The potential of Embodied Conversational Agents to stimulate adherence to eHealth interventions; steps towards a more dynamical approach

Mark Ronald Scholten

THE POTENTIAL OF EMBODIED CONVERSATIONAL AGENTS TO STIMULATE ADHERENCE TO EHEALTH INTERVENTIONS

STEPS TOWARDS A MORE DYNAMICAL APPROACH

Mark Ronald Scholten

THE POTENTIAL OF EMBODIED CONVERSATIONAL AGENTS TO STIMULATE ADHERENCE TO EHEALTH INTERVENTIONS

STEPS TOWARDS A MORE DYNAMICAL APPROACH

DISSERTATION

to obtain the degree of doctor at the Universiteit 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 Friday 12 November 2021 at 16.45 hours

by

Mark Ronald Scholten

born on the 4th of August, 1971
in Son en Breugel, The Netherlands

This dissertation has been approved by:

Supervisor:
prof. dr. J.E.W.C. van Gemert - Pijnen

Co-supervisor:
dr. S.M. Kelders

Cover design: Mark Scholten
Printed by: Gildeprint
ISBN: 978-90-365-5292-9
DOI: 10.3990/1.9789036552929

Graduation Committee:

Chair / secretary:	Prof. dr. T.A.J. Toonen
Supervisor:	Prof. dr. J. E. W. C. van Gemert-Pijnen (Universiteit Twente)
Co-supervisor:	Dr. S.M. Kelders (Universiteit Twente)
Committee Members:	Prof. dr. P.L.C. van Geert (Rijksuniversiteit Groningen)
	Prof. dr. D.K.J. Heylen (Universiteit Twente)
	Prof. dr. Max Louwerse (Universiteit van Tilburg)
	Dr. T.J.L. Van Rompay (Universiteit Twente)
	Dr. H.W. Steenbeek (Rijksuniversiteit Groningen)

Tempora mutantur, et nos mutamur in illis - Augustinus

(de tijden veranderen en wij veranderen met haar mee)

Contents

Chapter 1

General Introduction

Introduction

Embodied Conversational Agents (ECA's), animated talking heads on screen, are digital artefacts that have inspired scholars from various disciplines such as gaming, eHealth and eLearning. Their life-like appearance makes interaction with a computer a more engaging experience which has fueled the imagination of many researchers with regards to their applications.

Within this thesis, I have focused on the merits of ECA's as adjuncts to eHealth interventions. In order to do so, I will begin by taking a historical perspective on the onset and development of ECA's within this general introduction. From a similar chronological perspective, I will describe the eHealth field, the electronic delivery of healthcare services. I will do this with the objective of comparing what were initial expectations and where we stand today with ECA's in eHealth. As I will point out, both the ECA's and eHealth study fields hold promise but have not yet fully lived up to these expectations. Noteworthy, a large part of the promise dates back as far as the period 1994-2000. Within that era, important findings with the Human Computer Interaction (HCI) field were made, as I will describe. Furthermore, I will sketch how the eHealth and ECA research fields have contributed to each other since 2000. As I will argue, ECA's can potentially contribute to improving a tight issue in eHealth; the elevated non-adherence levels. Then again, expectations should be tempered, as renowned ECA meta-studies refer to ECA effects with caution, see e.g. "Nearing two decades of intense study of the topic researchers cannot say with much certainty the level of effectiveness one would expect to see .." (Schroeder & Gotch, 2015). Subsequently, I will touch upon the subject of the ECA and user behaving in synchrony and how this synchronous behavior can potentially contribute to the effectiveness of the ECA. Last, I will take a side-step into an adjacent study field; that of the Dynamical Systems Perspective (DSP). This study field typically examines how human behavior fluctuates over short time intervals (e.g. 10-40 seconds). In that perspective, I will describe how an Agent-Based Model (ABM) might be capable of simulating the shortly during cognitive-affective states of users during a psycho-education eHealth intervention. In addition, as I will point out, these simulation results can potentially contribute to a better understanding of when the ECA's supportive actions should be provided to eHealth users experiencing low motivation. Finally, I will state my research questions and outline the structure of this thesis.

1994-2000: Remarkable findings within the domain of Human Computer Interaction

25 years ago a series of important discoveries were made within the domain of Human Computer Interaction (HCI). Reeves and Nass (1996) described new ways in which humans respond to computers. At the time it was commonly accepted that computers were just machines, lifeless tools that were designed to carry out tasks for human users. Suddenly, Reeves and Nass (1996) reported on novel experimental effects. Computers that use flattery, or which praise rather than criticize their users were better liked by study participants. Furthermore, the authors found that computers that praise other computers are better liked than computers that praise themselves, and computers that criticize other computers are liked less than computers that criticize themselves. These effects became known as the media equation effect and the CASA (Computers as a Social Actor) effect. In essence, the CASA effect demonstrated that humans interpret computer actions as human actions while at the same time being well aware that computers are not human. Taking the CASA effect a step further, Nass, Fogg, and Moon (1996) ran an experiment, based on the hypothesis that the computer and human user could form a team together. The participants were primed as they were told that they would join forces with the computer placed in front of them. In addition, the experimental instructions mentioned that the users would be dependent on the computer's performance. By experimentally manipulating these user expectations, this study was able to demonstrate that humans display the same sorts of attitudes and behaviors as when collaborating with fellow humans. That is, compared to *control condition subjects* who were not told they were teaming up with the computer, *team subjects* reported they were more similar to the computer, and were more open to the computer's influence, all at significant levels. Further, team subjects thought the computer's information was of higher quality, and finally found the computer to be significantly friendlier.

Meanwhile, Fogg (1998) took the social aspect of the CASA effect, added new elements and coined another new HCI term *persuasive technology*. Fogg thought of persuasive technology as technology that is designed to change attitudes or behaviors of the users through persuasion and social influence, but not necessarily through coercion. Furthermore, Fogg (1998) proposed the *functional triad* as a classification of three basic ways that people view or respond to computing technologies: persuasive technologies can function as tools, media, or social actors – or as more than one at once. Through these terms, Fogg declared that computer technology should no longer be considered as expressions of passive code. Instead, the technology should be regarded as an active, dynamical instance. Not very long thereafter, in 2000, Picard published yet another seminal HCI study, titled 'Affective Computing'. Picard convincingly stated that computers should be endowed with affective capabilities. The term affective computing was defined as "computing that relates to, arises from, or influences emotion" (Picard, 2000, p. 1). Picard described how computers could be equipped with functionality to notice and respond to emotions as expressed by their users, in order to cater for a more natural form of human-computer interaction. Shortly before, Cassell (1999) had proposed to represent the computer as an actor with human-like characteristics such as a face and communication skills. The author baptized this digital artefact by the term *Embodied Conversational Agent* (ECA). ECA's were often described as animated talking heads, and colloquially referred to as 'robots on screen'. ECA's held promise and were soon deployed in several studies, especially in eLearning see e.g. Herman the Bug (Lester, Stone & Stelling, 1999), and Adele (Shaw, Johnson & Ganeshan, 1999). The publications of Cassell (1999) and Picard (2000) provide a hallmark within a blooming HCI period (1994-2000). During this period, the computer had transformed from a passive piece of electronical equipment to a conversational partner. The way had been paved for ECA research as a new discipline.

2001-2005: Further studies on ECA's

In 2005, Bickmore and Picard combined elements of persuasive technology and affective computing in their deployment of ECA's. They made a plea that ECA's should be deployed within eLearning and eHealth environments in order to support users. Furthermore, the authors underscored the importance of running studies with a longer duration (one to three months) in order to find out about their real potential. As they argued, a key aspect of any relationship (thus including the human-computer relationship) is its persistency and continuity. Yet, at the time, most ECA studies applied single-session experimental designs. In order to explore the longitudinal aspect of the human-computer relationship building process, Bickmore and Picard (2005) decided to develop and evaluate an ECA named Laura that could support a user during multiple interactions over an extended period of time. This way, Laura enriched the Fit Track behavior change system. Note that Fit Track was set up to motivate sedentary users to exercise; 30 minutes or more of moderate-intensity physical activity. Bickmore and Picard (2005) offered an intervention period of 30 days, followed by removal of the intervention. Finally, a follow-up measurement was done to check if the new behavior had become truly adopted. Within the control condition Fit Track was deployed without Laura. As the authors found, the relational agent Laura was generally liked by the study participants, but it did not result in significantly more exercise behavior compared to the control condition. To date, their longitudinal study is still relevant for this thesis, as one of the few studies describing an ECA in Health combined with ECA-user relationship building.

2006-2021: Proper design of ECA's appears to be and remain challenging

All the beforementioned developments (CASA effect, persuasive technology, affective computing, ECA's) gave rise to the idea that, within virtual environments, computers can take on roles that are normally taken by humans such as virtual support providers for patients in eHealth environments and as virtual tutors for students within electronic learning environments. So, back in 2001, it looked like it would be a matter of time for ECA's to become successful. However, reality turned out differently. To date, ECA's are not fulfilling their roles at a scale as one would expect. Compared to chatbots, disembodied dialogue-based artefacts that have close resemblances with ECA's, ECA's are not as often guiding visitors of medical or business websites. So, what happened?

Various studies have evaluated ECA research thereby including ECA studies as far back as 20 years. The meta study of Schroeder and Gotch (2015) on persisting issues in pedagogical agent research underscores that the effectiveness of including agent in a learning environment remains debatable. As part of their recommendations the authors advise treatment and control conditions that should not differ on more than a single dimension in order to precisely find out what ECA element contributes to what kind of outcome Furthermore, they promote the development and usage of low-cost ECA's as to stimulate adoption of ECA's as adjuncts in eLearning environments. Veletsianos and Russell (2014) equally state that ECA's in pedagogical contexts have not yet lived up to their promise. As an important cause they refer to the lack of both qualitative and interpretive studies, that prohibit gaining a deeper understanding of the ECA study field. Furthermore, the authors postulate that a multiplicity of variables, such as agent role, voice, and voice quality, interact in complex ways, making generalizations difficult. In addition, they recommend the deployment of agents in naturalistic contexts and open-ended environments. Finally, Veletsianos and Russell (2014) advocate the investigation of ECA's in long-term interventions, echoing the earlier statement of the seminal study of Bickmore and Picard (2005). In a similar vein, Weiss, Wechsung, Kühnel, and Möller (2015) evaluate ECA research. The authors take an interesting, contrary stance by stating that a speech dialogue with a computer is still far from self-evident. Direct manipulation, meaning clicking on buttons and icons has advantages compared to a speech dialogue, such as clear and predictable results. Within their ECA review, Johnson and

Lester (2016) refer to the lack of a generic technological ECA platform. As they state, earlier ECA work benefited from the availability of off-the-shelf Microsoft Agent platform. Indeed, since that platform has been discontinued in 2009, no comparable tool has taken its place. Logically, this has hindered the dissemination and implementation of ECA's on websites. Again making the comparison with chatbots, the claim of their success has indeed been reported as caused by the availability of platforms through the big technological companies Facebook, Google and Microsoft since 2016 (Brandtzaeg & Følstad, 2017). Other elements that stand out in ECA research are both the focus on specific design aspects such as gender and ethnicity (see e.g. Baylor, 2011), and the development of advanced ECA's (e.g. Gratch, Wang, Gerten, Fast & Duffy, 2007). See the lower part of Figure 1 for a timeline of inspirational studies within the ECA domain. Historically, relatively less attention has been paid to maximizing the ECA's effectiveness with an eHealth context. However, this has changed recently. A recent meta-study on the application of ECA's in clinical psychology (Provoost, Lau, Ruwaard & Riper, 2017) states that ECA piloting studies on the one hand show promising results with respect to usability and user acceptance, but on the other hand provide little hard evidence for their merits in clinical settings. Furthermore, the authors advocate a 'low-tech' ECA approach as it forces the field to think about the core attributes that can make the ECA effective. A second recent meta-study on ECA's and their lack of success has been conducted by ter Stal, Kramer, Tabak, op den Akker, and Hermens (2020). As the authors state in their review; the lack of a design standard is problematic for ECA's. As they argue, there are no clear guidelines with regards to the design and deployment of ECA features such as speech and/or textual output and facial and gaze expressions. As they conclude, consensus on design features of ECAs in eHealth is far from established. They therefore advise follow-up research that should focus on the modeling and formal definition of these design features. Finally, the authors repeat the stance of Bickmore and Picard (2005) and of Veletsianos and Russell (2014): ECA research should be conducted within both a long-term (6-12 weeks) and daily life setting. This safeguards that the ECA-user relationship building process is examined within a representative context for eHealth interventions.

From 2007 onwards: the creation of rapport as a cornerstone in ECA studies

The previous sections have sketched an image of ECA's that on the one hand hold promise in eHealth and eLearning and on the other hand fail to convincingly demonstrate this. Recommendations on ECA design such as more strictly formalization have been referred to. Noteworthy and although not part of any formal ECA design standard, the ECA research field has generally adopted the creation of a productive relationship as a de-facto design norm. Early ECA studies (e.g. Cassell, 1999; Reeves, 2000) primarily focused on the level of *engagement*, the enhancement of online experiences, that ECA's could instill amongst users. In addition, their pedagogical value (Lester, Stone & Stelling, 1999) was studied.

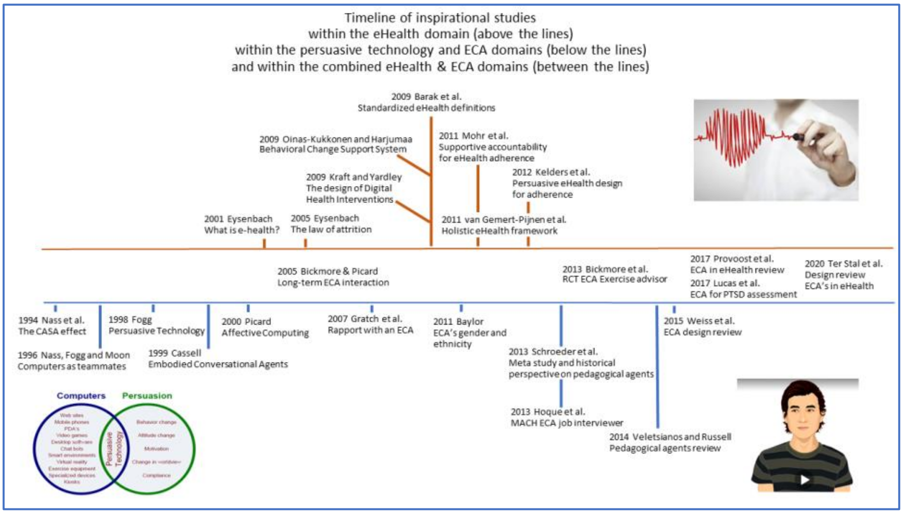


Figure 1. Timeline of inspirational studies within the eHealth domain (above the lines), ECA domain (below the lines), and both domains (between the lines).

Later studies (e.g. Bickmore & Picard, 2005) started to emphasized the relationship building project and framed their ECA as relational agent. Gratch, Wang, Gerten, Fast & Duffy (2007) took the partnership building process a step further, through an ECA that displayed contingent nonverbal behaviors indicating mutual attentiveness (e.g. mutual gaze), and coordination (e.g. postural mimicry and synchronized movements).

As the key relationship building outcome variable, the authors decided to put rapport into practice. Within human-to-human communication contexts, rapport is generally associated to terms as 'harmony, entrainment, fluidity, synchrony, and flow'. Those pleasant feelings are typically experienced when one is engaged in a good conversation with someone. As Gratch et al. (2007) mention; speakers seem tightly enmeshed in something like a *dance*. They rapidly detect and respond to each other's movements and together create a productive pattern. Gratch et al. (2007) defined highly relevant research questions: could an ECA effectively generate behavior that would engender feelings of rapport in human speakers? How would this compare to human generated contingent feedback? As a secondary goal they evaluated whether contingency (as opposed to frequency) of agent feedback was crucial for the creation of rapport. Their results indicated that contingency, the right timing of the ECA's non-verbal behavior, indeed mattered substantially when it came down to creating rapport. Remarkably, the authors also found that the agent generated behavior was as effective as the behavior of human listeners for the creation of rapport.

2001-2012; the eHealth period prior to the most important integrations with ECA's

Parallel to the developments within the ECA study field, the eHealth study domain commenced. That is, not long after the start of the development of the persuasive technology domain (Fogg, 1998), healthcare was considered as a suitable domain for application of new types of user-oriented technologies. User convenience, cost reduction and 'any place, anytime, anywhere' were considered as valuable arguments in favor of electronic healthcare delivery. In 2001, Eysenbach (2001), one of the pioneers, defined eHealth as an "emerging field in the intersection of medical informatics, public health and business, referring to health services and information delivered or enhanced through the Internet and related technologies" (p. 1).

See the upper part of Figure 1 for a timeline of inspirational studies within the eHealth domain. This part displays the early eHealth period (2001-2012) during which ECA's were not often involved. Note that the period after 2012 will be described in the following section. Early eHealth applications used in the clinical practice were teledermatology (van der Heijden, De Keizer, Bos, Spuls & Witkamp, 2011) and teleconsultation for diabetes care (Verhoeven, Tanja-Dijkstra, Nijland, Eysenbach & van Gemert-Pijnen, 2010). Despite the initial enthusiasm about eHealth, Eysenbach (2005) saw an issue coming up as early as in 2005: the high non-adherence rates amongst eHealth users. Note that non-adherence refers to the fact that not all participants use or keep using the intervention in the intended way. One of the causes, as was put forward by Kelders (2012), eHealth technology is often set up as a black box. That is; it is known what goes in (baseline measures) and what comes out (post-intervention measures), but limited attention has been paid to what happens inside the box. As to design transparent Health interventions, it had been postulated that their foundations should be bolstered. The design of eHealth interventions should be reinforced, according to Kraft and Yardley (2009). In addition, new and better definitions were needed for the eHealth field, as suggested by Barak, Klein, and Proudfoot (2009). Further standardization was proposed by van Gemert et al. (2011) with a holistic framework to design and develop eHealth interventions. In parallel, the relevance for eHealth interventions to communicate with their users became more apparent. Along these lines, Oinas-Kukkonen (2010) coined the term Behavior Change Support System (BCSS), "an information system designed to form, alter or reinforce attitudes, behaviors or an act of complying without using deception, coercion or inducements” (p. 6) .

As Kelders (2012) added: the communication function of BCSSs is key, which makes it basically more of a communication system than an information system. Oinas-Kukkonen and Harjumaa (2009) introduced a framework to classify the persuasive functions of BCSSs and eHealth intervention focusing on its communication functionalities. These communication functionalities were categorized within four main user support functions being; primary task support, dialogue support, social support and credibility support. Primary task support (e.g. reduction, tunneling, tailoring) provides a means to structure the eHealth program. Furthermore, dialogue support helps the user to achieve its objectives. This is done through providing feedback on the user’s actions and by improving the communication between the eHealth intervention and user, through praise, rewards, and similarity. Their framework can be used both as a frame of reference when developing an eHealth intervention and as an analysis-tool for existing eHealth interventions. Note that this framework is not only helpful in theory. Indeed there is evidence that communicative and supportive functionalities within eHealth interventions are truly effective. Webb, Joseph, Yardley, and Michie (2010) demonstrated that web-based interventions that include communication functionalities (through text messages) are more effective than web-based intervention that are void of this. Furthermore, Neff and Fry (2009) made clear that system-generated reminders increase the effect and adherence of web-based interventions.

2013-2021: partial evidence for ECA’s in eHealth

Shortly summarizing the previous sections: Adherence to self-guided eHealth interventions is low, especially in real-life settings. System-generated support provides a promising remedy. Furthermore, system-generated support within eHealth interventions has gained traction, which has been promoted by the design work of Oinas-Kukkonen and Harjumaa (2009) and the empirical evidence of e.g. Webb et al. (2010). ECA’s can personify these system-generated support actions. Despite the fact that ECA’s have been studied for more than twenty years, there is still uncertainty to what they are truly capable of. It is generally accepted that it is relevant for the ECA to first establish rapport with the user, as a pre-condition for being effective when support is provided. As can be concluded, the eHealth and ECA research domains can potentially be of value to each other. That is, self-guided eHealth has an ‘adherence demand’ for a technological solution for user support, that ECA’s can potentially provide. See the middle part of Figure 1, that displays that eHealth studies were started to be formally tested (e.g. by means of RCT’s) in combination with ECA's. Bickmore et al. (2013) ran an RCT on ECA applied as a motivating exercise coach for elderly people. In the simulated conversations, the ECA talked using synthetic speech and animated nonverbal behavior, and participants provided input by selecting what they wanted to say from a multiple-choice list of options on the touch screen. For two months, daily conversations with the ECA-coach were conducted. Furthermore, the ECA-coach set short-term and long-term exercise goals. As a follow-up on the two-month period, participants could interact with the ECA in a kiosk in their clinic waiting room for another 10 months. Control participants were given a control pedometer intervention that only tracked step counts for an equivalent period of time. Their results showed that ECA participants walked significantly more steps than control participants at two months, but this effect waned by 12 months. Moreover, intervention participants were highly satisfied with the program. In summary, the study demonstrated partial evidence. As second renowned eHealth-ECA study is provided by Lucas et al. (2017). This real-life study examined whether virtual human interviewers could increase disclosure of mental health symptoms among active-duty service members that just returned from a yearlong deployment in Afghanistan. Their ECA was based on the highly advanced Primer® platform (previously known as SimSensei®) that enabled the ECA to detect facial expressions of users and derive the user's emotional state. A virtual human interviewer conducted a semi structured screening interview via spoken language. The interviewer did the interview in three phases: phase one: rapport building, phase two: the clinical phase during which the interviewer asked a series of questions about symptoms, and finally phase three: rounding of the interview and bringing the interviewee back to a good mood, by asking questions such as: "What are you most proud of?". Service members reported more symptoms during a conversation with a virtual human interviewer than on the official Post-Deployment Health Assessment (PDHA) symptom checklist on paper. However, the results approached but did not reach statistical significance. This study demonstrates that an eHealth ECA can potentially be the better alternative compared to paper. Moreover, this study has been conducted within a psychologically highly sensitive context (user's fear of stigma), which makes it both ambitious and unique. However, as the ECA effects are not significant, the study does not provide hard evidence. Thus, again this study shows that the deployment of an ECA in eHealth is not a panacea.

The relationship between synchronous behavior and rapport building

The previous sections have described how ECA's have developed, especially within an eHealth context. Furthermore, it was laid out that -despite that they have been studied for more than two decades- their full potential has not really been demonstrated yet. Various studies (e.g. Bickmore & Picard, 2005; Veletsianos & Russell, 2014) have advised to run long-term ECA studies for the sake of getting a grip on the rapport (relationship) building process. But is that the whole story? Or are there alternative methods to create rapport? Indeed, there are complementary views on rapport building. That is, human communication studies have reported on synchronous movement rhythms leading to feelings of rapport, and resulting experiences of being part of one and the same social unity (Marsh, Johnston, Richardson & Schmidt, 2009; Tickle-Degnen & Rosenthal, 1990; Lakens & Stel, 2011). Moving in synchrony is argued to influence the degree to which individuals are perceived as a social unit (Marsh et al., 2009; Yzerbyt, Corneille, Seron & Demoulin, 2004). But also individuals themselves report experiences of being part of one and the same team. On a neural level this is explained by pathways that code for both action and perception (Overy & Molnar-Szakacs, 2009) which causes blurring of the self and the other. Altogether, based on the findings done within the human-human communication context, synchronous behavior is relevant for further exploration within the human-ECA context.

Agent-Based Models as used to simulate Human Computer Interaction

So far, primarily human-ECA studies have been described that use a rather traditional methodology. Typically, these studies mainly rely on post-experimental questionnaires that aim to capture all human-ECA interactions that take place during the experiment. However, in reality participants express a wealth of information, even during a small portion of a single experimental session. Facial and posture signals typically change on a short-term basis (10-40 seconds) and provide a rich source of information (D'Mello & Graesser, 2012). That is, on a moment-to-moment basis it can be assessed what emotions an eLearning or eHealth user is actually expressing. Moreover, the dynamics of these emotions (e.g. do they change from a certain state to its opposite state? If so, how often? Do they also change back? If so, how quickly?) can be studied by making use of the tools provided by the Dynamical Systems Perspective. Note that the Dynamical Systems Perspective is a class of mathematical equations that describe time-based systems with particular properties such as complexity and non-linearity and can be simulated through Agent-Based Models. An agent-based model (ABM) is a class of computational models for simulating the actions and interactions of autonomous agents (both inter-individual and intra-individual) with a view to assessing their effects on the system as a whole (Klein, Marx & Fischbach, 2018; Davis, O'Mahony, Gulden, Osoba & Sieck, 2018).

Research Questions

Altogether, based on the topics referred to in the previous sections, I have defined the following overarching question for this thesis: *is it advisable to add an ECA to a self-guided eHealth intervention for the purpose of eHealth user adherence and if so, what are the necessary conditions and guidelines?*

In more detail, I have defined the following supportive questions:

- RQ1: What does the scientific literature tell us about unaddressed eHealth user needs and about the capabilities of ECA's towards addressing these needs?
- RQ2: Do eHealth users appreciate ECA support in eHealth and if so, does the induction of experimental stress lead to higher appreciation levels?
- RQ3: Can rapport between user and ECA be built through synchronous speech?
- RQ4: Can an eHealth user's cognitive-affective states be computationally simulated and if so, can critically low user motivation states be repaired through ECA support?

Outline of the thesis

The following five studies will be described in this thesis:

Study one (Chapter two)

The first study (Chapter two) is a Scoped Review. Taking a meta-perspective, this study examines user needs towards support as expressed by eHealth users, support characteristics of ECA's and finally a mapping of user needs on the potential of ECA's to address them. The outcome of this study is used as the basis for the experimental design of study 2 (Chapter three), study four (Chapter five) and for the simulation study 3 (Chapter four).

Study 2 (Chapter three)

In order to investigate whether an ECA can effectively provide the types of support stemming from study one, study two (Chapter three) has been set up. An ECA is deployed as an adjunct to a self-guided positive psychology psycho-education intervention. The agent provides instructions and motivational support in between the online learning modules as to mitigate the risk of distraction. By deploying three versions of an ECA, varying the features of animation, speech, and visibility it is investigated whether users have a more positive experience than with a fourth text-only control condition.

Study three (Chapter four)

The lessons learned from study two are used for study three, being the second experimental ECA study. At the start of the experiment, stress is induced to one of the experimental conditions as a means to both increase the user need for external support and to make the experiment more life-like. A monologue-style ECA is deployed and compared with textual guidance as a control condition. The objective of this study is to find out whether stress induces a larger appreciation for ECA support amongst eHealth users.

Study four (Chapter five)

As a qualitative study, study four (Chapter five) explores the possibility to build rapport between user and ECA by simultaneously speaking out a series of phrases. It is known from human communication studies that doing the same thing at the same time (synchronizing) creates a bond. In daily life, this is experienced when carrying out the same dancing moves or when chanting together during rituals. A suitable synchronous task for human-ECA interaction would be to speak out textual phrases simultaneously. However, as far as we know, such as task has not yet been part of an experimental context. So in short, study five examines the effect of simultaneous speech between user and ECA with the objective of creating a bond. The rationale is that the existence of such a bond will make the ECA more effective when it is deployed as adjunct within an eHealth intervention.

Study five (Chapter six)

The fifth study takes a fundamental different angle than the previous study. Not just by using a different research method, a simulation, but especially as it focusses on the dynamical aspects of user motivation. An exploratory Agent-Based Model (ABM) on user motivation during eHealth psycho-education is designed, based on the literature of both motivational psychology and agent-based modeling. Simulations are run to find out whether critical points of user motivation can be found. These critical points such as persistent frustration are considered as the immediate precursors of non-adherence, that should either be avoided or repaired.

References

Barak, A., Klein, B., & Proudfoot, J. G. (2009). Defining internet-supported therapeutic interventions. Annals of behavioral medicine, 38(1), 4-17.

Baylor, A. L. (2011). The design of motivational agents and avatars. Educational Technology Research and Development, 59(2), 291-300.

Bickmore, T. W., & Picard, R. W. (2005). Establishing and maintaining long-term human-computer relationships. ACM Transactions on Computer-Human Interaction (TOCHI), 12(2), 293-327.

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.

Brandtzaeg, P. B., & Følstad, A. (2017, November). Why people use chatbots. In International conference on internet science (pp. 377-392). Springer, Cham.

Cassell, J. (1999). Embodied conversation: integrating face and gesture into automatic spoken dialogue systems.

Davis, Paul K., Angela O'Mahony, Timothy R. Gulden, Osonde A. Osoba, and Katharine Sieck, Priority Challenges for Social and Behavioral Research and Its Modeling. Santa Monica, CA: RAND Corporation, 2018. https://www.rand.org/pubs/research_reports/RR2208.html. Also available in print form.

D'Mello, S., & Graesser, A. (2012). Dynamics of affective states during complex learning. Learning and Instruction, 22(2), 145-157.

Eysenbach, G. (2001). What is e-health?. Journal of medical Internet research, 3(2), e20.

Fogg, B. J. (1998, January). Persuasive computers: perspectives and research directions. In Proceedings of the SIGCHI conference on Human factors in computing systems (pp. 225-232).

Gratch, J., Wang, N., Gerten, J., Fast, E., & Duffy, R. (2007, September). Creating rapport with virtual agents. In International workshop on intelligent virtual agents (pp. 125-138). Springer, Berlin, Heidelberg.

Hoque, M., Courgeon, M., Martin, J. C., Mutlu, B., & Picard, R. W. (2013, September). Mach: My automated conversation coach. In Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing (pp. 697-706).

Johnson, W. L., & Lester, J. C. (2016). Face-to-face interaction with pedagogical agents, twenty years later. International Journal of Artificial Intelligence in Education, 26(1), 25–36. doi:10.1007/s40593-015-0065-9

Kelders, S. M., Kok, R. N., Ossebaard, H. C., & Van Gemert-Pijnen, J. E. (2012). Persuasive system design does matter: a systematic review of adherence to web-based interventions. Journal of medical Internet research, 14(6), e152.

Klein, D., Marx, J., & Fischbach, K. (2018). Agent-based modeling in social science, history, and philosophy. an introduction. Historical Social Research/Historische Sozialforschung, 43(1 (163), 7-27.

Kraft, P., & Yardley, L. (2009). Current issues and new directions in Psychology and Health: What is the future of digital interventions for health behaviour change?.

Lakens, D., & Stel, M. (2011). If they move in sync, they must feel in sync: Movement synchrony leads to attributions of rapport and entitativity. Social Cognition, 29(1), 1-14.

Lester, J. C., Stone, B. A., & Stelling, G. D. (1999). Lifelike pedagogical agents for mixed-initiative problem solving in constructivist learning environments. User modeling and user-adapted interaction, 9(1), 1-44.

Lucas, G. M., Rizzo, A., Gratch, J., Scherer, S., Stratou, G., Boberg, J., & Morency, L. P. (2017). Reporting mental health symptoms: breaking down barriers to care with virtual human interviewers. Frontiers in Robotics and AI, 4, 51.

Marsh, K. L., Johnston, L., Richardson, M. J., & Schmidt, R. C. (2009). Toward a radically embodied, embedded social psychology. European Journal of Social Psychology, 39, 1217-1225.

Mohr, D., 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), e30.

Nass, C., Steuer, J., & Tauber, E. R. (1994). Computers are social actors. In Proceedings of SIGCHI ’94 Human Factors in Computing Systems (pp. 72–78). ACM. https://doi. org/10.1145/259963.260288

Nass, C., Fogg, B. J., & Moon, Y. (1996). Can computers be teammates?. International Journal of Human-Computer Studies, 45(6), 669-678.

Neff, R., & Fry, J. (2009). Periodic prompts and reminders in health promotion and health behavior interventions: systematic review. Journal of medical Internet research, 11(2), e16.

Oinas-Kukkonen, H., Harjumaa, M. (2009) Persuasive systems design: Key issues, process model, and system features. Communications of the Association for Information Systems 24, Article 28, 485–500.

Oinas-Kukkonen, H. (2010, June). Behavior change support systems: A research model and agenda. In *International Conference on persuasive technology* (pp. 4-14). Springer, Berlin, Heidelberg.

Overy, K., & Molnar-Szakacs, I. (2009). Being together in time: Musical experience and the mirror neuron system. Music Perception, 26(5), 489-504.

Picard, R. W. (2000). Affective computing. MIT press.

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), e151.

Reeves, B., & Nass, C. (1996). The media equation: How people treat computers, television, and new media like real people. Cambridge, UK: Cambridge university press.

Reeves, B. (2000). The benefits of interactive online characters. Center for the study of language and information, Stanford University.

Schroeder, N. L., Adesope, O. O., & Gilbert, R. B. (2013). How effective are pedagogical agents for learning? A meta-analytic review. Journal of Educational Computing Research, 49(1), 1-39.

Schroeder, N. L., & Gotch, C. M. (2015). Persisting issues in pedagogical agent research. Journal of Educational Computing Research, 53(2), 183-204.

Shaw, E., Johnson, W. L., & Ganeshan, R. (1999). Pedagogical agents on the web. In Proceedings of the third annual conference on Autonomous Agents (pp. 283-290).

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.

Tickle-Degnen, L., & Rosenthal, R. (1990). The Nature of Rapport and Its Nonverbal Correlates. *Psychological Inquiry, 1*(4), 285-293.

Van der Heijden, J. P., De Keizer, N. F., Bos, J. D., Spuls, P. I., & Witkamp, L. (2011). Teledermatology applied following patient selection by general practitioners in daily practice improves efficiency and quality of care at lower cost. British Journal of Dermatology, 165(5), 1058-1065.

van Gemert-Pijnen, J. E., Nijland, N., van Limburg, M., Ossebaard, H. C., Kelders, S. M., Eysenbach, G., & Seydel, E. R. (2011). A holistic framework to improve the uptake and impact of eHealth technologies. Journal of medical Internet research, 13(4), e111.

Veletsianos, G., & Russell, G. S. (2014). Pedagogical agents. In Handbook of research on educational communications and technology (pp. 759-769). Springer, New York, NY.

Verhoeven, F., Tanja-Dijkstra, K., Nijland, N., Eysenbach, G., & van Gemert-Pijnen, L. (2010). Asynchronous and synchronous teleconsultation for diabetes care: a systematic literature review. Journal of diabetes science and technology, 4(3), 666-684.

Webb, T., Joseph, J., Yardley, L., & Michie, S. (2010). Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. Journal of medical Internet research, 12(1), e1376.

Weiss, B., Wechsung, I., Kühnel, C., & Möller, S. (2015). Evaluating embodied conversational agents in multimodal interfaces. Computational Cognitive Science, 1(1), 1-21.

Yzerbyt, V., Corneille, O., Seron, E., & Demoulin, S. (2004). Subjective essentialism in action: Self-anchoring and social control as consequences of fundamental social divides. In V. Yzerbyt, C. M. Judd, & O. Corneille (Eds.), The psychology of group perception: Perceived variability, entitativity, and essentialism (pp. 101-124). New York: Psychology Press.

Chapter 2

Self-guided web-based interventions: a scoping review on user needs and on the potential of virtual agents to address them

Scholten MR, Kelders SM, Van Gemert-Pijnen JEWC
Self-guided web-based interventions: a scoping review on user needs and on the potential of virtual agents to address them

J Med Internet Res 2017;19(11):e383

Abstract

Background: Web-based mental-health interventions have evolved from innovative prototypes to evidence-based and clinically applied solutions for mental diseases such as depression and anxiety. Open-access, self-guided types of those solutions hold the promise of reaching and treating a large population against reasonable costs. However, a considerable factor that currently hinders the effectiveness of these self-guided web-based interventions is the high level of non-adherence. The absence of a human caregiver apparently has a negative effect on user adherence. However, it is unknown what it is of this support that yields higher adherence and effectiveness.

Objective: The objective of this paper is first to explore what is known in the literature about what support a user needs to keep motivated and engaged in an eHealth intervention. Second, the objective of the paper is to explore the current potential of Virtual Agents (VAs) to provide this support.

Methods: This study reviews and interprets the available literature on support within e-health interventions and the potential of VAs by means of a scoping review. The rationale for choosing a Scoping Review is that the subject is broad, diverse and largely unexplored which warrants a scoping review methodology in which it is sought to present an overview of a such potentially large and diverse body of literature pertaining to a broad topic.

Results: The results of the first part of this study suggest that during usage of self-guided online interventions there are user needs in terms of support and empathy that currently remain largely implicit and unaddressed. These support needs can be categorized as task related support and emotion related support. The results of the second part of this study suggest that VAs are capable of engaging and motivating users of IT applications in the domains of learning and behavioral change. However, especially longitudinal studies must be conducted to find out under what circumstances VAs can create and maintain a productive user relationship. Mapping the user needs on the VA capabilities suggest that VAs may provide a solution for improving the adherence levels.

Conclusion: Non-responsive VAs taking on an empathic role may be sufficient to create some positive impact on users. It is unclear, however, whether those type of VAs are competent enough and create sufficient believability amongst users to address the user's deeper needs for support and empathy. Responsive VAs may be

better suited for the job, but as they are costly to realize and maintain, further research should investigate whether this is a worthwhile path to take.

Keywords: eHealth; web-based intervention; embodied conversational agent; virtual agent; virtual humans; adherence; attrition

Introduction

Meta–analyses have demonstrated that web-based interventions for mental health have become reasonably successful treatments against common mental health problems such as depression and anxiety [1-3]. However, it is a consistent finding that human-supported web-based therapeutic interventions outperform self-guided interventions [4] (in which there is no support from a human). The mere online, sometimes remote presence of a human being, delivering informational support, emotional support or a therapeutic service results in significantly higher effect sizes [5]. In addition, human-supported interventions achieve higher rates of adherence, that is more participants use the intervention as intended, e.g. by completing all the lessons of an intervention [2, 3, 6]. Non-adherence is an important issue in web-based interventions for mental health [7] and becomes an even bigger problem when evidence-based therapies are deployed as free to access self-guided web-based therapeutic interventions [8]. In these interventions, adherence, defined as the percentage of users who complete all lessons, falls to a level as low as 1% [7] or even 0.5% [8]. The higher rates of adherence in human-supported interventions can be explained in favor of therapists, who prove to do an effective job in motivating clients during their change process [5]. However, also positive effects of electronic interventions have been found by features such as reminders and tailored advice [9]. Interestingly enough, Talbot [10] describes in her meta-study that a key converging finding is that the involvement of a professional support provider, a therapist, is not necessary. What is key is a minimal level of non-guiding human contact. Irrespective of whether this type of contact is provided by a layperson or a professional, it has equally large positive effects on intervention adherence. Moreover, scheduling support can already have an effect of itself on treatment effectiveness. A telephone contact scheduled at the start of the treatment to take place as soon as a self-help book has been read, yields surprisingly large completion rates and treatment outcomes [11]. This poses the question what this support is that is needed to achieve higher rates of adherence and effectiveness. A study of Cavanagh and Millings [12] provides evidence of built-in ‘common factors’ such an generating hope, empathy and warmth, collaboration and feedback, that increase the effectiveness of interventions. However, there is no common definition of the kind of support or ‘common factors’ that should be included in each intervention to be effective. The urgence of support is expressed by the statement of Kreijns et al. [13] who declare that the reason that digital learning environments fail is due to socio-emotional processes being “ignored, neglected, or forgotten”. As web-based health interventions share many characteristics with digital learning environments, it is a fair assumption that the same socio-emotional processes play a role and should be subject to study in relation to adherence.

Methods

This study was performed by means of structured data collection within the Web of Science and Scopus databases. As research method the Scoping Review has been chosen. A Scoping Review aims to map the existing literature in a field of interest in terms of the volume, nature, and characteristics of the primary research [14]. The rationale for choosing a Scoping Review for the subject of this paper, is that research on web-based interventions forms a large and diverse body of literature in which the role of support and its relationship to user motivation is barely explored and poorly understood. This is equally the case for system provided support provided by VAs within e.g. social learning contexts [15]. As far as to the best of our knowledge no studies have been conducted so far that systematically aimed to match user needs for web-based interventions to VA capabilities in order to find potential solutions for low adherence to the interventions. Having said that, seminal studies (e.g. [16] have suggested and indeed partly demonstrated that VAs have the potential to stimulate and motivate users which ultimately may have a positive effect on intervention adherence, which underscores the importance of the current study.

The study is divided into two parts:

- Part 1: a scoping review of meta studies on support in web-based interventions
- Part 2: a scoping review of the opportunities of virtual coaches to deliver support within web-based interventions for health or learning

Search strategy part 1: meta studies on support in web-based interventions
The Scopus database was searched with a combination of the concepts ‘support’, ‘web-based intervention’, and ‘review’. For each of the concepts, multiple key words were used (see Appendix A).

Inclusion criteria were:

- Papers had to address a web-based intervention for a mental and/or physical disorder in which support was subject of study
- Papers had to review multiple interventions/studies, or present ideas based on literature or earlier work

Exclusion criteria were:

- Papers that restricted themselves to a specific disease and/or intervention and did not generalize to eHealth within a broader context
- Papers that described the creation of a web-based intervention and did not take the empirical evaluation in scope
- Papers on social media and support solutions that were studied separate from the web-based intervention events
- Papers that did not describe support in functional terms (e.g. praise, reassurance) but only in technical delivery terms (e.g. SMS, e-mail)
- Papers that analyzed web-based interventions using high-level descriptive factors (e.g. “interactive component”, “supervision”, “tailored”) without going into more detail

The search resulted in 93 articles. Based on our in- and exclusion criteria, we selected 18 studies. By checking the references of these selected articles, we found another 4 relevant papers. Finally, 22 papers were included. See Figure 1 for the selection process.

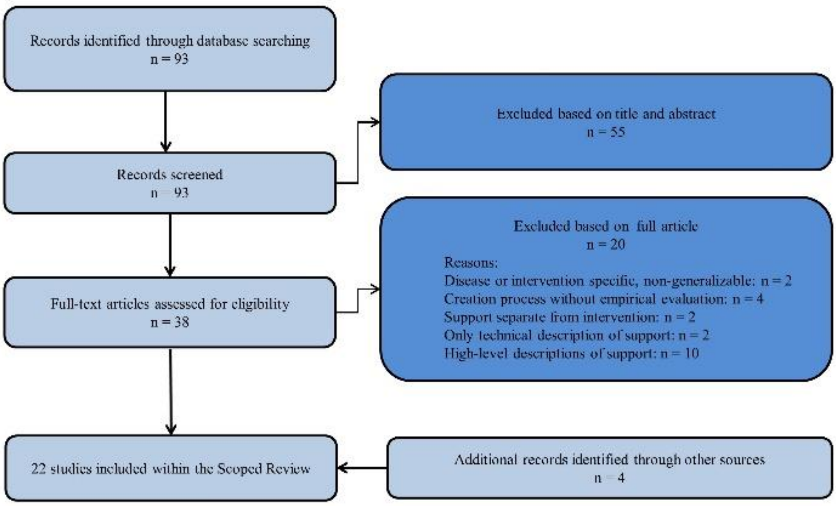


Figure 1. Flow diagram of the study selection of part 1 of the scoping review

Data extraction part 1

The entire content, including the introduction, discussion and references, of the 22 studies was checked for the user's perspective regarding usability and the needs they expressed. Subsequently, the specific needs with regards to support and motivation that –according to the authors- many self-guided web-based interventions were missing out on were listed and categorized within themes.

Search strategy part 2: opportunities of virtual coaches to deliver support within web-based interventions for health or learning

The search aimed to create a generic idea of the capabilities of VAs for supportive purposes. The Scopus and Web of Science databases were searched with a combination of the concepts 'virtual agents, 'web-based intervention', and 'support'. For each of the concepts, multiple key words were used (see Appendix A). As VAs are often used within a e-learning context, it was decided to include studies on Intelligent Tutoring Systems (ITS) as well. ITS was included as a key-word within the concept of 'web-based intervention'

Inclusion criteria were:

- Papers had to address VAs interacting with users or studies on VAs interacting with users

Exclusion criteria were:

- Papers that solely focused on the effects of VAs in Virtual Reality. VA studies that addressed VAs in VR but also on regular screens were not excluded.
- Papers that described computer simulations with agents/ during which interaction between human users and VAs were absent
- Papers that described a set-up of a VA but did not take the empirical validation in scope

The systematic search resulted in a limited number of studies (8). Moreover, these studies addressed a wide range of topics; from physical attributes [17], architecture [18], route planning [19], non-verbal behavior [20], virtual museum guide [21], empathy [22], to theoretical models [23] and articulation rates [24]. None of the studies provided a high-level picture of the capabilities of VAs with regards to support delivery. Therefore it was decided to expand the number of articles by means of hand search. We started the hand search by checking references within the 8 articles and searching on terms found within the 8 articles in Google Scholar.

The hand search had the following aims:

a) Finding synthesizing information on VAs within a health or pedagogical context with a focus on the delivery of support and motivating users. We started with the information found in [22] and additionally searched for meta studies on VAs.
b) Finding additional (founding) articles on the CASA effect as mentioned within [17] and [22].
c) Finding addition information on relationship building [24] and measures of relationship building as shortly described in [20, 24].
d) Finding additional information on theoretical models related to VAs as touched upon in [23].

The search procedure resulted in 53 included articles (Figure 2).

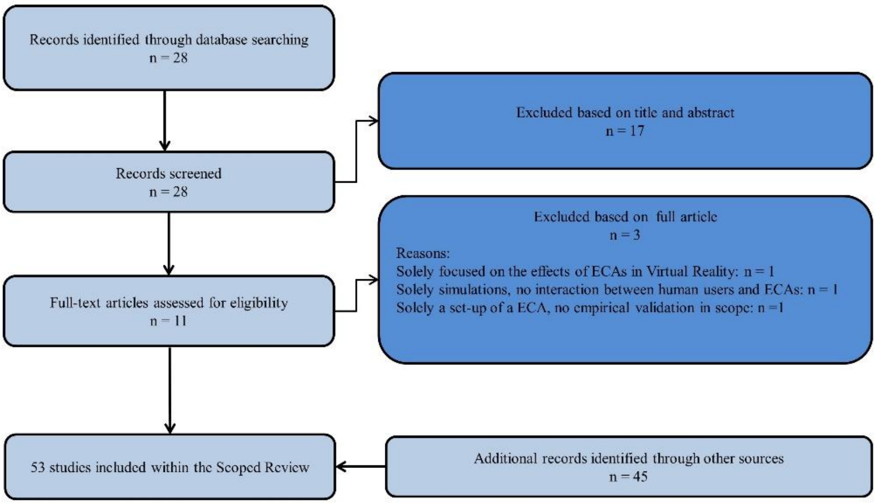


Figure 2. Flow diagram of the study selection of part 2 of the scoping review

Data extraction part 2

The entire content, including the introduction, discussion and references of the articles were analyzed on the presence of VA features. Subsequently, the various VA features were categorized within themes. The themes were chosen as a means to provide insight in the capacities of VAs first to communicate with users and secondly to motivate users.

Results

Part 1: results and themes found within the studies on the need for support in web-based interventions

The analyzed 22 studies suggest that a myriad of subtle interactions between users and computers play an important role in keeping a user motivated in continuing the web-based intervention. Although the elements of these interactions are very diverse, two common elements can be distinguished:

1. Users express the need for concrete feedback on their performance. Within the literature, this need is described as the principle of closure: the confirmation that an action has been successfully performed. This indicates that users of web-based interventions could benefit from task-related interaction and support (e.g. “Thank you for submitting your homework for this week. You sent it well on time.”). We call this *task related system support*.
2. Users express the need for interest and support for the issues they are dealing with. This suggests that users of web-based interventions could benefit from emotional support, that acknowledges both the user’s endeavors during the change program and the originating issue the user is dealing with. The dose and timing of the emotional support seems to have a large importance due to its intimate nature. Wrong timing could potentially harm the concept of user self-determination which is an appreciated feature of web-based self-guided interventions. We call this *emotion related system support*.

Table 1 shows the user needs that became apparent in the included papers and which we related to the two common elements mentioned above.

Table 1. User needs and issues and common user support mechanisms that can potentially fulfill these needs

User need or Issue	Common element	Source
1.Overcome user feelings of isolation	**Task related support** can fulfill this need by setting and reviewing login goals, positively reinforcing login and site use and answering questions regarding the functionality of the site. **Emotion related support** can fulfill this need by establishing a supportive relationship, In case login goals are not met or any other sign of diminished use of the intervention appears, the system can intercede and encourage use of the online intervention	[6, 25-28] [6, 25, 27, 28]
2.Interest in identity of user	**Emotion related support** can fulfill this need by providing the user with the opportunity to talk about the impact of the disease on their live and their idea on having become a patient.	[26, 29, 30]
3.Interest in concrete daily issues the user is struggling with	**Emotion related support** can fulfill this need by asking the user about their daily experiences and issues and responding by expressing empathy towards the user. If the user expresses a need for concrete, practical advice, the system could provide it accordingly or –for more complex questions- refer to a nurse or doctor connected to the system. This can be considered as **task related support in a broader context**.	[26, 31]
4.The ability for the user to refine the communication process	**Emotion related system support** provided alongside a more open interaction between user and system (e.g. by means of bi-directional free text or free speech) could potentially (and so far theoretically) increase the user feeling of contributing to its own change process	[32, 33]
5.The user need for encouragement	**Emotion related system support** could be delivered in terms of praising the user, delivering rewards and by other types of encouraging behavior	[6, 25, 30-32, 34-37]

6.Performance feedback mechanism for user responses	**Task related support** can fulfill this need by reviewing the user's contributions and by providing corrections in case the user made factual errors. This scenario is applicable for user performances that can be objectively graded (e.g. homework with factual information about an illness). In addition –preferably if opted so by the user- the user's achievements can be plotted against the achievements of the user's peer group. This scenario is applicable for e.g. user performances that can be measured in physical terms.	[6, 34, 38-40]
7.Users coping with experiences of negative affect during their change process	**Emotion related system support** can provide a dose of positive affect in case a phase of negative user affect that merits such a dose could be reliably distinguished.	[41]
8.Creating a setting of accountability towards the user	**Task related support** can play a positive role by objective goal setting, measuring the goals set, reminding the user of their goals set and by indicating which of these goals have (not yet) been met.	[27, 28, 31, 36, 37, 39, 42-45]

In the section below, the user needs from Table 1 will be discussed in more detail.

Theme 1. Overcome user feelings of isolation.

The anonymity of web-based interventions seems to play out as both strength and a weakness. Users feel encouraged to speak out, but sometimes also feel isolated due to its anonymous nature. Both task-related and emotion-related system support could potentially counteract feelings of isolation.

Theme 2. Interest in identity of user.

At a deeper level, users seem to expect (and probably need) a deeper interest in their identities. Knowles et al. [26] conclude as shortcomings found in 7 out of 8 studies: "sensitivity to 'Who I am' as a patient, including different clinical needs such as physical comorbidity." This is a case for emotion-related system support.

Theme 3. Interest in concrete daily issues the user is struggling with.
Users seem to have wish for a form of deeper interest in their concrete daily issues. Described by Knowles et al [26] as "sensitivity to 'How I Feel', recognizing the demands of depression on the user (such as emotional and motivational difficulties, and problems with concentration)." This is a case for emotion related system support. In case the user requests practical advice for daily issues, task related support in a larger domain than the intervention itself can also contribute.

Theme 4. The ability for the user to refine the communication process.
As reported by Donkin [32] users that were filling in questionnaires about how they felt, said that the questionnaire did not cover their feelings. Subsequently, these users had a strong wish to contextualize their answers. Indeed, an non-interactive tool as a questionnaire is perfectly fit for gathering experimental user data, but may be less suitable as a communication tool as it 'forces' answers to be put into a restricted set of categories. Emotion-related system support provided alongside a more open interaction between user and system (e.g. by means of bi-directional free text or free speech) could potentially (and so far theoretically) increase the user's feeling of contributing to its own change process.

Theme 5. The user need for encouragement.
As reported by Donkin et al. [32] and as quoted by Mohr [28] ".. patients want feedback on whether they are on the "right track" in their web-based intervention." Encouraging users during the intervention can likely be achieved by emotion-related system support.

Theme 6. Performance feedback mechanism for user responses.
Somewhat comparable to the statement of Donkin et al. [32], Helgadóttir [40] describes that many CCBT programs would benefit from a performance feedback mechanism for user responses. This would expand the system's ability to direct the user during their change program. By providing a direct task-related response "I have received your answers, thank you for your time and effort. Please allow me to comment on your answers" it would acknowledge the user's effort invested. By later analyzing the user responses and by providing feedback via e-mail, a second task-related support mechanism could be implemented.

Theme 7. Users coping with experiences of negative affect during their change process.
Kraft et al. [41] suggest that individuals should be assisted in coping with experiences of negative affect during their change process. They make a claim that many change program users struggle with the tension between their aspirations and their actual

status and behavior. During this struggle, the client's internal process of self-regulation is activated in order to alleviate the tension. Too much burden on the self-regulation process leads to ego-depletion [41], a status of a low level of mental energy. This status often results in increased relapse vulnerability of which therapy non-adherence can be considered as a special case. As a way to reverse this ego-depletion process, Kraft et al. [41] recommend a dose of positive affect, next to a period of rest for recovery. Emotion-related system support could provide such a dose of positive affect. The challenge would be to determine the moment that ego-depletion could be close.

Theme 8. Creating a setting of accountability towards the user.

In order to obtain adherence to web-based interventions that include human support, Mohr et al. [37] stress the importance of creating a setting of accountability towards the user, "the implicit or explicit expectation that an individual may be called upon to justify his or her actions or inactions". For such a setting, certain preconditions are necessary, such as participants that understand and agree with the benefits of their expected future behavior. Other preconditions are concrete goal setting and performance monitoring. Task-related machine support can play a positive role by reminding the user of their goal set and by indicating which of these goals have (not yet) been met. One should keep in mind that accountability might be harder to trigger amongst users who have been assigned to health interventions by their doctors and who did not primarily opt to participate by themselves.

Part 2: results and themes found within the studies on VAs with motivational capabilities

Table 2 below shows the results and the themes that were found in the selected studies.

Table 2. Themes on supportive VAs

Theme	Explanation	Sources
1. Computers As Social Actors (CASA)	Humans treat media in the same way as they treat other humans	Systematic search: [22] Hand Search: [46-50]
2. Open dialogue between user and computer	VAs have the ability to have an open verbal dialogue with users	Systematic search: [21] Hand Search: [51-53]
3.Visible conversation partner	Interaction with a 'talking face' leads to more trust and believability.	Systematic search: [17, 19, 22, 23] Hand Search:[54-64]
4. Human-Computer relationship	Interactions with an agent can lead to a relationship, which is important to keep users engaged over time	Systematic search: [24] Hand Search:[16, 65-71]
5. Measures of the Human-Computer relationship.	Human-VA relationship quality can be measured	Systematic search: [20] Hand Search: [16, 67, 72]
6. Responsive verbal and non-verbal communication	Computers should have the ability to notice and respond to verbally and non-verbally expressed emotions from their user, in order to create a more natural interaction	Systematic search: [22] Hand Search: [62, 73-80]
7. Impact of VAs on User motivation	There is evidence that VAs can motivate users, which is highly dependent on VA implementation, context, task etc.	Systematic search: [18] Hand Search: [56, 73, 81-83]
8. Methodological issues within VA research	Most experiments into VAs face similar methodological issues which have to be taken into account when interpreting the research.	Hand Search: [84-88]

Theme 1: Computers as Social Actors (CASA)

A large body of studies on VAs refer to the CASA effect [48, 50] as a cornerstone for studying human-computer interactions and especially human-VA interactions. The CASA effect demonstrates that humans treat media – in some respect- in the same way as they treat other humans. Various manifestations of this effect have been described such as:

- Computers that display flattery texts towards their users are preferred by their users compared to computers that do not display such texts
- Computers that textually praise other computers are better liked than computers that praise themselves, and computers that 'criticize' other computers are disliked compared to computers that criticize themselves
- Users who are partnered with an computer on basis of a color (e.g. the blue team) will have a more positive opinion on the computer and cooperate more with it than users who have to partner with a computer of the opposite, differently colored team

As an explanation of the CASA effect, it has been proposed that humans have a strong innate tendency to make social connections with other humans and other living creatures such as pets. This human tendency becomes real when objects such as personal computers demonstrate activities that could be socially interpreted by their users [50]. Although pc's can act socially, human users are logically aware of their non-social and non-living status. This seems a paradox: why would a human user socially respond to a pc while at the same time realizing that a pc does not warrant it? Nass and Moon [47] refer to 'mindless' (automatic, largely unaware) human behavior that the machine can trigger. This mindless behavior will be displayed as long as it remains socially acceptable. This phenomenon is also associated with the notion of 'suspension of disbelief', meaning that up to a certain point humans are willing to apply social rules to non-human yet communicative objects, irrespective of their non-living status.

Theme 2: Open Dialogue between user and computer

A following theme is the ability of computers and VAs to have an open verbal (textual or speech) dialogue with users. Within regular, day to day Human-Computer Interaction events, a user who interacts with their IT system will typically activate pre-defined menu options such as the 'save as' option within Microsoft Word. Subsequently, the computer will respond to the request by presenting a pop-up window which will enable the user to type in the file name of the document. In such a closed dialogue scenario, the interactions between user and software traditionally have a task-specific character (e.g. serve to reach a specific goal such as saving a document), have a short duration and are typically initiated by the user (and not by the computer). In contrast, VAs enable more open-ended and more relationship-oriented interactions. Interactions between VAs and users can span multiple question and answer pairs and can therefore be interpreted as a dialogue. The ELIZA study [53] described an early version of a textual psychotherapists that gave 'canned' responses to user questions as a result of quickly processing the input text provided and create a response out of it without realizing what the user had said (e.g. a question like: "Eliza, I feel miserable today" and an answer: "How often do you experience feelings of being miserable?"). Later studies create richer dialogue contexts to explore the capabilities of computers interacting with humans. Examples are first a study that has shown that a robot taking the role of museum guide who uses e.g. empathy and humor in his conversation style led to a more positive attitude towards the robot than the same robot without this enhanced conversation style [21]. A second study showed that a VA with high dialog capabilities reached more accurate answers when interviewing a subject than an agent with less dialog capabilities [51]. A third study [52] aimed to explore where open-dialogue options between users and VAs would lead to. The authors report that when learners are given opportunities to guide an open conversation, they especially ask off-topic questions. For example, learners often want to know about the agents' operating systems, design, purpose, and capabilities. Such conversations seem to serve the 'testing' of agents' abilities during which learners are attempting to discover the boundaries, limits, and capabilities of agents through 'game-like' inquiry.

Theme 3: Visible conversational partner

The following theme is the visibility of the conversational computer depicted as a (either static or animated) human face. According to Lisetti [59] the human face has a special status in human to human communication as it has often been identified as the most important channel for conducting trust and believability. As Lisetti states, the face as a communication channel has a higher status than bodily regions such as posture and gesture [55]. Multiple studies have supported this notion by demonstrating that users preferred to interact with a 'talking face' instead of a text only interface [63], an anthropomorphic agent together with a human voice has led to greater agent credibility [54], visible agents have led to greater positive motivational outcomes [62]) and task performance [64]. Besides empirical research, there are multiple theories that support this notion. The theories that were mentioned in the included sources are listed and explained in Table 3.

Table 3. Main theories and effects of visible VAs

Theory	Explanation	Source
Theory of Social Inhibition/ Facilitation	When in the presence of others, people perform learned tasks better and novel tasks worse. Empirical results have demonstrated that this principle also applies for the presence of VAs.	[64]
Social Agency Theory	By adding a visible VA as a screen tutor the social interaction schema is primed, which will cause the learner to try to understand and deeply process the computer delivered instructions	[60]
Social Modelling/ Social Learning Theory	Humans derive their knowledge, attitudes, behavior and goals by observing and imitating the surrounding social agents.	[17, 23]
Situational Dependency	Pedagogical agents are helpful when there is a need to increase companionship and decrease complexity	[56]
Social Exchange Theory	People prefer equitable relationships in which the contribution of rewards and costs are roughly equal. This equity principle also applies to human-computer relationships.	[57]
Persona Effect	The presence of a lifelike character in an interactive learning environment—even one that is not expressive— can have a strong positive effect on student's perception of their learning experience	[58]
Image principle	Image of a VA is not a key factor for learning, the level of animation of the VA is the key factor for learning.	[60]

Despite these positive experimental results and theoretical support for a visible, human-like personal computer, the visibility subject is somewhat controversial. Strong claims against the human face are provided by Norman [49] by his statement that a human face triggers false mental models and thus creates wrong user expectations. Other critique is provided by Rajan et al. [61] who demonstrated that it is first and foremost the voice (and not the visibility of the VA) that is responsible for positive learning effects.

Theme 4: Human-VA Relationship

A fourth theme is the concept that regular human-computer interaction events result in a relationship. Routine interactions between a user and their computer should be regarded as contributions to this human-computer relationship, as is argued by Bickmore et al. [16]. Although this relationship may be implicit, it has an impact on the user. The relationship plays a role even in case no relationship skills (e.g. empathy, humor) have been designed and built into the machine. The question arises whether a VA with a relationship-focused design could behave and be perceived as a competent social actor. This quality of the VA as a conversational partner is impacted by:

- *Interaction duration*. As described by Krämer et al. [70] getting people engaged with VAs is easy, but keeping then engaged over time is much more challenging. Bickmore et al. [16] (on physical activity) and Creed et al. [66] (on fruit consumption) conducted emotional virtual coach studies that spanned more than 28 days. They both found that deploying the emotional VA did not result in user behavior changes, but that users in general preferred to interact with the emotional virtual coaches.
- *Natural vs forced interaction*. Gulz [68] suggests that most VAs studies force the human-computer relationship too much. Users have no other option than to interact with the VAs they are confronted with.
- *User personality*. Von der Pütten et al. [71] make clear that it depends on the personality of the user how the human-computer relationship will develop. They demonstrated that 5 user personality factors were better predictors for the evaluation outcome of VAs than the actual behavior of the VA.

Theme 5: Measures of the Human-VA Relationship

The literature found mentions two regular measures with regards to the Human-VA Relationship.

- *Measure 1: Working Alliance*
  Working Alliance is a construct that originates from the psychotherapy literature and has been described as “the trust and belief that the helper and patient have in each other as team-member in achieving a desired outcome” [72]. Bickmore et al. [65] applied the working alliance inventory in their 30-day longitudinal study with a VA acting as an exercise coach. Participants who interacted with a VA with relational behavior enabled (empathy, social chat, form of address, etc.) scored the VA significantly higher on the Working Alliance Inventory compared to participants who interacted with the same VA with the relational behaviors disabled.
- *Measure 2: Rapport*
  A second important human-computer relationship measure is rapport. Rapport has been described as “the establishment of a positive relationship

among interaction partners by rapidly detecting and responding to each other's nonverbal behavior" [67]. Measurement of rapport has been conducted by Gratch et al. [67] in their evaluative VA study. Their results showed that the experience of rapport was of a comparable level compared to a face-to-face (i.e human interlocutor) condition.

Theme 6: Responsive verbal and non-verbal communication

Within human to human communication, the exchange of non-verbal information plays a key role. Social psychologists assert that more than 65% of the information exchanged during a person-to-person conversation is conveyed through the non-verbal band [74, 80]. The non-verbal channel is said to be especially important to communicate socio-emotional information. Socio-emotional content [75] is vital for building trust and productive human relationships that go beyond the purely factual and task-oriented communication. D'Mello et al. [75] describes the mutual impact of user and (synthetic) computer emotions as an affective loop which is pictured as follows:

- The user first expresses their emotion through verbal and physical interaction with the machine, e.g. through detectable gestures, usage of the keyboard or spoken language
- Then, the system responds by generating affective responses, through words, speech, animation and theoretically also colors and haptics
- This response affects the user in such a way that they become more involved in their further interaction with the computer

Concerning the importance of the affective loop, there are two stances:

- *Stance 1: Responsiveness of VAs (affective loop) is a critical condition for prolonged user interaction.* Doirado et al. [77] confirm the importance of the affective loop mechanism and state that a VA that lacks the capacity to understand the user and the capability to adapt its behavior (a non-responsive VA) will break the user's suspension of disbelief.
- *Stance 2: Autonomy of VAs (no affective loop) is a sufficient condition for prolonged user interaction.* Rosenberg-Kima et al. [62] deployed an autonomous (i.e. non-responsive) VA that introduced itself and provided a twenty-minute narrative about four female engineers, followed by five benefits of engineering careers. The VA was animated and its voice and lip movements were synchronized. The VA acted autonomously ; interaction between participants and VA was purely restricted to the user clicking on

the button for text topic. The results showed that the self-efficacy of the users and of their interest in the subject presented was significantly higher within the VA + voice condition compared to the voice-only condition. In support of these results, Baylor et al. [54] state that people are willing to interact with anthropomorphic agents even when their functionality is limited. As she indicates the mere visual presence and appearance will in some contexts be the determining factor and not so much its supportive, conversational or animation capabilities.

Theme 7: Impact of VAs on user motivation

Meta-studies and reviews [68, 80, 84, 85, 88] have reported on claims and evidence for positive VAs effects on learning, engagement and motivation. Schroeder et al. reviewed 43 studies and conclude that pedagogical agents have a small but significant effect on learning as ultimate outcome. Within their study, Schroeder et al. [80] did not make a distinction between responsive and non-responsive VAs. Specific research with regard to motivating users has also been conducted by deploying responsive VAs with the task to notice user frustration and empathically respond to it. Autonomous delivery of warmth and empathy by VAs towards users has shown positive effects, and studies show that this effect may be larger at the time the user experiences frustration [73, 85, 87]. All together the evidence for VAs that are capable of motivating users is mixed and inconclusive. VAs, whether they are non-responsive or responsive, provide a positive user experience as a result of their entertainment capabilities. Responsive VAs when specifically designed to detect user frustration and to empathically respond to it, have also empirically demonstrated positive effects on user attitudes. However, these positive effects have not yet been found in ecologically valid context but only within constrained contexts such as games with clear win and lose rules and as a result of system-generated moments of user frustration.

Theme 8: Methodological issues within VA research

The inconclusiveness regarding VA evidence as mentioned within the previous theme is claimed to be caused by methodological issues [85, 88]. Methodological issues make it difficult to compare study results and to draw generic conclusions. One of those issues is the difference in set-ups amongst VA studies. To name a few:

- Different modalities used for output: (synthesized or natural) speech or text
- Different levels of responsive emotional behavior; from textual responses projected alongside a static VA to fine-grained VA facial expressions intended to mirror the user’s facial expressions
- Different roles: tutor, peer, interviewer, coach

- Different implementations/ different computer code applied as Artificial Intelligence to steer the VA with code based on different behavioral theories

Many of these issues can be resolved by using a common, open research platform for VAs, such as the Virtual Human platform as provided by USCT [86]. Other issues can potentially be resolved by a common design framework for VAs as proposed by Veletsianos et al. with their EnALI framework [87]. Concerning the duration of the change programs several studies (e.g. [65, 66]) stress that the majority of virtual coaching studies concern short time spans of hours, which makes it difficult to study the development of the human-computer relationship and to realize effects on user behavior. Both Bickmore et al. and Creed et al. [65, 66] conducted emotional virtual coach studies that spanned more than 28 days. They both found that deploying the emotional VA did not result in user behavior changes, but that users in general preferred to interact with the emotional virtual coaches. Altogether Dehn and van Mulken [85] summarize the situation as follows: "… the simple question as to whether an animated interface improves human-computer interaction does not appear to be the appropriate question to ask. Rather, the question to ask is: what kind of animated agent used in what kind of domain influence what aspects of the user's attitudes or performance ".

Discussion

Part 1 of this Scoping Review aimed to explore what is known within the literature about what support a human needs to keep on being motivated and engaged during usage of an eHealth intervention. We found various user needs and issues related to support, which we divided into the following two main categories:

- *Task related system support*; concrete performance related feed-back
- *Emotion related system support*; support that had a empathic nature

It appeared that both task related support and emotion related support are regularly expressed user needs. Both needs therefore merit further attention in terms of research that aims to improve user adherence.

Part 2 of this Scoping Review aimed to give insight into the potential of VAs to deliver effective task related or emotion related support to humans.

On a high level, the following two kinds of VAs were distinguished:

- *Non-responsive (autonomous) VAs*. These VAs are not endowed with senses to 'see' or 'hear' the verbal or non-verbal signals that the user expresses, and logically also lack the capacity to interpret these signals. The VA is visually present to send out motivational messages intended to keep the spirits up. Advantages are that these kinds of VAs have demonstrated that they can engage users. Disadvantages are that forced presence of the VA runs the risk of annoying the user and can therefore become counter-productive.
- *Responsive VAs*. These VAs have the capability to capture and analyze the verbal and/or non-verbal signals sent by the user and emotionally respond to them. These VAs are set up with the intention to understand the user and to adapt their behavior accordingly. Advantages are that these VAs can tap into the rich sources of verbal and non-verbal information as spontaneously and freely provided by humans. However, disadvantages are that realizing a VA that does understand the user is a heavy task, requiring costly computational modeling of user BDI (Believe, Desire and Intentions) and affective loop facilities with a high chance of failure.

Table 4 below associates the themes from part 1 with the themes addressed within part 2, and indicates if responsive or non-responsive VAs can address the user need.

Table 4. User needs with supportive elements, associated VA features and the needed level of responsiveness of the VA

User need or issue	**Supportive Element**	**Associated VA Features**	**Needed responsive-ness**
1. Overcome user feelings of isolation	*Task related support* can fulfill this need by setting and reviewing login goals *Emotion related support* can fulfill this need by establishing a supportive relationship	- Computers as social actors - Visible conversation partner - Human-computer relationship	A non-responsive VA is sufficient
2. Interest in identity of user	*Emotion related support* can fulfill this need by providing the user with the opportunity to talk about the impact of the disease on having become a patient.	- Computers as social actors - Open dialogue - Visible conversation partner - Human-computer relationship - Responsive verbal and non-verbal communication	Not any VA is currently likely able to address this user need.
3. Interest in concrete daily issues the user is struggling with	*Emotion related support* can fulfill this need by asking the user about their daily experiences and issues.	- Computers as social actors - Open dialogue - Visible conversation partner - Human-computer relationship - Responsive verbal and non-verbal communication	A responsive VA is necessary, further research is advised.
4. The ability for the user to refine the communication process	*Emotion related support* provided alongside a more open interaction between user and system	- Open dialogue	A responsive VA is necessary, further research is advised.
5. The user need for encouragement	*Emotion related support* could be delivered in terms of e.g. praising the user	- Motivational effects	A non-responsive VA is sufficient
6. Performance feedback mechanism for user responses.	*Task related support* can fulfill this need by reviewing the user's contributions and by providing corrections in case the user made factual errors	- Computers as social actors - Visible conversation partner - Human-computer relationship	A non-responsive VA is sufficient
7. Users coping with experiences of negative affect during their change process	*Emotion related support* in the sense of providing a dose of positive affect at the right moment	- Responsive verbal and non-verbal communication - Motivational effects	A responsive VA is necessary, further research is advised.
8. Creating a setting of accountability towards the user	*Task related support* can play a positive role by objective goal setting	Computers as social actors Human-Computer Relationship	A non-responsive VA is sufficient

Principal Results

Non-responsive VAs

As described within table 4 above, non-responsive VAs are capable of helping out users with lighter/ more straightforward motivational tasks as described within the themes 1, 5, 6 and 8. Non-responsive VAs provide a likely solution to engage users which will de-isolate these users to some extent. A non-responsive VA can provide task-related support such as setting and reviewing login goals and emotion-related support by the delivery of supportive messages (theme 1). Non-responsive VAs are capable of motivating users by techniques such as praising (theme 5), delivering performance feedback (theme 6), and setting an expectation level towards the user (theme 8).

Responsive VAs

In contrast to non-responsive VAs, responsive VAs are capable of performing more complex motivational tasks as described within the themes 3, 4 and 7. First, responsive VAs are capable of having a dialogue with the user during which concrete daily issues the user is facing, can be effectively discussed (theme 3). Further research should focus on effective counter measures for users losing interest interacting with responsive VAs during longer-term interactions (e.g. 4-10 weeks with daily contact [65]. Logically, only with maintained user interest, concrete daily issues will be discussed and VAs can prove to be effective interlocutors. Second, during a dialogue the user can communicate what they are experiencing, which can serve as an alternative to filling in a questionnaire. This provides the user with the ability to refine the communication process (theme 4). Further research should focus on the accompanying technical and conversational complexities of such a refining dialogue. Thirdly, a responsive VA is capable of assisting users who cope with experiences of negative affect during their change process (theme 7). However, current experimental set-ups can only artificially create an subsequently mitigate a moment of frustration. Further research on VAs that detect and respond to spontaneous user emotion, should be conducted.

Not addressable by either responsive or non-responsive VAs

Dialogues between user and VA on deep, personal issues and identity related matters (theme 2) are currently technically too complex to realize. Smooth interactions are a necessary condition for VAs to become and remain a trustworthy counterpart. None of the VAs found is capable of truly meeting this condition of smoothness. These dialogues could therefore at present be best carried out by a human support provider.

Design factors for both responsive and non-responsive VAs

The VA literature of part 2 e.g. [55, 73] gave indications on successful design of VAs. Some design factors seems to be generically of value, irrespective of deploying either a responsive or non-responsive VA within an eHealth intervention. First, it is recommended [73] to communicate the *intention*, *capabilities* and *limitations* of the VA to the user. That is: The VA presents itself (e.g. as a coach, tutor or peer) before the start of the intervention and behaves according to its role and does so consistently. This way, the user will have clear expectations. Secondly, it is recommendable to provide users with the control over the presence of the VA, especially during longer term interactions. This will avoid annoyance amongst some users as reported by [16]. Thirdly, it is recommendable that the VA has short dialogues with the user that are restricted to "greasing the wheels" with regards to the user following the eHealth program. Systems that permit open dialogues with VAs often result in dialogues during which the VA is playfully tested [52]. By limiting the scope and length of the dialogues, the VA will more likely keep up its credibility as limited yet helpful interlocutor.

Limitations

This review has several limitations. No quantitative analyses were done and selection of the articles was done by interpretation of the researchers.

Conclusion

This study aimed to link 'demand' for support in web-based mental health interventions to the 'supply' of support from VAs.
Spontaneous empathy and the explicitly expressed intention of *non-responsive VAs* to deliver user support is likely to resolve some of the lighter motivational issues eHealth users are currently dealing with. *Responsive VAs* have even larger potential. However, they are more costly to realize and they create higher user expectations, which lead to a higher risk of failure. It may therefore be reasonable to first further explore the possibilities of non-responsive VAs and investigate what their added value may be in real world web-based mental health interventions. As a second step, it could then be explored if there is a need for responsive VAs and in what contexts. Krämer et al. [57, 70] suggest that humans prefer equitable human-computer relationships in which the contribution of rewards and costs are roughly equal. It is therefore an interesting hypothesis to empirically investigate whether such VA behavior can contribute to a more balanced human-computer relationship. Especially within a context during which users are asked to perform effort requiring tasks (such as learning or working on behavior change), a dose of positive affect may serve as an effective counterbalance to the user effort invested. Put differently; in case that the computer is not only *demanding* the user to perform tasks and invest time and effort but also actively *providing* support, the human-computer relationship may become more equipollent. Such an equipollent relationship will hypothetically last longer and stimulate web-based intervention adherence.

Conflicts of Interest
None declared

Abbreviations
VA: Virtual Agent

References

1. Andrews G, Cuijpers P, Craske MG, McEvoy P, Titov N. Computer Therapy for the Anxiety and Depressive Disorders Is Effective, Acceptable and Practical Health Care: A Meta-Analysis. PloS one 2010 Oct 13;5(10). ARTN e13196 10.1371/journal.pone.0013196
2. Cuijpers P, Donker T, Johansson R, Mohr DC, van Straten A, Andersson G. Self-guided psychological treatment for depressive symptoms: a meta-analysis. PloS one 2011;6(6):e21274. 10.1371/journal.pone.0021274
3. Richardson T, Richards D. Computer-based psychological interventions for depression treatment: a systematic review and meta-analysis. Clinical Psychology Review 2012;32(4):329-342.
4. Barak A, Klein B, Proudfoot JG. Defining Internet-Supported Therapeutic Interventions. Ann Behav Med 2009 Aug;38(1):4-17. 10.1007/s12160-009-9130-7
5. Andersson G, Carlbring P, Berger T, Almlöv J, Cuijpers P. What makes internet therapy work? Cognitive behaviour therapy 2009;38(S1):55-60.
6. Kelders SM, Kok RN, Ossebaard HC, Van Gemert-Pijnen JEWC. Persuasive System Design Does Matter: A Systematic Review of Adherence to Web-Based Interventions. J Med Internet Res 2012 Nov-Dec;14(6):17-40. ARTN e152 10.2196/jmir.2104
7. Schubart JR, Stuckey HL, Ganeshamoorthy A, Sciamanna CN. Chronic Health Conditions and Internet Behavioral Interventions A Review of Factors to Enhance User Engagement. Cin-Comput Inform Nu 2011 Feb;29:Tc9-Tc20. 10.1097/NCN.0b013e3182155274
8. Cugelman B, Thelwall M, Dawes P. Online Interventions for Social Marketing Health Behavior Change Campaigns: A Meta-Analysis of Psychological Architectures and Adherence Factors. J Med Internet Res 2011 Jan-Mar;13(1):84-107. ARTN e17 10.2196/jmir.1367
9. Fry JP, Neff RA. Periodic Prompts and Reminders in Health Promotion and Health Behavior Interventions: Systematic Review. J Med Internet Res 2009 Apr-Jun;11(2). ARTN e16 10.2196/jmir.1138
10. Talbot F. Client Contact in Self-Help Therapy for Anxiety and Depression: Necessary But Can Take a Variety of Forms Beside Therapist Contact. Behav Change 2012 Jun;29(2):63-76. 10.1017/bec.2012.4
11. Newman MG, Szkodny LE, Llera SJ, Przeworski A. A review of technology-assisted self-help and minimal contact therapies for anxiety and depression:

Is human contact necessary for therapeutic efficacy? Clinical Psychology Review 2011 Feb;31(1):89-103. 10.1016/j.cpr.2010.09.008
12. Cavanagh K, Millings A. (Inter) personal computing: the role of the therapeutic relationship in e-mental health. Journal of Contemporary Psychotherapy 2013;43(4):197-206.
13. Kreijns K, Kirschner PA, Jochems W. Identifying the pitfalls for social interaction in computer-supported collaborative learning environments: a review of the research. Comput Hum Behav 2003 May;19(3):335-353. Pii S0747-5632(02)00057-2 Doi 10.1016/S0747-5632(02)00057-2
14. Pham MT, Rajic A, Greig JD, Sargeant JM, Papadopoulos A, McEwen SA. A scoping review of scoping reviews: advancing the approach and enhancing the consistency. Res Synth Methods 2014 Dec;5(4):371-385. 10.1002/jrsm.1123
15. Kramer NC, Bente G. Personalizing e-Learning. The Social Effects of Pedagogical Agents. Educ Psychol Rev 2010 Mar;22(1):71-87. 10.1007/s10648-010-9123-x
16. Bickmore T, Gruber A, Picard R. Establishing the computer-patient working alliance in automated health behavior change interventions. Patient Educ Couns 2005 Oct;59(1):21-30. 10.1016/j.pec.2004.09.008
17. Jordine K, Wilson D-M, Sakpal R. What Is Age's Affect in Collaborative Learning Environments? Proceedings of the International Conference on Universal Access in Human-Computer Interaction; 2013.
18. Ieronutti L, Chittaro L. Employing virtual humans for education and training in X3D/VRML worlds. Comput Educ 2007 Aug;49(1):93-109. 10.1016/j.compedu.2005.06.007
19. Hofmann H, Tobisch V, Ehrlich U, Berton A. Evaluation of speech-based HMI concepts for information exchange tasks: A driving simulator study. Comput Speech Lang 2015 Sep;33(1):109-135. 10.1016/j.csl.2015.01.005
20. Novick D, Gris I. Building Rapport between Human and ECA: A Pilot Study. Human-Computer Interaction: Advanced Interaction Modalities and Techniques, Pt Ii 2014;8511:472-480.
21. Bickmore TW, Vardoulakis LMP, Schulman D. Tinker: a relational agent museum guide. Auton Agent Multi-Ag 2013 Sep;27(2):254-276. 10.1007/s10458-012-9216-7
22. Amini R, Lisetti C, Yasavur U, Rishe N. On-demand virtual health counselor for delivering behavior-change health interventions. Proceedings of the Healthcare Informatics (ICHI), 2013 IEEE International Conference on; 2013.

23. Apostol S, Soica O, Manasia L, Stefan C. Virtual Pedagogical Agents in the Context of Virtual Learning Environments: Framework and Theoretical Models. Elearn Softw Educ 2013:531-536.
24. Schulman D, Bickmore T. Modeling behavioral manifestations of coordination and rapport over multiple conversations. Proceedings of the International Conference on Intelligent Virtual Agents; 2010.
25. Kelders SM, Kok RN, Van Gemert-Pijnen JE. Technology and adherence in web-based interventions for weight control: a systematic review. Proceedings of the Proceedings of the 6th International Conference on Persuasive Technology: Persuasive Technology and Design: Enhancing Sustainability and Health; 2011.
26. Knowles SE, Toms G, Sanders C, Bee P, Lovell K, Rennick-Egglestone S, Coyle D, Kennedy CM, Littlewood E, Kessler D, Gilbody S, Bower P. Qualitative Meta-Synthesis of User Experience of Computerised Therapy for Depression and Anxiety. PloS one 2014 Jan 17;9(1). ARTN e84323 10.1371/journal.pone.0084323
27. Lehto T, Oinas-Kukkonen H. Persuasive Features in Web-Based Alcohol and Smoking Interventions: A Systematic Review of the Literature. J Med Internet Res 2011 Jul-Sep;13(3). ARTN e46 10.2196/jmir.1559
28. Mohr DC, Duffecy J, Ho J, Kwasny M, Cai X, Burns MN, Begale M. A Randomized Controlled Trial Evaluating a Manualized TeleCoaching Protocol for Improving Adherence to a Web-Based Intervention for the Treatment of Depression. PloS one 2013 Aug 21;8(8). ARTN e70086 10.1371/journal.pone.0070086
29. McClay CA, Waters L, McHale C, Schmidt U, Williams C. Online Cognitive Behavioral Therapy for Bulimic Type Disorders, Delivered in the Community by a Nonclinician: Qualitative Study. J Med Internet Res 2013 Mar;15(3):211-221. 10.2196/jmir.2083
30. Todd NJ, Jones SH, Lobban FA. What Do Service Users with Bipolar Disorder Want from a Web-Based Self-Management Intervention? A Qualitative Focus Group Study. Clin Psychol Psychot 2013 Nov;20(6):531-543. 10.1002/cpp.1804
31. Bradbury K, Dennison L, Little P, Yardley L. Using mixed methods to develop and evaluate an online weight management intervention. Brit J Health Psych 2015 Feb;20(1):45-55. 10.1111/bjhp.12125
32. Donkin L, Glozier N. Motivators and Motivations to Persist With Online Psychological Interventions: A Qualitative Study of Treatment Completers. J Med Internet Res 2012 May-Jun;14(3):262-273. ARTN e91 10.2196/jmir.2100

33. Nicholas J, Proudfoot J, Parker G, Gillis I, Burckhardt R, Manicavasagar V, Smith M. The Ins and Outs of an Online Bipolar Education Program: A Study of Program Attrition. J Med Internet Res 2010;12(5). ARTN e57 10.2196/jmir.1450
34. Cowpertwait L, Clarke D. Effectiveness of Web-based Psychological Interventions for Depression: A Meta-analysis. Int J Ment Health Ad 2013 Apr;11(2):247-268. 10.1007/s11469-012-9416-z
35. Jalil S, Myers T, Atkinson I. A Meta-Synthesis of Behavioral Outcomes from Telemedicine Clinical Trials for Type 2 Diabetes and the Clinical User-Experience Evaluation (CUE). J Med Syst 2015 Mar;39(3). ARTN 28 10.1007/s10916-015-0191-9
36. Kuijpers W, Groen WG, Aaronson NK, van Harten WH. A Systematic Review of Web-Based Interventions for Patient Empowerment and Physical Activity in Chronic Diseases: Relevance for Cancer Survivors. J Med Internet Res 2013 Feb;15(2). ARTN e37 10.2196/jmir.2281
37. Mohr DC, Cuijpers P, Lehman K. Supportive Accountability: A Model for Providing Human Support to Enhance Adherence to eHealth Interventions. J Med Internet Res 2011 Jan-Mar;13(1). ARTN e30 10.2196/jmir.1602
38. Cotter AP, Durant N, Agne AA, Cherrington AL. Internet interventions to support lifestyle modification for diabetes management: A systematic review of the evidence. J Diabetes Complicat 2014 Mar-Apr;28(2):243-251. 10.1016/j.jdiacomp.2013.07.003
39. Foster C, Calman L, Grimmett C, Breckons M, Cotterell P, Yardley L, Joseph J, Hughes S, Jones R, Leonidou C. Managing fatigue after cancer treatment: development of RESTORE, a web-based resource to support self-management. Psycho-Oncology 2015;24(8):940-949.
40. Helgadottir FD, Menzies RG, Onslow M, Packman A, O'Brian S. Online CBT I: Bridging the Gap Between Eliza and Modern Online CBT Treatment Packages. Behav Change 2009;26(4):245-253.
41. Kraft P, Drozd F, Olsen E. Digital therapy: Addressing willpower as part of the cognitive-affective processing system in the service of habit change. Lect Notes Comput Sc 2008;5033:177-+.
42. Gorlick A, Bantum EOC, Owen JE. Internet-based interventions for cancer-related distress: exploring the experiences of those whose needs are not met. Psycho-Oncology 2014;23(4):452-458.
43. Hamel LM, Robbins LB. Computer- and web-based interventions to promote healthy eating among children and adolescents: a systematic review. J Adv Nurs 2013 Jan;69(1):16-30. 10.1111/j.1365-2648.2012.06086.x

44. Ramadas A, Quek KF, Chan CKY, Oldenburg B. Web-based interventions for the management of type 2 diabetes mellitus: A systematic review of recent evidence. Int J Med Inform 2011 Jun;80(6):389-405. 10.1016/j.ijmedinf.2011.02.002
45. Sinclair C. Effectiveness and User Acceptance of Online Chronic Disease Management Interventions in Rural and Remote Settings: Systematic Review and Narrative Synthesis. Clinical Medicine Insights Therapeutics 2015;7:43.
46. Mori M, MacDorman KF, Kageki N. The uncanny valley [from the field]. IEEE Robotics & Automation Magazine 2012;19(2):98-100.
47. Nass C, Moon Y. Machines and mindlessness: Social responses to computers. J Soc Issues 2000 Spr;56(1):81-103. Doi 10.1111/0022-4537.00153
48. Nass CI, Brave S. Wired for speech: How voice activates and advances the human-computer relationship: MIT press Cambridge; 2005. 0262140926
49. Norman DA. Emotional design: Why we love (or hate) everyday things: Basic books; 2005. 0465051367
50. Reeves B, Nass C. How people treat computers, television, and new media like real people and places: CSLI Publications and Cambridge university press Cambridge, UK; 1996.
51. Conrad FG, Schober MF, Jans M, Orlowski RA, Nielsen D, Levenstein R. Comprehension and engagement in survey interviews with virtual agents. Front Psychol 2015 Oct 20;6. Artn 1578 10.3389/Fpsyg.2015.01578
52. Veletsianos G, Russell GS. What Do Learners and Pedagogical Agents Discuss When Given Opportunities for Open-Ended Dialogue? J Educ Comput Res 2013;48(3):381-401. 10.2190/Ec.48.3.E
53. Weizenbaum J. ELIZA—a computer program for the study of natural language communication between man and machine. Communications of the ACM 1966;9(1):36-45.
54. Baylor AL, Ryu J, Shen E. The effects of pedagogical agent voice and animation on learning, motivation and perceived persona. Proceedings of the Proceedings of World conference on educational multimedia, hypermedia and telecommunications; 2003.
55. Cowell AJ, Stanney KM. Embodiment and interaction guidelines for designing credible, trustworthy embodied conversational agents. Lect Notes Artif Int 2003;2792:301-309.
56. Kim CM, Baylor AL. A virtual change agent: Motivating pre-service teachers to integrate technology in their future classrooms. Educ Technol Soc 2008;11(2):309-321.

57. Krämer NC, von der Pütten A, Eimler S. Human-agent and human-robot interaction theory: similarities to and differences from human-human interaction. Human-Computer Interaction: The Agency Perspective: Springer; 2012:215-240.
58. Lester JC, Converse SA, Kahler SE, Barlow ST, Stone BA, Bhogal RS. The persona effect: affective impact of animated pedagogical agents. Proceedings of the Proceedings of the ACM SIGCHI Conference on Human factors in computing systems; 1997.
59. Lisetti C, Amini R, Yasavur U. Now all together: overview of virtual health assistants emulating face-to-face health interview experience. KI-Künstliche Intelligenz 2015;29(2):161-172.
60. Mayer RE. Cognitive theory of multimedia learning. The Cambridge handbook of multimedia learning 2014;43.
61. Rajan S, Craig SD, Gholson B, Person NK, Graesser AC, Group TR. AutoTutor: Incorporating back-channel feedback and other human-like conversational behaviors into an intelligent tutoring system. International Journal of Speech Technology 2001;4(2):117-126.
62. Rosenberg-Kima RB, Baylor AL, Plant EA, Doerr CE. The importance of interface agent visual presence: Voice alone is less effective in impacting young women's attitudes toward engineering. Lect Notes Comput Sc 2007;4744:214-222.
63. Sproull L, Subramani M, Kiesler S, Walker JH, Waters K. When the interface is a face. Hum-Comput Interact 1996;11(2):97-124. DOI 10.1207/s15327051hci1102_1
64. Zanbaka C, Ulinski A, Goolkasian P, Hodges LF. Effects of virtual human presence on task performance. Proceedings of the Proc. International Conference on Artificial Reality and Telexistence 2004; 2004.
65. Bickmore TW, Picard RW. Establishing and maintaining long-term human-computer relationships. ACM Transactions on Computer-Human Interaction (TOCHI) 2005;12(2):293-327.
66. Creed C, Beale R, Cowan B. The Impact of an Embodied Agent's Emotional Expressions Over Multiple Interactions. Interact Comput 2015 Mar;27(2):172-188. 10.1093/iwc/iwt064
67. Gratch J, Wang N, Gerten J, Fast E, Duffy R. Creating rapport with virtual agents. Intelligent Virtual Agents, Proceedings 2007;4722:125-+.
68. Gulz A. Benefits of virtual characters in computer based learning environments: Claims and evidence. International Journal of Artificial Intelligence in Education 2004;14(3, 4):313-334.

69. Kang S-H, Gratch J, Wang N, Watt JH. Does the contingency of agents' nonverbal feedback affect users' social anxiety? Proceedings of the Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems-Volume 1; 2008.
70. Kramer NC, Eimler S, von der Putten A, Payr S. Theory of Companions: What Can Theoretical Models Contribute to Applications and Understanding of Human-Robot Interaction? Appl Artif Intell 2011;25(6):474-502. 10.1080/08839514.2011.587153
71. von der Putten AM, Kramer NC, Gratch J. How Our Personality Shapes Our Interactions with Virtual Characters - Implications for Research and Development. Intelligent Virtual Agents, Iva 2010 2010;6356:208-221.
72. Horvath AO, Greenberg LS. Development and Validation of the Working Alliance Inventory. J Couns Psychol 1989 Apr;36(2):223-233. Doi 10.1037/0022-0167.36.2.223
73. Baylor AL, Rosenberg-Kima RB, Plant EA. Interface agents as social models: the impact of appearance on females' attitude toward engineering. Proceedings of the CHI'06 Extended Abstracts on Human Factors in Computing Systems; 2006.
74. Berry DC, Butler LT, de Rosis F. Evaluating a realistic agent in an advice-giving task. Int J Hum-Comput St 2005 Sep;63(3):304-327. 10.1016/j.ijhcs.2005.03.006
75. D’Mello S, Picard R, Graesser A. Towards an affect-sensitive autotutor. IEEE Intelligent Systems 2007;22(4):53-61.
76. D'Mello S, Olney A, Williams C, Hays P. Gaze tutor: A gaze-reactive intelligent tutoring system. Int J Hum-Comput St 2012 May;70(5):377-398. 10.1016/j.ijhcs.2012.01.004
77. Doirado E, Martinho C. I mean it!: detecting user intentions to create believable behaviour for virtual agents in games. Proceedings of the Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: volume 1-Volume 1; 2010.
78. Lisetti C, Amini R, Yasavur U, Rishe N. I can help you change! an empathic virtual agent delivers behavior change health interventions. ACM Transactions on Management Information Systems (TMIS) 2013;4(4):19.
79. Picard RW, Picard R. Affective computing. Vol 252: MIT press Cambridge; 1997.
80. Schroeder NL, Adesope OO, Gilbert RB. How effective are pedagogical agents for learning? A meta-analytic review. J Educ Comput Res 2013;49(1):1-39.

81. Brave S, Nass C, Hutchinson K. Computers that are care: investigating the effects of orientation of emotion exhibited by an embodied computer agent. Int J Hum-Comput St 2005 Feb;62(2):161-178. DOI 10.1016/j.ijhcs.2004.11.002
82. Klein J, Moon Y, Picard RW. This computer responds to user frustration: Theory, design, and results. Interact Comput 2002 Feb;14(2):119-140.
83. Moreno R. Software agents in multimedia: An experimental study of their contributions to students' learning. Human-computer interaction proceedings 2001:275-277.
84. Beale R, Creed C. Affective interaction: How emotional agents affect users. Int J Hum-Comput St 2009 Sep;67(9):755-776. 10.1016/j.ijhcs.2009.05.001
85. Dehn DM, van Mulken S. The impact of animated interface agents: A review of empirical research. Int J Hum-Comput St 2000 Jan;52(1):1-22. DOI 10.1006/ijhc.1999.0325
86. Hartholt A, Traum D, Marsella SC, Shapiro A, Stratou G, Leuski A, Morency L-P, Gratch J. All together now. Proceedings of the International Workshop on Intelligent Virtual Agents; 2013.
87. Veletsianos G, Miller C, Doering A. Enali: A Research and Design Framework for Virtual Characters and Pedagogical Agents. J Educ Comput Res 2009;41(2):171-194.
88. Veletsianos G, Russell GS. Pedagogical agents. Handbook of research on educational communications and technology: Springer; 2014:759-769.

Chapter 3

An empirical study to a pedagogical agent as an adjunct to an eHealth self-management intervention: what modalities does it need to successfully support and motivate users?

Scholten MR, Kelders SM, Van Gemert-Pijnen JEWC

An empirical study to a pedagogical agent as
an adjunct to an eHealth self-management intervention:
what modalities does it need to
successfully support and motivate users?

Frontiers in psychology, 2019, 10: 1063

Abstract

Background: Prior research has shown that the more patients know about their disease, health, and lifestyle the better the health outcomes are. Patients who are suffering from either physical diseases with mental consequences or from mental illnesses can independently contribute to their own feeling of mental well-being by following evidence-based online, self-guided therapeutic interventions. These self-guided therapeutic interventions during which there is no contact with a care provider have shown high effectiveness. However, users (patients) of self-guided eHealth interventions have difficulties fulfilling the entire trajectory as is mirrored in high non-adherence rates. Users have reported a need for support, that is traditionally provided by human care providers. This study investigates the opportunities from within the technology to increase its support level toward the user. It is known that Embodied Conversational Agents (ECA's) can provide such support towards users of eHealth interventions

Objective: The objective of this paper is to experimentally explore the potential of an Embodied Conversational Agent to support an user of an eHealth intervention.

Methods: We deployed a pedagogical agent acting as an adjunct to a self-guided positive psychology psycho-education intervention. This agent provided instructions and user support in between and explicitly not during the online learning modules as to avoid the risk of distraction. By deploying three versions of a pedagogical agent, varying the features of animation, speech, and visibility we investigated whether users felt more supported than by a fourth text-only control condition. All four conditions provided similar task-related support and emotion-related support to the user.

Results: The results of showed that our pedagogical agent made users feel guided and supported with respect to fulfilling their tasks. However, our pedagogical agents was not able to demonstrate effects of emotion-related support resulting in higher user motivation and an improved learning experience. Significant effects of visibility and voice were found, but animation of our pedagogical agent had no effect. On the feedback outcome variable we found a gender effect. Male participants graded the *visible* ECA's *higher* than female participants and graded the *non-visible* ECA *lower* than female participants.

Conclusion: Our experiment showed positive ECA effects when providing task-related support to users of a psycho-education environment. The ECA as a GUI seemed to make the task easier than text. However, our ECA was not capable of demonstrating effects as a result of its emotion-related support. This may be due to the friendly set-up of our experiment, that failed to bring users to a distressed, need-for-support mental state.

Keywords: eHealth; web-based intervention; embodied conversational agent; virtual agent; virtual humans; adherence; attrition

Introduction

eHealth is about the use of information and communication technology to reinforce health and health care. It refers to forms of prevention and education, diagnostics, therapy and care delivered through digital technology, independently of time and place. An important branch of eHealth consists of technological self-care solutions such as home telemonitoring applications that provide patients with direct insights through self-monitored data. Other self-care solutions focus on teaching indirect insights, leading to competence (disease knowledge) or disease management (making choices, acting responsibly) (Peeters et al, 2013). Research showed that the more patients know about their disease, health, and lifestyle the better the health outcomes are (Kennedy et al. 2017). Technological self-care (e.g. for chronic diseases) often goes hand in hand with self-management as a practice: the ability to actively participate in the management of health with the emphasis on physical and mental well-being. This involves medical management; changing, maintaining, and creating meaningful behaviors and dealing with the emotions of suffering from chronic disease(s) (Lorig and Holman, 2003). The question is whether self-management can be independently done by patients, that is without the help and support of a care provider. More precisely, the question is whether patients who are suffering from either physical diseases with mental consequences or from mental illnesses can independently contribute to their own feeling of mental well-being. Meta-analytic studies (Barak et al., 2008, Spek et al. 2007) have demonstrated the effectiveness of self-guided therapeutic interventions during which there is no contact with a care provider. Despite the effectiveness, patients show mixed opinions on these self-guided interventions. On the one hand patients report positive experiences with self-guided interventions (Walsh et al. 2018). However, disadvantages have also been reported by patients, such as the lack of human contact (Flynn et al., 2009).

Especially in case of self-guided e-mental Health interventions against depression, adherence can be low (Schubart et al., 2011). Low adherence is sub-optimal as greater exposure to website content is associated with increased benefit (REF Christensen Helen et al, 2004). Obvious follow-up questions are therefore *why* users do not adhere and even more *how* adherence can be stimulated. There seem to be no final answers to these questions but cues are certainly available.

A meta-analysis of the effectiveness of diabetes interventions suggests that participants' difficulties in understanding the use of Web-based interventions led to higher non-adherence rates (Lie et al., 2017)

In addition, some studies relate disease-specific effects such as severity to

adherence, with a high level of emotional distress leading
to early dropout (Davis and Addis, 1999).

In terms of solutions, the provision of support to enable patients to be confident and capable in managing health conditions is generally considered an important factor (de Silva, 2011, Wilkinson and Whitehead, 2009). In addition, there is empirical evidence that the *lack* of such a supportive relationship is associated with low levels of motivation to engage in self-care and may as such lead to *non-adherence*. (Bickmore, 2010, Drench et al. 2007).

In conclusion, user support is a relevant topic for user adherence. The next question is what kind of support users need. In order to answer this question, in an earlier study, we have analyzed (Scholten et al, 2017) studies on support needs as expressed by eHealth users. We found that users have a need to be encouraged (*emotion-related support*) but also value practical support (*task-related support*). Emotion-related support acknowledges both the user's endeavors during the change program and the originating issue the user is dealing with. It can be delivered in terms of praising the user, and by other types of encouraging behavior. In contrast, task-related support consists of actions such as setting and reviewing log-in goals of eHealth interventions, positively reinforcing log-in and intervention use and providing answers to users on questions regarding the functionality of the eHealth solution.

We suggested that fairly simple non-responsive Embodied Conversational Agents (ECA's) can provide a means for task-related support in order to make self-guided interventions a better experience. Embodied Conversational Agents are computer animations of faces or bodies, 'robots on screen'. They can enrich the mostly text and video based self-guided eHealth interventions with an interface that has stronger similarities with a human face. Furthermore, they personify the interface and can contribute to a feeling of trust in the system. (Andre and Pelachaud, 2010). ECA's are applied within various contexts; from computer games (Bostan at al, 2009), intelligent tutoring systems (D'Mello et al, 2007), museum guides (Kopp et al, 2005) conducting medical interviews (Kobori et al, 2018), and providing therapy for depression and anxiety (Fitzpatrick et al, 2017)

Within all this ECA variety, we focus in this paper on ECA's that take on the role of learning coach or tutor within e-learning environments as a) e-learning (psycho-education) is one of the cornerstones of self-guided e-health interventions and b) considerable progress in the application of ECA's within the scientific domain of e-learning has been made, which has created a solid basis for further research.

Current evidence for ECA's as tutors within e-learning

Indeed, promising ECA effects have been found on e-learning. Within the meta-study of Schroeder et al. (2013) on 43 studies including 3,088 participants, a small but significant effect was reported on learning. The participants learned more from a system with a pedagogical agent, than a system without one. Next to learning, positive effects of ECA's have been found on user motivation. The meta-study of Veletsianos and Russell (2014) reports on studies in which learner motivation and learning outcomes are promoted by pedagogical agents. However, the evidence is not equivocal as their meta-study also refers to studies in which the pedagogical agents did not demonstrate added value compared to text only conditions. They summarized these mixed results as a conundrum which is open for future research to resolve. A research topic that often goes together with the effectiveness of pedagogical agents is that of the modalities (eg speech, animation) of the agents used. The relevance of the modalities for learning is expressed by the social cue hypothesis (Domagck, 2010) that states that the presence of social cues cause learners to engage in sense-making processes and processing the learning material deeply. Social cues as represented by the ECA's modalities of e.g. visibility, speech and animation should -according to this hypothesis- have a positive impact on the learning process compared to a sheer textual environment.

Effects of ECA modalities of speech and visibility on e-learning

Equal to the effects of ECA's as a whole, the evidence for the effects of speech and visibility within ECA's is mixed. Atkinson (2002) found that an ECA using speech performed better than an e-learning environment that lacked an ECA. This positive effect was replicated by Lusk and Atkinson (2007) and also Graesser et al. (2004) came to the same conclusion. In contrast, Louwerse et al. (2005) report on studies in which pedagogical agents using speech had no additional effect compared to speech alone. Stated differently: those studies suggest that it is solely the speech that determines the learning effect and not the visual presence of the ECA. Schroeder (2013) found that speech-enabled ECA's provide a better solution on learning measures than 'ECA-less' learning environments. Schroeder therefore suggested that -contrary to the Louwerse et al statement- the ECA's visibility combined with their voice is more beneficial than voice alone. A potential way to reconcile these conflicting results on the ECA's visibility may be provided by the concept of *distraction*, which is also described as the *split-attention* principle (Louwerse et al., 2005). According to this principle, users are hindered to engage themselves in the learning process, when they are obliged to simultaneously interact with an ECA. Van Mulken, et al. (1998) found that ECA's are indeed distracting, but also motivating. Moreno and Mayer (2007) claim that the ECA's motivation should be greater than their distraction in order to facilitate learning. We are unaware of studied on design guidelines for optimizing user motivation while minimizing user distraction. However, there are solutions to the attention split. An ECA can provide motivating instructions in advance of an e-learning topic, somewhat comparable to a traditional teacher in a classroom. Then, the user is asked to start working on the topic, without further interference of the ECA. As a consequence, the user can both dedicate their full attention to first the ECA and then to their own learning process. Finally, a design argument in favor of a visible ECA, as a source for either speech or text, is provided by Cassell (2001). Cassell states that properly designed interfaces have affordances and visual clues that are in accordance with their role. Speech does not appear spontaneously; it therefore makes sense to present the ECA as its visible source.

Effect of the ECA modality of animation on e-learning

Technological advances have also made it easier to animate agents, instead of presenting them as a still image. However, limited knowledge is available on whether these animations have advantages. Baylor et al (2003) investigated the effects of pedagogical agent speech (human, machine-generated) and animation (present, absent) on learning and motivation, Animation gave somewhat contradictory results: participants learned significantly more but also reported that the agent was significantly less facilitative than when it was still. In addition, animation made the participants significantly less motivated about the topic. In contrast to these results, Schroeder et al reported that still ECA's produced a small but significant learning effect, whereas animated ECA's neither produced a positive nor a negative effect.

Expectations on ECA research within the domain of self-guided e-health interventions

In summary, the literature tells us that a visible speech-enabled or text-enabled ECA has e-learning benefits compared to no ECA at all. An important pre-condition is that the ECA will make a clear distinction between the moment it communicates to (or interacts with) the user and the moment they let the user learn. Whether an animated or still ECA is the better solution is open for further investigation. Within the context of this paper, we will concentrate on the e-learning domain within an e-mental health context, with patients as the targeted user group. Within this perimeter we will define what we can and should expect from an ECA. For this, please see the schematic picture of the research domain that we present below in figure 1.

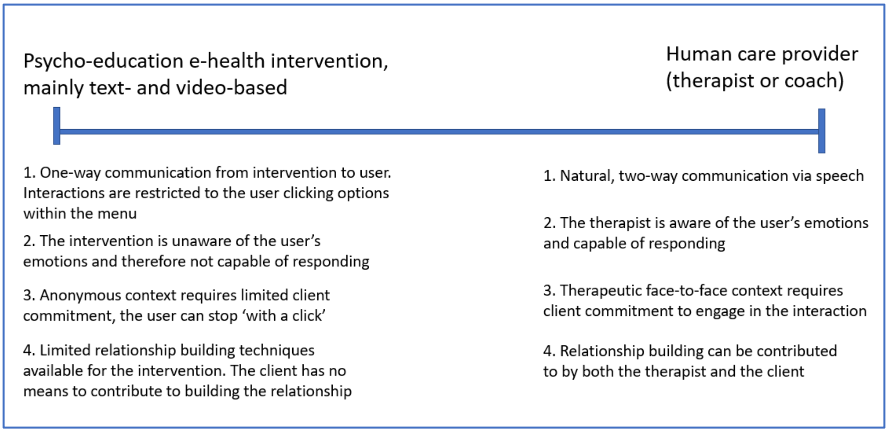


Figure 1. The left and right boundaries for ECA's as adjuncts in e-mental Health

On the left side we place a classic self-guided e-health intervention, such as MoodGym (https://moodgym.com.au/). Within this type of intervention, the user is typically asked to read information and do exercises in order to improve their mental being. The intervention does not 'see' or 'hear' how the user is doing, will not understand the user and will therefore not be capable of expressing personal interest. On the right side we position a (idealized) human care provider who can and will interact with the user. He/she can hear and see the user, will take their emotions into account and respond appropriately by eg expressing empathy. Within many experimental studies, ECA's are set up with the intention to simulate processes that hinge to the right side of the spectrum. These ECA's have the purpose of triggering social mechanisms that play a role within two-way human to human communication. Within this paper we opt for a different approach. Our aim is to find out whether we can make improvements on the left side: can we realize user experience improvements on a text- and video-based self-guided e-mental health intervention by adding an ECA that makes users more engaged and motivated? We choose this approach for the following reasons:

- Most evidence-based self-guided e-health interventions are text- and video-based and unaware of users' emotions. So, they are typical 'left-side' interventions. If we want to improve adherence to the present base of e-health solutions, and build upon the existing work done, we have to start left.
- By separating the therapeutic content from the user support aspects, existing evidence-based self-guided e-health interventions can remain unaltered. ECA's can be added as adjuncts for providing directions and user

support, without interfering with the functionality and evidence for the intervention.

- The ECA's we envision are widely available. If we would be able to find a positive ECA effect they can be fairly easily implemented within web-based environments.

The aim of our study is to investigate whether a straightforward, non-responsive ECA that delivers both task- and emotion-related support to users of a psycho-education intervention, will result in higher learning motivation amongst users as a remedy to enlarge adherence.

Material and Methods

Recruitment of participants

We started the recruitment process by adding the experiment as an option to the university of Twente e-health MOOC that is offered on the FutureLearn online course platform (https://www.futurelearn.com/). As the recruitment process of participants did not have the required pace, we decided to expand it. We recruited bachelor and master psychology students at the university of Twente]. In total 230 participants were included. As an inclusion criterion we set a high level of mastery of English. As an exclusion criterion we set participation in a pre-study with the ECA. The study protocol was reviewed and approved by the Twente Institutional Review Board.

Design

To investigate the differential outcome effects of ECA's with different modalities as mentioned within the ECA literature, speech, visual presence of the ECA, and the level of animation, we set up conditions with the following distinctive ECA features:

- The ECA is animated (1) vs the ECA is a still image (2) vs the ECA is not visible (3)
- The ECA expresses itself via speech (1) vs text (2)

Out of the six combinations, we left out animated, text (non-speech) as a key element of the ECA's animation consisted of the lip sync which we would lose without speech. In addition, we left out the option not visible, speech as a voice without a visible source would create an unusual set-up.

This way, we created the following 4 conditions:

1. AS = animated, speech (non-text),
2. SS = still, speech (non-text)
3. ST = still, text (non-speech)
4. TO (control condition) = text only

The study design was a between-subjects experiment with the before-mentioned four conditions to which participants were randomly assigned using randomization software; AS (58 participants; 44 female, 14 male) SS (58 participants; 46 female, 12 male) ST (55 participants; 49 female, 6 male) TO (59 participants; 43 female, 16 male).

Intervention

An e-learning intervention for making people knowledgeable about positive psychology was set up. Positive psychology focuses on the abilities of people and their potential to flourish. Positive psychology was chosen, being a relevant topic within the e-health domain; a number of treatments against depression are based on positive psychology principles (Hayes, 1999). In addition, positive psychology and happiness are subjects that are of general human interest. As we assumed, this would make it easier for participants to engage with our experimental set-up. The self-guided intervention was developed by analyzing the positive psychology topic (Gable, 2005) and creating a combination of theory and exercises, including the remunerated 'three good things exercise' (https://ggia.berkeley.edu/practice/three-good-things) and 'best possible self-exercise' (Renner et al, 2014)
A WordPress website (version 4.9.7) (https://wordpress.org) with 4 webpages was created, each representing a condition. The e-learning intervention on positive psychology was embedded as an online Microsoft PowerPoint presentation® and placed on the left side on each of the 4 webpages. On the right side of the 4 webpages the user support content was added, as to represent the 4 conditions. The user support consisted of task-related support (e.g. "within this experiment you will read about positive psychology and you will do some exercises") and emotional support (e.g. "well done!"). In addition, the user was stimulated to take advantage of the exercises outside of the experiment. An explicit distinction was made between the instruction as delivered by either the ECA or the text-only control condition on the one hand and the user learning activities on the other hand. This was done to avoid the split-attention effect (Louwerse et al., 2005). During instruction on the right side of the webpage, the user was told what learning modules would come next. Then the user was asked to click on the left side of the webpage and do the e-learning. When the e-learning module had come to an end, the user was asked to go

to the right side of the webpage for new instructions. For the animated (AS) condition an ECA was created through the Voki application (https://www.voki.com), see figure 2 below. For the other 3 conditions, a second Microsoft PowerPoint presentation was embedded on the right side of the page and added a still of the ECA (SS and ST conditions) and speech fragments (ST condition) or textual information (ST and TO conditions).

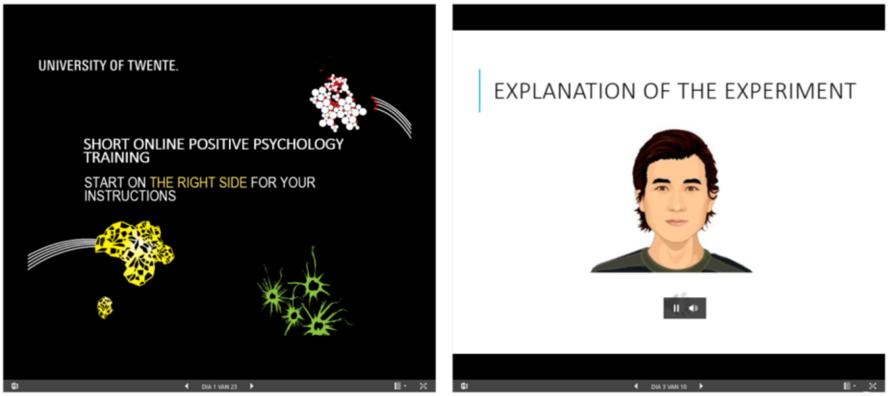


Figure 2. The e-learning intervention. On the left side of the webpage the educational content is displayed, on the right side the support condition with directions (task-related support) and encouragement (emotion-related support) in presented. The example shown is the AS condition; animated, speech.

Procedure

The webpages were put online and the study was run without human supervision to simulate the self-guided e-health intervention context. Users were provided with a URL that led to the Qualtrics system (https://www.qualtrics.com). A randomization software module redirected the users to one of four webpages. On right side of the webpage, the users received instructions through the ECA or instructional PowerPoint. They were asked to do the reading of the Positive Psychology PowerPoint on the left side and then to come back to the instructional side of the page for following instructions. This way, the users received instructions, performed an experimental task, received positive feedback and new instructions. After the introduction, this cycle was repeated twice. Then the users were redirected from the WordPress website to the Qualtrics environment to fill in the questionnaires.

Outcome measures

For the outcome measures, a variety of scales was selected. First, the EGameFlow scale (Fu et al, 2009) was selected, which measures learners' enjoyment of e-learning games. The developers of this scale refer to the application of flow theory within education (Whalen and Csikszentmihalyi, 1991) and argue that the flow experience is a pre-condition for successful e-learning. Autonomy and feedback have been implemented in the EGameFlow scale as a means to measure task-related support as provided by the Graphical User Interface (GUI) of the e-learning environment. This suited very well with the purpose of this experiment, in which there was an aim to measure the differential effects of ECA's as task-related and emotion-related support provider. From both the feedback and autonomy scale three items on the basis of validation were selected and on the basis of the distinctive formulation of the questions. Both scales use a seven-point Likert scale ranging from 'strongly disagree' to 'strongly agree'. Within the wording of the subscales the word "game" was replaced by "online training". Next, the Instructional Materials Motivation Survey (IMMS) was selected. This scale measures students' motivational reactions to self-directed instructional materials and is derived from the ARCS model (Keller, 1987). This model has been applied to ECA's in e-learning settings, e.g. (Shen, 2009). ARCS' A refers to gaining and keeping the learner's attention and stimulate their desire to learn. ARCS' R is about making the instruction relevant to the learners personal experience, needs and goals. The attention (12 items) and relevance (9 items) scales both use a five-point Likert scale ranging from 'not true' to 'very true'. Subsequently, Involvement was selected. The Personal Involvement Inventory (Zaichkowsky, 1994) is a context-free measure applicable to involvement with products, with advertisements, and with purchase situations. It has been applied before for measuring the effectiveness of environments with ECA's (Lo and Cheng, 2010). It was selected for this experiment to measure user motivation in general. The scale consists of 10 items and uses a seven-point Likert scale with varying category names such as 'appealing' versus 'not appealing' and 'means nothing' versus 'means a lot to me'. Last, the Rapport scale was selected. Rapport is an umbrella term for generic positive interactions between human counterparts, which as a term is also associated to terms as harmony, fluidity, synchrony and flow. Many studies have demonstrated that, when established, rapport facilitates a wide range of social interactions between humans including psychotherapy (Tsui and Schultz, 1985) teaching (Fuchs, 1987) and caregiving (Burns, 1984). Rapport has been used as outcome measure in studies with users interacting with an ECA (Gratch et al, 2007). Advanced ECA's that respond to the verbal and non-verbal behavior of the user in a contingent manner have indeed successfully created rapport. For this experiments' non-responsive ECA, we didn't expect effects of Rapport, but the outcome measure was added for exploratory and verification purposes. The Rapport scale (Cerekovic et al, 2014) consists of fifteen items and uses a five-point Likert scale ranging from ((1) – Disagree strongly to (5) - Agree strongly).

Analysis

Visibility, Speech and Animation as ECA modality features

As a first step, conditions on common features were categorized. The AS, SS and ST conditions were put together in the *Visible* ECA category (171 participants) and compared to the *Non-visible* category (59 participants) that solely consisted of the TO condition. Furthermore, the AS and SS conditions were put together in the *Speech* ECA category (116 participants) and were compared to the *Text* category (114 participants) that consisted of both the ST and TO condition. The Rapport outcome variable was only measured for the ST condition. Last, the AS represented the *Animated* ECA category (58 participants) and was compared to the *Non-Animated* ECA category (113 participants) that consisted of the SS, ST. Obviously, the TO condition was not part of this analysis as it did not contain a visible ECA. We used a two-way analysis of variance (two-way ANOVA) to calculate differential effects between the modality features and to calculate interaction effects on the ECA's feature and the gender of the participant. Although, prior to the analysis, we did not expect that gender would have an effect, a pre-analysis on gender showed differently.

Four conditions

Last, a two-way analysis of variance (ANOVA) on the four non-categorized conditions and their interaction with gender was conducted. This was done in order to look for effects of combinations of modalities, were combinations could be stronger (or less strong) than the individual modality effects. Additionally, t-tests were performed to look out for significant differences between individual conditions in combination with gender type.

Results

Visible vs non-visible ECA

The means and SD values of all outcome variables are shown in Table 1.

Table 1. Mean scores and standard deviation on the visibility- non-visibility distinction

	Visible ECA			Non-visible ECA		
	Female (n=139)	Male (n=32)	All (n=171)	Female (n=43)	Male (n=16)	All (n=59)
Feedback Mean (SD)	4.48 (1.20)**	4.90 (1.22)**	4.56 (1.21)*	4.50 (1.17)**	3.98 (0.98)**	4.36 (1.14)*
Autonomy Mean (SD)	5.39 (1.01)	5.66 (0.83)	5.44 (0.98)**	5.29 (1.00)	4.97 (0.96)	5.20 (0.99)**
Attention Mean (SD)	3.69 (0.62)	3.65 (0.46)	3.68 (0.59)	3.70 (0.62)	3.46 (0.57)	3.64 (0.61)
Relevance Mean (SD)	3.60 (0.67)	3.55 (0.59)	3.59 (0.66)	3.85 (0.57)	3.55 (0.50)	3.77 (0.57)
Involvement Mean (SD)	5.30 (1.12)	5.30 (1.07)	5.30 (1.10)	5.35 (0.97)	5.10 (0.63)	5.28 (0.90)
Rapport Mean (SD)	4.85 (0.69)	4.73 (0.74)	4.83 (0.70)	n.a.	n.a.	n.a.

*significant effect of p=.03, ** significant effect of p=.02

Comparing the visible and non-visible ECA, significant main effects were found on the outcome variables feedback* (F=4.64; p=.03), and autonomy** (F=5.17; p=.02); in both cases the visible ECA category resulted in significantly higher scores than the non-visible ECA.

No significant main effects were found for the other outcome variables: attention (F=0.65, p=0.42), relevance (F=1.14, p=0.29), involvement (F=0.15, p=0.70). Subsequently, the interaction between the visibility distinction and gender type was analyzed. A significant interaction effect between visibility*gender was found for the outcome variable feedback** (F=5.26, p=.02). The interaction effect is visually presented in figure 3 below; male participants graded the *visible* ECA *higher* than female participants but graded the *non-visible* ECA *lower* than female participants.

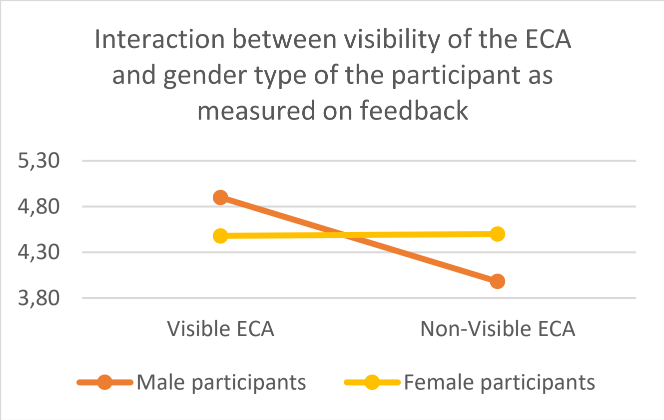


Figure 3. Feedback: interaction effect of visibility*gender type

Contrary to the feedback outcome variable, for the autonomy outcome variable a significant interaction effect between visibility and gender was *not* found (F=2.92, p=0.09) within the two-way ANOVA. In addition the two-way ANOVA showed that no significant interaction effects with gender were found for the other outcome variables; attention*gender: (F=0.86, p=0.36), relevance*gender (F=1.21, p=0.27), involvement*gender (F=0.45, p=0.50).

Text versus speech

For the means and SD values of the distinction of an ECA that communicates via speech or text, see table 2 below.

Table 2. Mean scores and standard deviation on the speech-text distinction

	Speech			Text		
	Female (n=90)	Male (n=26)	All (n=116)	Female (n=92)	Male (n=22)	All (n=114)
Feedback Mean (SD)	4.53 (1.22)	4.96 (1.28)	4,63 (1.24)*	4.44 (1.17)	4.17 (1.00)	4,38 (1.14)*
Autonomy Mean (SD)	5.37 (0.97)	5.65 (0.87)	5.43 (0.96)	5.36 (1.04)	5.16 (0.93)	5.33 (1.02)
Attention Mean (SD)	3.70 (0.64)	3.67 (0.45)	3.69 (0.60)	3.68 (0.60)	3.48 (0.55)	3.64 (0.60)
Relevance Mean (SD)	3.64 (0.63)	3.57 (0.42)	3.62 (0.59)	3.69 (0.68)	3.53 (0.70)	3.66 (0.68)
Involvement Mean (SD)	5.21 (1.10)	5.37 (1.02)	5.25 (1.08)	5.40 (1.06)	5.07 (0.83)	5.34 (1.03)
Rapport Mean (SD)	4.90 (0.07)	4.67 (0.14)	4,85 (0.70)	4.76 (0.10)	5.02 (0.29)	4,79 (0.72)

*significant effect of p=.02

A significant effect on feedback* (F=5,32, p=.02) was found; speech led to significantly higher scores than text. For the other variables no significant effects were found; autonomy (F=2.40, p=.12), attention (F=1.19, p=0.28), relevance (F=0.003, p=0.96), involvement (F=0.10, p=0.75), rapport (F=0.39, p=0.54). Subsequently, the interaction between the speech-text category and gender type was analyzed. No significant interaction effects were found. Feedback*gender (F=3.29, p=0.07), autonomy*gender (F=2.36, p=0.13), attention*gender: (F=0.79, p=0.38), relevance*gender (F=0.18, p=0.68), involvement*gender: (F=2.07, p=0.15), rapport*gender (F=2.14, p=0.15).

Animation vs no animation as categories

Subsequently, the effect of the modality of animation was analyzed. For the means and SD values of the distinction of an ECA that is animated or still, see table 5 below.

Table 3. Mean scores and standard deviation on the animated-still distinction

	Animated ECA			Still ECA		
	Female (n=44)	Male (n=14)	All (n=58)	Female (n=95)	Male (n=18)	All (n=113)
Feedback Mean (SD)	4.34 (1.23)	5.12 (1.25)	4.53 (1.27)	4.54 (1.19)	4.74 (1.20)	4.57 (1.19)
Autonomy Mean (SD)	5.28 (0.98)	5.82 (0.87)	5.41 (0.97)	5.44 (1.02)	5.53 (0.80)	5.45 (0.99)
Attention Mean (SD)	3.66 (0.69)	3.60 (0.49)	3.65 (0.64)	3.70 (0.59)	3.68 (0.45)	3.70 (0.57)
Relevance Mean (SD)	3.58 (0.61)	3.56 (0.34)	3.58 (0.56)	3.62 (0.70)	3.54 (0.74)	3.60 (0.70)
Involvement Mean (SD)	5.09 (1.16)	5.23 (1.10)	5.12 (1.14)	5.39 (1.08)	5.35 (1.08)	5.39 (1.08)
Rapport Mean (SD)	4.83 (0.11)	4.73 (0.19)	4.81 (0.72)	4.86 (0.07)	4.73 (0.17)	4.84 (0.70)

No significant effects of animation on any of the outcome variables was found: feedback (F=0.14; p=0.71), autonomy (F=0.13; p=0.72), attention (F=0.24, p=0.62), relevance (F=0.004, p=0.95), involvement (F=0.92, p=0.34), rapport (F=0.012, p=0.91).

No significant interaction effects between level of animation and gender type were found either. feedback*gender (F=1.42, p=0.24), autonomy*gender (F=1.30, p=0.26), attention*gender (F=0.03, p=0.86), relevance*gender (F=0.05, p=0.83), involvement*gender (F=0.17, p=0.68), rapport*gender (F=0.12, p=0.91). However, a gender effect on the variable feedback (F=4.15, p=0.04) was found, see figure 4 below. Male participants grade the ECA significantly higher that female participants.

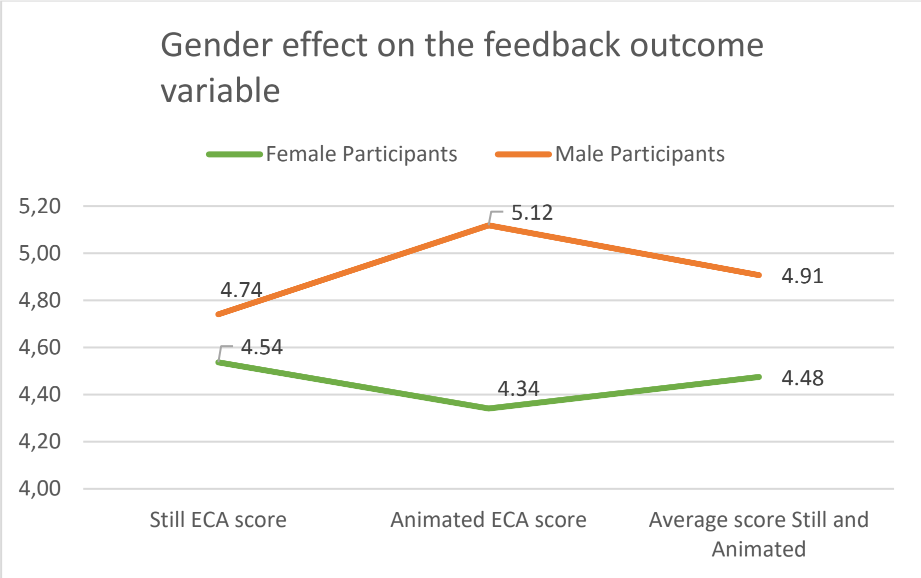


Figure 4. Gender effect on feedback

The reason this gender effect for feedback was solely found in the animation-still analysis is due to the text-only scores that were out of scope. This is in contrast to the speech-text and visibility-non-visibility analyses were text-only scores were in scope. Gender effects were not found for the other outcome variables; autonomy (F=2.573,0.111), attention (F=0.107, p=0.744), relevance (F=0.108, p=0.743), involvement (F=0.049, p=0.824), rapport (F=0.654, p=0.420).

Effects of Individual conditions

Last, the individual conditions were analyzed in order to look for differences between combinations of modalities.

Table 4. Mean scores and standard deviation of the four conditions

	AS	SS	ST	TO
Feedback Mean (SD)	4.72 (1.21)	4.53 (1.27)	4.41 (1.15)	4.36 (1.14)
Autonomy Mean (SD)	5.41 (0.97)	5.45 (0.95)	5.46 (1.04)	5.20 (0.99)
Attention Mean (SD)	3.65 (0.64)	3.74 (0.55)	3.65 (0.59)	3.64 (0.61)
Relevance Mean (SD)	3.58 (0.56)	3.67 (0.63)	3.54 (0.78)	3.77 (0.57)
Involvement Mean (SD)	5.12 (1.14)	5.37 (1.01)	5.40 (1.15)	5.28 (0.90)
Rapport Mean (SD)	4.81 (0.72)	4.89 (0.67)	4.79 (0.72)	n.a.

No significant effects of the conditions on the outcome variables were found; feedback (F=1.73; p=0.16); autonomy (F=1.70; p=0.17), attention (F=0.59, p=0.62), relevance (F=0.52, p=0.67), involvement (F=0.49, p=0.69), rapport (F=0.21, p=0.81). However, t-tests on the individual conditions revealed significant differences for the autonomy outcome variable between AS, the most feature rich condition and TO (p=0.04), the control condition. For the feedback outcome variable the differences between AS and TO (p=0.05) and SS and TO (p=0.05) both reached significance. Subsequently, the interaction between the four conditions and gender type was analyzed. For the means and SD values, see table 5 below.

Table 5: Mean scores and standard deviation of the four conditions*gender type

	AS		**SS**		**ST**		**TO**	
	Female (n=44)	Male (n=14)	Female (n=46)	Male (n=12)	Female (n=49)	Male (n=6)	Female (n=43)	Male (n=16)
Feedback Mean (SD)	4.34 (1.20)	5.12 (1.25)	4.71 (1.19)	4.78 (1.35)	4.37 (1.17)	4.67 (0.94)	4.50 (1.17)	3.98 (0.98)
Autonomy Mean (SD)	5.28 (0.98)	5.82 (0.87)	5.45 (0.97)	5.46 (0.87)	5.43 (1.08)	5.67 (0.68)	5.29 (1.00)	4.97 (0.96)
Attention Mean (SD)	3.66 (0.69)	3.60 (0.49)	3.74 (0.59)	3.76 (0.39)	3.66 (0.59)	3.53 (0.55)	3.70 (0.62)	3.46 (0.57)
Relevance Mean (SD)	3.58 (0.61)	3.56 (0.34)	3.69 (0.66)	3.57 (0.52)	3.54 (0.74)	3.48 (1.13)	3.85 (0.57)	3.55 (0.50)
Involvement Mean (SD)	5.09 (1.16)	5.23 (1.10)	5.33 (1.03)	5.53 (0.95)	5.45 (1.14)	4.98 (1.31)	5.35 (0.97)	5.10 (0.63)
Rapport Mean (SD)	4.83 (0.66)	4.73 (0.81)	4.97 (0.66)	4.59 (0.68)	4.76 (0.72)	5.02 (0.74)	n.a.	n.a.

No significant effects of the interaction between the conditions and gender type were found on any of the outcome variables feedback*gender (F=2.29, p=0.79), autonomy*gender (F=1.47, p=0.22), attention*gender: (F=0.35, p=0.79), relevance*gender (F=0.40, p=0.75), involvement*gender: (F=0.71, p=0.54), rapport*gender (F=1.46, p=0.23). However, independent sample t-tests with selections on male participants on AS vs TO as control condition showed significant effects on feedback (t=2.81, p=0.01) and autonomy (t=2.54, p=0.02). The independent sample t-tests on gender differences for AS showed that for feedback male participants (5.12) graded it significantly higher (t=2.06, p=0.04) than female participants (4.34). No other effects of independent sample t-tests were found.

Discussion

Principal results

Within this study we found that visibility of the ECA does have a positive effect on the outcome measures of feedback and autonomy. Furthermore, on feedback we found a gender effect. Male participants graded the *visible* ECA's *higher* than female participants and graded the *non-visible* ECA *lower* than female participants. This feedback effect was corroborated by gender analyses on animation and on the separate conditions, where male participants scored the ECA significantly higher than female participants. Speech communication by the ECA also had a positive effect on feedback, without differentiating between gender type. Animation did not show effects in this study.

Interpretation of the nature of the outcome variables

When interpreting these results, one of our first questions was: why were effects found on feedback and autonomy and not on the other outcome variables? We suspected that the nature of the outcome variables could play a role. As they measured different constructs, we decided to analyze their specific character and purpose in relation to our results. Figure 4 below depicts our experimental outcome variables, which we ranked according to the level of abstraction.

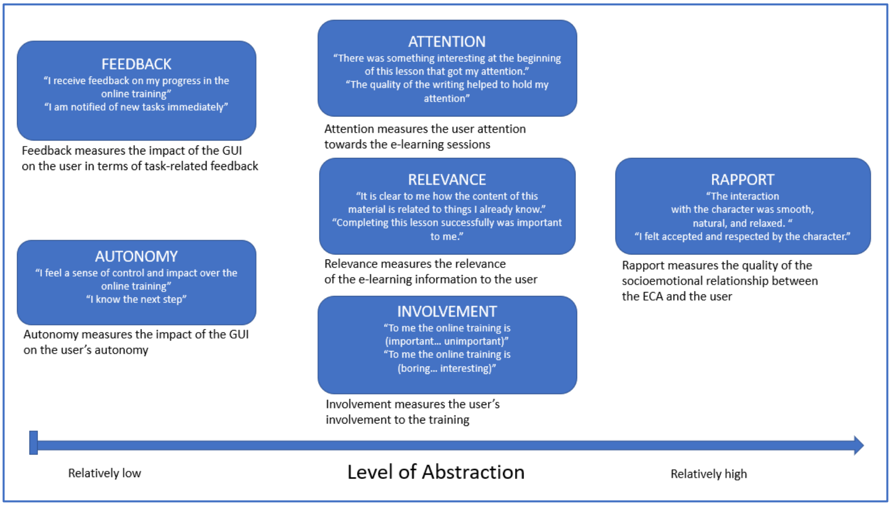


Figure 5. Sequence order of the outcome variables in terms of level of abstraction

As figure 5 shows, the task-related outcome variables feedback and autonomy are ranked lowest on level of abstraction. We will further discuss the figure, going from left to right.

Feedback and autonomy during the online training

These constructs address the way the GUI presents the user's tasks. The feedback and autonomy results demonstrate that when a user is doing the experiment, task-related support is more effective when delivered by a visible and speech-enabled ECA than by mere text. The social cue hypothesis that predicts deeper processing and higher personal relevance is therefore applicable to the modalities of visibility and speech, but not to animation. This is in accordance with the experimental result on animation of Lusk and Atkinson (2007) and with the stance that animation engages but also distracts users (Moreno et al, 2000). The engagement effect of animation seems to fit better with emotion-related support than with task-related support. However, our experiment did not demonstrate emotion-related effects of any kind, which we will discuss below in relation to the user state of distress. The explanation for the lack of an animation effect is further intricated by gender type; male participants graded animation significantly higher than female participants. This may be explained as a gender resemblance effect, (Baylor, 2011) but deserves further research.

Attention and relevance & involvement with the training

Attention, Relevance and Involvement question the user's learning experience. On these outcome variables, the visible and speech-enabled ECA did not induce effects. We interpret this as: although the users appreciated the feature-rich ECA providing task-related support (as demonstrated by the effects for feedback and autonomy), this effect did not *transfer* to the learning experience. In addition, the social cue hypothesis is not applicable to these outcome variables. We will expand on the reasons why this may be further below.

Rapport with the ECA

Most abstract is rapport, the relationship outcome variable. Rapport was measured on the three ECA conditions and not on the text-only condition. It measures the extent to which a relationship has been built between user and ECA. We added the variable for observation purposes. That is, we realized that it would be ambitious that signs of relationship would be found during a 30 minutes experiment where long-term interactions of e.g. 30 days are advised (Bickmore and Picard, 2005). This expectation was confirmed; no effects were found.

Comparison of our results to prior ECA studies and theories

Summarizing the results on the outcome variables, we found partial effects on feedback and autonomy. These constructs measure task-related support as provided by the GUI. No effects on learning experience and motivation were found, contrary to the results of the review study of Schroeder (2013). The implication of the social cue hypothesis of 'the more social cues, the more social effects' was therefore only partially confirmed. However, in line with the results of Schroeder (2013) we found an effect of the visibility of the ECA, but on a non-learning outcome variable: feedback. The feedback effect is in accordance with our expectation that users value practical support (*task-related support*) such as positively reinforcing log-in and intervention use when delivered by a simple, non-responsive ECA. This result for the feedback outcome variable also fits with Cassell's (2001) affordances that a visible ECA adds value as to make explicit *who* delivers the support. The *emotion-related support* of the ECA (positive confirmation after a lesson was done) seemed to have no effect on the learning experience.

Support, potentially only needed when in distress

The question is why the experiment did not show an ECA effect on learning experience. The answer may be found within the qualitative remarks of the participants, that generically stated it was a pleasant task. This makes it unlikely that a need for emotional support was induced. This probably made the social cues of the ECA superfluous. We further reason that users that experience episodes of distress (such as eHealth patients dealing with serious issues) have a greater need and indeed appreciation for support (Kraft et al, 2007). We envision a follow-up experiment during which users will carry out a mentally fatiguing pre-task, after which the effects of a supportive ECA will be assessed again. This concept is in line with the strength model (SM), a theory that describes that all acts of self-regulation rely on a common and limited energy source (Baumeister et al, 2007). According to this view, self-regulatory effort drains energy and leads to ego depletion (Baumeister et al, 2018) for which emotional support can provide a remedy (Kraft et al., 2007).

Additional measurement instruments

We started out by stating that ECA studies in general provide enigmatic results. Our results fit within this overall picture of ECA research. As an explanation, questionnaires as research tools may have their limitations measuring what users do and decide when interacting with ECA's. We envision a pre-experimental phase, during which users will shortly interact with both a text-only interface and an ECA interface. As a next step the user will be asked to choose their preferred interface for the core experiment. We wonder whether users will demonstrate a slight preference for ECA's compared to text-only solutions (as the present results suggest) or whether other results will appear. By continuing to use questionnaires at the end of the experiment, we may be able to cross-validate the users' prior decisions. Last, a remarkable result of our experiment is the *gender* effect that we found. Male participants valued our (male) ECA better in terms of feedback. We suspect this is an effect of gender similarity but it deserves further investigation. If we elaborate the action-driven method outlined above with a female ECA option, we will be able to test whether female participants choose female ECA's and whether they will score them higher on animation than they did within this experiment.

Limitations

Conclusions on ECA research are in general limited to their task and context. Concerning the task and context that were specific to our experimental set-up and could have influenced our results, we separated learning content (left part of the screen) from supportive content (right part of the screen). In addition, as learning content we used a positive psychology intervention. As supportive content we provided directions and gave positive feedback after a learning task was finalized by the user. This way we avoided distraction from the ECA towards the user, but we are not aware of similar set-ups in real life. The supportive content could be controlled by the user by using the click-through buttons, which provided user control, but which is unlike some other ECA set-ups that use vocal user input. Our feedback and autonomy outcome measures were both restricted to 3 items, more items would have been welcome. Our users were likely in a mental state of limited or no stress, which most likely did not induce a need for support.

Conclusion

Our experiment showed positive ECA effects when providing task-related support to users of a psycho-education environment. The ECA as a GUI seemed to make the task easier than text. However, our ECA was not capable of demonstrating effects as a result of its emotion-related support. This may be due to the friendly set-up of our experiment, that failed to bring users to a distressed, need-for-support mental state. Our hypothesis is that this disguises the true supportive potential of ECA's. Future research should aim to experimentally bring users to a mentally fatigued state within a long-term intervention in order to investigate whether emotional ECA support can be effective for user motivation. If indeed the ECA proves to be useful for users in such conditions, this provides a valuable argument for adding non-responsive ECA's to self-guided eHealth interventions for the sake of higher adherence and effect. We reckon that figure 1, describing a continuous line from support by the technology to support by human care providers, is relevant within the eHealth context. Our stance is that 'right-side' human support has its unique merits with which ECA's should not compete. The fact of the matter is that self-care technology has more potential than just providing tasks to users. The technology can be endowed with task-related and emotion-related supportive features from which users of self-guided interventions can benefit. We should not miss the opportunity to inform the 'left-side' technology to the support needs of our patients. As a means to realize this, we can add ECA's as a visible source of either supportive textual or (preferably) speech messages. In case we become successful at realizing support from within the technology itself, users of self-guided interventions will likely demonstrate higher adherence.

References

Peeters, J., Wiegers, T., & Friele, R. (2013). How technology in care at home affects patient self-care and self-management: a scoping review. International Journal of Environmental Research and Public Health, 10(11), 5541-5564.

Kennedy, B. M., Rehman, M., Johnson, W. D., Magee, M. B., Leonard, R., & Katzmarzyk, P. T. (2017). Healthcare Providers versus Patients' Understanding of Health Beliefs and Values. Patient experience journal, 4(3), 29.

Lorig, K. R., & Holman, H. R. (2003). Self-management education: history, definition, outcomes, and mechanisms. Annals of behavioral medicine, 26(1), 1-7.

Barak, A., Hen, L., Boniel-Nissim, M., & Shapira, N. A. (2008). A comprehensive review and a meta-analysis of the effectiveness of internet-based psychotherapeutic interventions. Journal of Technology in Human services, 26(2-4), 109-160.

Spek, V., Cuijpers, P. I. M., Nyklíček, I., Riper, H., Keyzer, J., & Pop, V. (2007). Internet-based cognitive behaviour therapy for symptoms of depression and anxiety: a meta-analysis. Psychological medicine, 37(3), 319-328.

Walsh, S., Szymczynska, P., Taylor, S. J., & Priebe, S. (2018). The acceptability of an online intervention using positive psychology for depression: A qualitative study. Internet interventions, 13, 60-66.

Flynn, D., Gregory, P., Makki, H., & Gabbay, M. (2009). Expectations and experiences of eHealth in primary care: a qualitative practice-based investigation. International journal of medical informatics, 78(9), 588-604.

Schubart, J. R., Stuckey, H. L., Ganeshamoorthy, A., & Sciamanna, C. N. (2011). Chronic health conditions and internet behavioral interventions: a review of factors to enhance user engagement. CIN: Computers, Informatics, Nursing, 29(2), 81-92.

Christensen, H., Griffiths, K. M., & Jorm, A. F. (2004). Delivering interventions for depression by using the internet: randomised controlled trial. Bmj, 328(7434), 265.

Lie, S. S., Karlsen, B., Oord, E. R., Graue, M., & Oftedal, B. (2017). Dropout from an eHealth intervention for adults with type 2 diabetes: a qualitative study. Journal of medical Internet research, 19(5).

Davis, M. J., & Addis, M. E. (1999). Predictors of attrition from behavioral medicine treatments. Annals of Behavioral Medicine, 21(4), 339-349.

de Silva, D. (2011). Evidence: helping people help themselves. The Health Foundation.

Wilkinson, A., & Whitehead, L. (2009). Evolution of the concept of self-care and implications for nurses: a literature review. International Journal of nursing studies, 46(8), 1143-1147.

Bickmore, T. (2010). Relational agents for chronic disease self management. Health Informatics: A Patient-Centered Approach to Diabetes, 181-204.

Drench, M. E., Noonan, A. C., Sharby, N., & Ventura, S. H. (2007). Psychosocial aspects of health care. Upper Saddle River, NJ: Pearson Prentice Hall.

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).

André, E., & Pelachaud, C. (2010). Interacting with embodied conversational agents. In Speech technology (pp. 123-149). Springer, Boston, MA.

Bostan, B., Kaplancali, U., Cad, K., & Yerlesimi, A. (2009). Explorations in Player Motivations: Game Mechanics. In GAMEON (pp. 5-11).

D'Mello, S., Picard, R. W., & Graesser, A. (2007). Toward an affect-sensitive AutoTutor. IEEE Intelligent Systems, 22(4).

Kopp, S., Gesellensetter, L., Krämer, N. C., & Wachsmuth, I. (2005, September). A conversational agent as museum guide–design and evaluation of a real-world application. In Intelligent virtual agents (pp. 329-343). Springer, Berlin, Heidelberg.

Kobori, Y., Osaka, A., Soh, S., & Okada, H. (2018). MP15-03 NOVEL APPLICATION FOR SEXUAL TRANSMITTED INFECTION SCREENING WITH AN AI CHATBOT. The Journal of Urology, 199(4), e189-e190.

Fitzpatrick, K. K., Darcy, A., & Vierhile, M. (2017). Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): a randomized controlled trial. JMIR mental health, 4(2).

Schroeder, N. L., Adesope, O. O., & Gilbert, R. B. (2013). How effective are pedagogical agents for learning? A meta-analytic review. Journal of Educational Computing Research, 49(1), 1-39.

Veletsianos, G., & Russell, G. S. (2014). Pedagogical agents. Handbook of research on educational communications and Technology içinde (ss. 759-769).

Domagk, S. (2010). Do pedagogical agents facilitate learner motivation and learning outcomes?. Journal of media Psychology.

Atkinson, R. K. (2002). Optimizing learning from examples using animated pedagogical agents. Journal of Educational Psychology, 94(2), 416.

Lusk, M. M., & Atkinson, R. K. (2007). Animated pedagogical agents: Does their degree of embodiment impact learning from static or animated worked examples?. Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition, 21(6), 747-764.

Graesser, A., Jackson, G. T., Ventura, M., Mueller, J., Hu, X., & Person, N. (2004). The impact of conversational navigational guides on the learning, use, and perceptions of users of a web site. In Agent-Mediated Knowledge Management (pp. 48-56). Springer, Berlin, Heidelberg.

Schroeder, N. (2013). Exploring Pedagogical Agent Use within Learner-attenuated System-paced Learning Environments.

Louwerse, M. M., Graesser, A. C., Lu, S., & Mitchell, H. H. (2005). Social cues in animated conversational agents. Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition, 19(6), 693-704.

Van Mulken, S., André, E., & Müller, J. (1998). The persona effect: How substantial is it?. In People and computers XIII (pp. 53-66). Springer, London.

Moreno, R., & Mayer, R. (2007). Interactive multimodal learning environments. Educational psychology review, 19(3), 309-326.

Cassell, J. (2001). Embodied conversational agents: representation and intelligence in user interfaces. AI magazine, 22(4), 67.

Baylor, A., Ryu, J., & Shen, E. (2003). The effects of pedagogical agent voice and animation on learning, motivation and perceived persona. In EdMedia: World Conference on Educational Media and Technology (pp. 452-458). Association for the Advancement of Computing in Education (AACE).

Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (1999). Acceptance and commitment therapy (p. 6). New York: Guilford Press.

Gable, S. L., & Haidt, J. (2005). What (and why) is positive psychology?. Review of general psychology, 9(2), 103.

Renner, F., Schwarz, P., Peters, M. L., & Huibers, M. J. (2014). Effects of a best-possible-self mental imagery exercise on mood and dysfunctional attitudes. Psychiatry research, 215(1), 105-110.

Fu, F. L., Su, R. C., & Yu, S. C. (2009). EGameFlow: A scale to measure learners' enjoyment of e-learning games. Computers & Education, 52(1), 101-112.

Whalen, S. P., & Csikszentmihalyi, M. (1991). Putting Flow Theory into Educational Practice: The Key School's Flow Activities Room. Report to the Benton Center for Curriculum and Instruction, University of Chicago.

Keller, J. M. (2006). Development of two measures of learner motivation. Unpublished Manuscript in progress. Florida State University.

Keller, J. M. (1987). Development and use of the ARCS model of instructional design. Journal of instructional development, 10(3), 2.

Shen, E. (2009). Effects of agent emotional support and cognitive motivational messages on math anxiety, learning, and motivation.

Zaichkowsky, J. L. (1994). The personal involvement inventory: Reduction, revision, and application to advertising. Journal of advertising, 23(4), 59-70.

Lo, S. K., & Cheng, M. W. (2010). The Effect of Online Agents on Advertising Effectiveness: The Presence Aspect. Management Review, 29, 99-102.

Tsui, P., & Schultz, G. L. (1985). Failure of rapport: Why psychotherapeutic engagement fails in the treatment of Asian clients. American Journal of Orthopsychiatry, 55(4), 561-569.

Fuchs, D. (1987). Examiner familiarity effects on test performance: implications for training and practice. Topics in Early Childhood Special Education, 7(3), 90-104.

Burns, M. (1984). Rapport and relationships: The basis of child care. Journal of Child Care.

Gratch, J., Wang, N., Gerten, J., Fast, E., & Duffy, R. (2007, September). Creating rapport with virtual agents. In International Workshop on Intelligent Virtual Agents (pp. 125-138). Springer, Berlin, Heidelberg.

Cerekovic, A., Aran, O., & Gatica-Perez, D. (2014, September). How do you like your virtual agent?: Human-agent interaction experience through nonverbal features and personality traits. In International Workshop on Human Behavior Understanding (pp. 1-15). Springer, Cham.

Moreno, R., Mayer, R., & Lester, J. (2000). Life-like pedagogical agents in constructivist multimedia environments: Cognitive consequences of their interaction. In EdMedia: World Conference on Educational Media and Technology (pp. 776-781). Association for the Advancement of Computing in Education (AACE).

Baylor, A. L. (2011). The design of motivational agents and avatars. Educational Technology Research *and Development*, *59*(2), 291-300.

Bickmore, T. W., & Picard, R. W. (2005). Establishing and maintaining long-term human-computer relationships. ACM Transactions on Computer-Human Interaction (TOCHI), 12(2), 293-327.

Kraft, P., Drozd, F., & Olsen, E. (2008, June). Digital therapy: Addressing willpower as part of the cognitive-affective processing system in the service of habit change. In International Conference on Persuasive Technology (pp. 177-188). Springer, Berlin, Heidelberg.

Baumeister, R. F., Vohs, K. D., & Tice, D. M. (2007). The strength model of self-control. Current directions in psychological science, 16(6), 351-355.

Baumeister, R. F., Bratslavsky, E., & Muraven, M. (2018). Ego depletion: Is the active self a limited resource?. In Self-Regulation and Self-Control (pp. 24-52). Routledge.

Chapter 4

Can an Embodied Conversational Agent effectively support eHealth users in distress?

Scholten MR, Kelders SM, Van Gemert-Pijnen JEWC

Can an Embodied Conversational Agent effectively support eHealth users in distress?

Heliyon, 2021, 7.3: e06509.

Abstract

Background: Stress is a prevalent issue amongst patients with chronic conditions. As eHealth interventions are gaining importance, it becomes more relevant to invoke the possibilities from the eHealth technology itself to provide support during experiences of stress. Embodied Conversational Agents (ECA's) also known as 'robots on screen' can potentially provide a remedy as support providers.

Objective: The objective of this paper is to investigate whether ECA support towards eHealth users is more readily accepted and appreciated by users who experience elevated levels of stress.

Methods: Within our eHealth experiment we applied a between-subjects design and experimentally studied the difference in appraisal of support as provided by either ECA's or textual guidance. The study was carried out amongst eHealth users of which half were deliberately put in a stressful pre-condition. The rationale was two-sided: we hypothesized that it would induce a need for external support and it would provide a fair representation of eHealth users in real life. The gender of the ECA was varied in order to investigate positive effects from a gender match between participant and ECA.

Results: The results show that the ECA did not demonstrate preferential effects compared to text as a control variable in any of the conditions. We suspect that the enduring visual presence of the ECA during task completion inhibited the users and led to the non-preferential effects.

Conclusion: our experimental results demonstrated that our ECA did not succeed in outperforming text, contrary to the results of our earlier study. The expected enlarged ECA support effect on users who experience stress, was not found., instead the ECA support effect vanished. This lack of evidence is not unprecedented in the ECA study field. As has been put forward within several ECA review studies; ECA research is multi-faceted and experimental studies regularly provide mixed and inconclusive results. We consider the results of our study as an affirmation of this phenomenon. Moreover, we realize that ECA research is challenging. The implementation of the ECA has to be spot-on for the participant to accept and prefer the ECA over textual guidance. If it is not implemented precisely right, the ECA will not yield preferential effects. In our study, the visibility of the ECA during task completion -despite its silent state- led to the absence of preferential effects.

Keywords: Stress; eHealth; support; Embodied Conversational Agent; ECA; persuasive technology

Introduction

It is well-established finding (see e.g. Vancamfort et al. 2016) that patients with chronic health conditions face elevated levels of stress. Stress is broadly defined as “a process by which a challenging emotional or physiological event or series of events result in adaptive or maladaptive changes required to regain homeostasis and/or stability” (Sinha and Jastreboff, 2013). Probably the most prominent physical cause of stress is pain (Abdallah and Geha, 2017, Rosenzweig et al., 2010). Stress amongst patients is also induced in more indirect ways such as the patient’s dwellings on his or her long-term prognosis. In Vancamfort et al.’s (2016) large epidemiological study on data from 229,293 adults living in 44 countries it is described in detail how chronic conditions lead to stress and reversely how stress worsens chronic conditions. Furthermore, the authors describe that stress can intensify the effect of chronic diseases such as asthma, arthritis, or diabetes as it increases experiences of pain and decreases adherence to medical treatment protocols. Within the eHealth domain, defined as ‘the use of information and communication technologies (ICT) for health’ (World Health Organization, 2016), stress is also referred to as a relevant factor. Leenen et al. (2016) describe eHealth patients’ stressful experiences in relation to their diseases in their study. As reported by the authors, carrying out eHealth self-management tasks is perceived by patients as an encounter with their physical and mental states. In a similar vein, Huygens et al., (2016) state that eHealth patients can become anxious from the information they find, particularly when reading information about complications that could occur at a later stage of their disease. But also carrying out seemingly innocent daily practical eHealth tasks can have unexpected stressful effects. Huygens et al. (2016) refer to a patient’s story measuring blood data as a routine, becoming more aware of his condition, and ultimately notifying this as a highly unpleasant experience. Another germane study (Kelders et al., 2013) reports on a group of users who dropped out from an intervention designed to reduce depressive complaints. This withdrawal occurred after a lesson that focused on the application of newly acquired skills in practice. Apparently, this lesson turned out to be too confrontational. Note that -from a treatment perspective- this lesson was as a key event for reaping the benefits from the eHealth intervention. Altogether, these studies suggest that eHealth self-management -although a sensible activity from a medical perspective- is often a daunting task from an emotional and personal perspective. In such as stressful situation, many patients lose motivation to continue using their eHealth interventions. Stated differently, intrinsic patient motivation starts to wane and external support has to be invoked.

1.1 Persuasive technology providing user support

A remedy to stimulate a patient's motivation is offered by persuasive technology. Persuasive technology is defined as 'computerized software or information systems designed to reinforce, change or shape attitudes or behaviors or both without using coercion or deception' (Oinas-Kukkonen & Harjumaa, 2009). According to van Gemert et al. (2018) persuasive technology is characterized by increased interactivity and engagement of users through modern information and communication technologies. A relevant instance of persuasive technology is the Embodied Conversational Agent, abbreviated as ECA. An ECA is a more or less autonomous and intelligent software entity with an embodiment used to communicate with the user (Ruttkay et al, 2004). Encouraging experimental set-ups have been realized with ECA's concerning the promotion of healthy behavior amongst patients (Sillice et al., 2018), training aspiring doctors for emotionally charged encounters with patients (Kron et al., 2017) and reaching out to a population that has an elevated PTSS profile but is avoiding mental healthcare (DeVault et al, 2014).

1.2 The present state of ECA study field

Although these ECA studies hold promise, Weiss et al. (2015) has convincingly outlined both the complexity and subtlety of the ECA study field. As Weiss et al. (2015) point out; depending on the application domain, different performance and quality aspects are important. That is, in a health literacy context, the ECA is required to engage the user. In contrast, in a care-taking situation, conveying empathy and provoking emotions are apt. With regards to their evidence, several meta-analyses have evaluated ECA effects, mostly within the eLearning domain. Within the meta-analysis of Schroeder et al. (2013) on 43 studies including 3,088 participants, a small but significant effect is reported on learning. The participants learned more from a system with an ECA, than a system without one. A second meta-analysis (Veletsianos and Russell, 2014) reports on studies in which both motivation and learning outcomes are promoted by ECA's. However, the authors also refer to studies on ECA's that failed to demonstrate added value compared to text-only conditions. Veletsianos and Russell (2014) summarized these mixed results as a conundrum and a challenge for new studies to take up.

1.3 This study as a successor of earlier positive ECA results

Within our earlier study (Scholten et al., 2019) we deployed a male ECA as an adjunct in an eHealth psycho-education intervention and compared its impact to a textual guidance control condition. We found a positive effect of the ECA's task-related, practical support. In contrast, we didn't find a user preference for the ECA because of its emotion-related, motivational capabilities. Following up on these results within this present study, we raise several topics.

As a first follow-up question on the Scholten et al., 2019 study: could the gender of the (male) ECA have played a role in the evaluation of the (mostly female) participants? Stated differently, could a match in gender between ECA and participant have contributed to a more positive user assessment? As reviewed in Baylor (2009), learners tend to be more influenced by an ECA of the same gender and ethnicity than agents who differ in those respects. Note that this phenomenon is similarly found in a human to human context; people are more readily persuaded by members of their in-group. We hypothesize that a female ECA in our new experimental set-up will result in enlarged support effects amongst female participants.

As a second follow-up topic, we hypothesize that study participants in distress are more in need of support than the participants in the original Scholten et al. (2019) study. In other words, stressed eHealth intervention users potentially value the supportive ECA better. Moreover, an experimental set-up including stress as a factor, makes it a more life-like eHealth intervention. Note however that empirical studies on ECA support for participants under stress are scarce. Prendinger et al. (2005) indicated that their affective ECA reduced the stress of participants as measured by galvanic skin response, and also led participants to experience a quiz as less difficult. Another study (Sanghoon et al., 2005) showed that the ECA's presence led to extra user stress. Thus, at first sight their ECA was counter-productive. However, as the authors concluded, the ECA's presence could still be considered as beneficial as it ultimately helped the user venting their stress experience. A last relevant study on reverse, adverse effects of ECA's (Blankendaal et al. 2015) has shown that an ECA can effectively create user frustration, but its impact is smaller than that of a human.As a third follow-up topic, the relationship between user and ECA needs to be further investigated. As we know from the literature (Bickmore et al., 2005), support that is provided by an ECA that has priorly established a relationship with the user has a much higher chance of being effective than support from an unfamiliar ECA. The quality of this user-ECA relationship is usually measured by the construct of rapport. Rapport has to do with a positive working relationship and being 'in tune' or 'click' with each other. The role of rapport in fostering effective social human interactions is well established. As reported by Gratch et al. (2013), rapport is underlying processes as diverse as social engagement (Tatar, 1997), success in teacher–student interactions (Bernieri, 1988), productive negotiations (Drolet and Morris, 2000), psychotherapeutic effectiveness (Tsui & Schultz, 1985). ECA's have been created that make use of small talk and humor as relationship building techniques (Bickmore, 2010). Some of these ECA's present themselves as mere speakers, thereby smartly avoiding the risk of falling short on their communication capabilities. In those cases, ECA's introduce themselves to the participant and explain their roles as support providers. This personal introduction -which is common practice within a human to human context- effectively creates a base of rapport between user and ECA (Bickmore, 2010). Note that such a personal introduction cannot be credibly provided through mere text, as there is not a visible sender as a source and point of reference. So, the visibility and personality of the ECA gives it distinctive qualities compared to textual guidance. For the purpose of this study we investigate whether this distinctive rapport creation ability will result in a user preference effect towards the supportive ECA. In addition, we aim to assess whether the ECA's rapport building activities will transfer to an overall positive eLearning experience. Moreover, we intend to investigate whether this effect will hold for users under stressful circumstances. Our underlying assumption is that stress will lead to an enlarged user need and appreciation for external support as provide by the ECA. As a precondition for effective support the literature tells us that ECA's should be capable of credibly presenting themselves as solutions for stress and avoid to be regarded as an additional source of user stress. As mentioned before, studies have shown that this can be achieved through the creation of a basic level of rapport with the user. However, it is an open question whether rapport will hold in stressful circumstances and whether the ECA will remain to be an effective support provider. We will therefore specifically investigate these matters within the present study.

1.4 This study

In this study we will include stress-induction on users, vary the ECA's gender, and stimulate the creation of rapport. We will verify the effects on the appreciation of the ECA. This brings us to the following research questions:

1) To what extent can we find preferential effects for the ECA compared to text, as to replicate the effect of the Scholten et al (2019) experiment?

2) To what extent does the experience of user distress positively affect the evaluation of the ECA?

3) To what extent do eHealth users provide higher ECA evaluations when interacting with an ECA of the same gender as compared to an ECA of different gender?

4) To what extent do positive user evaluations of the ECA lead to higher involvement of the user with the eHealth intervention?

Altogether, with the research questions we expect to find substantiation for ECA's as effective eHealth support providers. Stated differently, we aspire to find that the results corroborate the promise that ECA adjuncts provide a potential remedy for experiences of stress amongst users of self-guided eHealth interventions.

Materials and Methods

2.1 Recruitment of Participants

We recruited bachelor and master psychology and communication students at the University of Twente. As an inclusion criterion we set proficiency in English. As an exclusion criterion we set participation in a previous study with the ECA. The study protocol was reviewed and approved by the Twente Institutional Review Board. In total 106 participants were included. Participants were on average 20.4 years of age and represented 15 nationalities of which German (69%) and Dutch (18%) were most prominent. 80 participants were female (75.5%), 26 participants male (24.5%).

2.2 Design

To investigate the differential outcome effects of ECA's as a result of inducing stress and of a matching gender effect using a between-subjects design we set up the following pre-conditions and factors:

- Stressful versus non-stressful pre-condition (2 pre-conditions)
- Male ECA, female ECA, textual guidance (3 factors)

This resulted in 2*3 = 6 combinations to which participants were randomized.

The study design was a between-subjects experiment with two factors: the stress factor with 2 levels and the support factor with 3 levels. As portrayed in **Figure 1** below, randomization was done in two steps: during the first randomization, participants were either assigned to a stress or no stress pre-condition. During the second randomization, the participants were assigned to an e-learning intervention with as guidance either a female ECA, a male ECA, or text (control condition).

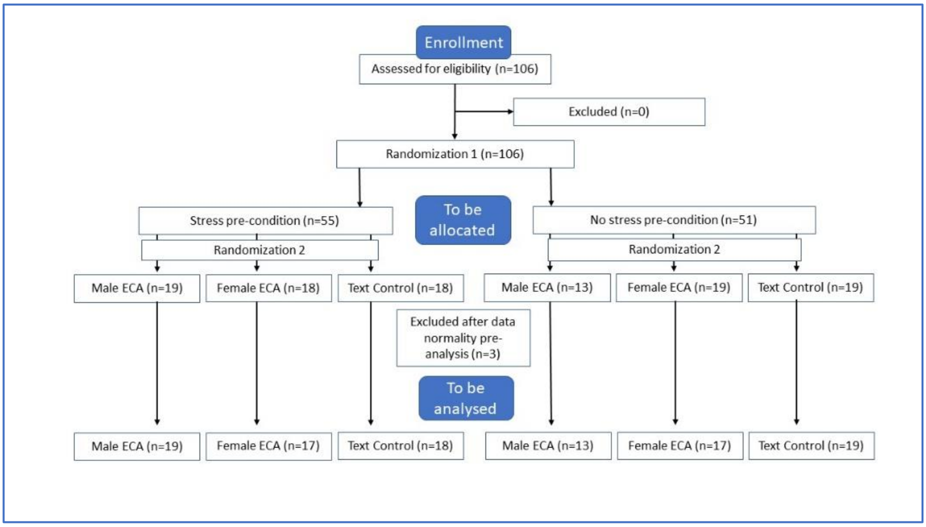


Figure 1. CONSORT flow chart for study participation.

2.3 Intervention

2.3.1 Pre-conditions

The pre-conditions were displayed on separate WordPress webpages (version 4.9.7) containing information on playing a Pac-Man game, see **Figure 2** below. The no stress webpage had a hyperlink to a regularly functioning Pac-Man version that had been uploaded to a GitHub site. The stress webpage contained a hyperlink to a second, invalidated Pac-Man version on GitHub. The invalidated Pac-Man version did not properly respond to the user's arrow key strokes in 30% of the occasions. Instead it went into another randomly chosen direction. This type of invalidation for the purpose of generating participant stress was inspired by the Affective Pac-Man solution from the study of Reuderink et al. (2009).

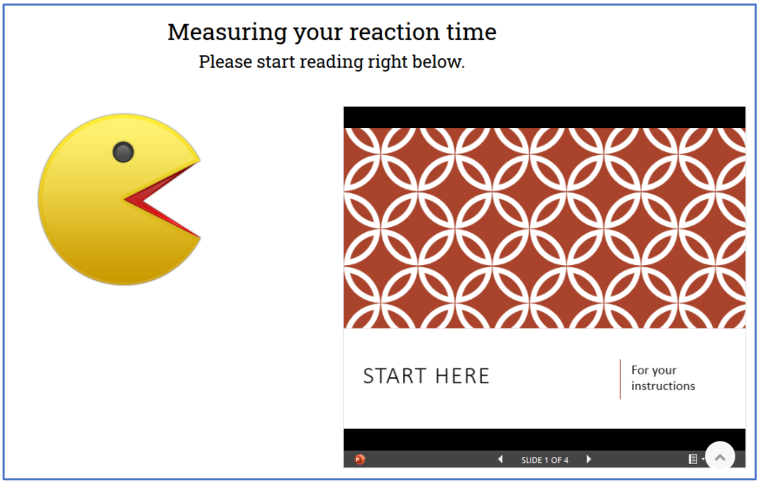


Figure 2. Webpage for the stressful and non-stressful pre-condition.

After the pre-conditional phase was rounded off, the participants were redirected to the main experiment.

2.4 Main experiment

This main experiment was run on a WordPress website (version 4.9.7) that contained the eHealth intervention on the left side of the webpage. The eHealth intervention was a PowerPoint® presentation with psycho-education material on positive psychology. The goal of the eHealth psycho-education intervention was to make users knowledgeable about positive psychology. Positive psychology focuses on the abilities of people and their potential to flourish. Several treatments against depression are based on positive psychology principles (Hayes et al., 1999). In addition, positive psychology and happiness are subjects that are of general human interest. As we reasoned, this topic would contribute to engage participants for our experiment. The self-guided eHealth intervention contained a combination of theory and exercises, including the remunerated “three good things exercise” and “best possible self-exercise” (Renner et al., 2014).

Figure 3. The eHealth psycho-education intervention. On the left side of the webpage the psycho-educational content is displayed, on the right side the support condition with guidance and directions (task-related support) and encouragement (emotion-related support) is presented. The example shown is the female ECA.

2.4.1 User support

As **Figure 3** displays, User guidance and support was provided on the right side of the webpage by either a female ECA, male ECA or text. The female and male ECA conditions were created through the Voki application. The ECA represented a virtual person in between 20 and 30 years of age, with Caucasian looks, acting as an informal (i.e. not medical) support provider. The female and male voices were provided by two Dutch speakers. The textual guidance condition was created using Microsoft PowerPoint®. All support conditions expressed the same guidance conveyed in English. The guidance was a combination of task-related support (e.g., “within this experiment you will read about positive psychology and you will do some exercises”) and emotional support (e.g., “well done!”). In addition, the user was stimulated to take advantage of the exercises in daily life.

An explicit separation was created between the instructional phase during which the ECA (or text) provided instructions and the learning phase, following those instructions. This was done to control for the split-attention effect (Louwerse et al., 2005). The effect contends that an ECA that is starting up conversations will distract the student when he is processing the e-learning material. Therefore, during the leaning phase, the ECA was silent.

2.4.2 Outcome Measures

We selected a variety of outcome measures, to measure both the practical benefit of support (as presented on the left side of **Figure 4**) and the socio-emotional benefit (as presented on the right side of **Figure 4**). Furthermore, we expected both the practical and socio-emotional aspects to impact the key outcome variable (as presented in the middle); the user's involvement with the eHealth intervention. We also decided to explore the application of an innovative graphical outcome measure, PrEmo. All outcome variables are discussed in further detail below, going from left to right in **Figure 4**.

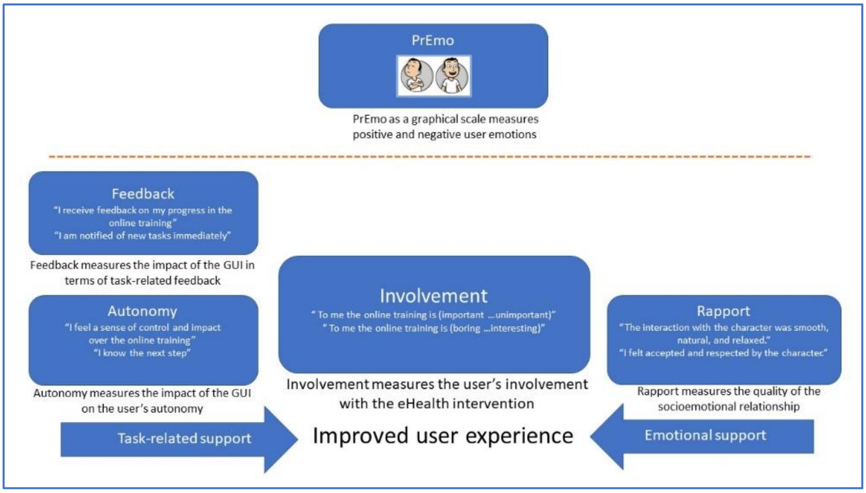


Figure 4. The ECA's task-related support (left side) and emotion-related support (right side) are hypothesized to have an impact on the user's involvement (middle) with the eHealth intervention

First, the autonomy and feedback dimensions of the larger EGameFlow scale (Fu et al., 2009) were selected as validated subscales that measure learners' enjoyment of e-learning games. Autonomy and feedback both represent the effects of task-related support. From both these scales three out of six items were chosen on the basis of validation and relevance to the experiment. Both scales use a seven-point Likert scale ranging from "strongly disagree" to "strongly agree." Within the wording of the subscales the word "game" was replaced by "online training."

Furthermore, as outcome variable involvement was selected. The Personal Involvement Inventory (Zaichkowsky, 1994) is a context free measure applicable to involvement with products, with advertisements, and with purchase processes. It has proven to be a useful outcome variable to evaluate environments with ECA's (Lo and Cheng, 2010, Scholten et al., 2019). In our case it assesses user motivation for the eHealth intervention. The scale consists of 10 items and uses a seven-point Likert scale with varying category names such as “important” vs. “not important” and “boring” vs. “interesting”. As indicated by the left and right arrows pointing to involvement, we expected that the involvement with the eHealth intervention would be positively impacted by both the practical and emotional benefits of the ECA’s support.

In addition, as outcome measure, PrEmo was chosen. PrEmo is a non-verbal self-report instrument that measures seven positive (further referred to as ‘PrEmoPos’) and seven negative emotions (further referred to as ‘PrEmoNeg’). It measures distinct emotions in a direct manner as it does not require the respondents to verbalize them, see **Figure 5** below**.**

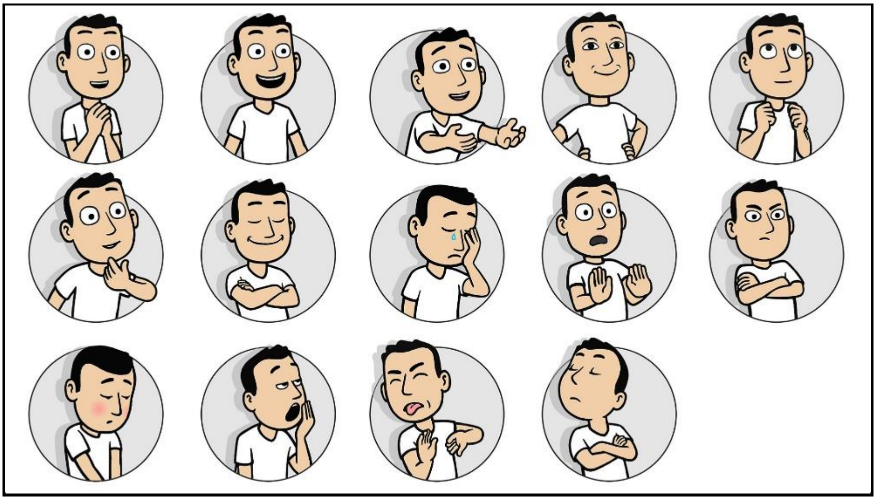


Figure 5: PrEmo visual outcome measure

Last, the Rapport scale was selected. Rapport is an umbrella term for generic positive interactions between human counterparts, which as a term is also associated to terms as

synchrony and flow. It has been used for ECA evaluations before (Brixey and Novick, 2019; Gratch, J., et al., 2013). The Rapport scale (Cerekovic et al., 2014) consists of fifteen items and we used a seven-point Likert scale ranging from [(1) – Disagree strongly to (7) – Agree strongly].

2.4.3 Procedure

The webpages were put online, and the study was run without human supervision to simulate self-guidance. Users were provided with an URL that led to the Qualtrics system. Users were directed to one of two pre-condition webpages representing the stressful and no stress pre-conditions. After carrying out the pre-condition, users were led to a Qualtrics environment where users were asked to fill in their Pac-Man high scores. In addition, they were asked to fill in a short version of the PrEmo questionnaire as a check on the effectiveness of the pre-conditions (manipulation check). After that, users were re-directed to the e-psycho-education environment. During instruction on the right side of the webpage, the user was told what learning module would come next. Then the user was asked to click on the left side of the webpage and follow up on the psycho-education tasks. When the psycho-education module had come to an end, the user was asked to go to the right side of the webpage for new instructions. When three psycho-education modules were done, the user was re-directed to the Qualtrics environment where the final questionnaires were presented.

Data analysis

2.5.1 Data distribution check

Before starting with the core statistical analysis, we first performed a check on the normality of the distributions of the outcome data. It appeared that the PremoNeg outcome variable was strongly left skewed (skewness=2.3, kurtosis =4.7). We tried to resolve this through re-normalization and by deleting outliers. However, both methods did not resolve the issue in a satisfactory way, so we decided not to involve PremoNeg in our main statistical analyses. Instead, we decided to run a sperate and specific non-parametric analysis on PremoNeg. In addition, the outcome variable Involvement showed some right skewness, (skewness=-1.2, kurtosis =1.8), which we resolved (skewness=-0.7, kurtosis =0.4) by deleting 3 outliers. As a result, the number of participants decreased from 106 to 103.

2.5.2 Stress pre-conditions and guidance conditions

As a first step, we did a manipulation check on the effect of the stress and no stress pre-conditions as measured after the start and after the end of the experiment by means of the outcome variable PrEmo. For the PrEmoNeg outcome variable we applied a non-parametric test that suited the non-gaussian distribution of the data. For the normally distributed PrEmoPos outcome variable we utilized the t-test. Secondly, we grouped all the cases and divided them according to the following three factors:

1) ECA with a gender that matches the gender of the participants
2) ECA with a gender that does not match with the gender of the participants

3) Textual guidance and support (control variable)

and analyzed the differential effects on our outcome variables using ANOVA. Thirdly, we analyzed the three guidance factors while taking stress into account using a two-way ANOVA. Furthermore, we performed a multiple regression test on the outcome variable involvement in order to find out to what extent external user support has an impact on user involvement.

RESULTS

3.1 Manipulation check: stress effects at start and end of the experiment

To check whether we succeeded in inducing a light level of stress amongst participants. We analyzed the PrEmo questionnaire as applied immediately after the pre-conditions. This questionnaire contained three items out of the six emotions for both PremoPos and PremoNeg. Furthermore, we analyzed the full PrEmoPos and PrEmoNeg questionnaire as applied after the experiment. The means, 95% Confidence Interval and SD values of the outcome variables are shown in Table 1.

Table 1. Mean scores and standard deviation of the effect of the pre-condition on the PrEmoNeg and PrEmoPos outcome variables, measured after the pre-conditions and after the experiment

	Stress pre-condition (n=54)	No stress pre-condition (n=49)
PrEmoNegStart (1-5)	2.2 (1.9-2.6; 0.2)*	1.9 (1.6-2.2; 0.2)**
PrEmoNegEnd (1-5)	1.6 (1.4-1.9; 0.1)*	1.6 (1.3-1.9; 0.1)**
PrEmoPosStart (1-5)	2.7 (2.4-3.0; 0.1)***	2.9 (2.6-3.2; 0.1)
PrEmoPosEnd (1-5)	3.1 (2.8-3.4; 0.1)***	3.1 (2.8-3.3; 0.1)

*significant evolvement effect of p=0.00

**significant evolvement effect of p=0.02

***significant evolvement effect of p=0.00

The independent samples median test on PrEmoNegStart showed no significant difference between the stress vs no stress pre-conditions but did show a tendency (p=0.10) in the direction of the stress pre-condition. For PrEmoNegEnd the independent samples median test clearly showed no significant difference between the stress vs no stress pre-conditions (p=0.94).

For PrEmoPosStart the t-test on the differences between the stress and no stress pre-condition showed no significant difference between the stress vs no stress pre-conditions (p=0.31). For PrEmoPosEnd the t-test on the differences between the stress and no stress pre-condition showed no significant difference (p=0.86) either. Altogether the 'start' results indicated that our stress manipulation led to a tendency effect on negative emotions. As a follow-up analysis, we investigated how the PrEmo measurements had evolved from start to end during the experiment. For both the stress (p=0.00) and no-stress (p=0.02) pre-condition, the start vs end difference of PrEmoNeg was significant with lower values at the end. This effect suggests that initially there was stress amongst users that had faded out at the end. For the PrEmoPos variable the end values were equal (3.1) for both the stress and no stress pre-conditions. Most importantly, for the stress pre-condition this represented a significant increase in positive emotions (p=0.00) whereas for the no-stress pre-condition this was not the case (p=0.27). Altogether the results on 'evolvement' of positive emotions indicated that the stress manipulation had been successful.

3.1.1 Stress vs. No stress

As a next step, we ran a one-way ANOVA to analyze the differences between the stress and no stress conditions. The means, 95% Confidence Interval and SD values of the outcome variables are shown in Table 2.

Table 2. Mean scores and standard deviation on the stress-no stress distinction.

	Stress pre-condition (n=54)	**No stress pre-condition (n=49)**
Feedback (1-7)	4.4 (4.2-4.6; 0.2)	4.7 (4.5-4.9; 0.2)
Autonomy (1-7)	5.4 (5.2-5.7;0.2)	5.3 (5.1-5.5; 0.2)
Involvement (1-7)	5.3 (5.2-5.4; 0.1)	5.4 (5.3-5.5;0.1)
PrEmoPos (1-5)	3.1 (3.0-3.2; 0.1)	3.1 (3.0-3.2; 0.1)
Rapport (1-7)	3.6 (3.4-3.8; 0.1) n=37	3.7 (3.5-3.9; 0.1) n=32

The one-way ANOVA revealed no significant differences on any of the outcome variables feedback (F=1.66; p=0.20), autonomy (F=0.23; p=0.63), involvement (F=0.27; p=0.61) premopos (F=0.03; p=0.86).

3.1.2 Stress and guidance

We sub-divided the stress and no stress pre-conditions into ECA guidance with a matching gender, ECA guidance with a not matching gender and finally textually guided section, in order to analyze potential differential effects. The means, 95% Confidence Interval and SD values of the outcome variables are shown in Table 3.

Table 3. Mean scores and standard deviation on the stress/not-stress pre-condition distinction, subdivided to the three support conditions

Table 3A Stress pre-condition (n=54)

	ECA matching gender (n=11)	ECA not matching gender (n=25)	Text (n=18)
Feedback (1-7)	4.6 (3.8-5.3; 0.4)	4.4 (3.9-4.9; 0.3)	4.2 (3.9-4.5; 0.3)*
Autonomy (1-7)	5.2 (4.5-5.8; 0.3)	5.5 (5.0-5.9;0.2)	5.4 (5.2-5.6; 0.2)
Involvement (1-7)	5.5 (4.9-6.1; 0.3)	5.0 (4.6-5.4; 0.2)	5.5 (5.2-5.8; 0.3)
PrEmoPos (1-5)	2.6 (2.0-3.1; 0.3)**	3.3 (2.9-3.7; 0.2)**	3.1 (2.9-3.3; 0.2)
Rapport (1-7)	3.7 (3.3-4.0; 0.2)	3.6 (3.54.1; 0.1)	n.a.

Table 3B No stress pre-condition (n=49**)**

	ECA matching gender (n=16)	ECA not matching gender (n=14)	Text (n=19)
Feedback (1-7)	4.5 (3.9-5.1; 0.3)	4.4 (3.7-5.0; 0.3)	5.1 (4.8-5.4; 0.3)*
Autonomy (1-7)	5.2 (4.5-5.8; 0.3)	5.5 (5.0-5.9; 0.2)	5.2 (4.9-5.5; 0.3)
Involvement (1-7)	5.5 (4.9-6.1; 0.3)	5.0 (4.6-5.4; 0.2)	5.4 (5.2-5.6; 0.2)
PrEmoPos (1-5)	2.6 (2.0-3.1; 0.3)	3.3 (2.9-3.7; 0.2)	3.0 (2.8-3.2; 0.2)
Rapport (1-7)	3.8 (3.5-4.1; 0.1)	3.6 (3.3-3.9; 0.1)	n.a.

*significant effect of p=0.04
**significant effect of p=0.04

Overall, the two-way ANOVA showed no significant effects; feedback (F=1.1; p=0.37), autonomy (F=0.21, p=0.96), involvement (F=0.74, p=0.59), premopos (F=0.96, p=0.45), rapport (F=0.47; p=0.7). However, significant effects were found on individual conditions: For the textual guidance, the difference between the stress pre-condition (4.2) and no stress pre-condition (5.1) was significant on the feedback outcome variable (p=0.04) in favor of the no stress pre-condition. For the stress pre-condition, the difference between the ECA with a matching gender (2.6) and with a not matching gender (3.3) was significant for the outcome variable premopos (p=0.04). Remarkably, the not matching gender ECA had the higher premopos score.

3.1.3 Effect of a gender match between ECA and participant

Next, we grouped the stress and no stress cases together and analyzed the effect of a match of gender between the participant and the ECA, using text as a control variable. The means, 95% Confidence Interval and SD values of the outcome variables are shown in Table 4.

Table 4. Mean scores and standard deviation on the gender match/mismatch distinction

	Matching gender (n=27)	Not matching gender (n=39)	Text (n=37)
Feedback (1-7)	5.0; 4.4-5.7; 0.2	4.5; 4.1-4.9; 0.2	4.7 (4.3-5.1; 0.2)
Autonomy (1-7)	5.5; 4.9-6.1; 0.2	5.5; 5.1-5.9; 0.2	5.3 (4.9-5.6; 0.2)
Involvement (1-7)	5.6; 5.0-6.1; 0.2	5.0; 4.7-5.4; 0.2	5.4 (5.1-5.8; 0.2)
PrEmoPos (1-5)	2.9; 2.3–3.4; 0.2	3.1; 2.8-3.4; 0.1	3.1 (2.8-3.3; 0.2)
Rapport (1-7)	3.8; 3.5-4.0; 0.2	3.7; 3.5-3.8; 0.2	n.a.

The one-way ANOVA revealed no significant differences on any of the outcome variables feedback (F=-0,10; p=0.66), autonomy (F=0.58; p=0.45), premopos (F=1.32; p=0.26) and rapport (F=0.64; p=0.43), although involvement (F=2.7; p=0.11) came close to a tendency. As a next step, we left out the textual condition and sub-divided the matching gender conditions in a female and male participant section, in order to analyze potential differential matching gender effects between ECA and user. We ran a two-way ANOVA. the means, 95% Confidence Interval and SD values of the outcome variables are shown in Table 5.

Table 5. Mean scores and standard deviation on the gender match/mismatch distinction subdivided to the participant's gender

	Matching gender (n=27)		Not matching gender (n=39)	
	A Female participant (n=23)	B Male participant (n=4)	C Female participant (n=27)	D Male participant (n=12)
Feedback (1-7)	4.3; 3.8-4.8; 0.3**	5.7; 4.5-7.0; 0.6**	4.2; 3-8-4.7; 0.2**	4.8; 4.1-5.5; 0.4
Autonomy (1-7)	5.1; 4.6-5.6; 0.2	5.8; 5.7-7.0; 0.6	5.3; 4.9-5.8;0.2	5.7; 5.0-6.3; 0.3
Involvement (1-7)	5.5; 5.1-5.9; 0.2	5.6; 4.6-6.6; 0.5	5.2; 4.8-5.6; 0.2	4.9; 4.3-5.5; 0.3
PrEmoPos (1-5)	2.9; 2.5-3.3; 0.2	2.8; 1.8-3.7; 0.5	3.3; 3.0-3.7; 0.2	2.9; 2.3-3.4; 0.3
Rapport (1-7)*	3.6; 3.5-3.8; 0.1***	4.4; 4.0-4.9; 0.2***	3.8; 3.6-3.9; 0.1***	3.3; 3.1-3.6; 0.1***

*significant ANOVA effect of p=0.00

**significant pairwise effects of p=0.02 (B vs C) and p=0.03 (A vs B)

***significant pairwise effects of p=0.00 (A vs B, B vs D, C vs D), p=0.01 (B vs C), p=0.05 (A vs D)

We found an effect from the ANOVA for rapport (F=10.53; p=0.00) indicating that the four conditions differed significantly from each other. For all the other outcome

variables we did not find an effect; feedback (F=1.89; p=0.18), autonomy (F=0.01; p=0.93), involvement (F=2.42; p=0.13) and premopos (F=0.59; p=0.45).
Pairwise comparisons showed for feedback significant differences between female participants with a matching gender (A) and male participants with a matching gender (B) (p=0.03) and between male participants with a matching gender (B) and female participants with a not matching gender (C) (p=0.02). Pairwise comparisons for rapport demonstrated significant differences between A and B (p=0.00), A and D (p=0.05), as well as between B and C (p=0.01), B and D (p=0.00) and finally C and D (p=0.00). The number of male participants of a matching gender (B) was very low (n=4) which made it difficult to draw firm conclusions. However, the high rapport scores for male participants with a matching gender (B) and low rapport scores for male participants with a not matching gender (D) suggest that male participants prefer a male ECA. For female participants the outcome was completely different. The rapport score (3.8) for the not matching gender C was even higher than the rapport score (3.6) for the matching gender A, but this difference did not reach significance (p=0.21).

3.1.4 The potential role of the participant's gender

As a next step, we analyzed the effect of the participant's gender (female versus male) as a stand-alone factor by a one-way ANOVA. Our objective was to find out whether ECA guidance had a differential effect on female versus male participants. The means, 95% Confidence Interval and SD values of the outcome variables are shown in Table 6.

Table 6. Mean scores and standard deviation on the participant's gender distinction.

	Female participant (n=50)	**Male participant (n=16)**
Feedback (1-7)*	4.3; 3.9-4.6; 0.2	5.0; 4.3-5.7; 0.3
Autonomy (1-7)	5.2; 4.9-5.6; 0.2	5.7; 5.1-5.6; 0.3
Involvement (1-7)	5.3; 5.1-5.6; 0.1	5.1; 4.5-5.6; 0.3
PrEmoPos (1-5)	3.1; 2.9–3.4; 0.1	2.8; 2.4-3.3; 0.2
Rapport (1-7)	3.7; 3.6-3.8; 0.1	3.6; 3.2-4.0; 0.1

*significant effect of p=0.03

For ECA guidance, we found a significant effect for feedback (F=4.74; p=0.03) in the direction of male participants. On the other variables no effects were found: autonomy (F=2.48; p=0.12), involvement (F=0.92; p=0.34), premopos (F=1.30; p=0.26), rapport (F=0.75; p=0.39).

3.1.5 The potential role of the ECA's gender

As a next step, we analyzed the effect of the ECA's gender (female versus male) as a stand-alone factor by a one-way ANOVA. Our objective was to find out whether guidance by either a female ECA or male ECA had a differential effect on participants. The means, 95% Confidence Interval and SD values of the outcome variables are shown in Table 7.

Table 7. Mean scores and standard deviation on the participant's gender distinction.

	Female ECA (n=35)	Male ECA (n=31)
Feedback (1-7)	4.5; 4.1-4.9; 0.2	4.4; 4.0-4.9; 0.2
Autonomy (1-7)	5.3; 4.9-5.7; 0.2	5.4; 5.0-5.8; 0.2
Involvement (1-7)	5.3; 4.9-5.6; 0.2	5.2; 4.9-5.6; 0.2
PrEmoPos (1-5)	2.9; 2.6–3.2; 0.2	3.2; 2.9-3.6; 0.2
Rapport (1-7)*	3.5; 3.4-3.7;0.1	3.9; 3.7-4.0;0.1

*significant effect of p=0.00

With respect to the gender of the ECA, we found a significant effect for rapport (F=8.74; p=0.00) in the direction of the male ECA. On the other variables no effects were found: feedback (F=0.04; p=0.85), autonomy (F=0.07; p=0.79), involvement (F=0.02; p=0.89), premopos (F=2.31; p=0.13).

3.1.6 User involvement and external support

Finally, we conducted a multiple regression test in order to check the extent to which user involvement was impacted by task-related support and emotion-related support. We took the outcome variable involvement as the dependent variable and feedback, autonomy and rapport as independent variables. The resulting regression was significant and explained 22% of variance (R-squared 0.22; p=0.00). Significant predictors in the model were autonomy (p=0.01) and rapport (p=0.01), but not feedback (p=0.73). Rapport was relatively most important contributing a R-squared value of 0.12, autonomy contributed a R-squared value of 0.09.

4. Discussion

4.1 Principal results

We didn't find the four suspected effects: the ECA guidance didn't yield higher scores than textual guidance in any situation; so, we didn't replicate the positive ECA guidance effect of the Scholten et al. (2019) study. Furthermore, our way of inducing stress did not positively impact the user's evaluation of the ECA. Thirdly, a main matching gender effect was not found. Instead, a gender-insensitive preference effect for the male ECA was found on rapport. Fourth, higher involvement with the eHealth intervention resulting from the deployment of the ECA was not found.

4.2 Interpretation of our effects

No ECA support effect compared to text as the control condition

Regarding guidance, we found that there was no preference for the ECA compared to text. A potential explanation is the remaining visibility of the ECA during the phase that the participants carried out their tasks. It is known from the social psychology literature that the attention of others fosters mastery of simple tasks but impairs mastery of complex and stressful tasks. This is also known as the theory of social facilitation and inhibition (Steinmetz and Pfattheicher, 2017; Zajonc & Sales, 1966). With the ECA study field, these impairment effects have been previously reported on; female participants were hindered by the presence of the ECA when they performed a task that was stressful due to its novelty (Zanbaka, 2004). Furthermore, Rickenberg and Reeves (2000) found that users felt more anxious when an ECA monitored their website work which led to a decrease in the user's task performance.

4.2.1 Stress did not induce higher ECA evaluations

In our experiment, our implementation of stress did not initiate the hypothesized causal chain of a higher user need for support and therefore more elevated levels of appreciation for the ECA. In order to understand the absence of this effect, we first checked whether we had been successful in inducing stress on participants in the first place. The results were mixed and inconclusive. The analysis on the evolvement of positive emotions showed a significant increase from start to end of the experiment in case of the stressful pre-condition. In contrast, the analysis on negative emotions only showed a tendency of an effect of the stressful pre-condition. In any event, the regular Pac-Man game provided a stressful baseline, that was hard for the invalidated Pac-Man to surpass. For all pre-conditions, the evolvement of all the PrEmo variables showed that users had lost their stress at the end of the experiment. This sketches an image of participants who are nervous at start of the experiment, just after they have played the pre-conditional Pac-Man game. During the experiment the participants recovered from the stress, possibly with some help of the Positive Psychology intervention. From other studies, we know that the duration of emotions including stress are highly variable, with emotions lasting anywhere from a few of seconds up to several hours, or even days (Verduyn et al., 2009). We reckon that either a stronger single dose of stress or multiple doses of stress during the experiment could have kept the participants in a more prolonged stressed state. We foresee a scenario of fluctuating stress levels during which the user is first stressed by an experimental stressful event and subsequently supported and comforted by the ECA, resulting again in lower stress levels. We find it plausible that such a fluctuating scenario will lead to higher user evaluations for the ECA. However, implementing such a scenario is challenging, stress as a side-effect impairs participants (Sänger et al, 2014). Moreover, inducing stress is not very common within eHealth experiments and is of course bounded by ethical guidelines.

4.2.2 No higher ECA evaluations when the ECA’s gender matches with the participant

With regards to a gender match between ECA and participant we did not find a main effect. A more fine-grained analysis on the rapport outcome variable demonstrated a gender specific effect; male participants had a preference for a male ECA. As an explanation for the absence of a main matching gender effect, the literature (e.g. ter Stal et al., 2019) mentions that although (gender) resemblance between participant and user is an important factor, multiple factors should be taken into consideration. This includes the ECA’s age, voice and its role. Within our set-up there was no difference between the female ECA and male ECA with regards to their ages and roles. In both cases we applied an ECA in its twenties, taking the role of a support provider. Still, the male ECA was preferred by all participants on the relationship (rapport) variable. Potentially, the voice or any aspect of the visual appearance of the male ECA was preferred, which outweighed the effect of the matching gender for the female participants. As an alternative explanation, the experimental role fitted the stereotype of the male ECA better. The experimental set-up with the Pac-Man pre-condition and ECA could be considered as somewhat technical. Baylor (2011) reports on gender stereotypes as applied by study participants within a technical context. As she reports, in the context of promoting young women’s motivation toward engineering as a career field, it was found that the male ECA’s were more influential than the female ECA’s. Other studies (Forlizzi et al., 2007; Zimmerman et al., 2005) also suggest that people prefer ECA’s along the lines of gender stereotypes and not along the line of matching genders.

4.2.3 Effect of the participants’ gender

For feedback (e.g. “I am notified of new tasks immediately”) the male participants graded the ECA (female and male ECA’s analyzed as one group) significantly higher than female participants did. This strongly resembled the results of Scholten et al. (2019) study with the male-only ECA. Furthermore, these results are in line with the results of Zhang et al (2013) who reported that male users graded the facilitating conditions (resources needed to achieve the behavior) of mobile eHealth technology as more important than female participants.

4.2.4 User involvement can partially be ascribed to user support

Our regression analysis showed that a portion of the user involvement could be ascribed to task-related and emotion-related support. We interpret this as an indication that, on the one hand, participants acted on an autonomous basis and were largely self-supportive. On the other hand, participants were stimulated by external support indeed. The relationship between external support and self-support is relevant in this sense. According to the Self-Determination (SDT) theory of Ryan and Deci (2003) the human needs of autonomy, competence and relatedness all require fulfillment. It is known (Legault, 2017) that external support can positively impact competence and have uplifting effects on relatedness, but result in lower experiences of autonomy. Application of SDT to the limited effect of external support that we found, can be interpreted in the light of participants who largely felt competent to carry out their experimental tasks. In case a stronger experimental dose of stress would have been provided we hypothesize that the level of competence would have been more strongly impaired. Our further hypothesis would be that in such a case the regression analysis would have portrayed stronger positive effects of external support in relation to user involvement. With respect to the *kind* of user support provided within the present experiment: in line with our expectations our results suggest that a mix of task-related and emotion-related support is advisable. Within the multiple regression analysis both types contributed, which a slightly higher contribution from emotional-related support compared to task-related support.

5. Conclusion

Our main ambition for studying the effectiveness of an ECA acting as an adjunct to a self-guided eHealth context was its potential to deliver higher evaluated user guidance and support than plain text. However, our experimental results demonstrated that our ECA did not succeed in outperforming text, contrary to the results of our earlier study. This lack of evidence is not unprecedented in the ECA study field. As has been put forward within several ECA review studies; ECA research is multi-faceted and experimental studies regularly provide mixed and inconclusive results. We consider the results of our study as an affirmation of this phenomenon. Moreover, we realize that ECA research is challenging. The implementation of the ECA has to be spot-on for the participant to accept and prefer the ECA over textual guidance. If it is not implemented precisely right, the ECA will not yield preferential effects. In our study, the visibility of the ECA during task completion -despite its silent state- led to the absence of preferential effects. With regards to this study's purpose, once more we emphasize the widely prevalence of patient stress and the potential from the eHealth technology itself to offer relief and support. Investigating the effectiveness of stress-relief by persuasive technology is relevant and in our eyes merits future research. The moment that we have succeeded in deploying effective stress-reducing technology, many eHealth users around the world will benefit and lead better lives.

Limitations

Conclusions on ECA research are specific to their task and context. Concerning the task and context that were used in our experimental set-up and that could have influenced our results; we separated learning content (left part of the screen) from supportive content (right part of the screen). In addition, as learning content we used a positive psychology intervention.

As supportive content we provided directions and gave positive feedback after a learning task was finalized by the user. This way we avoided direct distraction from the ECA toward the user. However, the visual presence of the ECA during task completion was not controlled for.

The supportive content could be controlled by the user by using the forward and backward buttons. This provided user control but deviated from other ECA set-ups that use vocal user input. Our intervention was a short-term, one-off intervention. It is not known how this can be translated to life interventions that typically span a period of 6–10 weeks and are used on a more frequent basis.

The manner we induced stress, the invalidated PacMan solely resulted in partial effects and is of course just an experimental representation of what chronic patients experience using eHealth solutions.

Directions for future work

Future research can be carried out by ECA's that are not visible during the episodes that the participants do their tasks, in order to avoid task inhibition. In addition, a stronger single dose or multiple doses of stress could be provided in order to prolong the stressful experience of participants. Then again this should be done cautiously in order not to overly frustrate participants. Furthermore, measurement of stress can be done by deploying smart devices such as wristbands (see e.g. Sevil et al., 2017) in order to track the stress' temporal dynamics. Subsequently, these measurements can be cross-validated with questionnaires on stress experiences as presented at the end of the experiment. With regards to the ECA's credibility and effectiveness, prior to the experiment the ECA can tell that there is a chance that there will be difficult episodes for the user. This can enlarge the credibility of the ECA, that can potentially be helpful at a later stage when the user truly experiences frustration and is subsequently supported. Last, we know what users are confronted with during the stressful PacMan and neutral PacMan pre-conditions. Without the need of measuring the user's emotions, the ECA can provide empathy towards the users who have just experienced the malfunctioning PacMan. All these possibilities mentioned, should be implemented with care. Within the ECA study field well-intending ECA's can unintentionally trigger amplification of stress, in case the ECA's implementation is not precisely right.

Ethics statement

This study was carried out in accordance with the recommendations BMS Ethics Committee of the university of Twente with informed consent from all subjects. All subjects gave informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the BMS Ethics Committee of the university of Twente.

References

Abdallah, C. G., & Geha, P. (2017). Chronic pain and chronic stress: two sides of the same coin? *Chronic Stress*, *1*, 2470547017704763.

Baylor, A. L. (2011). The design of motivational agents and avatars. *Educational Technology Research and Development*, *59*(2), 291-300.

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.

Bernieri, F. J. (1988). Coordinated movement and rapport in teacher-student interactions. Journal of Nonverbal behavior, 12 (2), 120–138.

Bickmore T. Relational agents for chronic disease self-management. In: Hayes BM, Aspray W, editors. Health Informatics: A Patient-Centered Approach to Diabetes. Cambridge, MA: MIT Press; 2010:181-204.

Bickmore, T. and Picard, R. 2005. Establishing and Maintaining Long-Term Human-Computer Relationships. ACM Transactions on Computer Human Interaction (ToCHI), 59(1): 21-49, 1–39. doi: 10.2190/ec.49.1.a

Blankendaal, R., Bosse, T., Gerritsen, C., De Jong, T., & De Man, J. (2015, May). Are aggressive agents as scary as aggressive humans? In Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems (pp. 553-561).

Brixey, J., & Novick, D. (2019). Building rapport with extraverted and introverted agents. In *Advanced Social Interaction with Agents* (pp. 3-13). Springer, Cham.

Cerekovic, A., Aran, O., and Gatica-Perez, D. (2014). "How do you like your virtual agent? Human-agent interaction experience through nonverbal features and personality traits," in Human Behavior Understanding. HBU 2014. Lecture Notes in Computer Science, eds H. S. Park, A. A. Salah, Y. J. Lee, L. P. Morency, Y. Sheikh, and R. Cucchiara (Cham: Springer), 8749.

Desmet, P. M., Porcelijn, R., & Van Dijk, M. B. (2007). Emotional design; application of a research-based design approach. *Knowledge, Technology & Policy*, *20*(3), 141.

DeVault, D., Artstein, R., Benn, G., Dey, T., Fast, E., Gainer, A., ... & Lucas, G. (2014, May). SimSensei Kiosk: A virtual human interviewer for healthcare decision support. In *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems* (pp. 1061-1068).

Drolet, A. L., & Morris, M. W. (2000). Rapport in conflict resolution: Accounting for how face-to-face contact fosters mutual cooperation in mixed-motive conflicts. Experimental Social Psychology, 36, 26–50.

Forlizzi, J., Zimmerman, J., Mancuso, V., & Kwak, S. (2007, August). How interface agents affect interaction between humans and computers. In *Proceedings of the 2007 conference on Designing pleasurable products and interfaces* (pp. 209-221).

Fu, F. L., Su, R. C., and Yu, S. C. (2009). EGameFlow: a scale to measure learners' enjoyment of e-learning games. Comput. Educ. 52, 101–112. doi: 10.1016/j. compedu.2008.07.004

Gratch, J., Wang, N., Gerten, J., Fast, E., and Duffy, R. (2007). "Creating rapport with virtual agents," in Intelligent Virtual Agents. IVA 2007. Lecture Notes in Computer Science, Vol. 4722, eds C. Pelachaud, J. C. Martin, E. André, G. Chollet, K. Karpouzis, and D. Pelé (Berlin: Springer).

Gratch, J., Kang, S. H., & Wang, N. (2013). Using social agents to explore theories of rapport and emotional resonance. Social Emotions in Nature and Artifact, 181, 181-197.

Hayes, S. C., Strosahl, K. D., and Wilson, K. G. (1999). Acceptance and Commitment Therapy. New York, NY: Guilford Press, 6.

Huygens MWJ, Vermeulen J, Swinkels ICS, Friele RD, van Schayck OCP, and de Witte LP. Expectations and needs of patients with a chronic disease toward self-management and eHealth for self-management purposes. BMC Health Serv Res. 2016; 16:232

Kelders SM, Bohlmeijer ET, van Gemert-Pijnen JE. Participants, usage, and use patterns of a web-based intervention for the prevention of depression within a randomized controlled trial. J Med Internet Res. 2013;15(8):e172. doi: 10.2196/jmir.2258. http://www.jmir.org/2013/8/e172/

Kron FW, Fetters MD, Scerbo MW, White CB, Lypson ML, Padilla MA, Gliva-McConvey GA, Belfore LA 2nd, West T, Wallace AM, Guetterman TC, Schleicher LS, Kennedy RA, Mangrulkar RS, Cleary JF, Marsella SC, Becker DM. Using a computer simulation for teaching communication skills: A blinded multisite mixed methods randomized controlled trial.
Patient Educ Couns. 2017 Apr;100(4):748-759. doi: 10.1016/j.pec.2016.10.024.

Leenen, L. A., Wijnen, B. F., de Kinderen, R. J., van Heugten, C. M., Evers, S. M., & Majoie, M. H. (2016). Are people with epilepsy using eHealth-tools? *Epilepsy & Behavior*, *64*, 268-272.

Legault L. (2017) Self-Determination Theory. In: Zeigler-Hill V., Shackelford T. (eds) Encyclopedia of Personality and Individual Differences. Springer, Cham

Lo, S. K., and Cheng, M. W. (2010). The effect of online agents on advertising effectiveness: the presence aspect. Manag. Rev. 29, 99–102.

Louwerse, M. M., Graesser, A. C., Lu, S., and Mitchell, H. H. (2005). Social cues in animated conversational agents. Appl. Cogn. Psychol. 19, 693–704. doi: 10.1002/acp.1117

Oinas-Kukkonen, H., & Harjumaa, M. (2009). Persuasive systems design: Key issues, process model, and system features. Communications of the Association for Information
Systems, 24 (1), 28.

Prendinger, H., Mori, J., Ishizuka, M.: Using human physiology to evaluate subtle expressivity of a virtual quizmaster in a mathematical game. International Journal of Human-Computer Studies 62(2) (2005) 231–245. doi:10.1016/j.ijhcs.2004.11.009

Renner, F., Schwarz, P., Peters, M. L., and Huibers, M. J. (2014). Effects of a best possible-
self-mental imagery exercise on mood and dysfunctional attitudes.
Psychiatry Res. 215, 105–110. doi: 10.1016/j.psychres.2013.10.033

Reuderink, B., Nijholt, A., & Poel, M. (2009, June). Affective Pac-Man: A frustrating game for brain-computer interface experiments. In *International conference on intelligent technologies for interactive entertainment* (pp. 221-227). Springer, Berlin, Heidelberg.

Rickenberg, R., & Reeves, B. (2000, April). The effects of animated characters on anxiety, task performance, and evaluations of user interfaces. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (pp. 49-56).

Rosenzweig, S., Greeson, J. M., Reibel, D. K., Green, J. S., Jasser, S. A., & Beasley, D. (2010). Mindfulness-based stress reduction for chronic pain conditions: variation in treatment outcomes and role of home meditation practice. *Journal of psychosomatic research*, *68*(1), 29-36.

Ruttkay, Z., Dormann, C., & Noot, H. (2004). Embodied conversational agents on a common ground: A framework for design and evaluation. In *From brows to trust: evaluating embodied conversational agents* (pp. 27-66). Kluwer.

Ryan, R. M., & Deci, E. L. (2003). On assimilating identities to the self: a self-determination theory perspective on internalization and integrity within cultures.

Sänger, J., Bechtold, L., Schoofs, D., Blaszkewicz, M., & Wascher, E. (2014). The influence of acute stress on attention mechanisms and its electrophysiological correlates. *Frontiers in behavioral neuroscience*, *8*, 353.

Sanghoon, A. L. B. D. W., & Roberto, P. E. S. (2005). The impact of frustration-mitigating messages delivered by an interface agent. Artificial intelligence in education: supporting learning through intelligent and socially informed technology, 125, 73

Scholten, M. R., Kelders, S., & Gemert-Pijnen, V. (2019). An empirical study of a pedagogical agent as an adjunct to an eHealth self-management intervention: what modalities does it need to successfully support and motivate users? *Frontiers in psychology*, *10*, 1063.

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), e383.

Schroeder, N. L., Adesope, O. O., and Gilbert, R. B. (2013). How effective are pedagogical agents for learning? A meta-analytic review. J. Educ. Comput. Res.

Sevil, M., Hajizadeh, I., Samadi, S., Feng, J., Lazaro, C., Frantz, N., ... & Cinar, A. (2017, May). Social and competition stress detection with wristband physiological signals. In *2017 IEEE 14th International Conference on Wearable and Implantable Body Sensor Networks (BSN)* (pp. 39-42). IEEE.

Sillice MA, Morokoff PJ, Ferszt G, Bickmore T, Bock BC, Lantini R, et al. Using Relational Agents to Promote Exercise and Sun Protection: Assessment of Participants' Experiences With Two Interventions. J Med Internet Res 2018 Feb 07;20(2):e48

Sinha, R., & Jastreboff, A. M. (2013). Stress as a common risk factor for obesity and addiction. *Biological psychiatry*, *73*(9), 827-835.

Steinmetz, J., & Pfattheicher, S. (2017). Beyond social facilitation: A review of the far-reaching effects of social attention. Social Cognition, 35(5), 585-599.

Tatar, D. (1997). Social and personal consequences of a preoccupied listener. Stanford, CA: Stanford
University

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*, 1-12.

Tsui, P., & Schultz, G. L. (1985). Failure of rapport: Why psychotherapeutic engagement fails in the treatment of Asian clients. American Journal of Orthopsychiatry,
55, 561–569.

Van Gemert-Pijnen, J. E. W. C., Kelders, S. M., Beerlage-de Jong, N., & Oinas-Kukkonen, H. (2018). Persuasive health technology. *eHealth research, theory and development, a multidisciplinary approach*, 344.

Vancampfort, D., Koyanagi, A., Ward, P. B., Veronese, N., Carvalho, A. F., Solmi, M., ... & Stubbs, B. (2017). Perceived stress and its relationship with chronic medical conditions and multimorbidity among 229,293 community-dwelling adults in 44 low- and middle-income countries. *American journal of epidemiology*, *186*(8), 979-989.

Veletsianos, G., and Russell, G. (2014). "Pedagogical Agents," in Handbook of Research on Educational Communications and Technology, 4th Edn, eds M. Spector, D. Merrill, J. Elen, and M. J. Bishop (Abingdon: Routledge), 759–769.

Verduyn, P., Van Mechelen, I., Tuerlinckx, F., Meers, K., & VanCoillie, H. (2009b). Intensity profiles of emotional experience over time. Cognition and Emotion, 23, 1427–1443.

Weiss, B., Wechsung, I., Kühnel, C., & Möller, S. (2015). Evaluating embodied conversational agents in multimodal interfaces. *Computational Cognitive Science*, *1*(1), 6.

Whalen, S. P., and Csikszentmihalyi, M. (1991). Putting Flow Theory into Educational Practice: The Key School's Flow Activities Room. Report to the Benton Center for Curriculum and Instruction. Chicago, IL: University of Chicago.

World Health Organization (2016)
Global diffusion of eHealth: making universal health coverage achievable. Report of the third global survey on eHealth. Geneva: World Health Organization; 2016. Licence: CC BY-NC-SA 3.0 IGO

Zaichkowsky, J. L. (1994). The personal involvement inventory: reduction, revision, and application to advertising. J. Advert. 23, 59–70. doi: 10.1080/ 00913367.1943.10673459

Zanbaka, C., Ulinski, A., Goolkasian, P., & Hodges, L. F. (2004, November). Effects of virtual human presence on task performance. In *Proc. International Conference on Artificial Reality and Telexistence 2004* (pp. 174-181).

Zhang, X., Guo, X., Lai, K. H., Guo, F., & Li, C. (2014). Understanding gender differences in m-health adoption: a modified theory of reasoned action model. *Telemedicine and e-Health*, *20*(1), 39-46.

Zimmerman, J., Ayoob, E., Forlizzi, J., & McQuaid, M. (2005). Putting a face on embodied interface agents.

Chapter 5

Synchronizing speech between a user and a Conversational Agent; can it establish rapport that is beneficial for eHealth interventions? An exploratory study.

Scholten MR, Kelders SM, Van Gemert-Pijnen JEWC
Synchronizing speech between a user and a Conversational Agent; can it establish rapport that is beneficial for eHealth interventions? An exploratory study.

Submitted

Abstract

Background: Self-guided eHealth, the use of technology to improve health, well-being and healthcare, provides autonomy to patients. However, as the usage statistics show, these type of solutions face elevated attrition rates. Embodied Conversational Agents ('robots on screen') as adjuncts to eHealth environments, have the capability to support patients during their endeavors. As such, their affective computing facilities can provide a potential remedy to the tight attrition issue. However, as an important pre-condition for ECA's to become effective support providers, they first have to build rapport with the user. Rapport building is an arduous task. It normally requires the deployment of a sophisticated and costly ECA that can respond to users in real-time, using both verbal and non-verbal communication channels.

Objective: The objective of this paper is to explore an alternative and novel way to build rapport; synchronous speech with a functionally modest, monologue-style ECA.

Methods: We set up an exploratory study during which we test a novel experimental task that serves to create speech synchrony. By means of a qualitative analysis we will test the synchronous speech task.

Results: The results show that users are fairly positive about speaking synchronously with the ECA. Nevertheless, users need to be priorly informed about the rhythm and pace of the ECA's speech.

Conclusion: Our explorative study results demonstrated promising effects from the synchronous speech task. Future studies can further test the enhanced synchronous speech task's practical value to ground ECA-originated support by building rapport with eHealth patients.

Keywords: eHealth, embodied conversation agent, rapport, synchrony, affective computing, HCI

Introduction

Worldwide, there is a growing demand for health services. Within the mental health domain, a disease such as major depression is the leading cause of years lived with disability. Within the physical health domain, chronical physical illnesses associated with an aging population equally lead to an increase of the demand for new types of health services. With this rise in demand, the future clinical workforce is expected to need the aid of smart technology; eHealth. eHealth has been defined in various ways. Within the context of this study, we follow the broad definition of van Gemert-Pijnen et al. (2018) “the use of technology to improve health, well-being and healthcare”.
An important subset of eHealth interventions consists of web-based interventions. According to Barak et. al (2009) a web-based intervention is defined as “a primarily self-guided intervention program that is executed by means of a prescriptive online program operated through a website and used by consumers seeking health- and mental health–related assistance.” Although the self-guidance has the advantage of providing autonomy to the patient, the usage statistics expose a salient issue with web-based interventions: their elevated attrition rates. Melville et al. (2010) reviewed internet-based treatment programs involving minimal therapist contact and their dropout rates. As they found, dropout ranged widely from 2% to 83%, with an overall average of 35%. In a similar vein, Buhrman et al. (2016) found considerable attrition levels ranging from 4% up to 54% within a review on internet interventions for chronic pain.
Undisputedly, web-based interventions are intended to be used from start to end. Hence, several studies have been examining root causes of early quitting and low usage. Relevant in this respect, Gemert-Pijnen et al. (2018) underscore the importance of the technology’s fit with the user and context. As the authors summarize, if the users feel like the eHealth technology does not match their personal needs and preferences, or cannot be embedded in their routines, it will not be used. So, are web-based interventions sufficiently attuned to user needs and the prevention of drop-out?
The literature suggests this is often not the case. As described in the scoped review of Scholten et al. (2017), users of self-guided eHealth interventions express they miss out on support, in a myriad of ways. First and straightforwardly, many patients experience a lack of *encouragement* after having fulfilled the tasks as demanded by the web-based intervention. In a broader health perspective, patients seek *acknowledgement* for daily issues they are struggling with, such as pain and sleeping difficulties. Obviously, the computer’s facilities for the detection of and providing support to these broader user needs are constrained. All the same, computers can

be endowed with user-oriented and supportive qualities (e.g., Scholten et al., 2020), according to the tenets of affective computing (Picard, 2003). By adding supportive functionality that is personal and visible (e.g., represented as an animated face and/or body) and natural language-based, a feeling of intimacy can be induced (Potdevin et al, 2020) amongst users. Such adjuncts are called Embodied Conversational Agents (ECA's). ECA's (also known as virtual agents, virtual humans, and relational agents) are animated, human-like figures simulating face-to-face conversation including verbal and nonverbal behavior (e.g., nodding and smiling).

1.1 Current evidence for ECA's

Functionalities of ECA's have been reviewed (e.g., Dehn and Van Mulken, 2000; Veletsianos and Russell, 2014) with the aim of finding evidence for what kind of ECA fits best within which context. As postulated by Dehn and Van Mulken (2000) and by a considerable number of later studies (e.g., Shamekhi et al., 2016, Baylor and Ryu, 2003) similarity between ECA and user (or user expectations) positively contributes to their approval. The similarity can have different origins. First, the similarity can be static, e.g., the user and ECA have the same gender, age or ethnicity. Next, the similarity can concern dynamical features such as similarity in the user's and ECA's behavior (e.g., same level of extraversion). Finally, contextual features are relevant, such as whether the ECA fulfills its role (e.g., student, tutor) comparable to how a typical human would carry it out. On a theoretical level, ECA-user similarity effects can be clarified by insights from social psychology. The similarity attraction theory (Byrne, 1997), the social identity theory (Tajfel, 1974) and the self-categorization theory (Turner and Reynolds, 2011) all support and explain the idea that people are attracted to, prefer and support relationships with similar others. Further building on their similarity potential, it has been suggested that ECA's can create a relationship with users. Apparently, their life-like appearance, their embodiment, exposing posture movements, eyes that are blinking, the visibility of a respiration rhythm, together with their capabilities to communicate via speech or text, instills a feeling of connection amongst users. Finally, attention from the ECA towards the user has been demonstrated to be an important feature. Heylen et al. (2002) proved that human-like gaze behavior of an ECA is evaluated more positively with regard to usability and involvement than more randomly determined gaze behavior. Comparable positive effects stemming from the ECA's gaze have been found by Bailenson et al. (2002).

1.2 Rapport building as a pivotal ECA concept

Within human-ECA studies, a universal concept for the productive ECA's attention as experienced by users, is *rapport*. Rapport has been studied across a range of scientific disciplines for its role in fostering emotional bonds and prosocial behavior. The role of rapport in stimulating effective social interaction is well established in the field of social psychological research. For example, rapport is argued to provide foundations for social engagement (Tatar, 1997), positive teacher–student interactions (Bernieri, 1988), success in negotiations (Drolet & Morris, 2000), psychotherapeutic effectiveness (Tsui and Schultz, 1985), and enhanced quality of child care (Burns, 1984). In their seminal article, Tickle-Degnen and Rosenthal (1990) equate rapport between humans with behaviors indicating positive emotions (e.g., head nods or smiles), mutual attentiveness (e.g., mutual gaze), and coordination (e.g., postural mimicry or synchronized movements) such as are expressed by the members of a well-functioning sports team in action. Likewise, within the human-ECA domain, agents have been created that effectively show attention and demonstrate positivity, leading to elevated levels of rapport as experienced by their users. In order to achieve this, ECA's effectively deploy two supplementary communication channels; the non-verbal and verbal channel.

First, ECA's can be endowed to timely respond to users in a non-verbal manner such as positive nodding and smiling. This type of responsive behavior is also known as non-verbal entrainment (Kenny et al., 2007) and can be utilized to stimulate the human interlocutor while speaking. Within an experimental human-ECA setting, Gratch et al., (2013) have successfully demonstrated how non-verbal entrainment behavior (i.e., positive ECA nodding) led to an elevated level of rapport amongst participants. However, later studies have i.e., argued (e.g., Weiss et al., 2015) that it is difficult for an ECA to show credible entrainment behavior during interactions with a longer time-span. The longer the interaction takes, the higher the chance that the ECA's timing will become faulty. As a result, the ECA's behavior will be evaluated as 'unreal' and disregarded by the user. As stated by Bickmore and Picard (2003): "it is an extremely challenging task to get the agent to maintain the illusion of human-like behavior over time; every aspect of the agent's appearance and verbal and non-verbal behavior must be correct or users will begin to discredit it."

In a pure verbal manner, an ECA can also create rapport, through autobiographical stories and small talk. Bickmore et al. (2010) have investigated the differential rapport building effects of ECA narratives either told by the ECA from the 1st-person perspective either as stories about a friend (3rd-person perspective). The purpose of their experiment was to assess the exercise (daily steps) stimulating behavior of an ECA taking on the role of exercise counselor for elderly people. As a mediating variable between the two ECA conditions and exercise behavior, the authors measured the level of user engagement. As the results showed, the rapport-building capabilities of the 1st-person condition was more effective than that of the 3rd-person condition. That is, the virtual exercise counselor speaking on its own behalf created

significantly higher system usage and user enjoyment. However, transfer effects of elevated user engagement on increased user walking behavior were not found. So, the results appeared to be somewhat contradictory. As an explanation the authors suggested that as a side-effect of the ECA's rapport-building and engagement techniques the pressure for the participants to exercise apparently diminished, ultimately leading to lower adherence to the exercise intervention.
Next to the existing non-verbal and verbal rapport building techniques readily deployed by ECA's, there may be another novel technique available; the creation of behavioral synchrony between ECA and user.

1.3 Synchrony informs rapport

Human communication studies have reported on synchronous movement rhythms leading to feelings of rapport, and resulting experiences of being part of one and the same social unity (Marsh et al., 2009; Tickle-Degnen & Rosenthal, 1990, Lakens and Stel, 2011). Moving in synchrony is argued to influence the degree to which individuals are perceived as a social unit (Marsh et al., 2009; Yzerbyt et al., 2004). But also individuals themselves report experiences of feeling to be part of one and the same team. On a neural level this is explained by pathways that code for both action and perception (Overy and Molnar-Szakacs, 2009) which causes blurring of the self and the other. As stated by Nozawa et al. (2019) practitioners and researchers in the field of education have recognized the effect of getting "in sync" with students, and have incorporated synchronous components such as singing and dancing as "warming-up" activities to facilitate better social and cognitive outcomes in the classroom. In a similar vein, there is evidence that a synchronized speech activity can create cooperation (Cummins, 2012) and rapport. In their study Reddish et al. (2013) showed that participants who were asked to read out loud a list of 54 English words in synchrony, cooperated more during a follow-up game. A second study yielding positive results on synchronized speech is provided by Harmon-Jones (2011). As experimental task she deployed three video's each containing multiple models that either sang or spoke out nonsense phrases during 75 seconds, in synchrony. Participants were assigned to either imitate, or merely listen to, the videos. Subsequently, participants rated their affiliation with the models. As Harmon-Jones found, both synchronous singing and speaking significantly increased the participants' affiliation with the models, with even higher scores for speaking than for singing. As social psychology principles and findings are often valid within human-ECA contexts, we expect that this synchronous speech task will equally positively impact user-ECA cooperation and rapport. To the best of our knowledge, such a synchronized speech task has not yet been proposed, let alone tested within human-ECA experiments. Within our experiment, we therefore aim to explore the user's view and acceptance of the novel experimental task that requires the user to speak out text lines simultaneously with the ECA. Secondly, we aim to find indications that this task creates rapport between study participants and ECA. In more detail, we intend to find marks for the most suitable condition. That is, we aim to examine whether phrases that contain information on teamwork and that are spoken out synchronously with the ECA will create an elevated level of rapport. In order to do so, we will apply a qualitative method for assessing the user's opinion on receiving motivational instructions from the ECA. Along these lines, we keep our ultimate objective in mind; a rapport-building ECA could become a remedy to non-adherence in future web-based interventions.

Methods

2.1 Design

To answer the research questions, an experiment was set up. Participants were asked to take part in an online course on positive psychology. At the start of the experiment, the participants were welcomed by Brian, the ECA. Brian was made in Voki® and presented on the right side of a WordPress webpage (version 4.9.7). Brian started by shortly introducing himself and explaining his role as support provider during the experiment. Then Brian asked the participants to speak out four phrases, that were displayed on the left side of the screen as part of a PowerPoint® presentation, see Figure 1 below.

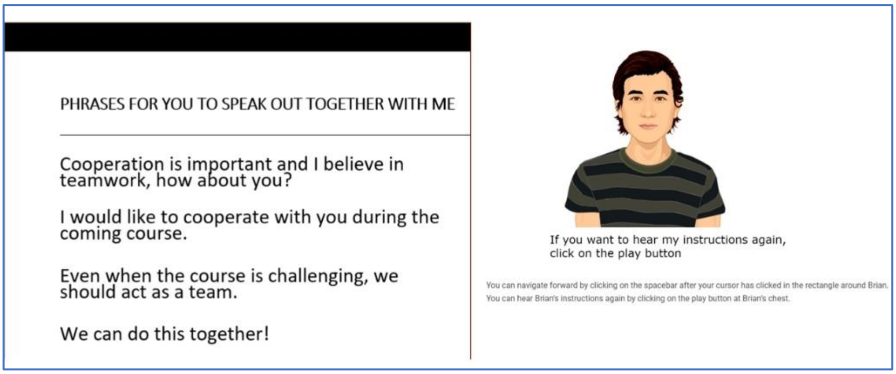


Figure 1. Webpage that displays four phrases on the left side and Brian on the right side.

Dependent on the speech condition the participant was part of, the participant either spoke out the phrases solely or synchronously with the ECA. Subsequently, the participants were asked to start with the online course. During the online course, Brian remained visibly present to support and guide the participants. The experiment took approximately 18 minutes for the participants to accomplish. See Figure 2 below for the sequence order of the experimental events.

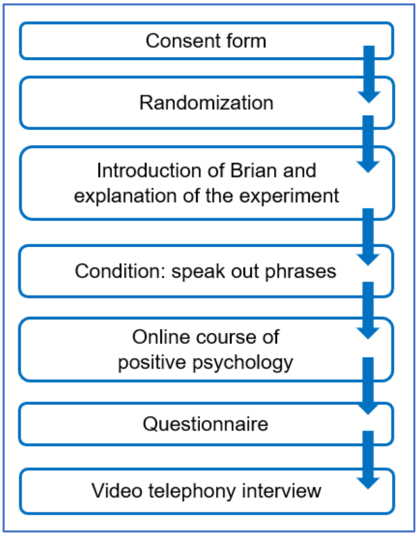


Figure 2: Flow chart of activities carried out by participants during the experiment

2.2 Speech conditions

Three different speech conditions were offered, all containing an equal number of 39 words put within 4 phrases;

Condition 1 (phrases about teamwork, synchronous speech). Within this condition the following phrases were presented visually to the study participant: *Cooperation is important and I believe in teamwork, how about you? I would like to cooperate with you during the coming course. Even when the course is challenging, we should act as a team. We can do this together!* For this condition the user was asked to speak the phrases aloud in the same rhythm as the ECA Brian did.

Condition 2 (neutral phrases not related to teamwork, synchronous speech). The following phrases were presented visually to the study participant: *Computers can be used to carry out various tasks for us. The personal computer has been introduced on august 2, 1981 by IBM. Next to Windows computers, there are Apple machines. Smartphones are both considered as telephones and computers.* Note that the phrases contained neutral information about computers and had the same length as the phrases in condition 1. For this condition the user was asked to speak the phrases aloud in the same rhythm as the ECA Brian did.

Condition 3 (neutral phrases not related to teamwork, non-synchronous speech, control condition). The same phrases as for condition 2 were presented visually to

the study participant. For this condition, the user was asked to speak the phrases aloud, by herself/ himself.
After the explanation of the speech procedure, ECA started a count-down and the participants started to speak. Then, Brian thanked the participants and the participants were asked to click through, in order to start with the online course. To these 3 combinations participants were randomized, see Figure 3 below. One participant was excluded from the analysis, after it became apparent that she had encountered too many technical difficulties.

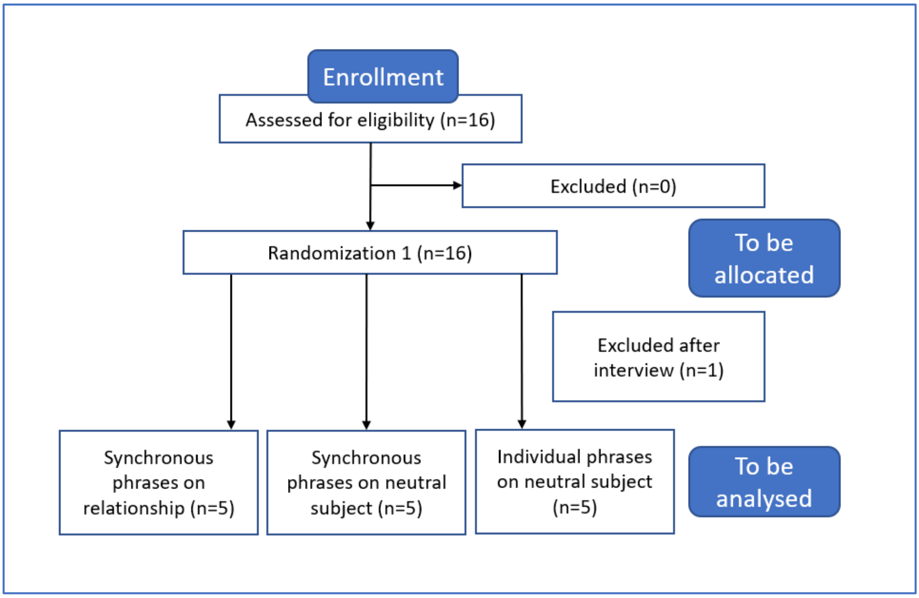


Figure 3. CONSORT Flow diagram of the experiment

2.3 Online course

The online course was run on a WordPress website (version 4.9.7) that contained the eHealth intervention on the left side of the webpage. The eHealth intervention was a PowerPoint® presentation with psycho-education material on positive psychology. The goal of the eHealth psycho-education intervention was to make users knowledgeable about positive psychology. Positive psychology focuses on the abilities of people and their potential to flourish. Several treatments against depression are based on positive psychology principles (Hayes et al., 1999). In addition, positive psychology and happiness are subjects that are of general human interest. As we reasoned, this topic would contribute to engage participants for our experiment. The self-guided eHealth intervention contained a combination of theory and exercises, including the renowned “three good things exercise” and “best possible self-exercise” (Renner et al., 2014).

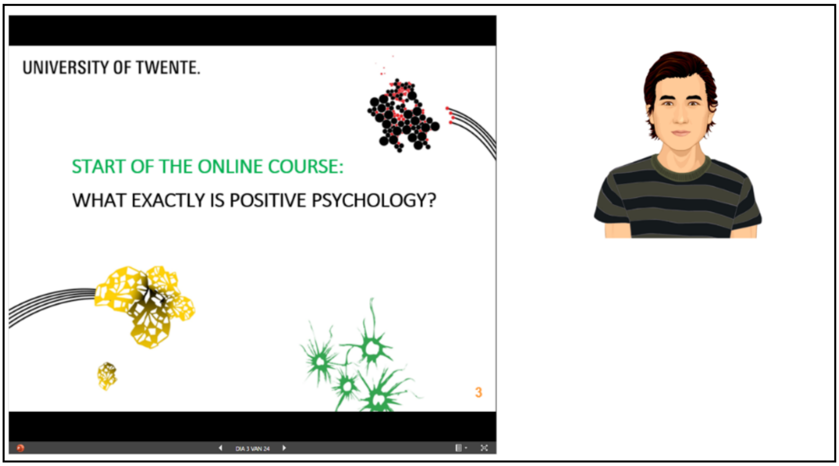


Figure 4 The eHealth psycho-education intervention. On the left side of the webpage the psycho-educational content is displayed, on the right side the motivating ECA Brian

As Figure 4 displays, user guidance and support were provided on the right side of the webpage by an ECA. The ECA condition was created through the Voki® application. The monologue-style ECA represented a virtual male in between 20 and 30 years of age, with Caucasian looks, acting as an informal (i.e., not medical) support provider. The user was asked to click on the ECA for the next voice segment to be spoken. The ECA showed lip synchronization and animation properties such as eyes blinking and chest breathing. Furthermore, the ECA's line of sight followed the cursor movements of the user. The ECA expressed guidance conveyed in English. After each piece of guidance and information, the ECA asked the user “Please click on the button to proceed.” The ECA’s displayed a combination of task-related support (e.g., “within this experiment you will read about positive psychology and you will do some exercises”) and motivational support (e.g., “So, let’s practice!”).

2.4 Recruitment of Participants

We recruited bachelor and master psychology and communication students at the University of Twente. As an inclusion criterion we set proficiency in English. As an exclusion criterion we set participation in a previous study with the ECA. The study protocol was reviewed and approved by the Twente Institutional Review Board with number 201009. In total 16 participants were included. Participants were on average 19.63 years of age and represented 4 nationalities of which Dutch (56.3%) and German (31.3%) were most prominent. Other nationalities were Lithuanian and

Spanish, all 6.3.%. 14 participants were female (87.5%), 2 participants were male (12.5%).

2.5 Procedure and Outcome Measures

At the end of the experiment, participants filled out a questionnaire in the Qualtrics® system. This questionnaire contained 10 questions, of which 4 were closed and 6 were open. The topic of the questions were the course as a whole, the role of Brian and the synchronous speech task. The questions were self-developed. Three out of the four open questions were about the relationship with Brian, one question was about the usefulness of the course. These questions were posed as a warming-up for the interview immediately thereafter. After filling in the questions, the users were asked to e-mail the researcher for the videotelephony interview to start. The interview was done through Microsoft TEAMS®. The objective of the interview was to find strengths and weaknesses of the ECA and to gather experiences on the user task of speaking aloud. Furthermore, the overall view of the participants on the online course was assembled. We asked the following questions during the interview:

- What was your experience on the online course?
- What was your experience on the task of speaking aloud?
- What is your opinion on Brian (the ECA)?
- Comparing this course with other online courses you have followed since the outbreak of Corona, what stands out for you?
- Did you experience a sense of *cooperation* with Brian?
- Did you experience a sense of *trust* with Brian?
- Did you experience a sense of *unity* of being part of the same team as Brian?
- Did you experience a sense of a *bond* with Brian?
- Do you have any other remarks?

After the interview, the participants were rewarded with a course credit

2.6 Data Analysis

On completion of the interviews, the interviewer transcribed the audio files verbatim. The content of the transcribed files was analyzed and subsequently grouped in themes together with the results of the open questions within the questionnaire using grounded theory (Thornberg et al., 2014). The main themes used for the analysis were defined *before* data collection, according to the objective of the experiment; the user's opinion about the speaking aloud task, about the ECA and finally about the online course. As most of the participant's input was provided on the ECA, we inductively decided to split the ECA main theme into a strength and weakness subtheme. Further subdivisions on themes were made inductively *after* data collection, based on the input data provided both within the questionnaire and during the interview. In detail, six steps were carried out performing our thematic analysis, following the descriptions of Braun & Clarke (2006). First, data gathered were familiarized by re-reading the transcripts for at least three times. Secondly, significant responses that directed to the research questions were highlighted and initial labels were attached. During the third phase, common labels abstracted from all the transcriptions were clustered together to form initial themes. Fourth, initial themes were reviewed to filter out instances that had minimal references. The filtering made use of categorizations that were based on the research questions and the literature as referred to within the introduction of this manuscript. Next, we put themes together that had a close semantic resemblance. The fifth stage involved the process of describing and labelling the final themes. Lastly, a report was written to describe the results and discussion and implemented in this manuscript.

3. Results

3.1 Qualitative results

Four main themes were applied deductively, in addition sub-themes emerged from the data, see **Table 1** below.

Table 1. Thematical topics found

Theme 1: The user experience of the task of speaking aloud	**Number of participants**
Positive	3 (spontaneous)
Mixed, but predominantly positive	9 (spontaneous)
Negative	3 (spontaneous)
The speech rhythm was a stumbling block	6 (spontaneous)
Theme 2: strengths of the ECA	**Number of participants**
Effect on participant: positive user experience	12 (spontaneous)
Effect on participant: autonomy during course	4 (spontaneous)
Effect on participant: clear structure of the course	4 (spontaneous)
ECA feature: the voice	5 (spontaneous)
ECA feature: the gazing behavior	2 (spontaneous)
Relational quality: cooperation in the sense of following	8 (prompted on cooperation)
Relational quality: trust on a rational level	4 (prompted on trust)
Theme 3: weaknesses of the ECA	**Number of participants**
One-sided communication	6 (spontaneous)
Possibility of bond	12 negative, 3 mixed (prompted)
Team experience	15 negative (prompted)
Theme 4: The experience on the course	**Number of participants**
Positive	12 (spontaneous)
Neutral	2 (spontaneous)
Negative	1 (spontaneous)

3.1.1. Theme 1: The mixed user experience on the task of speaking aloud

As our first theme, we asked participants about their experiences when carrying out the speak aloud task. The opinions varied. Some (3 out of 15) were straightforwardly **positive**; *I liked it. With these kinds of exercises, I'm quite an introvert, so speaking out loud gives another dimension to the course. [participant #13]. It went well. The virtual assistant was speaking slow and understandably. [participant #11].* Others (9 out of 15) were **mixed, but predominantly positive**; *Weird but also normal. I don't know how to explain. It is weird to speak it out loud with a virtual person, but I was more motivated to speak it out loud (an also in the same speed as Brian) than when they ask me to do it as a standard audio-recording. [participant #1].* Finally, some (3 out of 15) were **negative**; *I did not like this, it felt weird doing so and I do not really understand what the use of could be either. [participant #8] As very uncomfortable, but less uncomfortable than with a teacher. [participant #12].* Furthermore, following the speech rhythm of the ECA appeared to be a **stumbling block** as was spontaneously mentioned by the participants*; It was fine, although sometimes Brian's way of speaking was a little hard to follow. The breaks were quite irregularly timed. It was unfamiliar to do so and it was a little hard to try to talk in the same speech as Brian did because he pronounced some words or sentences different from how I would usually do. [participant #9].*

3.1.2 Theme 2: strengths of the ECA

Overall, the ECA provided a **positive user experience** on most (12 out of 15) participants: *I think that it was well done and the coach was very well-animated. I enjoyed the experience [participant #10].* With respect to other **effects** on participants; the ECA supported quite a few participants (4 out of 15) in their **autonomy**; *I knew that Brian was not actually waiting for me to complete my tasks so I did not feel pressured in any way and could fully concentrate on the tasks [participant #7].* Another ECA effect appeared that Brian provided a **clear structure** (on 4 out of 15 participants) with regards to following the course*… the instructions are so much clearer and I don't get distracted. It is easy to just scan a text and miss something important. It is also more motivating because you have to listen and because Brian is moving, I was really focused on him (does his mouth move at the right time when he spoke for example). [participant #3].* Some participants explained in detail which of the ECA's **features** created this positive user experience. The **voice** was specifically by quite a few (5 out of 15) participants. *He has a calm and relaxing voice. Even though he was animated I felt like someone was explaining things to me. [participant #13].* In addition, the ECA's **gazing behavior** was regularly referred to; *I really liked his eyes! The way they moved when you moved your cursor around him. [participant #9].* Finally, the participants replied on questions about the **relational qualities of the ECA**. When being asked whether the participants **cooperated** with the ECA, they predominantly (8 out of 15) indicated they followed the ECA's instructions; *I was doing something, he was doing something, and I was following it was more like a teacher student relationship than cooperation. [participant #15].* When we asked the participants whether they **trusted** the ECA, 4 participants answered by stating they trusted the ECA on a rational, academic level; *We're at the university and I think anything affiliated to the university is considered to be, a valid sort of credible source. So, considering Brian comes from that source, there is automatically an authoritative trust that has been implied. [participant #12].*

3.1.3 Theme 3: weaknesses of the ECA

Many (6 out of 15) participants mentioned spontaneously the **one-sided communication** as a drawback of the ECA; *A downside is of course that you cannot* *get any additional information or explanation besides what Brian is programmed to tell you. [participant #9].* When being asked whether the participants experienced an emotional **bond** with the ECA, this was mostly (12 out of 15 participants) denied and as a substantiation several types of answers were given. **If there would be a bond, it would be uni-lateral.** *What kind of bond can he have with me? [participant #7]. I was listening to him, but I didn't have the feeling that he knows me and accepts me. [participant #1].* **A bond might be created, but it would have a different nature than a human-human bond**; *Well, it's still quite inhumane. So that would definitely make it difficult, but I think you could form a relationship that is more similar to a relationship to a cartoon character. You can still form relationships because they still display quite humanlike features, but it's nothing like a relationship to real humans, I guess. [participant #11].* **A bond might be created, but more interaction time would be needed**; *Well, yeah, we do have a relationship, but I don't think that 20 minutes with an avatar is enough to form a bond, to be honest. [participant #14].* **A bond was not created as personal information could not be exchanged**; *I got, as I said earlier on the academic level, but not on a personal personal level. [participant #7].* **Power distance/ dependency on the ECA hindered a bond**; *… and I can't interfere in this whole process because I'm dependent on Brian. [participant #8]* Furthermore, when asked they had **a team experience** together with the ECA, this was denied by all the participants. No experience of belonging to the same team; *Because he was guiding me through it, I guess he was not like on the same level, like I was against guess because I had the feeling that he was like a step above me, like in a teacher's role because he was explaining it to me. [participant #15]*

3.1.4 Theme 4: the content of the course

Generally (12 out of 15), the participants referred to the content of the positive psychology course as **interesting**; *I really liked it, this topic really matches my mindset in daily life about getting the most out of your life/yourself and being positive. I probably will continue with these exercises. [participant #13].* However, one participant reported a **negative experience**; *I liked the first exercise. The second one was actually quite stressful for me. I'm a student so I don't really know how I want my future to look like. It definitely didn't give me a happy, positive feeling. [participant #14]*

3.2. Quantitative results

Although the qualitative approach was leading, we also conducted a small quantitative analysis on the four closed questions we posed on a 1-5 scale. Below, **in Table 2**, the means, 95% Confidence Interval and SD values of the outcome variables are presented. Note that these results are merely indicative.

Table 2. Mean scores and standard deviation for the three conditions

	Condition 1 (n=7)	Condition 2 (n=9)	Condition 3 (n=7)
Useful Course (1-5)	4.4; 3.9-4.9; 0.2	4.0; 3.6-4.4; 0.2	4.6; 4.1-5.1; 0.2
R: Useful advice of Brian (1-5)	4.4; 3.9-4.9; 0.2	4.6; 4.1-5.0; 0.2	4.6; 4.1-5.1; 0.2
R: Motivating Brian (1-5)	3.9; 3.3-4.4; 0.2	4.0; 3.5-4.5; 0.2	4.4; 3.9-4.9; 0.2
R: Feel accepted by Brian (1-5)	3.7; 3.4-4.0; 0.1	4.0; 3.8-4.2; 0.1	3.9; 3.6-4.1; 0.1

Table 2 shows the four measured variables. The ‘R’ indicates the rapport-related questions.

As Table 2 shows, both the course and the ECA were evaluated positively (all mean scores of 3.7 or higher). Apparently, Brian was valued for its advice and somewhat less for its motivational qualities and its ability to make the participants feel accepted. Remarkably, for all four questions, the scores for condition 1 were lower than for condition 3, the control condition. However, the one-way ANOVA and LSD post-hoc test for pairwise comparisons did not demonstrate any significant differences.

4. Discussion

4.1 The user experiences with the speech task, ECA, and online course

With respect to the user experiences, we applied as four main themes: the mixed user experience on the task of speaking aloud, the strengths of the ECA, the weaknesses of the ECA, and finally the content of the course. With respect to the strengths and weaknesses of the ECA we applied as subcategories ECA features: what aspects of the ECA’s look and feel were brought up by the study participants, ECA effect: what effect did the ECA have on the study participants, and finally the ECA’s relational qualities/ short-comings: what relationship building features (and the lack of them) were mentioned by the study participants.

4.2 The experimental task of speaking aloud needs finetuning

The interview demonstrated that the task of speaking aloud was a positive event for most of the participants. Nonetheless, a number of participants indicated that the way the ECA would pronounce the phrases was unpredictable. Therefore, the pace and rhythm of the ECA’s speech became somewhat hard to synchronize. Furthermore, for the participants the rationale of speaking out loud was missing. These findings were in line with the results of our exploratory quantitative analysis.

This analysis provided no indications that differential levels of rapport were created as a result of synchrony or content manipulations. As an explanation, we reckon that our synchronous speech task has not fully matured yet and needs further refinement. We foresee that several improvements are possible. First, the participants could be given the opportunity **to warm up and practice**. When they were asked to speak synchronously with the ECA, they hadn't heard the ECA speaking out the phrases before and it was immediately for real. If the participants would once have heard the ECA's speech before speaking it out, they could get acquainted to the ECA's speaking rhythm. As a result, their speech rhythm would be better attuned to the ECA's speech rhythm. Consequently, the participants would have experienced a more successful completion of the task, which would likely have made the task more effective. An earlier study on synchronous speech involving two humans (Cummins, 2011) concluded that -remarkably-, practice at the task did not lead to markedly better performance. As the author stated, it was never the case that one speaker was consistently leading and the other lagging behind. Rather, the speech of the two speakers seemed to fuse, with only minimal leading or lagging, and no consistent leader. In contrast, in our human-ECA set-up, this fusion was not reported on. This absence may be explained by the fact that the pre-recorded speech of the ECA was not adapting to the participant's speech. In other words, it all came down to the participants to create synchrony, which is known to be a harder task. In fact, as Konvalinka et. al (2010) found, participants can effectively synchronize with an unresponsive computer. However, according to the authors, performance deteriorates if the computer additionally acts irregularly, as was the case according to our participants. Second, **the rationale for doing the task** should have been accounted for by the ECA more explicitly. As stressed by Klein et al. (2006), explanation of events through the use of mental models increases task performance when users interact with computers (see e.g., Ziefle and Bay, 2004; Xie et al, 2017). For a future experiment, we therefore foresee a more elaborate introduction such as; "I would like to bond with you. As means, I would like you to speak out several phrases aloud together with me, as a to warm up for the experiment and to start cooperating with you." Likely, this would have provided sense to this activity. Furthermore, it would have led to better integration with the other cooperative activities with the ECA later in the experiment. Note that our initial decision not to disclose the rationale for the task was to prevent any participant's bias towards cooperation with the ECA. Although we still consider that as a valuable argument, we now reckon this is of lesser importance than clearly explaining the purpose of the speech task. As a third improvement, it is important to keep in mind that, in daily practice, speaking and listening are **purposeful social activities**. Stated differently; speech is normally heard by an audience that generally will respond in one way or another. See e.g., Prinz (1990) "the ability to produce language is of no use when there is no one to listen, and the ability to understand language is of no use when there is no one to produce it" (pp. 177). In a similar vein, Bickmore and Cassell (2005) underscore the importance of an ECA that properly responds to the user. Within our experimental set-up, the ECA had no listening abilities and the participants only received verbal feedback by the ECA after the participants had clicked on the forward button ("thank you for speaking out the phrases"). So, in short, the participants lacked the experience of a listening ECA. Consequently, as an improved set-up, the participant's speech could be recorded and subsequently played out by the ECA in order to indicate hearing of the participant's speech; "You spoke out the following phrases:' (playing recorded phrases). Thank you for doing so."

4.3 The ECA's relational qualities and shortcomings

With respect to its features, the ECA's voice and gazing behavior were mentioned by the study participants as positive. This is in accordance with earlier studies that reported on voice (e.g., Scholten et al., 2017) and gaze (e.g., Heylen et al., 2002). In addition, participants generally asserted that the content of the course was interesting and relevant. However, during the interview the participants elaborately addressed the ECA's relational qualities and shortcomings. On the one hand, the ECA managed to support the study participants in their autonomy when following the course. Moreover, participants reported they felt at ease with the ECA, meaning that the ECA didn't put any sort of pressure on them. In addition, the ECA provided structure to the course. It created a more vivid user experience compared to mere text reading as some participants spontaneously asserted. On the other hand, users expressed the urge to verbally interact, a need the ECA obviously could not fulfill. As participants were literally missing something essentially reciprocal, this one-sided communication had an effect on their experienced relationship with the ECA. Not surprisingly therefore, participants mentioned as relational short-coming that they basically had no choice than to follow the ECA's instructions. If there was something like a relationship, it was framed as a 'dependent' and a 'teacher-student' relationship, devoid of involvement from the ECA's side.

4.4 Ways to improve the ECA's relational shortcomings

Interestingly, participants referred to examples from the film industry such as human actors and cartoon figures to point out that a true relationship with the ECA could not be created, as by definition. However, contrary to these broadcasting examples, a unique ECA feature is its one-on-one relationship with its interlocutor. This one-on-one relationship can be utilized for the revelation of the ECA's involvement. As previous studies have demonstrated (e.g., Bickmore et al, 2010), an ECA that exposes itself to an eHealth user by means of an autobiographical story creates rapport. Note that within our experimental set-up the ECA only briefly introduced itself with its name and purpose ("I am Brian, your virtual coach. I am here to guide you through the training'). With hindsight and based on the participant's feedback that the ECA was poor in involvement, we could have equipped the ECA with a more disclosive autobiographical story. We think of something like "I was made in 2019 at the university of Twente. Since then, I have been deployed in several experiments to support study participants. Now I am here to guide you through this course. I expect I can help you by providing you some extra information and guidance. The scientists who created me gave me a name and you can call me Brian." This way the ECA would move away from the instructor role towards more of a companion-like role, take away social distance and create rapport.

5. Limitations

Conclusions on ECA research are specific to their task and context. Concerning the task and context that were used in our experimental set-up and that could have influenced our results; we separated learning content (left part of the screen) from supportive content (right part of the screen). In addition, as learning content we used a positive psychology intervention.

As supportive content we provided directions and gave positive feedback after a learning task was finalized by the user. This way we avoided direct distraction from the ECA toward the user.

6. Conclusions and future work

We conducted a qualitative explorative study on a novel experimental task to create rapport between ECA and user. Participants provided positive feedback on the ECA's supportive role leading to positive effects on their experiences of autonomy. However, instilling a feeling of bonding or a team experience could not be accomplished by our ECA. Moreover, as our study showed, the synchronous task will have to be improved in order to become practical and verifiable. Future studies with an enhanced version of the synchronous speech task can therefore truly explore its potential. This way, we can assess whether speaking in unison with an ECA is not just to be taken literally, but can also deliver practical results. If deployment of the synchronous speech task will ultimately provide evidence for rapport building, it will have been substantiated that speaking in unison with an ECA is not restricted to figure of speech. Moreover, an important hurdle will have been taken for supportive ECA's as valuable adjuncts to web-based interventions. Then again, we cannot of course not expect that these ECA's will address user needs as well as human caregivers. Nonetheless, it is realistic to expect that such ECA's will perform better than the current baseline. This baseline is set by web-based interventions that face considerable attrition levels, partly due to not sufficiently meeting human needs that naturally and understandably go along with health issues.

Acknowledgements

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of interest

On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

Bailenson, J.N., Beall, A.C., Loomis, J., Blascovich, J., Turk, M.: Transformed Social Interaction: Decoupling Representation from Behavior and Form in Collaborative Virtual
Environments. PRESENCE: Teleoperators and Virtual Environments, 13 (2002) 428-441

Barak, A., Klein, B., & Proudfoot, J. G. (2009). Defining internet-supported therapeutic interventions. Annals of behavioral medicine, 38(1), 4-17.

Baylor A. L., Ryu J. 2003. The effects of image and animation in enhancing pedagogical agent persona. J. Educ. Comput. Res. 28, 373–395

Baylor, A. L. (2011). The design of motivational agents and avatars. *Educational Technology Research and Development*, *59*(2), 291-300.

Beall, A. C., Bailenson, J.N., Loomis, J., Blascovich, J., Rex, C.: Non-Zero-Sum Mutual Gaze in Collaborative Virtual Environments. Proceedings of HCI International, Crete (2003).

Bernieri, F. J. (1988). Coordinated movement and rapport in teacher-student interactions. Journal of Nonverbal behavior , 12 (2), 120–138.

Beskow, J., & McGlashan, S. (1997). *Olga - A Conversational Agent with Gestures.* Paper presented at the IJCAI Workshop on Animated Interface Agents.

Bickmore, T., & Cassell, J. (2005). Social dialogue with embodied conversational agents. In Advances in natural multimodal dialogue systems (pp. 23-54). Springer, Dordrecht.

Bickmore, T., Schulman, D., & Yin, L. (2010). Maintaining engagement in long-term interventions with relational agents. *Applied Artificial Intelligence*, *24*(6), 648-666.

Bickmore, T., & Picard, R. (2003, April). Subtle expressivity by relational agents. In Proceedings of the CHI 2003 Workshop on Subtle Expressivity for Characters and Robots.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative research in psychology, 3(2), 77-101.

Burns, M. (1984). Rapport and relationships: The basis of child care . Journal of Child Care , 2 , 47–57 .

Byrne, D. (1997), “An overview (and underview) of research and theory within the attraction paradigm”, Journal of Social and Personal Relationships, Vol. 14 No. 3, pp. 417-431.

Cassell, J., Pelachaud, C., Badler, N., Steedman, M., Achorn, B., Becket, T., et al. (1994). *Animated conversation: rule-based generation of facial expression, gesture & spoken* 43 *intonation for multiple conversational agents*. Paper presented at the Proceedings of the 21st annual conference on Computer graphics and interactive techniques.

Cassell, J., & Thorisson, K. R. (1999). The Power of a Nod and a Glance: Envelope vs. Emotional Feedback in Animated Conversational Agents. *Applied Artificial Intelligence, 13*, 519-538.

Cummins, F. (2011). Periodic and aperiodic synchronization in skilled action. Frontiers in Human Neuroscience, 5, 170.

Dehn, D. M., & Van Mulken, S. (2000). The impact of animated interface agents: a review of empirical research. *International journal of human-computer studies*, *52*(1), 1-22.

Drolet, A. L., & Morris, M. W. (2000). Rapport in conflict resolution: Accounting for how face-to-face contact fosters mutual cooperation in mixed-motive conflicts . Experimental Social Psychology , 36 , 26–50 .

Gratch, J., Kang, S. H., & Wang, N. (2013). Using social agents to explore theories of rapport and emotional resonance. *Social emotions in nature and artifact*, *181*, 181-197.

Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (1999). Acceptance and commitment therapy (p. 6). New York: Guilford press.

Harmon-Jones, C. K. (2011). Does musical behavior promote affiliation?. Texas A&M University.

Heylen, D., Es, I. van, Nijholt, A., Dijk, B.: Experimenting with the gaze of a conversational
agent, in: van Kuppevelt, J., Dybkjaer, L., Bernsen, N. (Eds.), Proceedings of the International CLASS Workshop on Natural, Intelligent and Effective Interaction with Multimodal Dialogue Systems. Kluwer Academic, New York (2002)

Kenny, P., Hartholt, A., Gratch, J., Swartout, W., Traum, D., Marsella, S., & Piepol, D. (2007, November). Building interactive virtual humans for training environments. In Proceedings of i/itsec (Vol. 174, pp. 911-916).

Klein, G., Moon, B., & Hoffman, R. R. (2006). Making sense of sensemaking 1: Alternative perspectives. IEEE intelligent systems, 21(4), 70-73.

Konvalinka, I., Vuust, P., Roepstorff, A., & Frith, C. D. (2010). Follow you, follow me: continuous mutual prediction and adaptation in joint tapping. Quarterly journal of experimental psychology, 63(11), 2220-2230.

Lakens, D., & Stel, M. (2011). If they move in sync, they must feel in sync: Movement synchrony leads to attributions of rapport and entitativity. Social Cognition, 29(1), 1-14.

Marsh, K. L., Johnston, L., Richardson, M. J., & Schmidt, R. C. (2009). Toward a radically embodied, embedded social psychology. European Journal of Social Psychology,
39, 1217-1225.

Melville, Katherine M., Casey, Leanne M., & Kavanagh, David J. (2010) Dropout from internet-based treatment for psychological disorders. British Journal of Clinical Psychology, 49(4), pp. 455-471.

Nozawa, Takayuki, et al. "Prior physical synchrony enhances rapport and inter-brain synchronization during subsequent educational communication." Scientific reports 9.1 (2019): 1-13.

Overy, K., & Molnar-Szakacs, I. (2009). Being together in time: Musical experience and the mirror neuron system. Music Perception, 26(5), 489-504.

Picard, R. W. (2003). Affective computing: challenges. International Journal of Human-Computer Studies, 59(1-2), 55-64.

Potdevin, D., Clavel, C., & Sabouret, N. (2021). Virtual intimacy in human-embodied conversational agent interactions: the influence of multimodality on its perception. Journal on Multimodal User Interfaces, 15(1), 25-43.

Prinz, W. (1990). A common coding approach to perception and action. In O. Neumann & W. Prinz (Eds.), Relationships between perception and action (pp. 167-201). Berlin, Germany: Springer-Verlag

Rabinowitch, T.-C. & Meltzoff, A. N. Synchronized movement experience enhances peer cooperation in preschool children. J Exp Child Psychol 160, 21–32 (2017).

Ramseyer, F. & Tschacher, W. Nonverbal synchrony of head- and body-movement in psychotherapy: different signals have different associations with outcome. Front Psychol 5, 979 (2014).

Reddish, P., Fischer, R., & Bulbulia, J. (2013). Let’s dance together: synchrony, shared intentionality and cooperation. *PloS one*, *8*(8), e71182.

Renner, F., Schwarz, P., Peters, M. L., & Huibers, M. J. (2014). Effects of a best-possible-self mental imagery exercise on mood and dysfunctional attitudes. Psychiatry research, 215(1), 105-110.

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), e383.

Scholten, M. R., Kelders, S. M., & van Gemert-Pijnen, J. (2020). Applying an Agent-based Model to Simulate Just-In-Time Support for Keeping Users of eLearning Courses Motivated. Nonlinear Dynamics, Psychology, and Life Sciences, 24(4), 403-429.

Shamekhi, A., Czerwinski, M., Mark, G., Novotny, M., & Bennett, G. A. (2016, September). An exploratory study toward the preferred conversational style for compatible virtual agents. In *International Conference on Intelligent Virtual Agents* (pp. 40-50). Springer, Cham.

Tatar, D. (1997). Social and personal consequences of a preoccupied listener . Stanford, CA : Stanford University

Tajfel, H. (1974), “Social identity and intergroup behavior”, Information (International Social Science
Council), Vol. 13 No. 2, pp. 65-93.

Thornberg, R., Perhamus, L., & Charmaz, K. (2014). Grounded theory. Handbook of research methods in early childhood education: Research methodologies, 1, 405-439.

Tickle-Degnen, L., & Rosenthal, R. (1990). The Nature of Rapport and Its Nonverbal Correlates. *Psychological Inquiry, 1*(4), 285-293.

Tsui, P., & Schultz, G. L. (1985). Failure of rapport: Why psychotheraputic engagement fails in the treatment of Asian clients . American Journal of Orthopsychiatry , 55 , 561–569

Turner, J.C. and Reynolds, K.J. (2011), Self-categorization Theory, Handbook of Theories in Social
Psychology, 1st ed., pp. 399-417.

van Gemert-Pijnen, L. J., Kip, H., Kelders, S. M., & Sanderman, R. (2018). Introducing ehealth. In eHealth research, theory and development: a multi-disciplinary approach (pp. 3-26). Routledge.

Van Pinxteren, M. M., Pluymaekers, M., & Lemmink, J. G. (2020). Human-like communication in conversational agents: a literature review and research agenda. *Journal of Service Management*.

Veletsianos, G., and Russell, G. (2014). “Pedagogical Agents,” in Handbook
of Research on Educational Communications and Technology, 4th Edn, eds
M. Spector, D. Merrill, J. Elen, and M. J. Bishop (Abingdon: Routledge),
759–769.

WEI-ERN, J. W. (2012). *Establishing Rapport with Conversational Agents: Comparing the Effect of Envelope and Emotional Feedback* (Doctoral dissertation).

Weiss, B., Wechsung, I., Kühnel, C., & Möller, S. (2015). Evaluating embodied conversational agents in multimodal interfaces. Computational Cognitive Science, 1(1), 1-21.

Xie, B., Zhou, J., & Wang, H. (2017). How influential are mental models on interaction performance? exploring the gap between users' and designers' mental models through a new quantitative method. Advances in Human-Computer Interaction, 2017.

Yzerbyt, V., Corneille, O., Seron, E., & Demoulin, S. (2004). Subjective essentialism in action: Self-anchoring and social control as consequences of fundamental social divides. In V. Yzerbyt, C. M. Judd, & O. Corneille (Eds.), The psychology of group perception: Perceived variability, entitativity, and essentialism (pp. 101-124). New York: Psychology Press.

Ziefle, M. and Bay, S., "Mental models of a cellular phone menu.
Comparing older and younger novice users," in Proceedings of the International Conference on Mobile Human-Computer Interaction, pp. 25–37, International, Berlin, Germany, 2004.

Chapter 6

Applying an Agent-based Model to Simulate Just-In-Time Support for Keeping Users of eLearning Courses Motivated

Scholten MR, Kelders SM, Van Gemert-Pijnen JEWC, Steenbeek HW

Applying an Agent-based Model
to Simulate Just-In-Time Support
for Keeping Users
of eLearning Courses Motivated

Nonlinear dynamics, psychology, and life sciences,
2020, 24.4: 403-429.

Abstract

Background: Persuasive technology can support users of self-paced eLearning courses during critical moments of low motivation. Agent-based models (ABMs) – a relatively unfamiliar phenomenon within the persuasive technology and eLearning domains- offers a potentially relevant methodology to understand when the support should be delivered. Using ABMs, the dynamics of motivational user states can be simulated. Subsequently, emerging user patterns can be traced that can potentially provide insight in the ebb and flow of motivation.

Objective: The objective of this study is first to simulate trajectories towards critically low motivation of eLearning users and second to simulate supportive repairment actions with the purpose of reversing the non-productive motivational trend.

Methods: For the purpose of this study, we designed an exploratory ABM on motivation based on the mental energy notion of which the foundations can be found both within the literature of motivational psychology and agent-based modeling.

Results: During the simulations we succeeded in generating moments of critically low user motivation. In addition, we were able to simulate the positive impact of external user support at those critical moments.

Conclusion: Our results suggest that it is plausible to put further energy in developing ABM models with the ultimate goal of feeding persuasive technology with the ability to deliver just-in-time user support during eLearning.

Key Words: agent-based model, simulation, self-learning motivation, eLearning, persuasive technology

Introduction

Motivation and Support in eLearning

eLearning (Moore, Dickson-Deane, & Galyen, 2011) is an approach to teaching and learning that is based on the use of electronic media and devices as tools for improving access. In case users participate in self-paced eLearning modules, during which there is no contact with a care provider or tutor, many users quit, resulting in high non-adherence rates of 40% - 80% (Bawa, 2016). In his review, Bawa (2016) mentions low motivation and a lack of support as precursors of discontinuation. Thus, the question becomes how the strength of self-guidance can be combined with mitigating the drop-out risk due to user experiences of low motivation.

An encouraging approach is to provide automated support from the 'e' (IT) system itself. It has been found that system generated (automated) support conditions performed equally well as human support conditions (Kelders, Bohlmeijer, Pots, & van Gemert-Pijnen, 2015). In a similar vein, positive effects of electronic environments using features such as reminders and tailored advice have been reported on (Neff & Fry, 2009). These examples fall under the umbrella of persuasive technology. Persuasive technology (Fogg, 1998) is technology that is designed to change attitudes or behaviors of the users through persuasion and social influence and avoids coercion. This technology can be added to eLearning environments with the purpose of motivating the online students. Embodied Conversational Agents (e.g. DeVault, Artstein, Benn, Dey, Fast, Gainer, & Lucas, 2014) mostly described as animated computational artefacts or briefly 'robots on screen', provide an illustration. Another example of persuasive technology is gamification software, also known as the use of game design elements in non-game contexts (Deterding, Dixon, & Khaled, 2011).

Agent-based Models for Dynamical User Information

Empirical studies have shown that motivation during eLearning is a variable process (Conati, 2002; Kapoor, Mota, & Picard, 2001; Kort, Reilly, & Picard, 2001) with episodes of high and low motivation. In order to grasp these variations, many data points have to be measured. Note that most experiments in the fields of eLearning and persuasive technology (e.g. Scholten, Kelders, & van Gemert-Pijnen, 2019) measure average values of user motivation using post-experimental questionnaires. As a result, fluctuations of motivational states become flattened out within the aggregated outcome levels (Enfield, 2014; Steffensen, & Pedersen, 2014). As an alternative, gaining insight in the dynamical states (e.g. of users) can be done through agent-based models (Wilensky & Rand, 2015). An agent-based model (ABM) is a dynamical computational system with a set (system) of interconnected elements

that undergo change because of inter-element influences. The elements within the computational system can represent units at various levels; from neurons (in neural systems) and thoughts (in cognitive-affective systems) to individuals (in relationships) to groups (in societies). The goal of running simulations with an agent-based model is to search for explan-atory insights into the collective behavior of agents. That is, the interactions of the individual agents at the micro level create system behavior at the macro level. The system can demonstrate macro characteristics, that are hard – if not impossible – to predict from the features at the micro level. This phenomenon is also referred to as emergence (Bassett & Gazzaniga, 2011; Sawyer, 2005). Within the neuroscience domain, emergence has been associated with mental states, the relationship between mind and the brain and to human consciousness in general (Eberlen, Scholz, & Gagliolo, 2017).
Next to displaying emergence, the second main strength of agent-based models (ABMs) is their ability to provide insight into the temporal dynamics of processes. It concerns processes of very different kinds, as is demonstrated within the domains of physics and biology (Jackson, Rand, Lewis, Norton, & Gray 2017). Closer to the contexts of eLearning motivation, the Dynamical Systems Perspective (DSP) to which ABMs belong, has rendered valuable insight in the domain of social psychology (Steenbeek, van der Aalsvoort, & van Geert 2014; Vallacher, Read, & Nowak, 2017). Knowing that self-paced eLearning programs are characterized by varying users' emotional states and enlarged drop-out rates, we want to exploit the dynamical and temporal aspects of ABMs. Stated differently, the ambition of our study is to design an ABM for simulation of the temporal dynamics of eLearning motivation. In addition, we intend to simulate external support effects as a means to reverse episodes of critical low user motivation towards productive emotional states.

Literature Review

We followed the guidelines of Wilensky and Reisman (2006) on how to create an ABM model. We reviewed three sources of literature with the purpose of grounding our ABM model, with regards to motivational psychology, ABM models, and motivational student states during eLearning.

Motivational Psychology

As a first step, we performed a literature search in the domain of motivational psychology. We encountered several studies (e.g. Hufford, Witkiewitz, Shields, Kodya, & Caruso, 2003; Scott, 2016) that refer to the ABM-related notion of tipping points, moments were the situation changes drastically (instead of linearly). This has also been described as the *catastrophe principle*, and has been successfully computationally modeled for motivation within organizations (Guastello, 1987). We consider the fundamental decision to quit an eLearning course as a potential candidate of a catastrophe. However, the studies we found either applied different techniques than agent-based modeling (Guastello, 1987) either did not translate their theoretical notions into a computational model at all (e.g. Hufford et al., 2003; Scott, 2016). We, therefore, built our own model and shifted our search strategy towards theoretical and empirical notions on motivation that we could implement as computational variables. We encountered such a notion; *mental energy* (e.g. Alexander & Winne, 2012; Sevincer, Busatta, & Oettingen, 2014). Mental energy mobilization fuels effort and performance toward attaining the desired future (Oettingen et al., 2009; Oettingen, Marquardt, & Gollwitzer, 2012) such as obtaining a learning or therapeutic objective. In a similar vein, Simpson and Balsam (2015) refer to motivation in relation to energy according to their definition *motivation as the energizing of behavior in pursuit of a goal*. Finally, as a physiological representation of energy, Domenico and Ryan (2017) refer to the dopaminergic system that is underpinning motivation. Altogether these studies suggest a pivotal role of mental energy with regards to motivation.

Agent-Based Modeling

As a second step, we investigated the literature on existing ABM models on mental energy. This did not lead to an outcome. We therefore decided to take a model from an entirely different domain into consideration; the predator-prey model (Wilensky & Reisman, 2006). This model simulates an eco-system, grounded on the notion of energy. That is, the model simulates the wandering agents sheep and wolf and the stationary agent grass, present in an imaginary space and each possessing an amount of energy. In case agents of different kinds encounter each other, the weaker agent disappears (grass is eaten by sheep, sheep are eaten by wolves) and the energy level of the stronger agent increases. This way the stronger agent can survive; the energy boost counterbalances the natural decay of energy for each wandering agent. Despite the seemingly simple structure of the predator-prey model, it has demonstrated interesting dynamical patterns within the field of ecology. Because of the excellent reputation of this model and following the TAPAS guidelines (Take A Previous model and Add Something, see Frenken, 2004), we decided to use this model as our basis for our simulation study. Our agents would not represent organisms like wolf or sheep or smaller physical units like neurons. Instead, our agents would represent intra-individual units of mental energy on an abstract level, comparable to notions known from the Artificial Intelligence (AI) field, being sub-symbolic units (Asai & Fukunuga, 2018), an approach that also has been referred to as connectionism (see e.g. the learning simulation study of Dovgopoly & Mercado, 2013). Like the original predator-prey model, our agents wander around in imaginary space. Contrary to the predator-prey model our agents do not destroy each other but vanish from the simulation when their energy has dropped to zero. Furthermore, obviously, our space does not symbolize an eco-system but an eLearning student. For substantiating our model, we analyzed several more studies in the field of biological psychology and physiology that referred to mental energy as a pivotal notion. Benton, Parker and Donohoe (1996) make a connection between glucose levels in the blood and cognitive functioning. Likewise, Fairclough and Houston (2004) relate metabolism to the effort spent on mental tasks. As these models describe: doing a challenging task leads to a drain of energy, which could be added as an energy-lowering mechanism to our ABM model. We encountered the computational model of Read and Miller (2019) that embedded the approach and avoidance principles (e.g. Gray, 1987) as two separate motivational mechanisms. Inspired by their study, we could make a distinction in our ABM model between agents with a high level of mental energy (representing the approach principle) and agents with a low level of mental energy (representing the avoidance principle).

Motivational States during eLearning

As a third and last step, we investigated the literature on motivational user states during eLearning to find theoretical and empirical ground for our ABM simulations. Various eLearning studies have shed a light on the affects users experience during their activities. Both incidence and persistence of various affects (measured per time unit of 20 seconds) during eLearning have been empirically investigated (Baker, D'Mello, Rodrigo, & Graesser, 2010). The authors referred to a model to ground affect categories, the two axes of displeasure/pleasure (axis 1) and deactivation/activation (axis 2) (Russell, 2003). Baker et al. (2010) took the affective states of boredom, frustration, confusion, engaged concentration, delight and surprise into account, as being most relevant in relation to eLearning results. Engaged concentration was the most common state, followed by confusion. On persistence, boredom scored highest and even came close to be a non-transitory mood. Stated differently, once a learner had become bored, it became difficult for them to become engaged again. Other experiments have studied the relationship between user affects and learning outcomes. Post-test scores of physics understanding decreased as function of negative affect during learning (Linnenbrink & Pintrich, 2002).

Engagement

With regards to positive effects during eLearning, flow (Ceja & Navarro, 2009; Csikszentmihalyi, 2020) and engagement have been mentioned (D'Mello et al., 2008). Flow is regarded as a highly desirable positive affective state that is beneficial to learning. Logically most eLearning programs would want to promote and prolong the state of flow. However, as mentioned by D'Mello et al. (2008) any intervention on the part of the tutor runs the risk of adversely interfering with the flow experience. Instead, the authors decided to focus on measuring negative states such as frustration and boredom. Therefore, we follow the same approach and focus on reversing negative cognitive-emotional states into productive states for learning. We also address negative motivational states in the following sections.

Frustration

Despite its notorious reputation, it is argued (Mentis, 2007) that frustration is a natural learning experience that may not need remediation. Frustration among users of information systems is only of concern if it is associated with events that are outside of the user's locus of control, such as a program bug. A frustrating event of this nature interrupts the user's cognitive flow and ambition to fulfill their computerized tasks. In other situations, such as a video game that has been designed for invoking user frustration due to difficult episodes that the user must overcome (Gee, 2012), frustration is not even a burden but an indispensable part. Frustration mitigation would be counter-productive with regards to fulfilling the goals of the electronic environment. In other words, this type of frustration is largely productive. From a learning theory perspective, both frustration and confusion have a positive connotation as they are associated with the principle of learning through cognitive disequilibrium (Graesser & Olde, 2003; Piaget, 1952). According to this principle, a student is an engaged, steady learner up to the point they encounter a learning element that they do not yet understand. This first leads to confusion. In case the lack of comprehension endures, the student can become frustrated. It is not until the student understands the new knowledge element that a new equilibrium is created. As of that moment, understanding on a higher level is created, which is generally accompanied with positive affect.

Boredom

As a second unproductive motivational state for eLearning, boredom is commonly referred to. As subtle behavioral indicators of boredom pupils going off-task and users zoning out are described by D'Mello and Graesser (2012). Boredom is a likely predictor of non-adherence within a self-paced eLearning context. Empirical studies have demonstrated that loss of attention pertaining to boredom can be overcome with outside help. This was exhibited by a study on an Intelligent Tutoring System (ITS) on biology that aimed to promote engagement and learning by dynamically detecting and responding to student's boredom and disengagement (D'Mello, Olney, Williams, & Hays, 2012). As a means, the system made use of an eye tracker to monitor the student's gaze patterns and identified when the user was no longer looking at the screen. If the longitude of the gaze away period exceeded a threshold, an animated pedagogical agent displayed the motivational intervention, "please pay attention" to reengage the user. Their results showed that the motivational intervention succeeded to bring the students' attention back to the pedagogical agent; statistically significant decreases were found for off-screen gaze behavior.

Motivational User States as ABM Output

Motivational states and their transitions have been empirically tested in relation to eLearning activities on computer literacy (D'Mello & Graesser, 2012; study 2). Although the authors did not make any reference to ABM models, their study described notions such as "states" and "transitions" that have a remarkable semantic resemblance with the ABM vocabulary. Furthermore, their study relates well-defined theoretical learning notions (e.g. equilibrium hypothesis) to empirical motivational states of eLearning users. We therefore used their empirical results on motivational user states as target states for our simulations.

D'Mello and Graesser (2012) used video technology during their experiment to assess the participants' motivational states. After the experiment, the participants self-evaluated their motivational states, using a pre-defined list of main affective states to choose from. Patterns of transitions (trajectories) that occurred significantly above chance level were found which are depicted as main motivational states (ellipses) and transitions (arrows) in Figure 1. Note that the authors related the state transitions to learning theories which gave their study a well-grounded basis. The highlights of their study:

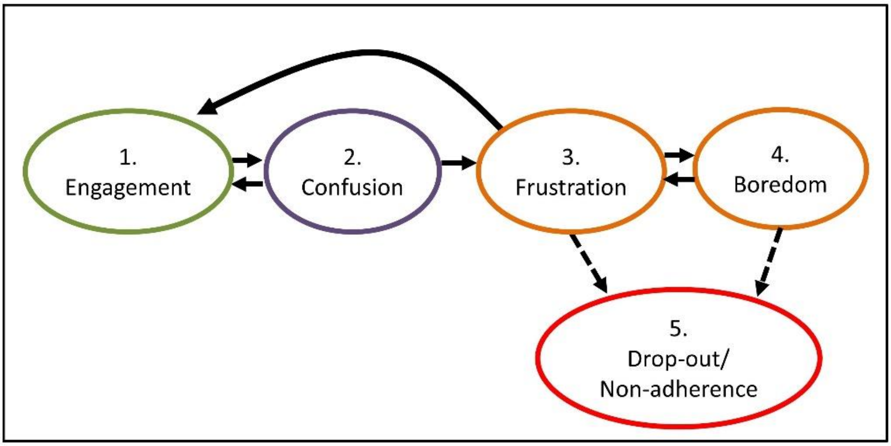


Figure. 1. Foundations of our simulation study. Relevant motivational states (elipses) and trajectories (arrows) during eLearning as the foundation of our simulation study.

The *disequilibrium hypothesis* was confirmed; learners who are in an engaged state (state 1 of Figure. 1), detect an impasse and transition into cognitive disequilibrium (a state of confusion, state 2 of Figure 1). Impasses were defined as contradictions, anomalies, system breakdowns or errors.

The *productive confusion* hypothesis was confirmed; learners who can master the information that relate to the impasse and can integrate the information into existing knowledge schema's go back from confusion (state 2 of Figure 1) to engagement/flow (state 1 of Figure. 1).

The *hopeless confusion* hypothesis was confirmed; not all learners can resolve all the impasses they encounter. In case they fail to resolve the information, their learning goals are blocked, and they experience a sense of failure. As a result, they go from confusion (state 2 of Figure. 1) to frustration (state 3 of Figure 1).

The *disengagement* hypothesis was confirmed; learners who cannot get themselves out of the episode of frustration (state 3), will eventually stop trying and transition into the state of boredom (state 4 of Figure 1).

We decided to simulate their four main states and added state 5, drop-out, as an extra state (i.e. not in scope of D'Mello & Graesser, 2012) that is relevant for the success of self-paced eLearning programs and which we expect to follow from *persistent* frustration or boredom.

In summary, we aimed to investigate the topic of variable user motivation by conducting simulations with an ABM. Figure 2 portrays the scope of this study. We conducted simulations with an agent-based model as depicted on the left side. We also simulated various motivational states of D'Mello and Graesser (2012), of which persistent boredom and frustration had our special interest as we suspected them to lead to non-adherence (depicted as the ellipse at the bottom right side). Furthermore, we purported to simulate engagement (flow) as a motivational state, as we considered it as a predictor of adherence. In addition, we intended to simulate confusion as a motivational state that is expected to reside in between flow and frustration. Furthermore, we aimed to find trajectories of motivational states when simulating the empirical results of D'Mello and Graesser (2012; study 2). Finally, our objective was to repair unproductive student's states by bringing them back to engagement through an act of external support. This type of support is provided by persuasive technology that is capable of sending out suitable motivational messages towards the eLearning student.

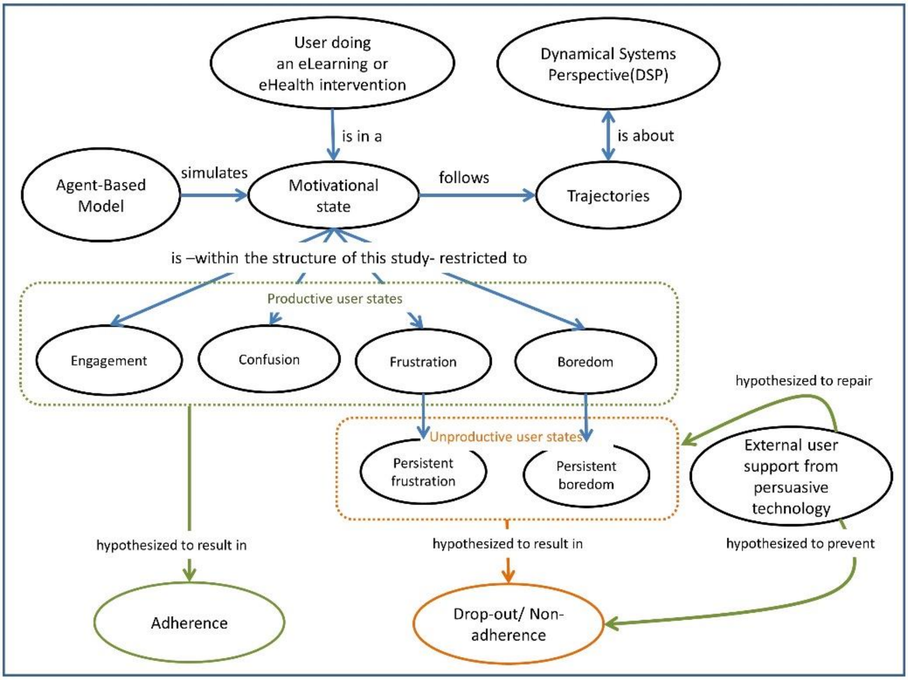


Figure 2. Structure of our study.

Material and Methods

Principles of the ABM model

Inspired by the common notion of energy within the motivational psychology and ABM literature, we decided to use energy as the main agent's characteristic. We implemented the notions of D'Mello and Graesser (2012) on boredom and engagement, using mental energy as a single variable. We generalized the study's principles of engagement, confusion and frustration to "high energy" and of boredom to 'low energy. Furthermore, our ABM model was based on the swarm principles of the predator-prey model, with agents that move within an imaginary space and encounter each other. Within our ABM low-level agents would interact with each other and create high-level, emerging patterns of motivation at the eLearning student level.

Mechanisms and Variables within the ABM Model

We designed and built our ABM model, using Netlogo version 6.0.1. Within this section we elaborate further on the model's mechanisms and variables.

Mechanisms Inducing Mental Energy Fluctuations

We implemented three supplementary mechanisms for increasing men-tal energy at the agent's level; as a first autonomous mechanism, high-energy agents can transfer their energy to their low-energy neighbor as to represent an internal transmission effect of motivation. This mechanism does not affect the number of agents, but does affect the energy level per agent. As a second autonomous mechanism, two agents who meet each other in the imaginary space (i.e. they are at adjacent places) can create a child agent that receives half the energy of its parents. This offspring mechanism was borrowed from the original predator-prey model and mainly affects the number of agents. Thirdly, as a researcher-controlled mechanism, agents receive a dose of mental energy. This dose is determined by the variable support-effect-on-energy. The occurrence of being supported/ motivated is determined by the variable support-chance. For decreasing mental energy, we created a single mechanism: during the simulation, each cycle leads to subtraction of the agent's energy level as high as the decay variable. This represents the effort the student puts in following the eLearning course.

Variables

Our motivation simulation model contains the following energy-related variables:

- The variable *initial-level-of-energy*, values ranging from 0-10 represen-ting the student's level of mental energy at the start of the eLearning course.

Prior to the simulation, when the set-up phase of the model is carried out, energy is distributed over the agents with initial-level-of-energy as the average value.

- The variable *energy-decay*, values ranging from 0.2-2.0. When the simulations runs, each cycle leads to subtraction of the agent's energy level as high as the energy-decay value. In case the energy level of the agent has dropped to zero, the agent disappears from the simulation. In case all agents have disappeared, the simulation stops.
- The variable *support-chance* ranging from 0.05-0.50, the chance that an individual agent receives a supportive message. This variable is a pre-condition for the variable support-effect-on-energy to become effective.
- The variable *support-effect-on-energy*, values ranging from 0-10. This energy dose is added to the current energy of the agent during each simulation cycle. When the models runs autonomously, this variable represents a process of self-motivation. When this variable is enlarged during the simulation, the surplus represents external support.

In addition the model contained the following two control variables:

1. The variable *initial-number-of-agents* had values ranging from 25-100. This value determines how many agents are put into place during the set-up phase, just before the simulation starts to run. This is a typical and generic variable for a model of this kind.

2. The variable *low-high-energy-ratio* had values ranging from 3-12. The boundary value for the ratio of Low Energy Agents divided by High Energy Agents. If there are substantially more low energy agents and this boundary value is reached, the user has come to the state of mental fatigue and will drop out of the eLearning course, so the simulation stops.

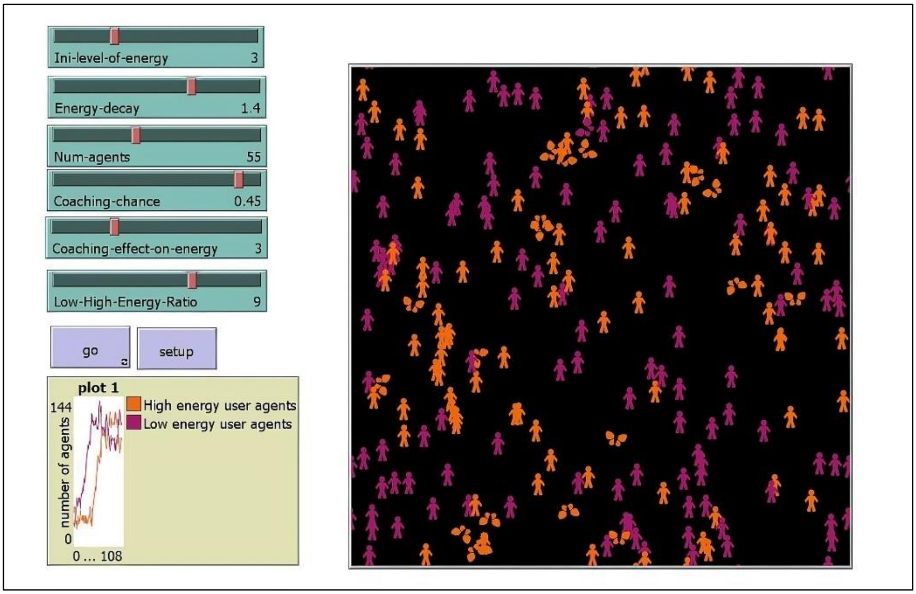


Figure 3. Netlogo ABM model. Lighter agents are HEA. Darker agents are LEA.

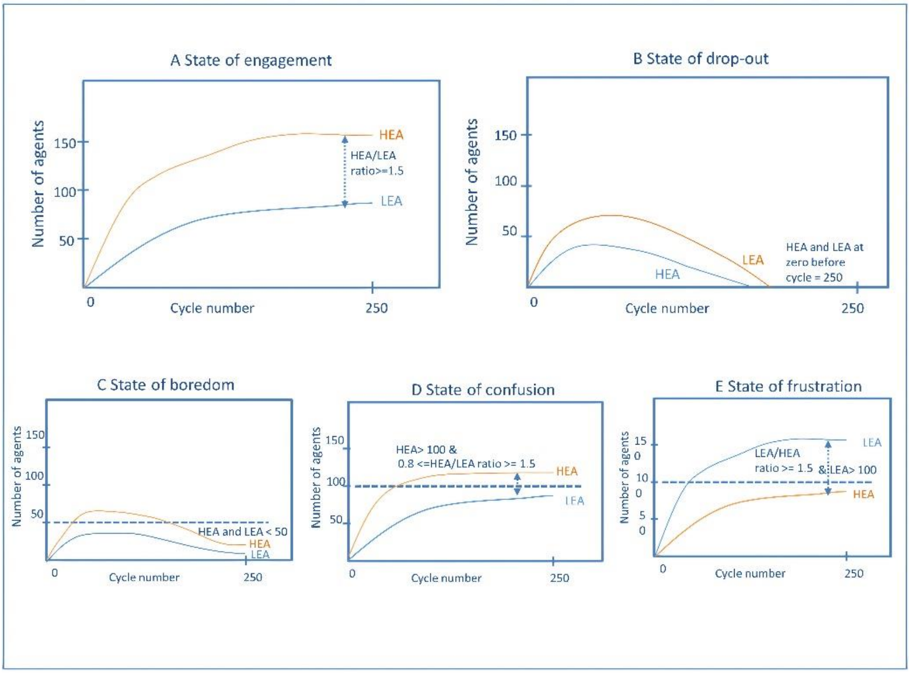


Figure 4. Graphical representations of the five motivational states A to E. See text for explanation.

Static Boundaries

Furthermore, the model contains several static boundaries. As a first static boundary, the energy level of an agent has been maximized at a value of 10. As a second static boundary we made the model stop at 250 cycles, as pre-simulations had shown that the model's patterns did not change once this number had been reached. As a third static boundary, the number of agents has been maximized at 4 times the number of initial agents as pre-simulations showed that a ceiling value was needed to prevent overcrowding of the simulation space which made the simulations useless.

After the pre-simulation phase, we set a criterion for what we considered high energy agents and low energy agents. We defined Low Energy Agents (LEA) as having an energy level of 6 or lower. We defined High Energy Agents (HEA) as having an energy level of 7 or higher. We defined a standard configuration set of variables around the average (middle) values of the model's variables. Note that, in our case, high and low energy agents are not opposing forces but agents that are gradually distinctive. HEA can become LEA during the simulation (due to energy decay) and LEA can over time become HEA (resulting from support-chance and support-effect-on-energy).

We defined the five states we aimed to simulate in Figure. 4. In Figure 4, motivation states A-E are depicted as *expected* graphs of High Energy Agents (HEA, solid line) and Low Energy Agents (LEA, dotted line). From above to below and from left to right: A. State of engagement. Criteria: HEA>100 and HEA to LEA ratio >=1.5. B. State of drop-out. Criterion: HEA and LEA = 0 before cycle = 250. C. State of boredom. Criteria: HEA and LEA both < 50 but > 0 and cycle 250 is reached. D. State of confusion. Criteria: HEA > 100 and HEA to LEA ratio < 1.5. E. State of frustration. Criteria: LEA>100 and LEA to HEA ratio >=1.5. (Note that this is equal to A. state of engagement with HEA replaced by LEA).

Table 1. Standard Values of the variables we used for ABM simulations.

Parameter	Initial Value
Initial-level-of-energy	5
Energy-decay	1.2
Number-of-agents	70
Support-chance	0.35
Support-effect-on-energy	4
High-low-energy-ratio	9

Standard Configuration Set

We created a standard configuration set for the ABM variables (Table 1). This set always resulted in a state of engagement, but was at the same time very sensitive to adaptations, as pre-simulations had shown. Therefore, this standard set was used as a starting point for our simulations.

Objective of the Simulations

The objective of our simulations was first to find whether we could simulate main motivational states (D'Mello & Graesser, 2012) being engagement, confusion, frustration and boredom and drop-out as our added fifth motivational state. Our second objective was to simulate motivational state transitions of D'Mello and Graesser (2012). Our third objective was to find out whether we could reverse a trend towards drop-out by simulating effects of external support.

We ran simulations with various levels for the model's variables, applying single-variable (i.e. changing one variable at a time) and multi-variable (changing two or more variables at a time) simulations. Last, we simulated external user support for the purpose of bringing users back to a state of engagement.

Results

We ran single-variable, multi-variable and trajectory simulations. After each simulation, the Netlogo model generated a csv file with the simulation results containing: cycle number, number of HEA, number of LEA. Subsequently, we categorized the end state of the simulation into one of the five categories as presented in Figure. 1. The results of the various types of simulations are discussed in three separate sections below.

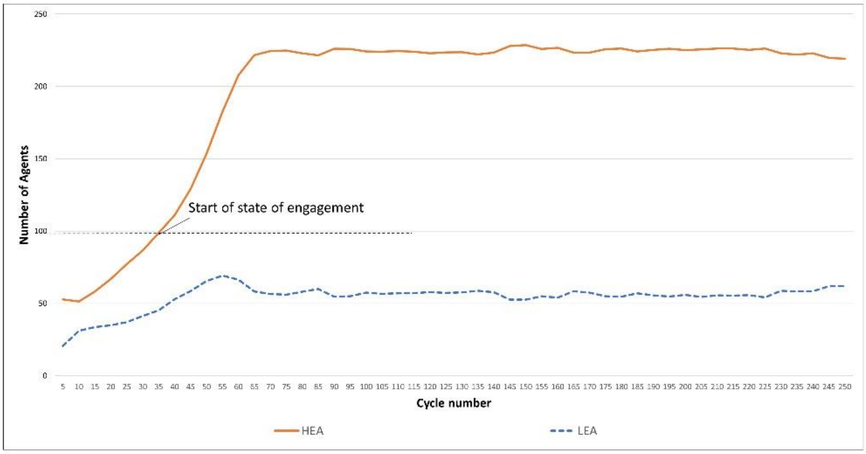


Figure 5. Single-variable simulation result for the variable initial-level-of-energy set to the maximum value 7.

Single-Variable Simulations

We started with manipulating a single variable at a time. Subsequently, we ran simulations to get a first understanding of the dynamics of the model. We

Table 2. Summary of the Effects and Motivational States Resulting from the Single-Variable Simulations.

Variable	Effect	Number of motivational end states simulated
Initial-level–of-energy	Non-determining as a stand-alone factor. Irrespective of the value chosen, there is always an engaged motivational state simulated.	1 simulated state for values 0, 5, 7: engagement (3 * 10 times)
Energy-decay	Determining as stand-alone factor in case of very high values. A 1.8 value results in early simulation stops, which represents an early drop-out of the user.	1 simulated state for values 1.0, 1.2, 1.4: engagement (3 * 10 times). 2 states for value 1.8: drop-out (7 times) and boredom (3 times)
Number-of-agents	Non-determining as a stand-alone factor. Irrespective of the value chosen, there is always an engaged motivational state simulated.	1 simulated state for values 50, 70, 80: engagement (3 * 10 times)
Support-chance	Stand-alone, it is a determining factor for the success of the simulation. Simulations done with the very low values (0.05, 0.10, 0.15) all represent drop-out. Simulations done with the low (0.25), medium/standard (0.35) and high level (0.45) all represent engagement.	1 simulated state per value 0.05, 0.10, 0.15: drop-out (3*10 times) 0.25, 0.35, 0.45: engagement (3*10 times)
Support-effect-on-energy	Stand-alone it is a determining factor. With values of 1 and 2, the simulation does not reach the end of 250 ticks and drop-out is simulated. With values 3 and 4, the value of 250 ticks is reached and there are substantially more high energy agents than low energy agents and engagement is simulated.	1 simulated state per value 1: dropout (10 times) 2: drop-out (10 times) 3: engagement (10 times) 4: engagement (10 times)

ran 10 simulations with each value set. We varied one variable at a time, starting from the set of standard values as presented in Table 1. We ran single-variable simulations manipulating all the variables, being initial-level-of-energy, energy-decay, number-of-agents, support-chance and support-effect-on-energy. Figure 5 portrays an example of a single-variable simulation.

As Figure. 5 depicts, initially and up to cycle 65, HEA and LEA both grow, although the growth rate of HEA is much higher as depicted by the steeper solid orange line. Then the growth stops and the number of HEA and LEA agents both remain stable, as can be seen by the flat lines. Relating this graph to a motivational state; this is a clear case of an engaged learner who remains motivated up to the end.

As the summary of the single-variable simulations (Table 2) demonstrates, the support variables support-chance and support-effect-on-energy are both determining factors for (dis)engagement. This is valuable information in the light of our ambition to simulate the positive effects of support for disengaged users. We present these support results within the results section on user support.

Multi-variable Simulations

After our simulations with single variables, we ran multi-variable simulations during which we changed two or more variables at a time. Our aim was to look out for configurations that would lead to fundamentally different outcomes (e.g. one simulation ending in engagement, another simulation resulting in early drop-out), while departing from the same configuration. Discovering the sensitivity of the model, could provide us with deeper insight in how motivation develops. In order to come to a proper set-up of variables, we chose values with conflicting directions such a high energy-decay (during the single- variable simulations leading to early drop-out) combined with a high support-chance (during the single-variable simulations leading to prolonged engagement). We selected configurations as diverse as possible on the criteria of (a) different end states between configurations (e.g. drop-out, engagement) and (b) different end states within configurations. This led to our set of seven configurations, as presented in Table 3 below.

-

Table 3. Configuration of Variables used for the Multi-variable Simulations.

Configuration number	Initial level of energy	Energy-decay	Number of agents	Support-chance	Support effect on energy	High-Low Energy Ratio
1	5	1	50	0.45	3	9
2	5	1	55	0.45	3	9
3	5	1.2	55	0.45	3	9
4	5	1.4	55	0.45	3	9
5	5	1.4	60	0.45	3	9
6	6	1.4	60	0.45	3	9
7	5	1.4	65	0.45	3	9
Standard set (reference values)	5	1.2	70	0.35	4	9

Using the configurations in Table 3 we ran 10 multi-variable simulations with each of the seven configuration sets. As an example, Figure 6 depicts a multi-variable simulation based on configuration 3 of Table 3.

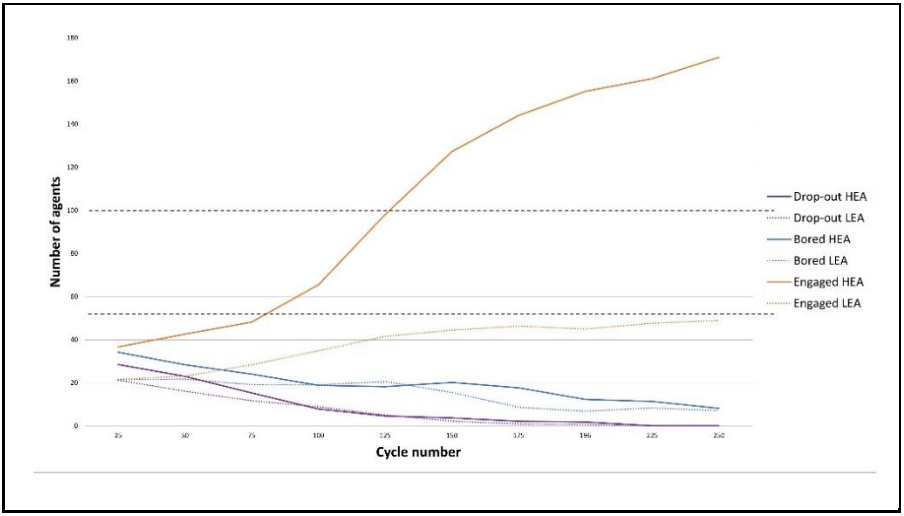


Figure 6. Multi-variable result for configuration 3 of Table 3. The solid lines within Fig. 6 represent the High Energy Agents (HEA), the dotted line the Low Energy Agents (LEA) of the three categories being drop-out, boredom and engagement.

As depicted in Fig. 6; shortly after start of the simulations, at about cycle 25, HEA lines start to diverge from their LEA counterparts, and go into directions that do not change anymore. Six simulations (lines with circles) showed engagement, with high numbers of HEA and a HEA/LEA ratio far above 1.5. Two simulations (lines with triangles) showed disengagement (boredom) with both a low number of HEA and LEA (<50). Two simulations (lines with squares) led to a drop-out.

Table 4 summarizes of the results of the multi-variable simulations using all seven configurations. Table 4 portrays that we were able to simulate three out of four motivational states of D'Mello and Graesser (2012; study 2); engagement, confusion, boredom, *but not frustration*. We succeeded in simula-ting drop-out as our added fifth state.

Regarding frustration, we decided to simulate more specifically why we were not able to find it as an end state. See Fig.7 for the result of this simulation. The right side of Fig. 7 displays the trajectory from frustration towards drop-out. We took as configuration: number-of-agents = 80, initial-level-of-energy = 4, energy-decay = 1.8, support-chance = 0.35, support-effect-on-energy = 4, low-high-energy-ratio = 9.

Table 4. Results of the Multi-variable Simulations Using Table 3.

Configuration	Motivational end states found
Configuration 1	3 states: Boredom (5 times), drop-out (3 times), engagement (2 times)
Configuration 2	1 state: Engagement (10 times)
Configuration 3	3 states: Boredom (2 times), drop-out (2 times), engagement (6 times)
Configuration 4	2 states: Confusion (1 times), drop-out (9 times)
Configuration 5	2 states: Engagement (8 times), drop-out (2 times)
Configuration 6	2 states: Confusion (6 times), drop-out (4 times)
Configuration 7	2 states: Confusion (8 times), drop-out (2 times)

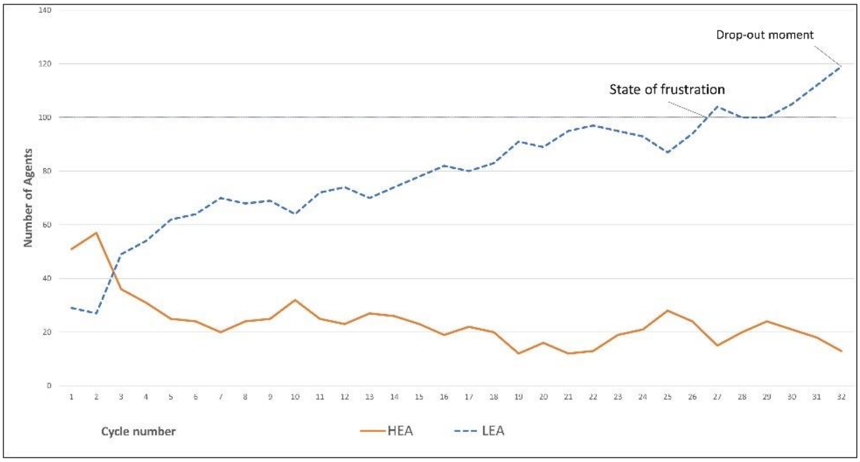


Figure 7. Frustration as a short-lived state in our model.

As can be seen, from cycle 2 onwards the dotted LEA line continues to go up and the solid HEA line continues to go down. Not long thereafter, the state of frustration is reached (LEA > 100 and LEA to HEA ratio >= 1.5), shortly followed by the drop-out state as determined by the Low-High-Energy-Ratio. In short, this figure visualizes that frustration is a so-called *repellor* state, a state the system strongly tends to move away from, namely towards drop-out.

External User Support

For the preparation of the simulations on external support, we took our standard configuration set (see Table 1) and enlarged energy-decay to a very high value of 1.8 as previous simulations had shown (see Table 2, variable energy-decay) that this configuration would have a high chance of an early drop-out. We could repair the trajectory towards drop-out by providing a dose of support. We set the simulation speed at a low value and started the simulation. See the results in Figure 8.

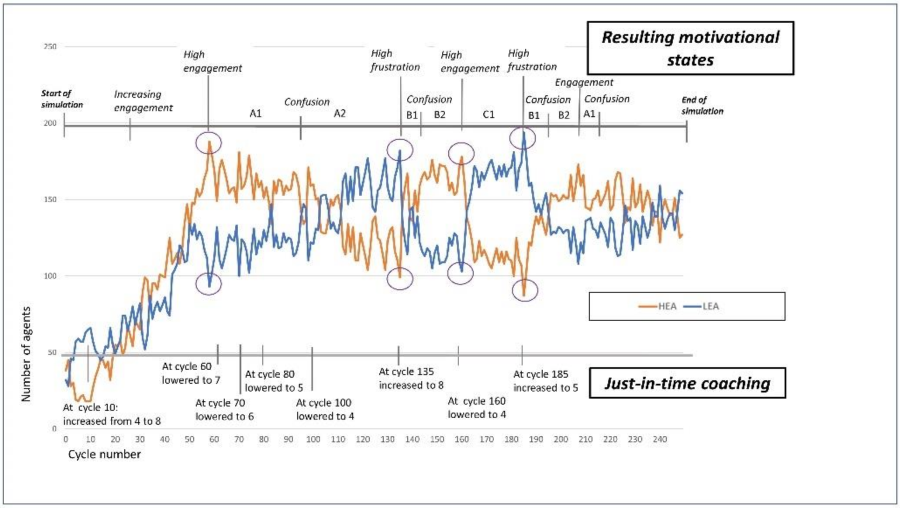


Figure 8. ABM simulation results in terms of motivational states (e.g. engagement, confusion) and trajectories (e.g. T1, T2) as portrayed on the upper horizontal line, resulting from just-in-time supportive acts as indicated on the lower horizontal line.

Common to all our simulations, our simulation started off with a low number of HEA and LEA, both less than 50. At cycle 10 we saw a strong upward trend of LEA and a strong downward trend of HEA. We counter-acted the trend towards drop-out by manually increasing the support-effect-on-energy from 4 to 8 as to simulate stronger support (supportive) messages. As a result, both the LEA and HEA went up, and the model reached a highly engaged state at cycle 60. We lowered support-effect-on-energy step by step and (following trajectory T1) the system first transitioned into confusion and then (following trajectory T2) into frustration. As to reverse the frustration at cycle 135, we repeated the cycle 10 intervention and increased the support-effect-on-energy from 4 to 8. The U1 trajectory was created and the system shortly landed in the state of confusion and (after U2) directly came into engagement again. At cycle 160 we lowered support-effect-on-energy in one go to 4, as opposed to our previous step by step approach. The system went directly to frustration, following trajectory V1. We increased support-effect-on-energy by a minimal amount from 4 to 5 at cycle 185. The system went into confusion at cycle 200 and engagement at cycle 210, and from then on, the swings decreased in size and the system ended in the state of confusion at cycle 250.
After the Figure 8 simulation, we still lacked a simulation of the “emotional repair” trajectory from boredom to engagement resulting from external support. We used configuration 1 as that would give us the largest chance (5 out of 10) of encountering boredom. As Figure 9 depicts, we increased support-effect-on-energy at cycle 68

from 3 to 6 and the system indeed transitioned from a state of boredom into a state of engagement at cycle 230. So we succeeded here in simulating a positive support effect, amending the user's boredom into engagement.

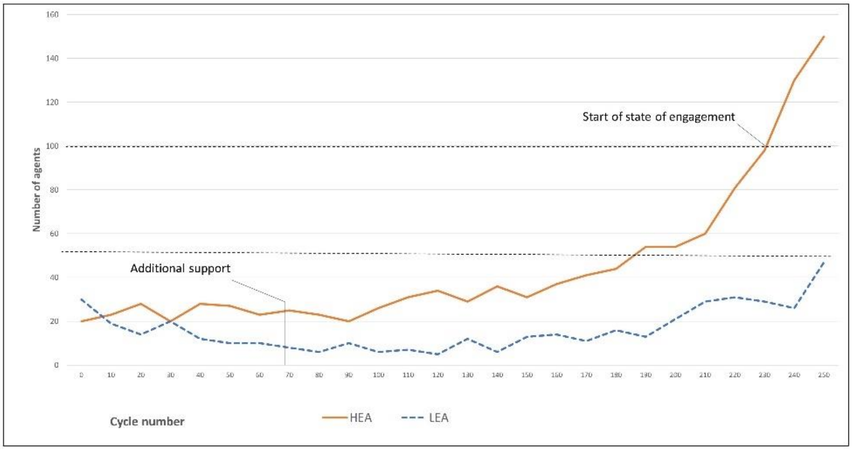


Figure 9. From boredom to engagement resulting from external support (additional support).

Motivational Trajectories

To make the connection between the introduction and the simulation results explicit, we plotted the resulting simulation results of Figures 6, 7 and 8 with the motivational states of D'Mello and Graesser (2012; study 2) and created Figure 10. Figure 10 therefore exhibits the trajectories we were able to find (as indicated by the check sign) and failed to find (as indicated by the cross sign).

As indicated by the check signs: we found forward (T1) and backward (U2) trajectories from engagement to confusion. Furthermore, we found a forward (T2) and backward (U1) trajectory from confusion to frustration.

As indicated by the cross signs, we could not simulate the trajectories between frustration and boredom. In our model frustration did not lead to boredom and boredom did not result in frustration. Remarkably enough, these trajectories were not hypothesized prior to their empirical discovery (D'Mello & Graesser, 2012) either. In addition, we were not able to simulate a trajectory that went from frustration directly to engagement (D'Mello & Graesser, 2012); we encountered confusion as a short-living intermediary state. Again, this direct trajectory was (D'Mello & Graesser, 2012) not hypothesized prior to their experimental finding (D'Mello & Graesser, 2012). With respect to notifying a drop-out risk, we found trajectories from frustration to drop-out (see Figure7) and from boredom to drop-out (see Figure 6).

With respect to reversing trajectories towards drop-out by providing external support: we could bring a bored user back in engagement (see Fig. 9) and a frustrated user back to confusion (U1) and subsequently to engagement (U2). So, altogether we succeeded in simulating bringing eLearning users who are at the verge of dropping out back to a productive state.

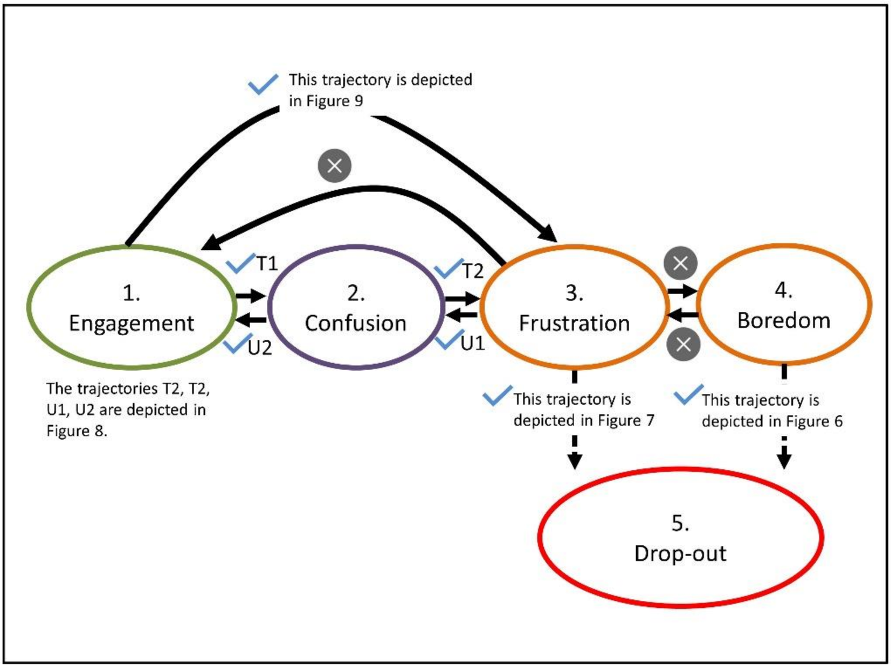


Figure 10. Summary of our ABM simulation results plotted on the motivational states of D'Mello and Graesser (2012; study 2).

Discussion

We had a three-fold objective: first to find whether we could simulate four main motivational states (D'Mello & Graesser, 2012; study 2) being engagement, confusion, frustration and boredom and drop-out as our added fifth motivational state. Our simulation results showed that we were able to simulate three out of four states as end states, we found frustration only as a short-lived repellor state towards drop-out.

Our second objective was to simulate motivational state transitions (D'Mello & Graesser, 2012). We found most motivational state transitions, but we couldn't simulate the two trajectories connecting the states of frustration and boredom. Furthermore, we couldn't replicate the trajectory from frustration directly to engagement, skipping the state of confusion.

Our third objective was to find out whether we could reverse a trend towards drop-out by simulating positive effects of external support. Our simulations showed that by enlarging the support part at critical moments, we were able to bring users back from both frustration and boredom to engagement.

Just-In-Time Support

The single-variable simulations that lacked the additional just-in-time support interventions, showed that departing from a low value for support-chance a high drop-out risk resulted. Reversely, these single-variable simulations showed that with a stable and high support-chance engagement was established. The single-variable simulation results for support-effect-on-energy showed the same pattern; low values resulted in dropout, high values led to user engagement.

During the additional support simulations, we were able to counteract the loss of motivation effect by manually increasing the variable support-effect-on-energy just-in-time and we succeeded in bringing the user back to a state of engagement. Thus, external support was highly relevant to user motivation in our model.

When a User's Expectations Are Not Met

Furthermore, we simulated a user expectations effect. An eLearning course that apparently disappoints the user (high energy-decay) and during which limited external support (medium support-chance and medium support-effect-on-energy) is available, led to high drop-outs rate during the single-variable simulations.

Predictors of Drop-Out

In the introduction we described frustration and boredom as predictors of drop-out, and our model's simulations frustration did not lead to boredom and boredom did not result in frustration. Despite their similarities, our results suggest that boredom and frustration are still fundamentally different motiva-tional states. The trajectory from frustration to drop-out is in our eyes an example of an emotion-driven, highly energetic decision to stop. In contrast, the trajectory from boredom to drop-out represents prolonged ennui (D'Mello & Graesser, 2012) that is associated with low mental energy.

Limitations and Future Research

We started our study with the low adherence levels during eLearning programs and mentioned the dynamics of user states that cannot be captured by post-experimental questionnaires. As an alternative we proposed a combination of empirical user data (e.g. gained through video observations) together with simulations. Our results provide further support for the methodological perspec-tive that it does not make sense to regard towards a user as residing in an average motivational state. Instead, our study provides further support for the concept of a dynamical user transitioning from a set of multiple motivational states and of which unproductive transitions need external action. Using a scientific methodology that is in accordance with this finding is therefore advisable. A challenge remains of course how an eLearning can reliably detect that a user is in a state of persistent boredom or persistent frustration. In our study we visually assessed the simulated curves of HEA and LEA agents. In reality, an eLearning course should make use of genuine real-time user motivation data. See the next session for our vision on those matters. Future research can be done within a real-life e-learning setting. Data collection can take place on behavioral log data of students (e.g. number of online exercises done, time spent per e-learning unit). These data -with the students' permission- can be collected on a continuous basis during the course. Subsequently, this type of information can be used to monitor the students' level of participation. In case the patterns within the data suggest an elevated risk for drop-out the e-learning institution can send a motivational e-mail to the students.

Furthermore, the collection of behavioral log data can be combined with the collection of physiological data as is enabled by several technological developments. Both within experimental settings (e.g. affective computing, see e.g. Garbarino, Lai, Bender, Picard, & Tognetti, 2014) as in personal contexts (e.g. quantified self, see e.g. Lee 2014) physiological signals of different kinds (e.g. heartbeat, heart rate variability, sleep patterns) are being measured. As a consequence, it is likely that in the coming years more short-term user state data will become available for analysis and interpretation.

In addition, the capacity to automatically analyze these data is also becoming more readily available. Technological advancements within the domains of A.I. and more specifically Machine Learning provide means to analyze these types of user data and to respond to users in real-time. Combining the delivery with the capabilities of AI to process real-time behavioral and physiological user data can lead to a new generation of user-adaptive eLearning programs that will likely positively affect user adherence (see e.g. Nahum-Shani et al., 2017; Spruijt-Metz & Nilsen, 2014). This suits a research methodology that is familiar with analyzing the dynamical aspects of these data and our stance is that ABMs have a valuable offer to make here. By combining dynamical, short-term empirical outcomes (e.g. video recordings, physiological measurements), with simulated dynamics and the more stationary empirical outcomes of student questionnaires, scholars have more measurement instruments at their disposal. These will enable them to assess the user's motivational state from different angles. These assessments can be made computational and implemented in the eLearning programs so that the persuasive technology knows what kind of support it should provide when.

Then again, we realize that a simulation-based approach has still to live up to its promise of offering both an insightful and representative view on mental processes of students within a broader scientific context. Only a limited number of eLearning scholars is currently familiar with dynamical systems modeling. Devising a proper simulation model of psychological processes is challenging and choosing the assumptions requires some dauntlessness. Crossing the chasm of physiological to psychological user states is equally ambitious. For good reasons these challenges have been called "wicked problems" (Davis, O'Mahony, & Gulden 2017). Nevertheless and concluding we think that even wicked problems can be tackled and ABMs are likely to contribute.

Conclusion

Our ABM model demonstrated that a significant part of the empirical motivational data of D'Mello and Graesser (2012) could be replicated during the simulations. In addition, we succeeded in simulating just-in-time support effects. However, we are just at the beginning. Future research should be done to find support for the idea that this approach can ultimately inform eLearning programs to deliver user support at suitable moments for the sake of larger user adherence.

References

Alexander, P. A., & Winne, P. H. (2012). Handbook of educational psychology. Routledge.

Asai, M., & Fukunaga, A. (2018, April). Classical planning in deep latent space: Bridging the subsymbolic-symbolic boundary. Paper presented to the Thirty-second AAAI Conference on Artificial Intelligence. New Orleans, LA.

Baker, R. S., D'Mello, S. K., Rodrigo, M. M. T., & Graesser A. C. (2010). Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive–affective states during interactions with three different computer-based learning environments. International Journal of Human-Computer Studies, 68, 223-241.

Bassett, D. S., & Gazzaniga, M. S. (2011). Understanding complexity in the human brain. Trends in Cognitive Sciences 15, 200-209.

Bawa, P. (2016). Retention in online courses: Exploring issues and solutions—A literature review. Sage Open, 6(1), 2158244015621777.

Benton, D., Parker, P. Y., & Donohoe, R. T. (1996). The supply of glucose to the brain and cognitive functioning. Journal of Biosocial Science, 28, 463–479.

Ceja, L., & Navarro, J. (2009). Dynamics of flow: A nonlinear perspective. Journal of Happiness Studies, 10, 665-684.

Csikszentmihalyi, M. (2020). Finding flow: The psychology of engagement with everyday life. Hachette UK: Basic Books.

Conati, C. (2002). Probabilistic assessment of user's emotions in educational games. Journal of Applied Artificial Intelligence, 16, 555-575.

Davis, P. K., O'Mahony, A., & Gulden, T. (2017). Challenges for Social and Behavioral Research and Its Modeling. In Proceedings of the 2017 international conference of the Computational Social Science Society of the Americas (pp. 13-28). Santa Fe, NM: ACM.

D'Mello, S., Jackson, T., Craig, S., Morgan, B., Chipman, P., White, H. & Graesser, A. (2008). AutoTutor detects and responds to learners affective and cognitive states. In B. Woolf, E. Aïmeur, R., Nkambou, & S. Lajoie (Eds.). Intelligent tutoring systems: 9th international conference on intelligent tutoring systems, Montreal, Canada, June 23-27, 2008, proceedings (pp. 306-308). New York, NY: Springer.

D'Mello, S., & Graesser, A. (2012). Dynamics of affective states during complex learning. Learning and Instruction, 22, 145-157.

D'Mello, S., Olney, A., Williams, C., & Hays, P. (2012). Gaze tutor: A gaze-reactive intelligent tutoring system. International Journal of Human-Computer Studies, 70, 377-398.

Deterding, S., Dixon, D., & Khaled, R. (2011). Gamification: Toward a definition. The ACM CHI Conference on Human Factors in Computing Systems 2011 (pp. 12-15). Vancouver, BC, Canada: Association for Computing Machines.

DeVault, D., Artstein, R., Benn, G., Dey, T., Fast, E., Gainer, A., & Lucas, G. (2014). SimSensei Kiosk: A virtual human interviewer for healthcare decision support. In Proceedings of the 2014 international conference on autonomous agents and multi-agent systems (pp. 1061-1068). Paris, France: International Foundation for Autonomous Agents and Multiagent Systems.

Di Domenico, S. I., & Ryan, R. M. (2017). The emerging neuroscience of intrinsic motivation: A new frontier in self-determination research. Frontiers in Human Neuroscience, 11, 145. doi: 10.3389/fnhum.2017.00145

Dovgopoly, A., & Mercado, E. (2013). A connectionist model of category learning by individuals with high-functioning autism spectrum disorder. Cognitive, Affective, & Behavioral Neuroscience, 13, 371-389.

Eberlen, J., Scholz, G., & Gagliolo, M. (2017). Simulate this! An introduction to agent-based models and their power to improve your research practice. International Review of Social Psychology, 30, 149-160.

Enfield, N. P. (2014). Causal dynamics of language. In N. J. Enfield, P. Kockelman & J. Sidnell (Eds.), Cambridge handbook for linguistic anthropology (pp. 319-335). Cambridge, UK: Cambridge University Press.

Fairclough, S. H., & Houston, K. (2004). A metabolic measure of mental effort. Biological Psychology, 66, 177-190.

Fogg, B. J. (1998). Persuasive computers: perspectives and research directions. In Proceedings of the SIGCHI conference on Human factors in computing systems (pp. 225-232). ACM Press/Addison-Wesley Publishing Co.

Frenken, K., (2004). History, state and prospects of evolutionary models of technical change: A review with special emphasis on complexity theory. Utrecht University, The Netherlands. Retrieved August 31, 2020 from http://www.narcis.nl/publication/RecordID/oai:library.tue.nl:656114.

Garbarino, M., Lai, M., Bender, D., Picard, R. W., & Tognetti, S. (2014). Empatica E3—A wearable wireless multi-sensor device for real-time computerized biofeedback and data acquisition. In 2014 4th international conference on wireless mobile communication and healthcare-transforming healthcare through innovations in mobile and wireless technologies (MOBIHEALTH) (pp. 39-42). Athens, Greece: Institute of Communication and Computer Systems.

Gee, J. P. (2012). Situated language and learning: A critique of traditional schooling. New York, NY: Routledge.

Graesser A. C., & Olde B. A. (2003). How does one know whether a person understands a device? The quality of the questions the person asks when the device breaks down. Journal of Educational Psychology, 95, 524-536.

Gray, J. A. (1987). The psychology of fear and stress (Vol. 5). Cambridge, UK: Cambridge University Press.

Guastello, S. J. (1987). A butterfly catastrophe model of motivation in organizations: Academic Performance. Journal of Applied Psychology, 72, 165-182.

Hufford, M. R., Witkiewitz, K., Shields, A. L., Kodya, S., & Caruso, J. C. (2003). Relapse as a nonlinear dynamic system: Application to patients with alcohol use disorders. Journal of Abnormal Psychology, 112, 219-227.

Jackson, J. C., Rand, D., Lewis, K., Norton, M. I., & Gray K. (2017) Agent-based modeling: A guide for social psychologists. Social Psychological and Personality Science, 8, 387-395.

Kapoor, A., Mota, S., & Picard, R. W. (2001). Towards a learning companion that recognizes affect. MIT Media Laboratory Affective Computing Technical Report No. 543. Retrieved March 15, 2009 from http://vismod.media.mit.edu/ pub/tech-reports/TR-543.pdf.

Kelders, S. M., Bohlmeijer, E. T., Pots, W. T., & van GemertPijnen, J. E. (2015). Comparing human and automated support for depression: fractional factorial randomized controlled trial. Behaviour Research and Therapy, 72, 72-80.

Kort, B., Reilly, R., & Picard, R. W. (2001). An affective model of interplay between emotions and learning: Reengineering educational pedagogy-building a learning companion. Proceedings of the IEEE international conference on advanced learning technologies (43-46). Los Alamitos, CA: IEEE Computer Society Press.

Lee, V. R. (2014). What's happening in the "quantified self" movement? In ICLS 2014 proceedings (pp. 1032-1036). Boulder, CO: International Society of the Learning Sciences.

Linnenbrink, E. A., & Pintrich, P. R. (2002). The role of motivational beliefs in conceptual change. In M. Limón & L. Mason (Eds.) Reconsidering conceptual change: Issues in theory and practice (pp. 115-135). Dordrecht: Springer.

Liu, B., & Sundar, S. S. (2018). Should machines express sympathy and empathy? Experiments with a health advice chatbot. Cyberpsychology, Behavior, and Social Networking, 21, 625-636.

Mentis, H. M. (2007). Memory of frustrating experiences. In Information and Emotion: The Emergent affective paradigm in information. Behavior Research and Theory, 1, 197-210.

Moore, J. L., Dickson-Deane, C., & Galyen, K. (2011). e-Learning, online learning, and distance learning environments: Are they the same? The Internet and Higher Education, 14(2), 129-135.

Nahum-Shani, I., Smith, S. N., Spring, B. J., Collins, L. M., Witkiewitz, K., Tewari, A., & Murphy, S. A. (2017). Just-intime adaptive interventions (JITAIs) in mobile health: Key components and design principles for ongoing health behavior support. Annals of Behavioral Medicine, 52, 446-462.

Neff, R., & Fry, J. (2009). Periodic prompts and reminders in health promotion and health behavior interventions: Systematic review. Journal of Medical Internet Research, 11(2), e16. doi: 10.2196/jmir.1138

Oettingen, G., Marquardt, M. K., & Gollwitzer, P. M. (2012). Mental contrasting turns positive feedback on creative potential into successful performance. Journal of Experimental Social Psychology, 48, 990-996.

Oettingen, G., Mayer, D., Sevincer, A. T., Stephens, E. J., Pak, H., & Hagenah, M. (2009). Mental contrasting and goal commitment: The mediating role of energization. Personality and Social Psychology Bulletin, 35, 608-622.

Piaget, J. (1952). The origins of intelligence in children. (M. Cook, Transl.). New York, NY: International Universities Press.

Read, S. J., & Miller, L. C. (2019). A neural network model of motivated decision-making in everyday social behavior. Social-Behavioral Modeling for Complex Systems, 1, 145-162.

Russell, J. A. (2003). Core affect and the psychological construction of emotion. Psychological Review, 110, 145-172.

Sardi, L., Idri, A., & Fernández-Alemán, J. L. (2017). A systematic review of gamification in e-Health. Journal of Biomedical Informatics, 71, 31-48.

Sawyer, R. K. (2005). Social emergence: Societies as complex systems. NY: Cambridge University Press.

Scholten, M. R., Kelders, S., & Gemert-Pijnen, V. (2019). An empirical study of a pedagogical agent as an adjunct to an eHealth self-management intervention: What modalities does it need to successfully support and motivate users? Frontiers in Psychology, 10, 1063. doi: 10.3389/fpsyg.2019.01063

Scott, B. (2016). Cybernetic foundations for psychology. Constructivist Foundations, 11, 509-517.

Sevincer, A. T., Busatta, P. D., & Oettingen, G. (2014). Mental contrasting and transfer of energization. Personality and Social Psychology Bulletin, 40, 139-152.

Simpson, E. H., & Balsam, P. D. (2015). Behavioral neuroscience of motivation. Cham, Switzerland: Springer.

Spruijt-Metz D, & Nilsen W. (2014). Dynamic models of behavior for just-in-time adaptive interventions. IEEE Pervasive Computing, 13(3), 13-17.

Steenbeek, H., van der Aalsvoort, D., & van Geert, P. (2014). Collaborative play in young children as a complex dynamic system: Revealing gender related differences. Nonlinear Dynamics, Psychology, and Life Sciences, 18, 251-276.

Steffensen, S., & Pedersen, S. B. (2014). Temporal dynamics in human interaction. Cybernetics & Human Knowing, 21, 80-97.

Vallacher, R. R., & Nowak, A. (1997). The emergence of dynamical social psychology. Psychological Inquiry, 8, 73-99.

Vallacher, R. R., Read, S. J., & Nowak, A. (Eds.). (2017). Computational social psychology. London, UK: Routledge.

Wilensky, U., & Rand, W. (2015). An introduction to agent-based modeling: Modeling Natural, Social, and Engineered Complex Systems with NetLogo. Cambridge, MA: MIT Press.

Wilensky, U., & Reisman, K. (2006). Thinking like a wolf, a sheep, or a firefly: Learning biology through constructing and testing computational theories—an embodied modeling approach. Cognition and Instruction, 24, 171-209.

Chapter 7

General Discussion

Discussion

Within this general discussion, first the findings of the studies presented in this thesis regarding the deployment of ECA's as adjuncts to eHealth interventions, will be summarized. Second, the main results will be discussed. Third, the main implications of these findings will be elaborated upon.

Chapter two: Scoped Review on currently unaddressed eHealth user needs versus ECA's capabilities for support

Study one, chapter two, set the baseline for this thesis. The scoped review described unaddressed needs of users of eHealth interventions and the capabilities of ECA's to provide support (see RQ one). Furthermore, chapter two set forth how the ECA's supportive capabilities can be mapped on the eHealth user needs. As main categories for user needs, we described task-related support for guidance on practical issues and emotion-related support for deeper, personal matters often pertaining to the state of being a patient and the associated loss of health. With regards to ECA's we defined non-responsive and responsive ECA's. Non-responsive ECA's are fairly easy to implement as adjuncts in eHealth interventions and not costly. They can utter speech fragments, yet are not capable of capturing and responding to emotionally expressed user needs. A typical example is provided by the Voki® ECA we used in our experiments. Responsive ECA's, on the other hand, are more sophisticated and far more expensive. They have the capability to detect and process verbal and nonverbal information as expressed by users. Furthermore, they have the capacity to respond to users in real-time. A typical example is provided by the Primer® ECA platform as e.g., used by Lucas et al. (2017).

Bringing the topics of user needs and ECA's together; non-responsive ECA's are likely to be effective providers in case of less personal, task-related guidance and support. A typical example is factual feedback on a patient's homework. A more personal issue such as the need for encouragement is more suitable for a responsive ECA to handle, as it can process some of the user's emotional signals, so far mostly in lab situations. Finally, very personal issues such as a patient coping with experiences of negative affect during e-psychotherapy can be most appropriately handled by a human caregiver.

Within this thesis we made use of non-responsive ECA's as they can serve as easy-to-implement adjuncts to existing eHealth interventions. In case it is proven that they are effective these ECA's can be added with limited hassle to eHealth interventions that are evidence-based but suffer from elevated attrition levels.

Chapter three: empirical evidence for the eHealth user's appreciation of ECA support

Study two, in chapter three, empirically assessed whether a supportive non-responsive ECA can be a valuable adjunct to an eHealth psycho-education intervention (see the initial part of RQ two). During the experiment, the added value in terms of task-related support was measured through the deployment of the autonomy and feedback outcome measures. Autonomy aims to measure the extent to which the user feels enabled and supported by the eHealth intervention to carry out the experimental tasks. Feedback measures the extent to which the study's participants feel informed by the eHealth intervention on the coming actions. In addition, the effect of emotion-related support resulting in a relationship was measured through the rapport outcome variable. In addition, the study tested multiple modalities; speech, text, animation. The ultimate objective for deployment of the ECA was to achieve elevated user motivation and involvement for the eHealth intervention. The user motivation and involvement served as a proxy for user adherence in a real-life eHealth setting. Study two found positive effects on feedback and autonomy induced by the ECA, as compared to the textual control condition. The speech-based ECA outperformed the text-based ECA on one of the outcome variables; feedback. The ECA's ultimate objective was not achieved; no differential effects were found for motivation and involvement due to the ECA's supportive actions. These partial effects are in accordance with the results of the meta-study of Schroeder, Adesope, and Gilbert (2013); ECA's as adjuncts in eLearning generate mixed positive and negative effects in terms of learning and motivation.

Noteworthy on study two is that emotional support seemed hardly needed, as participants expressed they overall had a positive learning experience. We would expect a larger effect of the ECA's (both task-related and emotion-related) support in case users are under stress. This was the reason to start up study three (chapter four) as a follow-up study.

Chapter four: the eHealth user's stress level and the hypothesized enlarged appreciation of the ECA's support

In study three, chapter four, we provided half of the participants with an invalidated, stressful Pac Man game. We hypothesized that this would induce elevates stress levels amongst study participants, as defined in the second part of RQ two. Furthermore, we reasoned that elevated stress levels would induce a higher need and evaluation of the ECA's supportive actions during the eHealth psycho-education intervention. Moreover, such an experiment context would provide a fair representation of experiences of health patients in real-life. That is, chronic conditions such as asthma, arthritis and diabetes lead to stress. Altogether, stress plays a central patient affecting role in many real-life eHealth interventions. As chapter four showed, stress induction at the beginning had been successful, but stress had vanished at the end. Surprisingly, stress had a *negative* impact on the evaluation of the ECA. Instead of an *enlargement* of the study two effect, the ECA no longer outperformed textual support on any of the outcome measures. As an explanation we posited that ECA's can also produce adverse effects (Louwerse, Graesser, McNamara & Lu, 2009), due to e.g. distraction (Weiss, Wechsung, Kühnel, & Möller, 2015) or being too confrontational (Rickenberg & Reeves, 2000). The multiple regression analysis showed that rapport was a significant predictor for involvement with the eHealth intervention, and showed that autonomy and feedback were not. This implies that rapport building between the ECA and user is crucial for promoting adherence. It further strongly suggests that task-related support within an eHealth psycho-education context is of lesser importance towards user adherence. Providing support for the validity of our experimental set-up, the no-stress condition of chapter four replicated the chapter three outcomes. That is, for the no-stress solution the ECA outperformed text on the outcome variables autonomy and feedback, exactly like the ECA did in (the no-stress) study two as described in chapter three. Thus, contrary to our expectations, our ECA was solely positively evaluated in a low-stress context and not in a stressful context.

Chapter five: rapport building between ECA and user through synchronized speech
Study four in chapter five explored a novel, speech synchrony option to establish rapport between our non-responsive ECA and user, as defined in RQ3. That is, there is ample evidence (e.g. Gratch, Wang, Gerten, Fast & Duffy, 2007) that for ECA's to become effective support providers, it is important they first build rapport with the user. As mentioned in the previous section, rapport was a significant predictor for involvement with the eHealth intervention and showed that autonomy and feedback were not. Rapport building between ECA and user is easier said than done. It normally requires sophisticated *responsive* ECA's that can react to users in real-time, using both verbal and non-verbal communication channels, for both capturing and responding to user signals. However, it is known from the literature that humans who have moved in synchrony (dance, march) or spoken in synchrony cooperate more and experience feelings of rapport towards each other. The same mechanism may work between ECA and human. During our qualitative study, we therefore explored an alternative and novel way to build rapport; synchronous speech with a non-responsive, also called monologue-style ECA. As our study results show, users were fairly positive about speaking synchronously with the ECA. Nevertheless, the

results did not demonstrate that rapport was built. Based on our results, we can say that there are a series of practical improvements possible. First, users need to become more aware of the rhythm and pace of the ECA's speech in order to be able to synchronize it. Second, the rationale of the experimental task, especially the parts of speaking out loud and speaking in synchrony with the ECA, require a more profound explanation. The rationale is both a pre-condition for participants to take the task seriously and to regard the task as useful. Third, after the experimental task has been carried out, users should receive clearer feedback. That is, a more explicit feedback mechanism on the words spoken by the user ("thank you for speaking out the following words: …) will diminish the chance of thinking the words were spoken in vain. In conclusion, it requires follow-up research to verify the validity of the hypothesized causal relationship between speech synchrony and user-ECA rapport.

Chapter six: simulations of unproductive cognitive-affective states of eHealth users and repairment actions through ECA support

Whereas the previous studies took an empirical approach towards the delivery of support through an ECA, study five in chapter six, used simulation as a research method for studying the influence of support on user motivation and as defined in RQ4. Study five explored unproductive cognitive-affective states, predictive for non-adherence within a simulated eHealth psycho-education environment. Previous studies (e.g. D'Mello et al., 2010) had described the concept of affective just-in-time interventions as to repair a learner's unproductive learning state. Within an earlier study D'Mello, Picard, and Graesser (2007) had summarized this principle as the affective loop. In short, the affective loop describes that the eLearning or eHealth system detects that the user is bored or frustrated. Subsequently, the system intervenes in order to bring the user back to a more productive cognitive-affective state, more suitable for learning. In a similar vein, studies (e.g. Andrew, Borriello, & Fogarty, 2007) refer to the term *kairos*, 'the right moment' to act. The key questions for self-guided eHealth interventions would of course be how the system is able to determine that 'right' moment and what a suitable repairment act should look like. In the D'Mello and Graesser (2012) study the hopeless confusion pattern was identified as potential moment for a repairment act. Learners who fail to meet what the eLearning environment demands, transition from a state of confusion to a state of frustration. Furthermore, the authors encountered the disengagement pattern; learners who cannot get themselves out of a state of frustration transition into a state of boredom. Note that the state of persistent boredom represented the least productive user state within their experiment. In real-life, those persistently bored learners plausibly would have stopped doing eLearning, making this a relevant proxy state for non-adherence.

We devised, built and run an Agent-Based Model (ABM) to simulate D'Mello and Graesser's (2012) cognitive-affective states and the trajectories connecting these states. See Figure 1 below of a graphical representation of these states, using the ball in the valley metaphor. The left part shows a cognitive-affective state, the right part demonstrates a transition of cognitive-affective states.

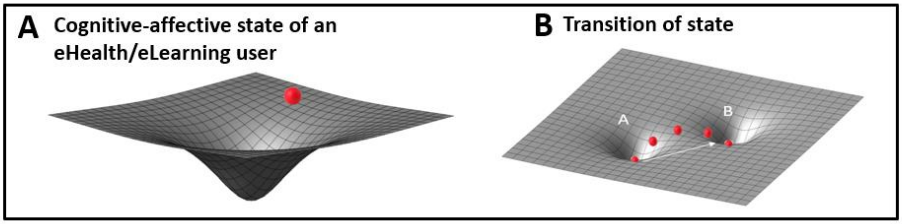


Figure 1 The cognitive-affective state (left, A) and a transition of state (right, B).

The results showed that a significant part of the user states and trajectories could be successfully simulated by the ABM. The hopeless confusion pattern and disengagement pattern were both demonstrated during our simulations. Most importantly, the capability to repair these unproductive states by means of external support at the right moment was equally successfully demonstrated. The non-productive user patterns could be reversed and users could be brought back to state of engagement through supportive actions.

The implications of these results are two-fold:

First, if it can be *simulated*, it can also be *predicted* through simulation. In very practical terms it means that an eHealth system with an ECA as adjunct, could -so far theoretically- be equipped with a predictive ABM that could warn the ECA when a tipping point towards non-adherence would be approaching. As a next step, the ECA would send out a supportive act towards the user in order to bring him back into a more productive cognitive-affective state and as such prevent non-adherence. There is evidence for such emotional repairment acts during eLearning (Woolf et al., 2009).

Second, these results provide strong indications that the domains of affective computing, eHealth, ECA and DSP have relevant intersections. Basically, affective computing has the outlook of making the eHealth intervention emotionally aware of the user's present cognitive-affective state. ECA's hold the promise of personifying automated support. Finally, DSP can connect the affective computing, ECA, and eHealth domains as it holds the promise of understanding the dynamics of user states and therefore the capability to determine the 'right' moment for the delivery of support towards the eHealth/eLearning user. By combining these four domains, it becomes more plausible that the deployment of ECA’s for eHealth adherence will succeed.

Implications for Research and Development

Within the introduction of this thesis, the following research questions were formulated, in order to find out whether ECA could contribute to the self-guided eHealth intervention adherence challenge.
RQ1: What does the scientific literature tell us about unaddressed eHealth user needs and about the capabilities of ECA’s towards addressing these needs?
RQ2: Do eHealth users appreciate ECA support in eHealth and if so, does the induction of experimental stress lead to higher appreciation levels?
RQ3: Can rapport between user and ECA be built through synchronous speech?
RQ4: Can an eHealth user’s cognitive-affective states be computationally simulated and if so, can critically low user motivation states be repaired through ECA support? Taking all these research questions together, it became apparent in study one that users of self-guided eHealth interventions have needs for support, that have largely been implicit and therefore unaddressed by these interventions. Or as firmly stated by Kreijns, Kirschner, and Jochems (2003) the reason that digital learning environments fail is because of socioemotional processes that are being “ignored, neglected, or forgotten”.
Following up on the study one result, resolving unfulfilled user needs through computerized support is certainly not a straightforward task. Generically speaking, ECA’s have been regarded as digital artefacts with a lot of potential for two decades, but their evidence for motivating users is still modest. This generic picture was confirmed by the results from studies two and three: we found positive effects resulting from the ECA’s task-related support, but our ECA was not capable of involving and motivating the participants as proxies for adherence. These limited effects could be explained by the equally limited rapport-building facilities (i.e. through the delivery of effective emotion-related support) of our monologue-style ECA. As a next step, we aimed to circumvent the ECA’s limitations through the synchronized speech task we devised and explored in study four. Study four provided useful guidelines for improvement of the synchronous speech task. It requires a follow-up study to find out whether the improved synchronized speech task can indeed effectively induce rapport. Finally, our simulation study five demonstrated how unproductive user states could be repaired through the delivery of supportive actions. By detecting unproductive states and their patterns, the study made it plausible that non-adherence can be prevented.
So, based on these five studies, what can we conclude on the deployment of ECA’s as a means to enlarge eHealth adherence? In first instance we repeat the earlier

conclusions: 'ECA's have potential, but the potential still has to materialize.' But more important than this conclusion is the question: what are opportunities for future ECA studies to truly untap the ECA's potential for support in self-guided eHealth? In this note, three important topics came up during the five studies; first the *time-dependent character of eHealth user states*, second the *integrated multi-modal aspect of eHealth user expressions* and third the *two-sided* and *real-time character of the ECA-user mirroring process*.

The time-dependent character of eHealth user states; stress and user motivation

First, within study three, the time-dependent (temporal) character of eHealth user stress became apparent. By applying the PrEMO questionnaire on positive and negative emotions at the start and at the end of the study, study three demonstrated that the eHealth user was initially stressed and that stress had vanished at the end of the experiment. Furthermore, the study three results demonstrated no preference for the ECA's support compared to mere text amongst stressed participants. So, as the study results suggest, stress may inhibit a positive user evaluation of ECA support. In other words, the well-intended ECA support, may be perceived as confrontational. Such effects are not uncommon, adverse effects of ECA's amongst users under stress have been found before (Zanbaka, 2004). In a related note, Rickenberg and Reeves (2000) found that users felt more anxious when an ECA monitored their website work which led to a decrease in the user's task performance. These effects can be explained by the social psychology literature that states that the attention of others fosters mastery of simple tasks but impairs mastery of complex and stressful tasks. This is also defined as the theory of social facilitation and inhibition (Steinmetz and Pfattheicher, 2017; Zajonc & Sales, 1966). However, stress will not likely entirely block the effectiveness of ECA support. Rapport seems to play a pivotal mediating role here. Tsui & Schultz (1985) describe that rapport is a key determinant for effective psychotherapy, known for its stressful episodes. Even more specific to the notion of stress Kim, Roth, and Wollburg (2015) found that the patient-rated therapeutic alliance (a concept close to rapport) is a key factor for the effectiveness of breathing therapies against anxiety and panic. Thus, as a pre-condition for user stress reduction, the eHealth system should be capable of building rapport and of tracking the eHealth user's stress level. At higher stress levels, the ECA as adjunct to the eHealth system can (and often should) intervene and support the user. This is done with the objective to bring the user back to a motivated state and to promote adherence. Of course, this should be done with caution as elevated user stress levels can also result in counter-productive effects as described by the theory of social facilitation and inhibition (Steinmetz &

Pfattheicher, 2017).
Within study five, a second user state that changes over time was encountered; user motivation. Remember that study five computationally simulated how a user's motivational level fluctuates in time during eLearning as demonstrated by D'Mello and Graesser (2012). Using video-material of users, this study showed that during a single experimental session, users can feel motivated, bored and frustrated. Furthermore, as was exhibited, these states can be short-lived but can also reach a more permanent status. Next, the fluctuations between these states provide valuable information as well. A user that goes from motivation to frustration and back on a regular basis is likely having a different experience than a user that follows that trajectory just once. Important to note, both the states and the fluctuations can only become apparent to the eHealth/eLearning intervention, in case they are noticed. This seems hardly surprising, yet to date is only done within specific experiments (e.g. Woolf et al., 2009). Thus, as a recommendation for truly supporting eHealth users, the eHealth intervention needs to measure their cognitive-affective states on a continuous basis. A negative cognitive-affective user state runs the risk of becoming persistent and should be noticed in time. Especially since a persistent negative cognitive-affective user state is an even better predictor for non-adherence than the before mentioned user stress level (Baker, D'Mello, Rodrigo & Graesser, 2010). That is to say, the cognitive-affective user state reflects the impact of *all events* (including the stressful ones) on the user, which makes it (at least hypothetically) a more precise and reliable predictor for non-adherence.
Stated in methodological terms, applying a questionnaire *after* the experiment will provide valuable user data. However, these questionnaire data do not provide the full picture of what the user has been experiencing *during* the experiment. For that, we need short-term measurements of the cognitive-affective states of eHealth users. Worthwhile to mention, short-term measurements are done within more recent eHealth studies; gauging log data (Sieverink, Kelders, Poel & van Gemert-Pijnen, 2017) or physiological data (Nelson, Verhagen & Noordzij, 2016). Thus, these studies in their present set-ups, already represent a major development in how eHealth and ECA studies are conducted. Further steps can be taken, however, that will be further elaborated upon, within the coming sections.

Human user and ECA use two-sided mirroring techniques over integrated modalities

Next to the time-based aspect of eHealth user states, a second topic that became apparent during this thesis, is the *two-sided* and *integrated multi-modal* character of human-ECA communication. This topic came up, when preparing study five on synchronous behavior between human user and ECA. Several studies were reviewed, in order to substantiate the set-up of the experiment. As found, various human-to-human interaction studies (e.g. Louwerse, Dale, Bard & Jeuniaux, 2012). have demonstrated that mirroring and synchrony are found across multiple modalities on a two-sided basis. That is, during the interaction, humans start to use the same words, start to display equal facial expressions, and show comparable gestures. Important to note, some, but not all mirroring is done at the same time (synchronously). Within conversations, people mirror each other with both short delays (facial mirroring with delays of seconds) and longer delays (lexical mirroring with delays of several seconds up to minutes). Then again, postural mirroring (e.g. two people both leaning forward) can be observed at the same point in time. So, when interpreting the mirroring data, these different delays should be taken into account. Important to note, these different modalities work together as a package to sustain the conversation. In other words, for understanding how smooth (or difficult) the human-to-human conversation is going, the degree of mirroring across all modalities provides valuable information. This suggests that it makes sense to apply a two-sided and integrated multi-modal approach to human-ECA research as well. Based on lexical, facial, gestural and postural expressions of human and ECA, the suggestion would be that an *overall mirroring score* is calculated during user-ECA interactions for both the user and ECA. Statistically this can be realized by means of cross-recurrence quantification analysis (CRQA) (Cox, van der Steen, Guevara, de Jonge-Hoekstra & van Dijk, 2016). CRQA is considered as a powerful nonlinear time-series method to study coordination and cooperation between people. In essence, this method captures (delayed) mirroring of behavior when two interlocutors interact. In case, during the conversation, people start to 'lend' words of each other and copy facial expressions, the CRQA score will go up.

Note that this approach, is fundamentally different from the recommendation of Schroeder et al. (2013) who proclaim a single-variable approach for ECA research. Their recommendation is understandable from a control perspective. Start with one variable (e.g. facial expressions of an ECA) and only take another variable (e.g. tone of voice of an ECA) into consideration when the first variable is fully understood. However, the Schroeder et al. (2013) approach goes beyond the point that humans by definition use all modalities they have on a simultaneous basis, for telling their environment how they are doing. Humans intrinsically expect their interlocutor to do the same as has been demonstrated in human-robot studies (Tsiourti, Weiss, Wac & Vincze, 2019). The authors demonstrated that incongruous emotional information across the auditory (vocal prosody) and visual (whole-body expressions) modalities decreased the observers' understanding of the robot's emotions. In case a robot or ECA displays supportive intentions through just a single modality (e.g. verbally, but not through facial expressions) the message become incongruous and will therefore lead to lower understanding by the user.

Human-ECA rapport, as the basis for a successful ECA, can be assessed in real-time

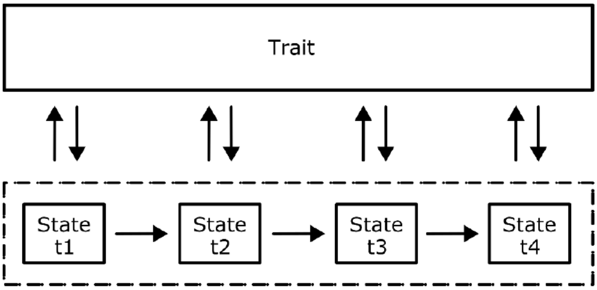
As mentioned before, short-term measurements are helpful to track the user's cognitive-affective state. Furthermore, the application of both an overall ECA and user mirroring score is recommended. As a third suggestion, it seems advisable that these overall mirroring scores are calculated on a real-time basis. As such, they can provide valuable information about the rapport-building process during the interaction. The user mirroring score indicates whether the human user really portrays feelings of rapport (or the absence thereof) to the ECA on a real-time basis. This has practical benefits. When there is a situation of low rapport, the ECA can decide to repair the situation in time. As such, the ECA can rebuild rapport by means of techniques such as small talk. If later, the user is in need of support, the ECA has gained sufficient 'rapport credits' to provide it effectively.

Note that this approach is different from recommendations made by renowned scholars in earlier ECA days. As described within the introduction, Bickmore and Picard (2005) recommend the usage of *long-term studies* (e.g. 6 weeks or more) with daily interactions between user and ECA in order to realistically assess the process of relationship building. However, such long-term experimental designs have serious practical challenges. Participants are often harder to recruit and the costs are elevated. Furthermore, so long as the user's cognitive-affective states are not measured, it is still not precisely known what is happening in terms of relationship building. In addition, note that this two-sided and real-time set-up is also different from the rapport building agent as used by Gratch, Wang, Gerten, Fast, and Duffy (2007). In their study, the ECA used mirroring techniques to build rapport with the user, but the mirroring expressed by the user was not measured. In other words, the authors took a one-sided, ECA-centered approach towards rapport building. Moreover, the authors did not measure rapport from a process perspective but restricted themselves to an outcome perspective, by deploying the rapport questionnaire at the end of the experiment.

The three suggestions of this thesis in a broader perspective

Within the previous sections, three suggestions were provided for the way ECA research can be conducted. It was proposed to take the *short-term, integrated multi-modal and two-sided and real-time* relationship building perspective in scope, especially when conducting laboratory experiments. Typically, within a laboratory situation multi-modal cognitive-affective user data are readily available for short-term and real-time measurements (e.g. video recording to capture a user's facial expressions and speech). Within this section these suggestions will be put in a broader scope and related to current developments within the psychology field. As mentioned throughout this thesis, rapport building between ECA and user stands (and will likely remain standing) central to ECA research. Within this section it is recommended to update the measurement of the key variable, rapport, on the basis of the beforementioned three suggestions. That is, by combining the post-experimental rapport questionnaire data with the short-term and real-time cognitive-affective measures, scholars can gauge the rapport construct from two different angles. The *rapport score* can be assessed through the rapport questionnaire as a one-off. In contrast, the *rapport building process score* can be assessed continuously during the experiment, through the short-term behavioral measurements and by continuously calculating overall mirroring scores through CRQA. Such a combined longer-term, short-term approach has been recently described by De Ruiter, van der Gaag, Jeronimus, and Kunnen (2019). referring to the *state* and the *trait* measurements of psychological constructs, see the picture below. Note that the state and trait principle is a psychological theoretical concept that has a long history, especially within the personality field (see e.g. Allport and Odbert, 1936). Recent insights and technical developments (e.g. computational modelling, time series analysis) have provided new opportunities for empirical research on state and trait. Not surprisingly therefore, relevant empirical and simulation studies on state and trait (e.g. De Ruiter, Van Geert, & Kunnen, 2017) have been underscoring that a psychological concept can (and probably should) be analyzed from both perspectives.

Figure 1.3



Self-organization. Higher order structures, such as individual traits, emerge from the iterative sequence of system states over time. In turn, the individuals' states are constrained by the higher order trait. This means that the states of an individual are stabilized as a result of traits that have developed. Figure from "Self-organization" by M.A.E. van der Gaag, 2018. Copyright 2018 by CC BY 4.0.

Figure 2 State and trait. Reprinted from Psychosocial Development in Adolescence: Insights from the Dynamic Systems Approach (p. 18). De Ruiter, van der Gaag, Jeronimus, and Kunnen (2019). Copyright 2018 by Creative Commons. Reprinted with permission.

Applying the state and trait to motivation in eHealth, the *trait* is the construct that is typically measured by questionnaires on involvement and motivation. It is the more stationary instantiation of a human trait such as motivation in general or motivation to follow an eHealth program. The *state* is the varying, shortly during construct, the eHealth user's motivation that is applicable at a *specific moment in time*. Obviously, the two are related but not equal. Repetitive changes in state, will affect the trait in the long run, working from the bottom up. Reversely, a fundamental new insight can change the trait and will consequently have an impact on the experienced states, working top-down. Fundamental to this model and the before mentioned Dynamical Systems Perspective (DSP) is that the trait should not be considered as the only relevant and measurable experimental construct. Neither should the mere average of states as measured by a post-experimental questionnaire. In contrast, according to the tenets of DSP, the *user states, their fluctuations* and finally *the patterns* they form, all hold important experimental state information. In human dyadic studies DSP principles have been successfully applied on interaction patterns between children (see e.g. Steenbeek and van Geert, 2005). It is therefore plausible to assume that the same principles can be successfully applied to human-ECA studies within an eHealth context.

Finally, within the beginning of the introduction section it was stated (see e.g. study one and study two) that a non-responsive, monologue-style ECA likely provides a minimalistic, yet suitable solution for a substantial part of ECA research. This stance is not in accordance with the three suggestions made. That is, for measuring states of user and ECA, a responsive, dialogue-style ECA is required. So how useful are non-responsive ECA's? For practical and cost reasons, monologue-style ECA's can still be beneficial, especially if the updated synchronous speech task would succeed in building rapport. However, for truly understanding the interaction between human and ECA, responsive ECA's should be chosen.

Strengths and Limitations

A limitation of the studies concerns the participants who are self-selected, mostly healthy and young students. As such they don't represent the general eHealth target group. This means that the results are somewhat harder to generalize. Then again, for this kind of (predominantly exploratory) research, such an experimental set-up seems sufficient. Moreover, the general principles we encounter (e.g. the role of stress and rapport) will also hold for patients in real-life settings. A second limitation concerns the usage of an online ECA study amongst subjects, mostly participating from home. Other than in a laboratory setting with face-to-face contact between researcher and participant, a lower level of control could be maintained by the researcher. Possibly, some participants could have used a more prolonged period for doing the experimental tasks, but then again, we expect that this has had limited consequences. Note that, on the other hand, the online set-up chosen offered a practical solution for running the experiments during the Corona-pandemic. Moreover, the experimental online set-up was representative for how most eHealth applications are used in daily life, namely at home and self-directed. Third, our initial objective was to evaluate the effectiveness of our ECA for eHealth adherence. However, during the thesis' period, we got new insights (e.g. on the role of stress, on real-time rapport, and on the dynamical systems perspective). This led to a change of the follow-up studies, fitting within the frame of agile science. As a limitation of this approach, the ABM simulations and synchronous speech became subject of study relatively late. Then again, it can also be considered as a strength as we needed the experience of the earlier study results to come up with these new ideas.

Conclusion

ECA studies have been conducted for more than two decades, since Cassell formally introduced the term in 1999. After all these years, ECA's hold the promise of more productive HCI, but to date their results are still mixed and inconclusive. Within our experiments we found that the ECA's practical guidance (task-related support) scored significantly higher than textual guidance, but no rapport effect (emotion-related support) was found as a reputable indicator for an effective ECA. Moreover, no involvement effect, serving both as our ultimate experimental outcome and as a proxy for larger user adherence, was found. When participants were put under stress, as to create a more real-life like experimental condition, the ECA's practical merits disappeared. So, although our ECA's support provided some positive effects, it remains debatable whether it would contribute to larger user adherence in a real-life, self-guided eHealth setting.

In short, our results correspond with most previous ECA studies: slightly positive, yet not fully convincing. However, our results also conform to many previous ECA studies in another, more promising way; rapport building seems key for an ECA to be effective. As recommended, measuring the rapport building process and rapport outcome can be fruitfully combined. By applying a short-term, integrated multi-modal and two-sided and real-time perspective, the relationship building process can be measured on a fine-grained basis. By combining this with outcome measurements from questionnaires, the best of both worlds is used. Furthermore, it is recommended that the more recent developments in eHealth and ECA's using log data (Sieverink et al., 2017), physiological data (Nelson et al., 2016) and cognitive-affective data (D'Mello and Graesser, 2012) can benefit from the Dynamical Systems Perspective. Along the lines of DSP, both the user states and their fluctuations can be measured and analyzed.

For eHealth adherence purposes, the user's motivational state provides relevant information to assess whether he is in need of support. Furthermore, by comparing behavioral states of eHealth user and ECA, and assessing their degree of mirroring across multiple modalities, the state of the rapport building process can be gauged. Finally, the fluctuations in both the user's motivational state and the user-ECA rapport state can be assessed. Such fluctuations will likely provide new insights in the intra-experimental dynamics of the experiences of users. It is expected that this dynamical approach will give the combined ECA and eHealth fields a new impetus and that the evidence for effective ECA's can gain strength in order to live up to their promise.

References

Abdallah, C. G., & Geha, P. (2017). Chronic pain and chronic stress: two sides of the same coin?. Chronic Stress, 1, 2470547017704763.

Allport, G. W., & Odbert, H. S. (1936). Trait names: A psycho-lexical study. Psychological Monographs, 47(1, Whole No. 211).

Andrew, A., G. Borriello, and J. Fogarty (2007). Toward a systematic understanding of suggestion tactics in persuasive technologies. In Proceedings of the 2nd international conference on Persuasive technology, PERSUASIVE'07, Berlin, Heidelberg, pp. 259–270. Springer-Verlag.

Baker, R. S., D'Mello, S. K., Rodrigo, M. M. T., & Graesser, A. C. (2010). Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive–affective states during interactions with three different computer-based learning environments. International Journal of Human-Computer Studies, 68(4), 223-241.

Bickmore, T. W., & Picard, R. W. (2005). Establishing and maintaining long-term human-computer relationships. ACM Transactions on Computer-Human Interaction (TOCHI), 12(2), 293-327.

Cox, R. F., van der Steen, S., Guevara, M., de Jonge-Hoekstra, L., & van Dijk, M. (2016). Chromatic and anisotropic cross-recurrence quantification analysis of interpersonal behavior. In Recurrence plots and their quantifications: expanding horizons (pp. 209-225). Springer, Cham.

De Ruiter, N.P., van der Gaag, M.A.E., Jeronimus, B.F., Kunnen, E.S. (2019). What we can gain from a dynamic systems approach to psychosocial development in adolescence. In book: Psychosocial Development in Adolescence: Insights from the Dynamic Systems Approach (Editors: Kunnen, E.S., de Ruiter, N.M.P., Jeronimus, B.F., van der Gaag, M.A.), chapter 1. London: Routledge Psychology.

De Ruiter, N. M., Van Geert, P. L., & Kunnen, E. S. (2017). Explaining the "how" of self-esteem development: The self-organizing self-esteem model. Review of General Psychology, 21(1), 49-68.

D'Mello, S., Picard, R. W., & Graesser, A. (2007). Toward an affect-sensitive AutoTutor. IEEE Intelligent Systems, 22(4), 53-61.

D'Mello, S., Lehman, B., Sullins, J., Daigle, R., Combs, R., Vogt, K. & Graesser, A. (2010, June). A time for emoting: When affect-sensitivity is and isn't effective at promoting deep learning. In International conference on intelligent tutoring systems (pp. 245-254). Springer, Berlin, Heidelberg.

D'Mello, S., & Graesser, A. (2012). Dynamics of affective states during complex learning. Learning and Instruction, 22(2), 145-157.

Gratch, J., Wang, N., Gerten, J., Fast, E., & Duffy, R. (2007, September). Creating rapport with virtual agents. In International workshop on intelligent virtual agents (pp. 125-138). Springer, Berlin, Heidelberg.

Kim, S., Roth, W. T., & Wollburg, E. (2015). Effects of therapeutic relationship, expectancy, and credibility in breathing therapies for anxiety. *Bulletin of the Menninger Clinic*, *79*(2), 116-130.

Kreijns, K., Kirschner, P. A., & Jochems, W. (2003). Identifying the pitfalls for social interaction in computer-supported collaborative learning environments: a review of the research. *Computers in human behavior*, *19*(3), 335-353.

Louwerse, M. M., Graesser, A. C., McNamara, D. S., & Lu, S. (2009). Embodied conversational agents as conversational partners. Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition, 23(9), 1244-1255.

Louwerse, M. M., Dale, R., Bard, E. G., & Jeuniaux, P. (2012). Behavior matching in multimodal communication is synchronized. Cognitive science, 36(8), 1404-1426.

Lucas, G. M., Rizzo, A., Gratch, J., Scherer, S., Stratou, G., Boberg, J., & Morency, L. P. (2017). Reporting mental health symptoms: breaking down barriers to care with virtual human interviewers. Frontiers in Robotics and AI, 4, 51.

Nelson, E. C., Verhagen, T., & Noordzij, M. L. (2016). Health empowerment through activity trackers: An empirical smart wristband study. Computers in Human Behavior, 62, 364-374.

Rickenberg, R., & Reeves, B. (2000, April). The effects of animated characters on anxiety, task performance, and evaluations of user interfaces. In Proceedings of the SIGCHI conference on Human factors in computing systems (pp. 49-56).

Schroeder, N. L., Adesope, O. O., & Gilbert, R. B. (2013). How effective are pedagogical agents for learning? A meta-analytic review. Journal of Educational Computing Research, 49(1), 1-39.

Sieverink, F., Kelders, S., Poel, M., & van Gemert-Pijnen, L. (2017). Opening the black box of electronic health: collecting, analyzing, and interpreting log data. JMIR research protocols, 6(8), e156.

Steenbeek, H., & van Geert, P. (2005). A dynamic systems model of dyadic interaction during play of two children. European Journal of Developmental Psychology, 2(2), 105-145.

Treede, R. D., Rief, W., Barke, A., Aziz, Q., Bennett, M. I., Benoliel, R., ... & Wang, S. J. (2019). Chronic pain as a symptom or a disease: the IASP Classification of Chronic Pain for the International Classification of Diseases (ICD-11). Pain, 160(1), 19-27.

Tsiourti, C., Weiss, A., Wac, K., & Vincze, M. (2019). Multimodal integration of emotional signals from voice, body, and context: Effects of (in) congruence on emotion recognition and attitudes towards robots. International Journal of Social Robotics, 11(4), 555-573.

Vallacher, R. R., Van Geert, P., & Nowak, A. (2015). The intrinsic dynamics of psychological process. Current Directions in Psychological Science, 24(1), 58-64.

Weiss, B., Wechsung, I., Kühnel, C., & Möller, S. (2015). Evaluating embodied conversational agents in multimodal interfaces. Computational Cognitive Science, 1(1), 1-21.

Woolf, B., Burleson, W., Arroyo, I., Dragon, T., Cooper, D., & Picard, R. (2009). Affect-aware tutors: recognising and responding to student affect. International Journal of Learning Technology, 4(3-4), 129-164.

Epilogue

Most of all, I think that the statement of Gratch et al. (2007), referring to rapport as a productive human-ECA interaction that can be described as a *dance*, is most accurate. That is, the term dance is in accordance with a more traditional research methodology consisting of post-experimental questionnaires that provide an appraisal to the dance as a whole. From the questionnaire perspective the dancing couple has just finalized their performance and stands in front of the jury waiting for their one-off judgement.

Moreover, dance is a suitable metaphor for the DSP approach towards human behavior as a process that unfolds through time. Following the DSP principles, the dancing couple is video-recorded and the performance of the dancing couple is assessed by studying consecutive dancing states within short time intervals. Out of all the states, an overall impression emerges, which we also refer to as the quality of the dance.

Samenvatting
(summary in Dutch)

Samenvatting

Al ruim twee decennia wordt onderzoek gedaan naar Embodied Conversational Agents (ECA's; 'robots on screen'), wanneer gerekend wordt vanaf het moment dat Justine Cassell deze term introduceerde in 1999.
Na al deze jaren onderzoek, zijn ECA's nog altijd een belofte en bieden ze geen zekerheid voor een betere gebruikservaring. Doordat een ECA een menselijk gezicht geeft aan een computer en zich via spraak uit, wordt de computer meer zoals wij, menselijke gebruikers.
Tegelijkertijd, kijkend naar hun evidentie, zijn de resultaten niet eenduidig positief te noemen. ECA's laten geregeld positieve effecten zien op bijvoorbeeld leerprocessen, maar zeker niet in alle omstandigheden. Meta-studies die geprobeerd hebben dit gebrek aan overtuigende effecten te verklaren, hebben verschillende oplossingsrichtingen geopperd, zoals een strikter gebruik van ontwerpstandaarden en het consequenter toepassen van standaard ECA-platformen. Maar ook de meta-studies lijken er niet helemaal uit te komen. Voor ECA-onderzoek lijkt het daarom raadzaam om nieuwe wegen te verkennen.
In dit proefschrift is onderzocht of ECA's kunnen bijdragen aan het verhogen van de adherentie van gebruikers van zelfstandig gevolgde eHealth interventies. Dat wil zeggen: van eHealth interventies waarbij de gebruikers geen (menselijke) hulp van buitenaf krijgen. Immers: (psychologische) zelfhulp interventies zijn er in overvloed, maar zonder externe ondersteuning, die wellicht ook ervaren wordt als stok achter de deur, is de adherentie vaak laag. Voor het uitvoeren van het onderzoek volgden we in eerste instantie de reguliere aanpak met experimenten en vragenlijsten.
Onze experimentele resultaten lieten zien dat de praktische hulp die onze ECA's konden bieden (task-related support) inderdaad significant hoger scoorden dan dezelfde hulp aangeboden via tekst. We vonden geen effecten van emotionele ondersteuning vanuit de ECA; dit type ondersteuning leidde niet tot een sterkere emotionele klik (rapport) van de gebruiker met de ECA.
Daarnaast vonden we geen effecten op betrokkenheid van de gebruiker (involvement) bij de eHealth interventie, als indicator voor adherentie. Bovendien, wanneer gebruikers in een stressvolle situatie werden geplaatst, om daarmee een meer realistische eHealth situatie te creëren, bleken de positieve praktische ondersteuningseffecten van de ECA te verdwijnen.
Al met al, hoewel onze ECA positieve effecten liet zien, is het toch onzeker of onze ECA in staat zou zijn om hogere adherentie te bewerkstelligen in een real-life eHealth context. Of te wel: onze resultaten komen overeen met veel voorgaande studies: enigszins positief, maar toch niet echt overtuigend.

Echter, onze resultaten kwamen ook op een andere manier overeen met voorgaande studies; het werken aan een emotionele klik met de gebruiker (rapport) is een belangrijke sleutel tot succes voor de ECA. Merk op dat rapport gaat over het ervaren van een verbinding en om harmonie tussen gesprekspartners. Rapport wordt ook wel in verband gebracht met het 'in synchronie' zijn met elkaar.

Gaandeweg werd tijdens het onderzoekstraject duidelijker dat het belangrijk is om zowel het bouwen aan rapport als het ervaren van rapport te meten. Door een korte-termijnperspectief te nemen (10 tot 40 seconden) en de verschillende modaliteiten te meten (gebruik van woorden door gebruikers, gebruik van gezichtsuitdrukkingen, gebruik van andere lichaamsuitdrukkingen) kan real-time, met video-opnames, bepaald worden in welke mate de interactiepartners (gebruiker en ECA) in synchronie zijn met elkaar. Hoe meer de ECA en de gebruiker elkaars communicatie-uitingen overnemen (dezelfde woorden, simultane gezichtsuitdrukkingen), hoe meer synchronie er gemeten wordt. Hoe meer sychronie er ontstaat, hoe sterker de band tussen gebruiker en ECA wordt. Hoe sterker de onderlinge band, hoe effectiever de ECA en -naar verwachting- hoe beter de ECA in staat zal zijn om de eHealth adherentie van de gebruiker positief te beïnvloeden.

Door deze analyses op lichaamstaal te combineren met vragenlijsten aan het einde van de interactie, kan zowel vanuit een dynamisch perspectief als statisch perspectief gemeten worden hoe succesvol de interactie is voor het creëren van rapport met de ECA, zoals ervaren door de gebruiker.

Bovenstaande aanbeveling lijkt sterk op de methodiek die een gerenommeerd ECA-onderzoeker als Jonathan Gratch toepast, maar gaat een stap verder door methodische inzichten vanuit een ander onderzoeksgebied hierbij te betrekken; het Dynamical Systems Perspective (DSP). Volgens de principes van DSP kunnen zowel de gebruiker statussen als de fluctuaties tussen deze statussen relevante informatie bevatten. Door voort te bouwen op recente ontwikkelingen in de ECA en eHealth domeinen; het gebruik van log data, fysiologische data en cognitief-affectieve data en deze vanuit een DSP-perspectief te analyseren, ontstaan nieuwe onderzoeksmogelijkheden.

Om dit concreet te maken in termen van motivatie en ondersteuning tijdens het gebruik van eHealth en eLearning interventies; de gebruikers motivatie kan tijdens het experiment kort-cyclisch gemeten worden en op basis hiervan kan ingeschat worden of de gebruiker externe ondersteuning door de ECA nodig heeft. Ook kan -door meerdere modaliteiten hierbij te betrekken- kort-cyclisch gemeten worden hoe sterk de emotionele band is die de ECA en de gebruiker op dat moment met elkaar bereikt hebben. Als gebruiker en ECA elkaars woorden gaan gebruiken, dezelfde lichaamshouding hebben en op hetzelfde moment glimlachen, dan weet je dat het goed zit met die band. Tenslotte kunnen ook de fluctuaties qua gebruiker motivatie en qua emotionele band tussen gebruiker en ECA in kaart gebracht worden. In veel onderzoeken worden deze fluctuaties niet gemeten en dus ook niet gebruikt voor nadere analyses. Het traditionele onderzoek naar ECA's legt veel nadruk op het gebruik van vragenlijsten aan het einde van het experiment. Echter, volgens de methodiek van DSP, zouden we die fluctuaties juist wèl moeten meten en kunnen ze ons nieuwe inzichten geven. Zo is het aannemelijk dat eHealth gebruikers langere perioden van motivatie afwisselen met korte periodes waarin de motivatie ontbreekt. Pas als blijkt dat een dergelijk patroon doorbroken wordt en de demotivatie significant toeneemt, moet het leveren van ECA-ondersteuning overwogen worden. Als de ECA intussen een voldoende sterke band met de gebruiker opgebouwd heeft, is de kans groot dat de ECA daarin effectief zal zijn. Het is dus de kunst om op basis van kort-cyclische metingen te achterhalen of je de gebruiker -zelfs als deze licht gefrustreerd is- even met rust moet laten, of dat het moment daar is om de gebruiker op te peppen. Daarnaast gaat het om het verzamelen van kort-cyclische gegevens over de band tussen de gebruiker en de ECA en het ontdekken van eventuele patronen daarin. Dit verzamelen van gegevens gebeurt allemaal *tijdens* het experiment.

Door de patronen uit de dynamische data *tijdens* het experiment te vergelijken met de meer statische inzichten uit vragenlijsten die *na* het experiment worden afgenomen, kan kruisbestuiving plaatsvinden.

Mijn verwachting is dat een dergelijke gecombineerde statisch-dynamische benadering een nieuwe kans biedt voor onderzoek naar de bruikbaarheid van ECA's binnen eHealth interventies. Ik verwacht daarom dat een dergelijke aanpak eraan bij zal dragen dat ECA's daadwerkelijk laten zien waarvoor ze in potentie geschikt zijn; effectieve ondersteuners van eHealth gebruikers.

Dankwoord

Volgens Simone, mijn vrouw, had ik het al over promoveren, kort nadat ik afgestudeerd was voor mijn masterstudie cognitiewetenschap. Als ik voor mijzelf de film terugdraai; toen ik in 1995 afstudeerde aan de Radboud Universiteit in Nijmegen, had ik echt het gevoel iets bijzonders over de wereld geleerd te hebben. Ik had kennis gemaakt met AI, neurale netwerken en bewustzijn. Ik maakte uitgebreid kennis met het begrip reductionisme, het idee dat de waarheid te vinden is in de kleinste elementen, vaak de neuronen of de genen. Ik had kennis gemaakt met begrippen als de onvolledigheidsstelling van Gödel, die je vertelt dat je zelfs in de wiskunde niet alles kunt bewijzen, maar sommige zaken voor waar moet aannemen. Aha, dus zelfs de exacte wetenschap doet uiteindelijk ook aannames. Eenmaal op zoek naar een baan in de ICT gaven recruiters aan cognitiewetenschap óók een bijzondere studie te vinden, want eentje vroeg waarom ik niet *gewoon* economie of bedrijfskunde was gaan studeren …

En blijkbaar plantte die opmerking een zaadje in me, want mede dankzij ondersteuning van mijn toenmalige werkgever in de telecom-industrie, ben ik in 2002 gestart met mijn MBA aan de Rotterdam School of Management. Toen ik mijn EMBA in 2004 afrondde, had ik voor de tweede keer het gevoel dat een studie nieuwe vensters op de wereld voor mij had geopend. Mijn MBA leerde me waar het om ging als je zaken wilde doen of leiding gaf.

Mijn honger naar kennis bleek echter nog niet gestild. In 2014 begon het weer te kriebelen en kreeg ik opnieuw de gedachte ‘snap ik nu waar het in deze wereld om draait?’. Die vraag werd vervolgens: ‘Wat als ik eens zou gaan promoveren? Zou ik daarmee mijn kennis een nieuwe boost kunnen geven?’

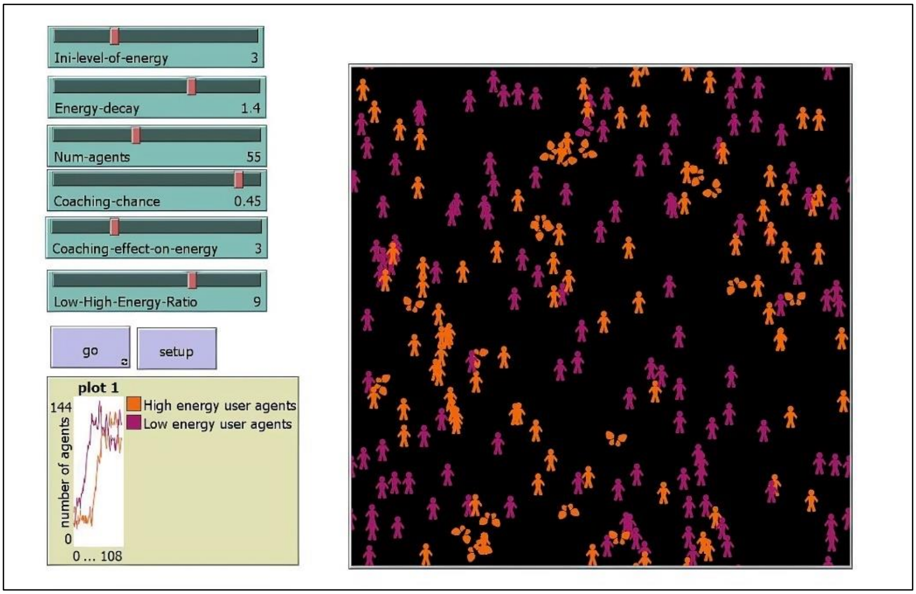
En zo geschiedde. Ik startte mijn promotie-traject aan de UT met Embodied Conversational Agents in eHealth en kwam uiteindelijk vooral uit op het Dynamical Systems Perspective. En voor de derde keer heb ik het idee dat een studie mij iets fundamenteels over de wereld heeft geleerd. Ik hoop dat dit leerproces wederkerig is. M.a.w., ik hoop dat dit proefschrift uiteindelijk niet alleen van waarde blijkt te zijn geweest voor mijn persoonlijke ontwikkeling, maar ook voor de wetenschap. Of belangrijker nog: de gemeenschap.

Mijn grootste dank gaat naar mijn vrouw Simone. Haar steun en het luisteren naar mijn verhalen over waarom het Dynamical Systems Perspective en Complexity Theory zo waardevol is om de wereld te begrijpen, hebben mij ontegenzeglijk heel veel geholpen. Simone, je bent naar mij toe afwisselend motiverend en positief-kritisch geweest. Hopelijk is dat ook precies hoe ik naar jou toe ben. Ik ben zo ontzettend blij met jou als levenspartner.

Vervolgens gaat mijn dank naar Isa en Maud, mijn dochters, inmiddels jonge vrouwen geworden. Geduldig hebben ze geluisterd naar mijn verhalen over mieren die bruggen bouwen over rivieren en over wat dat zegt over hoe de wereld werkt. Intussen dachten ze natuurlijk na over wat ze zo-even nog op hun mobieltjes hadden gelezen ….
Dan wil ik mijn copromotor Saskia bedanken. In voor jou drukke tijden ben je voor mij altijd bereikbaar gebleven voor input en feedback. Je feedback was altijd nuchter en pragmatisch, heerlijk! Dat is dus óók wetenschap. Zonder jou had ik deze promotie niet kunnen doen.
Mijn promotor Lisette wil ik bedanken. Ik herinner me de avond dat je mij belde, nadat we elkaar overdag hadden ontmoet op een andere universiteit waar ik in eerste instantie zou gaan promoveren. Het werd een steal. Je zag in mij dat ik graag de gebaande paden wilde verlaten en gaf aan dat de universiteit Twente en specifiek de PGT vakgroep daar een goede plek voor is. Dat is gebleken. Je hebt me de ruimte gegeven, tot aan het einde toe, toen ik je mijn stellingen voorlegde. Ga je gang. Ik mocht buiten de lijntjes kleuren. Academische vrijheid. Ik waardeer dat ten zeerste.
Ik bedank familieleden en vrienden, die me steevast hebben gemotiveerd. In het bijzonder: tante Annemieke, oom Koos, Marc (tevens paranimf), Menco (ook paranimf) en Karin, Marty en Agnes, mijn neef Robert, mijn nicht Katinka. Ik bedank mijn loop- en praatvrienden Bertil en Miriam. Ik bedank de mensen die op mijn pad gekomen zijn en mij hebben geïnspireerd; mijn Zen-leraren Ingrid en Sandra. De gesprekken met Paul over het houden van je eigen koers. Ik bedank mijn schoonouders Eildert en Roely voor het lekkere eten dat ze hebben gekookt en de ontelbare keren dat ze op Isa en Maud hebben gepast. Ik bedank mijn schoonzus Margit. Ik bedank de Eckartvriendinnen Floor, Nelleke en Veronique. We kennen elkaar nu zo'n 38 jaar, we zien elkaar niet veel en zijn toch dichtbij voor mij.
Ik bedank mijn broer Jeroen voor de mooie momenten die we samen hebben, al wandelend of in de sauna. Ik bedank mijn zus Claudia, voor haar openheid en de inzichten die ze me geeft. Ik bedank mijn bonus-moeder Annette voor de wandelingen in de Treek en de familie-momenten samen. Ik bedank mijn broer Jan-Paul voor het mij betrekken bij de thuiswerktafel: top dat we samen het ondernemerspad zijn opgegaan. Zo leren we elkaar op een nieuwe manier kennen.

Tenslotte bedank ik mijn ouders; Johan en Anneke. Zij konden dit helaas vandaag niet meemaken, maar zouden bijzonder trots op me zijn. Wat betreft het creatieve, vooruitstrevende en analytische was mijn vader Johan mijn voorbeeld. Ten aanzien van het liefdevolle, spirituele (al zou ze zelf dat woord nooit gebruiken) en literaire heb ik heel veel kunnen leren van mijn moeder Anneke.

Pa en ma, dank jullie daarvoor. Dat mijn ouders minstens zo goed konden dansen als een Embodied Conversational Agent en ik, laat onderstaande foto goed zien.



Studies on Embodied Conversational Agents (ECA's; 'robots on screen') have been conducted for more than two decades since the term was formally introduced in the late nineties of the previous century. Even after all these years, ECA's hold the promise of more productive Human-Computer Interaction, but to date their results are still mixed and inconclusive.

Within this thesis it has first been investigated whether ECA's can contribute to improving user adherence to self-guided eHealth interventions. Secondly, main methodological issues with regards to ECA research have been addressed and an alternative approach has been proposed.

To do so, this thesis used insights from a different domain; the Dynamical Systems Perspective (DSP). DSP is a class of mathematical equations that describe time-based systems with properties such as complexity and non-linearity and can be simulated through Agent-Based Models. The principle of rapport is common to both the ECA and DSP domain and created the bridge we needed. Note that rapport is the experience of harmony between conversational partners. It is associated to the notion of 'being in sync'.

This thesis demonstrated that the simulations run with the Agent-Based Model (ABM) were associated to fluctuations with regards to user motivation. The eHealth experiments run with the ECA showed small positive effects due to the ECA's task-related support.

By comparing behavioral states of eHealth user and ECA and assessing their degree of mirroring across multiple modalities (e.g., facial, lexical, postural), the state of the rapport building process can be gauged on a short-term and real-time basis. Noteworthy, these measurements can be taken *during* the eHealth intervention. Likewise, the *fluctuations* in both the user's motivational state and the user-ECA rapport state can be assessed. As an approach, it is recommended to compare the dynamical patterns that arise from these short-term data taken *during* the experiment with the static questionnaire data taken *after* the experiment.

It is expected that this combined dynamical-statical approach will give the ECA in eHealth domain a new impetus. This way, the evidence for effective ECA's can gain strength in order to live up to their long-standing promise.

ISBN 978-90-365-5292-9