-365-2881-8
25. Kottink-Hutten, A.I.R. Assessment of a two-channel implantable peroneal nerve stimulator post-stroke. February 2010, ISBN 978-90-365-2959-4
26. van Weering, M.G.H. Towards a new treatment for chronic low back pain patients. Using activity monitoring and personalized feedback. May 2011, ISBN 978-90-365-3305-8
27. Gulmans, J. Crossing boundaries. Improving communication in cerebral palsy care. February 2012, ISBN 978-90-365-3305-8
28. Molier, B.I. Influence of augmented feedback on learning upper extremity tasks after stroke. March 2012, ISBN 978-90-365-3293-9
29. Dubbeldam, R. Towards a better understanding of foot and ankle kinematics in rheumatoid arthritis. The effects of walking speed and structural impairments. October 2012, ISBN 978-90-365-3407-9
30. Evering, R.M.H. Ambulatory feedback at daily physical activity patterns. A treatment for the chronic fatigue syndrome in the home environment? April 2013, ISBN 978-90-365-3512-0
31. Malhorta, S. Does spasticity interfere with functional recovery after stroke? A novel approach to understand, measure and treat spasticity after acute stroke. November 2013, ISBN 978-90-365-2567-0
32. Tabak, M. Telemedicine for patients with COPD. New treatment approaches to improve daily activity behaviour. Roessingh Research and Development. February 2014, ISBN 978-94-6108-590-0
33. Trompetter, H.R. ACT with pain. Measurement, efficacy and mechanisms of Acceptance and Commitment Therapy. September 2014, ISBN 978-90-365-3708-7

34. op den Akker, H. Smart Tailoring of Real-Time Physical Activity Coaching Systems. October 2014, ISBN 978-90-365-3762-9
35. Jansen-Kosterink, S.M. The added value of telemedicine for physical rehabilitation. December 2014, ISBN 978-90-823196-0-6
36. Velstra, I.M. Advanced insights in upper limb function of individuals with cervical spinal cord injury. December 2015, ISBN 978-90-365-3929-6
37. Kloosterman, M.G.M. Keep on rolling. Functional evaluation of power-assisted wheelchair use. June 2016, ISBN 978-90-365-4299-9
38. Prinsen, E.C. Adapting to change. Influence of a microprocessor-controlled prosthetic knee on gait adaptations. December 2016, ISBN 978-90-365-4206-7
39. Wolvers, M.D.J. A coach in your pocket. On chronic cancer-related fatigue and physical behavior. March 2017, ISBN 978-90-365-4299-9
40. Cabrita, M. Active and Pleasant Ageing supported by Technology. November 2017, ISBN 978-90-365-4407-8
41. Haarman, J.A.M. TIBAR. Therapist Inspired Balance Assisting Robot. November 2017, ISBN 978-90-365-4407-8
42. Nijenhuis, S.M. Roll up your sleeves! Technology-supported arm and hand training at home after stroke. April 2018, ISBN 978-90-365-4510-5
43. Cranen, K. Acceptance of telerehabilitation in chronic pain: the patients' perspective. June 2018, ISBN 978-90-365-4555-6
44. Boerema, S.T. Sensing human activity to improve sedentary lifestyle. September 2018, ISBN 978-90-365-4604-1
45. Radder, B. The wearable hand robot. Supporting impaired hand function in activities of daily living and rehabilitation. November 2018, ISBN 978-90-365-4658-4
46. Krabben T. A reaching hand. Towards an Active Therapeutic device for the upper extremity following stroke. December 2018, ISBN 978-90-365-4660-7
47. Timmerman, J.G. Cancer rehabilitation at home. The potential of telehealthcare to support functional recovery of lung cancer survivors. January 2019, ISBN 978-90-4701-7
48. Nikamp-Simons, C.D.M. The sooner the better?! Providing ankle-foot orthoses in the rehabilitation after stroke. May 2019, ISBN 978-90-365-4747-5
49. Achterkamp, R. Towards a balanced and active lifestyle. June 2019, ISBN 978-94-6323-656-0
50. Engbers, C. Keep Cycling. How Technology can Support Safe and Comfortable Cycling for Older Adults. September 2019, ISBN 978-90-365-4848-9
51. Ommeren, A.L. Offering a helping hand. Getting a grip on needs and preferences of stroke patients regarding soft-robotic technology supporting hand function. October 2019, ISBN 978-90-365-4835-9
52. ter Stal, S. Look Who's Talking. Appearance of Embodied Conversational Agents in eHealth. March 2021, ISBN 978-90-365-5126-7

