Similarity in action with an Embodied Conversational Agent: can synchronous speech yield higher levels of rapport? An exploratory eHealth study.

Mark R. Scholten[1,2] , Saskia M. Kelders[1,3] and Julia E. W. C. Van Gemert-Pijnen[1,4]

[1] Department of Psychology, Health and Technology, Centre for eHealth and Wellbeing Research, University of Twente, Enschede, Netherlands
[2] HU University of Applied Sciences, Utrecht, Netherlands
[3] Optentia Research Focus Area, North-West University, Vanderbijlpark, South Africa
[4] Institute for Biomedical Engineering, University of Waterloo, Canada
m.r.scholten@utwente.nl

Abstract. Self-guided eHealth has the benefit of providing autonomy to patients. However, the autonomy comes with a cost; elevated attrition rates. Embodied Conversational Agents ('robots on screen'), have technological capabilities to build rapport with eHealth users and to support them, but are costly to realize and their evidence is inconclusive. We investigated a novel and low-technological method to build rapport. eHealth users synchronized their speech with a monologue-style ECA, a method for which there exists evidence within the human-to-human communication domain. User experiences were investigated using predominantly qualitative methods. As our study results show, users are fairly positive about speaking synchronously with the ECA. However, the experimental task needs refinements. Users need to priorly hear, at least once, the pace of their artificial interlocutor in order to synchronize. Future studies can further examine the refined synchronous speech task and its potential for more widely accessible rapport-building ECA's aiming for eHealth adherence.

Keywords: Human computer interaction, embodied, conversational agent, virtual human, chatbot, synchrony, eHealth, E-learning, dynamical systems, rapport

1 Introduction

Worldwide, there is a growing demand for health services. Within the mental health domain, diseases such as major depressions are the leading causes of a substantial number of years lived with disability. Within the physical health domain, chronical illnesses, associated with an aging population, equally lead to an increase of the demand for 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 as definition [1] "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. [2] a web-based intervention is "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 self-guided web-based interventions have the advantage of providing autonomy to the patients, their usage statistics expose a salient issue: elevated attrition rates. A review on internet-based treatments for psychological disorders [3], involving minimal therapist contact, found dropout ranges from 2% to 83%, with an overall average of 35%. A review focusing on internet interventions for chronic pain [4] found attrition levels ranging from 4% up to 54%.

Undisputedly, web-based interventions have been designed for usage from start to end. Hence, several studies have been examining root causes of drop-out and low usage. Relevant for the prevention of drop-out is the technology's fit with the user and context [1]. If the users feel the eHealth technology does not meet their personal needs and preferences, or cannot be embedded in their daily routines, it will not be used [1]. So, are web-based interventions currently sufficiently attuned to the needs of users they are aiming to help?

The literature suggests this is often not the case. As described in their scoped review [5], users of self-guided eHealth interventions express they miss out on support, in various ways. First and straightforwardly, many patients experience a lack of *encouragement* after having fulfilled the tasks the web-based intervention demanded them to do. Furthermore, patients miss out on *acknowledgement* for daily issues they are struggling with, such as pain and sleeping difficulties.

All the same, computers can be endowed with user-oriented and supportive qualities (e.g., [6]), according to the tenets of affective computing [7]. By adding supportive functionality that is personal and visible (e.g., represented as an animated face and/or body) and based on natural language, a feeling of intimacy can be induced amongst users [8]. 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).

Relevant for the subject of this paper, ECA's are suitable candidates to provide support to users of web-based interventions and can potentially provide a remedy to elevated attrition rates [9].

1.1 Current Evidence for ECA's, Similarity Is a Cornerstone

Functionalities of ECA's have been reviewed [10, 11] with the aim of finding evidence for what kind of ECA fits best within which context. As already postulated back in 2000 [11] and in more recent studies [12, 13]; similarity between ECA and user consistently lends to their approval.

The similarity can have various origins. First, the similarity can be static, e.g., the user and ECA can have the same gender, age or ethnicity. Next, the similarity can concern dynamical features such as users and ECA's behaving in similar ways (e.g.,

same level of extraversion). Finally, contextual similarity is relevant, such as an ECA fulfilling its role (e.g., student, tutor) comparable to how a human would typically fulfill it. ECA-user similarity effects can be clarified through social psychology theories. The similarity attraction theory [14], the social identity theory [15] and the self-categorization theory [16] all support and explain the idea that people are attracted to, prefer and support relationships with similar others. Further building on their similarity characteristics, it has been suggested that ECA's can create relationships with users. Their life-like appearance, as exposed by their embodiment, posture movements, blinking eyes, and the visibility of their respiration rhythm, together with their capabilities to communicate via speech or text, instills feelings of connection amongst users.

Finally, mutual visual attention between ECA and user has been demonstrated to be a productive relational feature. An ECA showing human-like gaze behavior is evaluated more positively regarding usability and involvement than an ECA showing randomly determined gaze behavior [17]. Reversely, users demonstrate visual attention towards the ECA, aligned with social interaction [18].

1.2 Building Rapport as a Pre-condition for Successful Human-ECA Interactions

Within dyadic human-ECA studies, a universal concept resulting from the productive ECA's attention, is rapport. Rapport is a smooth, positive interpersonal interaction [19], a harmonious form of relationship building [20]. Rapport has been studied across a range of scientific disciplines for its role in fostering emotional bonds and prosocial behavior.

Specifically, within the field of social psychological research, the role of rapport in mediating ultimate outcomes on social interaction is well established. That is, rapport is argued to provide foundations for social engagement [21], positive teacher–student interactions [22], success in negotiations [23], psychotherapeutic effectiveness [24], and enhanced quality of childcare [25].

In their seminal article, Tickle-Degnen and Rosenthal [26] equate rapport between humans with behaviors indicating positive emotions (e.g., head nods and smiles), mutual attentiveness (e.g., mutual gaze), and coordination (e.g., postural mimicry and synchronized movements). For studies on human-ECA communication Embodied Conversational Agents have been created that effectively show attention and demonstrate positivity [27], fostering elevated levels of rapport as measured amongst their users. For this purpose, ECA's effectively express both non-verbal and verbal information, serving as complementary communication channels, see the coming section.

First, ECA's are endowed with the capacity to timely respond to users using non-verbal language such as positive nodding and smiling [28]. This type of responsive behavior is also known as non-verbal *entrainment* [29] and stimulates the human interlocutor while speaking.

Within an experimental human-ECA setting, it has been successfully demonstrated how non-verbal entrainment behavior (i.e., positive ECA nodding) led to elevated levels of rapport amongst participants [30]. In a similar vein, the importance of similar non-verbal behavior, being hand gestures and head nods, during dyadic human-ECA

communication have been investigated [31]. Within their study the authors assessed the effect of the amplitude of non-verbal behaviors expressed by an ECA across two conversational sessions, in time separated by at least one day. As their results showed, non-verbal entrainment created a significant level of rapport, but increasing the entrainment amplitude during the second session did not further increase rapport. It has been argued [32] 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 summarized by Bickmore and Picard (p. 2) [33] "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.".

ECA's have a second instrument to create rapport with the user, utilizing pure *verbal* information, such as autobiographical stories and small talk. Indeed, differential rapport building effects of ECA narratives have been empirically found [34].

Importantly, there is evidence for user-ECA rapport resulting in lower levels of attrition in eHealth. First, as stated in [35] ECA's that interact with the user can build trust and rapport, leading to companionship and long-term, continual use. Reversely, it has been reported that the *lack* of user-ECA rapport yields the opposite effect: lower levels of adherence [36]. Within an adjacent field, E-learning, equal positive causal effects of user-ECA rapport on user adherence levels have been found. As stated in [37]; rapport building ECA's result in higher learner motivation which leads to lower dropout ratios.

1.3 ECA Research Stalemate Requires New Technological and Empirical Directions

Despite the beforementioned positive ECA effects, reviews [10, 11, 37, 38] on the effectiveness of ECA's concluded that, generically speaking, their evidence is mixed and inconclusive. In 2016, this conclusion was formulated [10] as: "Nearing two decades of intense study of the topic, researchers cannot say with much certainty the level of effectiveness one would expect to see ...". An evaluation of the study field of conversational agents that lack human-like embodiments (e.g., Siri®, Alexa®) [38] led to an even bolder conclusion: "like having a really bad PA".

Along with the acknowledgement of the lack of strong and convincing evidence for conversational agents, the way ECA studies are set up, has been analyzed. With the start of each new ECA research project, the design and build of the ECA platform is started anew, which requires considerable budget and efforts [40]. As a solution, a generic technological platform for the study of ECA's has been suggested [40]. Indeed, several universal ECA study platforms have been put into practice [40–42], but use has been limited so far.

A recent review [34] puts emphasis on the lack of design standards for ECA's. As the authors state, consensus on design features of ECAs in eHealth is far from established and hinders progress in the ECA study field.

Meanwhile, optimism with regards to technology is a primary factor within many ECA studies [43–45]. Confidence is expressed with regards to ECA's that will profit from

conversational AI developments, even though such agents are capital-intensive to realize and create new design challenges [44]. Moreover, reviews on this new generation of AI-based ECA's have so far and again reported on mixed and inconclusive results on their effectiveness, usability, and satisfactoriness [46, 47].

A pragmatic review [48] therefore advises to take a technological step back. The authors advocate the usage of low-tech ECA's as best in class for routine clinical eHealth contexts.

The low-tech ECA as a category refers to an ECA that mainly sends out information. A low-tech ECA barely receives, let alone processes user state information. Stated differently, this kind of ECA does not use affect sensing and entrainment functionalities which would enable the ECA to connect with users. For the contrast with a high-tech ECA, see the description in section 1.2. In concrete terms, a typical low-tech ECA consists of a non-personalized automated textual feedback mechanism accompanied by a static photograph of a clinician [48]. Low-tech ECA's have the advantage of their wide availability, low cost and fairly easy implementation. These advantages can have considerable positive implications for the study field as a whole. Low-tech ECA's create a condition for replication of ECA studies and of comparison of ECA study results [11, 49].

Summarized, after almost twenty-five years of study, the ECA study field still holds promise. However, despite the ongoing technological progress made, the study domain consistently demonstrates challenges in providing univocal evidence. As technological progress has not done the trick yet, it makes sense to involve less technology-driven views on ECA's and to explore adjacent scientific domains. Next to the existing technology-driven rapport building techniques as described in the previous section, there is an adjacent rapport building technique available: the task-based inducement of behavioral synchrony between ECA and user.

1.4 Interpersonal Synchrony Informs Rapport

Human-to-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 [20, 50, 51]. Within mundane contexts, synchronous vocalization rituals, like group singing during birthday parties and sport team yelling on the pitch, are culturally well-embedded rituals to create feelings of cohesiveness.

Moving in synchrony (i.e., with overlap in time and form) is argued to influence the degree to which individuals are externally perceived as a social unit [51, 52] . But also, individuals themselves report experiences of belonging to one and the same team. As is said, the individuals become embedded within a new social unit, which creates a reality of its own. On a neural level this effect is explained by pathways that code for both action and perception [53] which causes blurring of the self and the other. Overall, meta-analyses on interpersonal synchrony [54, 55] report on positive effects, both on pro-social behavior and pro-social attitudes. Within the review of Mogan et al. [55], the pro-social effects are specified as social bonding, social cognition and positive affect, all with significant outcomes. On a higher descriptive level these synchrony effects between individuals are also referred to as dynamically coupled systems [56], associated to emergence [57] and non-linear mathematics [58], pertaining to the field

of the Dynamical Systems Perspective [59] with high relevance to the psychology field [60].

Finally, and important to note, setting a common goal before and during syncing enables people to predict each other's cooperation in the future [54] and bolsters the synchrony effect [54, 55].

1.5 Interpersonal Synchrony Applied within a Human-ECA Context

Can the positive effects of interpersonal synchrony be applied to human-ECA communication? Intuitively, such an approach looks encouraging. A synchronous activity is also referred to as *similarity in action* [61], semantically adjacent to *similarity* between humans and ECA's with an existing evidence base [12, 13]. But *how* can this interpersonal synchrony be empirically induced by conversation agents? In human-robot studies, asking participants to move together with the robot seems to be an acceptable activity [62, 63]. In contrast, human-ECA environments do not usually provide such a shared physical space. Therefore, speaking together seems to be the more practical alternative. Synchronous speech has already shown evidence within human-to-human studies [64, 65], specifically for the outcome measures cooperation and rapport [66, 67].

In addition, the synchronized speech task has been recommended for human-ECA communication [68]. Moreover, there is empirical evidence for synchronizing speech with humans recorded on video. Harmon-Jones (2011) [69] deployed three videos each containing multiple models that either sang or spoke out nonsense phrases for 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, synchronous speaking significantly increased the participants' affiliation with the artificial models, compared to mere listening. Then again, although empirical progress has been made, a task and user evaluation study of the synchronous speech task within a human-ECA context has not yet been carried out, to the best of our knowledge.

The present study aims to fill this gap. It will apply the human-to-human synchronous speech task [64–66] within a human-ECA context and evaluate the task and its acceptability. Secondly, our study aims to find indications that this speech task creates rapport between study participants and ECA. Thirdly, the study aims to examine whether phrases that semantically contain teamwork information (based on the common goal evidence as reported on in the literature [54, 55]), will lead to an elevated level of rapport.

To do so, we will apply a predominantly qualitative method for assessing the user's opinion on receiving motivational instructions from an ECA that has been added as an adjunct to an eHealth psycho-education intervention. Psycho-education is a branch within eHealth that makes patients aware of their situation and stimulates coping and acceptance [70, 71]. Within our experiment, we will offer psycho-education about positive psychology [72]. Positive psychology, 'the study of happiness' is a subject of general interest which makes it suitable for a broad range of psycho-education experiments.

Along these lines, we keep our ultimate objective in mind; an easily implementable rapport-building ECA can potentially provide a practical remedy for non-adherence to web-based eHealth interventions.

2. Methods

2.1 Design

To answer the beforementioned three research questions (i.e. firstly on user acceptance of the synchronous speech task, secondly on indications that the synchronous speech task fosters rapport, thirdly examining a common goal effect), an experiment was set up. Sixteen 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 monologue-style ECA, see figure 1 below with Brian uttering "I am Brian your virtual coach, welcome to the experiment."

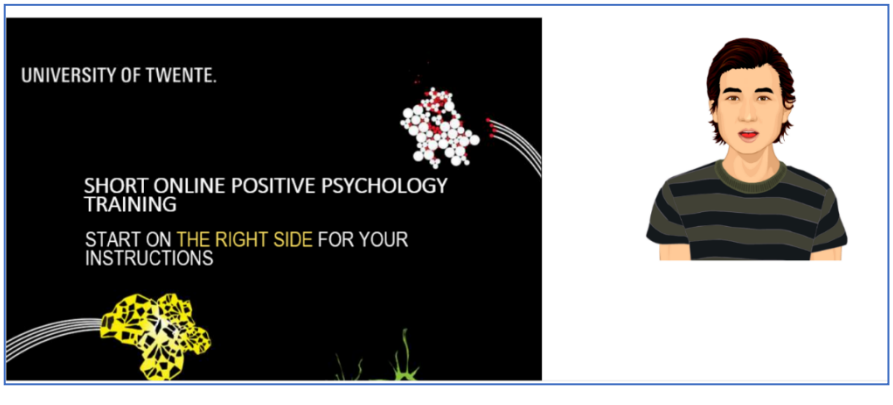

Fig. 1. Webpage that displays the online course on the left side and the welcoming ECA Brian on the right side of the screen

Then the participants were asked to speak out phrases. These phrases pertained to three speech conditions, of which the details will be described in section 2.2. Subsequently, the participants started with the online course. For the course details, see section 2.3. The entire experiment took approximately 18 minutes for the participants to accomplish. See figure 2 below for an overview of the experimental events in sequence order.

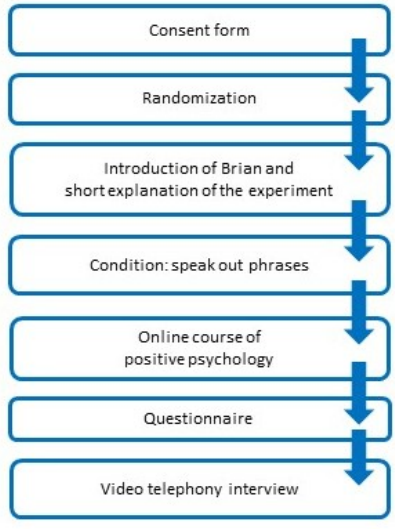

Fig. 2. Flowchart of activities carried out by participants during the experiment

2.2 Speech Conditions

After the introduction of Brian, the study participants were asked to speak. There were three different speech conditions, to which participants were randomized, as depicted in figure 3 below. One participant was excluded from the analysis, after it became apparent that they had encountered too many technical difficulties resulting in (the user) not being able to click through the flow of the online course All speech conditions contained an equal number of 39 words put within four phrases.

Condition 1 (phrases containing teamwork semantics, 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 (semantically neutral phrases, 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 (semantically neutral phrases, 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.

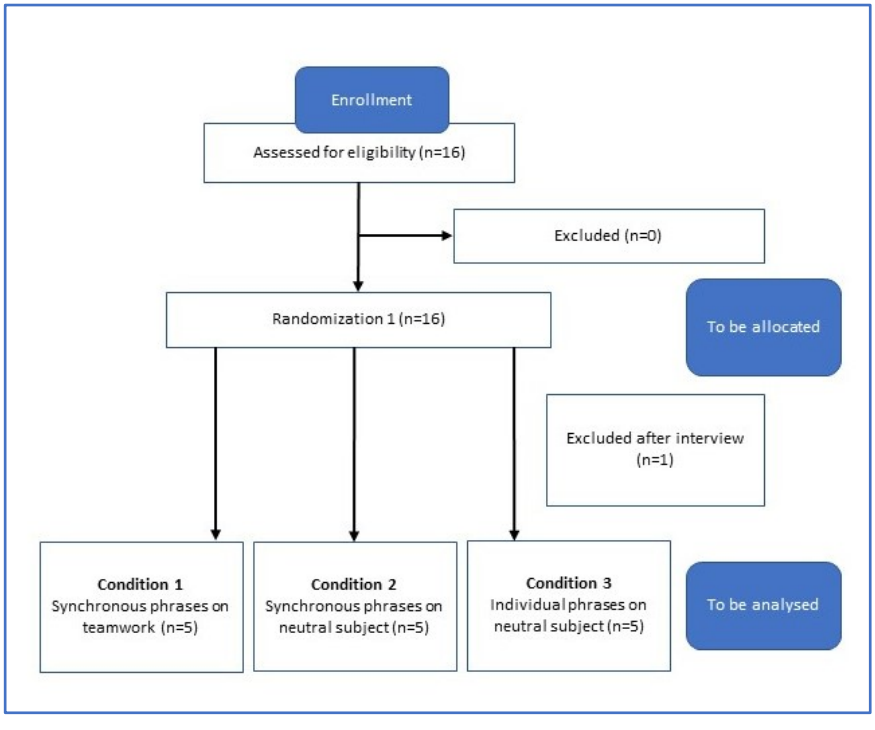

Fig. 3. CONSORT Flow diagram of the experiment including the randomization of the speech conditions

Fig. 4. Webpage that displays four phrases pertaining to the speech task on the left side and supportive ECA Brian on the right side (condition one phrases are shown)

For all these three speech conditions the speech procedure was explained by Brian. Subsequently, Brian started a count-down after which the participants were asked to speak out loud their phrases. Figure 4 provides an example of the phrases as presented on the left side of the screen and the supportive ECA Brian on the right side of the screen.

After speaking out the phrases, the participants were thanked and then clicked through, to start the online course.

2.3 Online Course and User Guidance

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 [72]. 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* [73].

User guidance and support were provided on the right side of the webpage by the ECA Brian. The ECA condition was created through the Voki® application. As depicted in **Figure 1** and **Figure 4** above, 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 ECA showed lip synchronization and animation properties such as blinking eyes and a breathing chest. Furthermore, the ECA's line of sight followed the cursor movements of the user and as such demonstrated visual attention towards the user. In addition, the ECA expressed guidance conveyed in English.

After each piece of guidance and information, the ECA asked the user “Please click on the spacebar 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., “Well done!”). Note that this set-up has been used within a previous human-ECA eHealth study [74].

2.4 Recruitment of Participants

We recruited bachelor and master psychology and communication students at our university. 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 university’s Institutional Review Board with number 201009. In total 16 participants were included. Participants were on average 19.63 years of age and represented four 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%), two participants were male (12.5%).

2.5 Procedure and Outcome Measures

As mentioned, the study's objective was to assess three topics, being the experience with a) the synchronous speech task b) with the ECA Brian, and c) with the psycho-education intervention. Furthermore, the study had two measurement objectives, first and foremost a qualitative evaluation to gain new fundamental insights, and secondly a short quantitative evaluation. To cover these mixed objectives, we decided to self-develop a questionnaire consisting of an open/qualitative and closed/quantitative section.

For the open part, we formulated questions about the user experience of the course and the ECA. For the closed part we borrowed insights from existing questionnaires in the ECA study field, being on feeling accepted as a user by the ECA [75], on usefulness of the ECA [76] and usefulness of the course [77], and finally on motivational capabilities of the ECA [78].

At the end of the experiment, participants filled out our questionnaire in the Qualtrics® system. In concrete, this questionnaire contained the following 10 questions:

1. What did you think about the online course? (open)
2. To what extent is the course's content useful for you? (closed, 1-5)
3. How did you feel doing the online exercises? (open)
4. What is your opinion on doing a course like this online? (open)
5. What did you think of Brian? (open)
6. Did Brian give useful advice? (closed, 1-5)
7. To what extent was Brian a motivating virtual person? (closed, 1-5)
8. To what extent did you feel accepted by Brian? (closed, 1-5)
9. We have asked you to speak out the phrases aloud. How did you experience this? (open)
10. If you would be asked to do this during a true online course, would you do it? (open)

After filling in the questions, the users were asked to e-mail the researcher for the interview to start. As the study was carried out during Covid-19, the interview was done through Microsoft TEAMS®. The objective of the interview was to gather experiences on the user task of speaking aloud and to find strengths and weaknesses of the ECA at formulated at a more personal level than in our questionnaire.

With regard to the open questions: we asked about the course, Brian and the speaking aloud task. With regards to the closed questions; we used *cooperation*, *trust*, *unity/team* which we found as measures within studies with synchronizing humans [50, 67]. In addition, we used *bond* as a measure that we found both within synchronous movement studies [67, 79] as in studies about relationship building with a health context [80].

This way, the overall view of the participants on the online course was assembled. We asked the following questions during the interview:

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

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 thematic analysis [81].

The four main themes used for the analysis were defined along with the research questions of the experiment. See **Table 1** below for the research questions and the themes of this study.

Table 1. The research questions and the themes of this study

Research Questions	Related themes
RQ 1: Acceptability of the synchronous speech task	Theme 1: the user experience of the task of speaking aloud Theme 4: the experience with the course
RQ 2: Indications that the synchronous speech task fosters rapport (and consequently results in a more positive evaluation of the ECA)	Theme 2: strengths of the ECA Theme 3: weaknesses of the ECA
RQ 3: Examining a common goal effect through semantics of speech (and consequently results in a more positive evaluation with the course)	Theme 4: the experience with the course

Further divisions on themes were made after data collection, based on the input data provided both within the questionnaire and during the interview. For example, as strengths of the ECA (theme 2) users spontaneously mentioned as sub-themes: 'positive user experience', 'autonomy', 'clear structure', 'voice', 'gazing behavior'.

In detail, six steps were carried out performing our thematic analysis, following the descriptions of [82]. 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 subthemes. Fourth, initial subthemes 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 subthemes together that had a close semantic resemblance. The fifth stage involved the process of describing and labelling the final subthemes. 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, in addition sub-themes emerged from both the questionnaire and interview data, see **Table 2** below. Note that the prompted answers relate to the closed questions and the spontaneous answers to the open questions and the elaborations provided by the user on the closed questions.

Table 2. Themes and sub-themes 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 with the course	**Number of participants**
Positive	12 (spontaneous)
Neutral	2 (spontaneous)
Negative	1 (spontaneous)

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 (three out of 15) were straightforwardly positive; *I liked it. With these kinds of exercises, I am 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 (nine 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 at the same speed as Brian) than when they ask me to do it as a standard audio-recording [Participant #1].* Finally, some (three 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 pace as Brian did because he pronounced some words or sentences different from how I would usually do this [Participant #9].*

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 (four 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 four out of 15 participants) with regard 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 must 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 mentioned by quite a few (five 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 (eight 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].*

Theme 3: Weaknesses of the ECA. Many (six 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 unilateral**. *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 make it difficult, but I think you could form a relationship that is more like 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 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. 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].*

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 didn't give me a happy, positive feeling [Participant #14].*

3.2. Quantitative Results

Although we conducted our study taking an explorative and qualitative perspective, we ran a small quantitative analysis on a 1-5 scale on the four closed questions we posed. Below, in figures 5, 6 and 7, the means, per condition, of the four closed questions are presented.

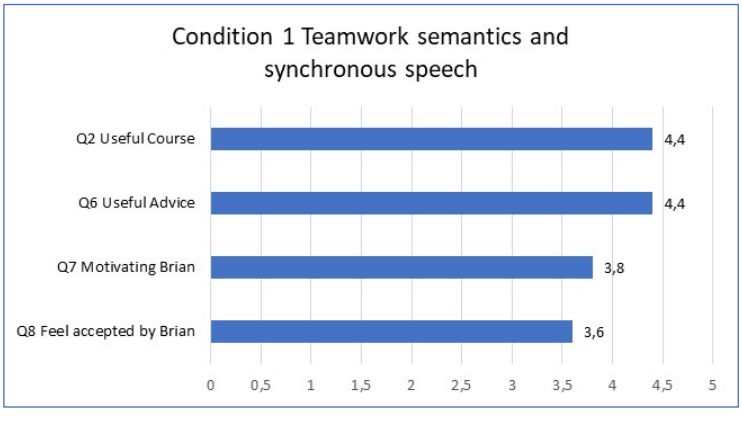

Fig. 5. Condition 1 Teamwork semantics and synchronous speech: the mean values of the four closed questions

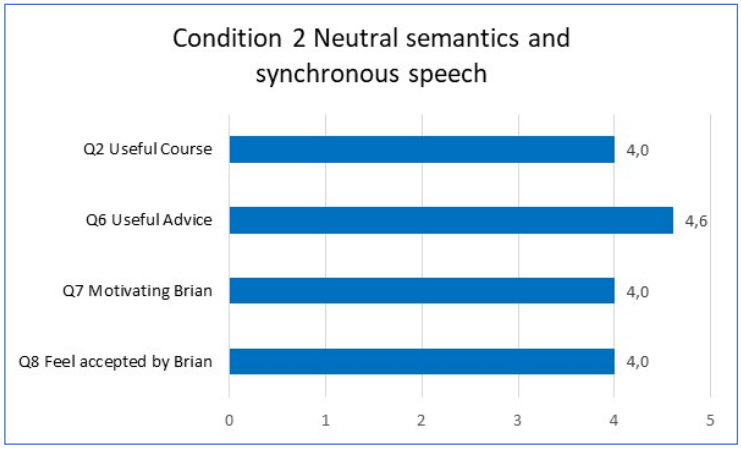

Fig. 6. Condition 2 Neutral semantics and synchronous speech: the mean values of the four closed questions

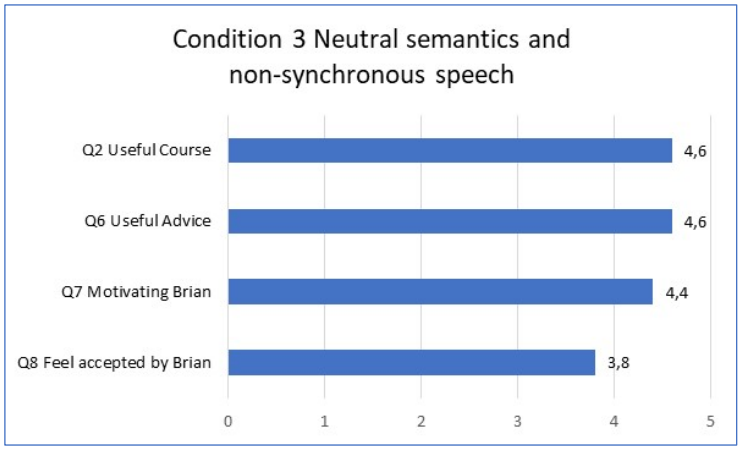

Fig. 7. Condition 3 Neutral semantics and non-synchronous speech: the means of the four closed questions

As the data were on an ordinal scale we conducted a Kruskal-Wallis test to examine the differences on appreciation of Brian and the course. No significant differences (Chi square = .99, p = .61, df = 2) were found among the three conditions.

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/ shortcomings: what relationship building features (and the lack thereof) were brought up 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, several participants indicated that the way the ECA would pronounce the phrases was unpredictable. The pace and rhythm of the ECA's speech became somewhat hard to synchronize; the participants had to improvise. Furthermore, for the participants the rationale of the task of speaking out loud was missing. These findings were in line with the results of our exploratory quantitative analysis. This quantitative analysis provided no indications that differential levels of rapport were created because of synchrony or teamwork semantics. Altogether, these results demonstrate that our synchronous speech task has not fully matured yet and needs several refinements.

First, the participants should be given the opportunity **to warm up and practice speaking**. When they were asked to speak synchronously with the ECA, they hadn't heard the ECA speaking out the phrases before. The participants should get acquainted to the ECA's speaking pace and rhythm by listening to it, at least once. Consequently, future participants will experience successful completion of the task. We consider this as a necessary pre-condition for the rapport effect we seek. An earlier study on synchronous speech involving two humans [66] 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. Note that, in contrast, in our human-ECA set-up, this fusion was never mentioned by the participants. This absence may be explained by the fact that the pre-recorded speech of the ECA could not adapt itself 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 found [83], participants can effectively synchronize with a non-adaptive computer. However, according to the authors, performance *does* deteriorate if the computer additionally acts in irregular and unpredictable ways, which was the case in our study according to our participants.

Second, **the rationale for doing the task** should have been accounted for by the ECA more explicitly. As stressed by [84] explanation of events using mental models

increases task performance when users interact with computers (see e.g., [85, 86]). For a future experiment, we therefore foresee a more elaborate introduction such as "I would like to bond with you. As a means, I would like you to speak out several phrases aloud together with me, as to warm up for the experiment and to start cooperating with you." Likely, this would have provided sense to this activity towards the participants. 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 will respond in one way or another. See e.g., [87] "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" (p. 177). In a similar vein, Bickmore and Cassell underscore [88] 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 summary, the participants had missed the experience of an ECA demonstrating listening behavior. Consequently, as a practical improvement of the set-up, the participant's speech could be recorded and subsequently played out by the ECA to signal 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 Limitations

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 [74] and gaze [17]. In addition, participants 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 limitations. 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 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 failed to fulfill. As participants were literally missing something reciprocal, this one-sided communication influenced their experienced relationship with the ECA. Not surprisingly therefore, participants mentioned that they 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 genuinely experienced involvement from the ECA's side.

4.4 Ways to Work Around the ECA's Relational Limitations

Participants referred to examples from the film industry such as human actors and cartoon figures to point out that a two-sided 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 smartly utilized for the revelation of the ECA's involvement. As previous studies have demonstrated [33], an ECA that exposes itself to an eHealth user by means of an autobiographical story, will create 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."). Based on the participant's feedback, future ECA's can be equipped with a longer autobiographical story. Practically, we think of an introduction like "I was made in 2019 at this university. 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. Please call me Brian." This way the ECA would move away from the instructor role towards more of a companion-like role and as such take away some social distance.

5. Limitations

Conclusions on ECA research are known to be 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 the user finalized a learning task. This way we avoided direct distraction from the ECA toward the user.

Concerning the study's conditions, we decided to use two conditions to cater for a semantic and a synchronization effect and compared them to a third control condition. We decided not to apply a fourth condition, teamwork semantics combined with non-synchronized speech, that theoretically would isolate the semantic effect. We did this for methodological reasons. That is, such a condition would have led to a participant talking about teamwork to a passive ECA. As an event, this would represent rapport-building by the participant and not by the ECA, which would not fit within the scope of our study.

Concerning the size of our study; we included 16 participants. A future, non-exploratory study will include more participants (around 15-20 participants per condition) to cater for a more robust quantitative analysis.

6. Conclusions and Future Work

We conducted a user-evaluation study on the synchronous speech task, novel to the ECA study field, to create rapport between a monologue-style 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.

The participants provided important insights with regards to the task to speak synchronously with an ECA. As our study results showed, the experimental task needs several practical refinements. With these refinements, the potential of cost-efficient and functionally modest ECA's, can be genuinely assessed within eHealth contexts. Although we cannot expect that these ECA's will support users as well as human caregivers can, it seems fair to hypothesize that web-based interventions equipped with these ECA's will perform better than the current baseline. This baseline is set by self-guided web-based interventions that face considerable attrition levels, partly due to not sufficiently providing user support.

On a higher descriptive level, we crossed scientific domains. We applied recent insights on rapport building from the human-to-human communication and dynamical systems domains to the human-ECA domain, as to supplement the current high technology-based ECA's.

Our stance is that the refined synchronous speech task warrants further investigation during future human-ECA studies. Much alike the eHealth study field, the human-ECA field shows promise, but faces notable challenges in materializing. Fresh ideas are recognized to prevent a too strong an entanglement with high-tech ECA's, having relative low accessibility, and still yielding modest results. A promising new avenue is the synchronized speech task, similarity in action, having demonstrated positive empirical effects within the human-to-human communication domain. Using a refined synchronous speech task, it can be assessed whether speaking in unison with an accessible ECA is not merely figure of speech but can genuinely provide evidence for user-ECA rapport and adherence to eHealth interventions.

CRediT author statement. Mark Scholten: conceptualization, investigation, methodology, data curation, and writing - original draft. **Saskia Kelders:** conceptualisation, methodology, writing - review and editing. **Julia E. W. C. Van Gemert-Pijnen:** conceptualisation.

References

1. van Gemert-Pijnen L. (J. E.W.C.), Kip H., Kelders S.M., Sanderman R.: Introducing eHealth eHealth Research, Theory and Development. pp. 3-26. Routledge (2018) https://doi.org/10.4324/9781315385907-1
2. Barak A., Klein B., Proudfoot J.G.: Defining Internet-Supported Therapeutic Interventions Annals of Behavioral Medicine, 38, pp. 4-17 (2009) https://doi.org/10.1007/s12160-009-9130-7 PMid:19787305
3. Melville K.M., Casey L.M., Kavanagh D.J.: Dropout from Internet-based treatment for psychological disorders British Journal of Clinical Psychology, 49, pp. 455-471 (2010) https://doi.org/10.1348/014466509X472138 PMid:19799804

4. Buhrman M., Gordh T., Andersson G.: Internet interventions for chronic pain including headache: A systematic review Internet Interventions, 4, pp. 17-34 (2016) https://doi.org/10.1016/j.invent.2015.12.001 PMid:30135787 PMCid:PMC6096254
5. 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. https://doi.org/10.2196/jmir.7351 PMid:29146567 PMCid:PMC5709656
6. Scholten, M. R., Kelders, S. M., & Van Gemert-Pijnen, J. E. (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. https://doi.org/10.3389/fpsyg.2019.01063 PMid:31143152 PMCid:PMC6520839
7. Picard R.W.: Affective computing: challenges International Journal of Human-Computer Studies, 59, pp. 55-64 (2003) https://doi.org/10.1016/S1071-5819(03)00052-1
8. Potdevin D., Clavel C., Sabouret N.: Virtual intimacy in human-embodied conversational agent interactions: the influence of multimodality on its perception Journal on Multimodal User Interfaces, 15, pp. 25-43 (2021) https://doi.org/10.1007/s12193-020-00337-9
9. Abdulrahman, A., & Richards, D. (2021). In Search of Embodied Conversational and Explainable Agents for Health Behaviour Change and Adherence. Multimodal Technologies and Interaction, 5(9), 56. https://doi.org/10.3390/mti5090056
10. DEHN D.M., van MULKEN S.: The impact of animated interface agents: a review of empirical research International Journal of Human-Computer Studies, 52, pp. 1-22 (2000) https://doi.org/10.1006/ijhc.1999.0325
11. Veletsianos G., Russell G.S.: Pedagogical Agents Handbook of Research on Educational Communications and Technology. pp. 759-769. Springer New York, New York, NY (2014) https://doi.org/10.1007/978-1-4614-3185-5_61
12. Baylor A.L., Ryu J.: The Effects of Image and Animation in Enhancing Pedagogical Agent Persona Journal of Educational Computing Research, 28, pp. 373-394 (2003) https://doi.org/10.2190/V0WQ-NWGN-JB54-FAT4
13. Shamekhi A., Czerwinski M., Mark G., Novotny M., Bennett G.A.: An Exploratory Study Toward the Preferred Conversational Style for Compatible Virtual Agents Presented at the (2016) https://doi.org/10.1007/978-3-319-47665-0_4
14. Byrne D.: An Overview (and Underview) of Research and Theory within the Attraction Paradigm Journal of Social and Personal Relationships, 14, pp. 417-431 (1997) https://doi.org/10.1177/0265407597143008
15. Tajfel H.: Social identity and intergroup behaviour Social Science Information, 13, pp. 65-93 (1974) https://doi.org/10.1177/053901847401300204
16. Turner J.C., Reynolds K.J.: Self-Categorization Theory Handbook of Theories of Social Psychology. pp. 399-417. SAGE Publications Ltd, 1 Oliver's Yard, 55 City Road, London EC1Y 1SP United Kingdom (2011) https://doi.org/10.4135/9781446249222.n46
17. Heylen D., van Es I., Nijholt A., van Dijk B.: Controlling the Gaze of Conversational Agents Presented at the (2005) https://doi.org/10.1007/1-4020-3933-6_11
18. Louwerse M.M., Graesser A.C., McNamara D.S., Lu S.: Embodied conversational agents as conversational partners Applied Cognitive Psychology, 23, pp. 1244-1255 (2009) https://doi.org/10.1002/acp.1527
19. Abbe, A., & Brandon, S. E. (2014). Building and maintaining rapport in investigative interviews. Police practice and research, 15(3), 207-220. https://doi.org/10.1080/15614263.2013.827835
20. Tickle-Degnen L., Rosenthal R.: The Nature of Rapport and Its Nonverbal Correlates Psychological Inquiry, 1, pp. 285-293 (1990) https://doi.org/10.1207/s15327965pli0104_1
21. Tatar D.: Social and personal consequences of a preoccupied listener Unpublished doctoral dissertation, Department of Psychology, Stanford University, Stanford, CA, (1997)

22. Bernieri F.J.: Coordinated movement and rapport in teacher-student interactions Journal of Nonverbal Behavior, 12, pp. 120-138 (1988) https://doi.org/10.1007/BF00986930
23. Drolet A.L., Morris M.W.: Rapport in Conflict Resolution: Accounting for How Face-to-Face Contact Fosters Mutual Cooperation in Mixed-Motive Conflicts Journal of Experimental Social Psychology, 36, pp. 26-50 (2000) https://doi.org/10.1006/jesp.1999.1395
24. Tsui P., Schultz G.L.: Failure of rapport: Why psychotherapeutic engagement fails in the treatment of Asian clients. American Journal of Orthopsychiatry, 55, pp. 561-569 (1985) https://doi.org/10.1111/j.1939-0025.1985.tb02706.x PMid:4073227
25. Burns M.: Rapport and relationships: The basis of childcare. Journal of Child Care, 2, pp. 47-57 (1984)
26. Morbini F., DeVault D., Georgila K., Artstein R., Traum D., Morency L.-P.: A Demonstration of Dialogue Processing in SimSensei Kiosk Proceedings of the 15th Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL). pp. 254-256. Association for Computational Linguistics, Stroudsburg, PA, USA (2014) https://doi.org/10.3115/v1/W14-4334
27. Huang L., Morency L.-P., Gratch J.: Virtual Rapport 2.0 Presented at the (2011) https://doi.org/10.1007/978-3-642-23974-8_8
28. Kenny P., Hartholt A., Gratch J., Swartout W., Traum D., Marsella S., Piepol D.: Building interactive virtual humans for training environments Proceedings of i/itsec. vol. 174. pp. 911-916 (2007)
29. Gratch J., Kang S.-H., Wang N.: Using Social Agents to Explore Theories of Rapport and Emotional Resonance Social Emotions in Nature and Artifact. pp. 181-197. Oxford University Press (2013) https://doi.org/10.1093/acprof:oso/9780195387643.003.0012
30. Novick D., Gris I.: Building Rapport between Human and ECA: A Pilot Study Presented at the (2014) https://doi.org/10.1007/978-3-319-07230-2_45
31. Weiss B., Wechsung I., Kühnel C., Möller S.: Evaluating embodied conversational agents in multimodal interfaces Computational Cognitive Science, 1, pp. 6 (2015) https://doi.org/10.1186/s40469-015-0006-9
32. Bickmore T., Picard R.: Subtle expressivity by relational agents Proceedings of the CHI 2003 Workshop on Subtle Expressivity for Characters and Robots (2003)
33. Bickmore T., Schulman D., Yin L.: MAINTAINING ENGAGEMENT IN LONG-TERM INTERVENTIONS WITH RELATIONAL AGENTS Applied Artificial Intelligence, 24, pp. 648-666 (2010) https://doi.org/10.1080/08839514.2010.492259 PMid:21318052 PMCid:PMC3035950
34. ter Stal S., Kramer L.L., Tabak M., op den Akker H., Hermens H.: Design Features of Embodied Conversational Agents in eHealth: a Literature Review International Journal of Human-Computer Studies, 138, pp. 102409 (2020) https://doi.org/10.1016/j.ijhcs.2020.102409
35. Bickmore T., Schulman D., Yin L. (2010). Maintaining engagement in long-term interventions with relational agents. Appl. Artif. Intell. 24 648-666. 10.1080/08839514.2010.492259 [PMC free article] [PubMed] [CrossRef] [Google Scholar] https://doi.org/10.1080/08839514.2010.492259 PMid:21318052 PMCid:PMC3035950
36. Fitton, I. S., Finnegan, D. J., & Proulx, M. J. (2020). Immersive virtual environments and embodied agents for e-learning applications. PeerJ Computer Science, 6, e315. https://doi.org/10.7717/peerj-cs.315 PMid:33816966 PMCid:PMC7924662
37. Schroeder N.L., Gotch C.M.: Persisting Issues in Pedagogical Agent Research Journal of Educational Computing Research, 53, pp. 183-204 (2015) https://doi.org/10.1177/0735633115597625
38. Johnson W.L., Lester J.C.: Face-to-Face Interaction with Pedagogical Agents, Twenty Years Later International Journal of Artificial Intelligence in Education, 26, pp. 25-36 (2016) https://doi.org/10.1007/s40593-015-0065-9

39. Luger E., Sellen A.: "Like Having a Really Bad PA" Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. pp. 5286-5297. ACM, New York, NY, USA (2016) https://doi.org/10.1145/2858036.2858288
40. Huang H.-H., Masuda T., Cerekovic A., Tarasenko K., Pandzic I.S., Nakano Y., Nishida T.: Toward a Universal Platform for Integrating Embodied Conversational Agent Components Presented at the (2006) https://doi.org/10.1007/11893004_28
41. Takahashi T., Takeda H.: Proposal of a script language for embodied conversational agents as asynchronous conversational media Proceedings of the first international joint conference on Autonomous agents and multiagent systems part 3 - AAMAS '02. p. 1387. ACM Press, New York, New York, USA (2002) https://doi.org/10.1145/545056.545136
42. Rojc M., Mlakar I., Kačič Z.: The TTS-driven affective embodied conversational agent EVA, based on a novel conversational-behavior generation algorithm Engineering Applications of Artificial Intelligence, 57, pp. 80-104 (2017) https://doi.org/10.1016/j.engappai.2016.10.006
43. Masche J., Le N.-T.: A review of technologies for conversational systems International conference on computer science, applied mathematics and applications. pp. 212-225. Springer (2017) https://doi.org/10.1007/978-3-319-61911-8_19
44. Allouch M., Azaria A., Azoulay R.: Conversational Agents: Goals, Technologies, Vision and Challenges Sensors, 21, pp. 8448 (2021) https://doi.org/10.3390/s21248448 PMid:34960538 PMCid:PMC8704682
45. McTear M.: Conversational AI: Dialogue Systems, Conversational Agents, and Chatbots Synthesis Lectures on Human Language Technologies, 13, pp. 1-251 (2020) https://doi.org/10.2200/S01060ED1V01Y202010HLT048
46. Milne-Ives M., de Cock C., Lim E., Shehadeh M.H., de Pennington N., Mole G., Normando E., Meinert E.: The Effectiveness of Artificial Intelligence Conversational Agents in Health Care: Systematic Review Journal of Medical Internet Research, 22, pp. e20346 (2020) https://doi.org/10.2196/20346 PMid:33090118 PMCid:PMC7644372
47. Schachner T., Keller R., v Wangenheim F.: Artificial Intelligence-Based Conversational Agents for Chronic Conditions: Systematic Literature Review Journal of Medical Internet Research, 22, pp. e20701 (2020) https://doi.org/10.2196/20701 PMid:32924957 PMCid:PMC7522733
48. Provoost S., Lau H.M., Ruwaard J., Riper H.: Embodied Conversational Agents in Clinical Psychology: A Scoping Review Journal of Medical Internet Research, 19, pp. e151 (2017) https://doi.org/10.2196/jmir.6553 PMid:28487267 PMCid:PMC5442350
49. Heidig, S., & Clarebout, G. (2011). Do pedagogical agents make a difference to student motivation and learning?. Educational Research Review, 6(1), 27-54. https://doi.org/10.1016/j.edurev.2010.07.004
50. Wiltermuth, S. S., & Heath, C. (2009). Synchrony and cooperation. Psychological science, 20(1), 1-5. https://doi.org/10.1111/j.1467-9280.2008.02253.x PMid:19152536
51. Marsh K.L., Johnston L., Richardson M.J., Schmidt R.C.: Toward a radically embodied, embedded social psychology European Journal of Social Psychology, 39, pp. 1217-1225 (2009) https://doi.org/10.1002/ejsp.666
52. Yzerbyt V., Estrada C., Corneille O., Seron E., Demoulin S.: Subjective Essentialism in Action The psychology of group perception: Perceived variability, entitativity, and essentialism, pp. 101-124 (2004)
53. Overy K., Molnar-Szakacs I.: Being Together in Time: Musical Experience and the Mirror Neuron System Music Perception, 26, pp. 489-504 (2009) https://doi.org/10.1525/mp.2009.26.5.489
54. Rennung M., Göritz A.S.: Prosocial Consequences of Interpersonal Synchrony Zeitschrift für Psychologie, 224, pp. 168-189 (2016) https://doi.org/10.1027/2151-2604/a000252 PMid:28105388 PMCid:PMC5137339

55. Mogan R., Fischer R., Bulbulia J.A.: To be in synchrony or not? A meta-analysis of synchrony's effects on behavior, perception, cognition and affect Journal of Experimental Social Psychology, 72, pp. 13-20 (2017) https://doi.org/10.1016/j.jesp.2017.03.009
56. Butler E.A.: Temporal Interpersonal Emotion Systems Personality and Social Psychology Review, 15, pp. 367-393 (2011) https://doi.org/10.1177/1088868311411164 PMid:21693670
57. Miles L.K., Nind L.K., Macrae C.N.: The rhythm of rapport: Interpersonal synchrony and social perception Journal of Experimental Social Psychology, 45, pp. 585-589 (2009) https://doi.org/10.1016/j.jesp.2009.02.002
58. Wiggins S., Wiggins S., Golubitsky M.: Introduction to applied nonlinear dynamical systems and chaos, Springer, (2003)
59. de Jonge-Hoekstra L., van der Steen S., van Geert P., Cox R.F.A.: Asymmetric Dynamic Attunement of Speech and Gestures in the Construction of Children's Understanding Frontiers in Psychology, 7, (2016) https://doi.org/10.3389/fpsyg.2016.00473 PMid:27065933 PMCid:PMC4814764
60. Vallacher R.R., van Geert P., Nowak A.: The Intrinsic Dynamics of Psychological Process Current Directions in Psychological Science, 24, pp. 58-64 (2015) https://doi.org/10.1177/0963721414551571
61. Valdesolo P., DeSteno D.: Synchrony and the social tuning of compassion. Emotion, 11, pp. 262-266 (2011) https://doi.org/10.1037/a0021302 PMid:21500895
62. Mörtl A., Lorenz T., Hirche S.: Rhythm Patterns Interaction - Synchronization Behavior for Human-Robot Joint Action PLoS ONE, 9, pp. e95195 (2014) https://doi.org/10.1371/journal.pone.0095195 PMid:24752212 PMCid:PMC3994045
63. Lehmann H., Saez-Pons J., Syrdal D.S., Dautenhahn K.: In Good Company? Perception of Movement Synchrony of a Non-Anthropomorphic Robot PLOS ONE, 10, pp. e0127747 (2015) https://doi.org/10.1371/journal.pone.0127747 PMid:26001025 PMCid:PMC4441426
64. Cummins F.: Practice and performance in speech produced synchronously Journal of Phonetics, 31, pp. 139-148 (2003) https://doi.org/10.1016/S0095-4470(02)00082-7
65. Cummins F.: Measuring synchronization among speakers reading together Presented at the November 1 (2019) https://doi.org/10.36505/ExLing-2006/01/0020/000020
66. Cummins F.: Periodic and Aperiodic Synchronization in Skilled Action Frontiers in Human Neuroscience, 5, (2011) https://doi.org/10.3389/fnhum.2011.00170 PMid:22232583 PMCid:PMC3248675
67. Reddish P., Fischer R., Bulbulia J.: Let's Dance Together: Synchrony, Shared Intentionality and Cooperation PLoS ONE, 8, pp. e71182 (2013) https://doi.org/10.1371/journal.pone.0071182 PMid:23951106 PMCid:PMC3737148
68. Delaherche E., Chetouani M., Mahdhaoui A., Saint-Georges C., Viaux S., Cohen D.: Interpersonal Synchrony: A Survey of Evaluation Methods across Disciplines IEEE Transactions on Affective Computing, 3, pp. 349-365 (2012) https://doi.org/10.1109/T-AFFC.2012.12
69. Harmon-Jones C.K.: Does musical behavior promote affiliation?, Texas A&M University, (2011)
70. Whittemore, R., Jaser, S. S., Faulkner, M. S., Murphy, K., Delamater, A., Grey, M., & TEENCOPE Research Group. (2013). Type 1 diabetes eHealth psychoeducation: youth recruitment, participation, and satisfaction. Journal of medical Internet research, 15(1), e2170. https://doi.org/10.2196/jmir.2170 PMid:23360729 PMCid:PMC3636279
71. Lally, R. M., Kupzyk, K. A., Bellavia, G., Hydeman, J., Gallo, S., Helgeson, V. S., ... & Brown, J. K. (2020). Caring Guidance™ after breast cancer diagnosis eHealth psychoeducational intervention to reduce early post-diagnosis distress. Supportive Care in Cancer, 28, 2163-2174. https://doi.org/10.1007/s00520-019-05028-0 PMid:31414245 PMCid:PMC7083810

72. Hayes S.C., Strosahl K.D., Wilson K.G.: Acceptance and commitment therapy: The process and practice of mindful change, Guilford press, (2011)
73. Renner F., Schwarz P., Peters M.L., Huibers M.J.H.: Effects of a best-possible-self mental imagery exercise on mood and dysfunctional attitudes Psychiatry Research, 215, pp. 105-110 (2014) https://doi.org/10.1016/j.psychres.2013.10.033 PMid:24252218
74. Scholten, M. R., Kelders, S. M., & Van Gemert-Pijnen, J. E. (2021). Can a monologue-style ECA more effectively motivate eHealth users in initial distress than textual guidance?. Heliyon, 7(3), e06509. https://doi.org/10.1016/j.heliyon.2021.e06509 PMid:33842700 PMCid:PMC8020434
75. Cerekovic A., Aran O., Gatica-Perez D.: Rapport with Virtual Agents: What Do Human Social Cues and Personality Explain? IEEE Transactions on Affective Computing, 8, pp. 382-395 (2017) https://doi.org/10.1109/TAFFC.2016.2545650
76. Wechsung, I., Weiss, B., Kühnel, C., Ehrenbrink, P., & Möller, S. (2013, August). Development and validation of the conversational agents scale (CAS). In INTERSPEECH (pp. 1106-1110). https://doi.org/10.21437/Interspeech.2013-298
77. Schnall, R., Cho, H., & Liu, J. (2018). Health Information Technology Usability Evaluation Scale (Health-ITUES) for usability assessment of mobile health technology: validation study. JMIR mHealth and uHealth, 6(1), e8851. https://doi.org/10.2196/mhealth.8851 PMid:29305343 PMCid:PMC5775483
78. Birk, M. V., Atkins, C., Bowey, J. T., & Mandryk, R. L. (2016, May). Fostering intrinsic motivation through avatar identification in digital games. In Proceedings of the 2016 CHI conference on human factors in computing systems (pp. 2982-2995). https://doi.org/10.1145/2858036.2858062
79. Fujiwara, K., Kimura, M., & Daibo, I. (2020). Rhythmic features of movement synchrony for bonding individuals in dyadic interaction. Journal of Nonverbal Behavior, 44(1), 173-193. https://doi.org/10.1007/s10919-019-00315-0
80. Andrusyna, T. P., Tang, T. Z., DeRubeis, R. J., & Luborsky, L. (2001). The factor structure of the Working Alliance Inventory in cognitive-behavioral therapy. The Journal of psychotherapy practice and research, 10(3), 173.
81. Vaismoradi, M., Turunen, H., & Bondas, T. (2013). Content analysis and thematic analysis: Implications for conducting a qualitative descriptive study. Nursing & health sciences, 15(3), 398-405. https://doi.org/10.1111/nhs.12048 PMid:23480423
82. Braun V., Clarke V.: Using thematic analysis in psychology Qualitative Research in Psychology, 3, pp. 77-101 (2006) https://doi.org/10.1191/1478088706qp063oa
83. Konvalinka I., Vuust P., Roepstorff A., Frith C.D.: Follow you, Follow me: Continuous Mutual Prediction and Adaptation in Joint Tapping Quarterly Journal of Experimental Psychology, 63, pp. 2220-2230 (2010) https://doi.org/10.1080/17470218.2010.497843 PMid:20694920
84. Klein G., Moon B., Hoffman R.R.: Making Sense of Sensemaking 1: Alternative Perspectives IEEE Intelligent Systems, 21, pp. 70-73 (2006) https://doi.org/10.1109/MIS.2006.75
85. Xie B., Zhou J., Wang H.: 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, pp. 1-14 (2017) https://doi.org/10.1155/2017/3683546
86. Ziefle M., Bay S.: Mental Models of a Cellular Phone Menu. Comparing Older and Younger Novice Users Presented at the (2004) https://doi.org/10.1007/978-3-540-28637-0_3
87. Prinz W.: A Common Coding Approach to Perception and Action Relationships Between Perception and Action. pp. 167-201. Springer Berlin Heidelberg, Berlin, Heidelberg (1990) https://doi.org/10.1007/978-3-642-75348-0_7
88. Bickmore, T., & Cassell, J. (2005). Social dialogue with embodied conversational agents. Advances in natural multimodal dialogue systems, 23-54. https://doi.org/10.1007/1-4020-3933-6_2