

# Embodied Conversational Agents in eHealth: How Facial and Textual Expressions of Positive and Neutral Emotions Influence Perceptions of Mutual Understanding

SILKE TER STAL<sup>1,2,\*</sup>, GERBRICH JONGBLOED<sup>3</sup> AND MONIQUE TABAK<sup>1,2</sup>

<sup>1</sup>*eHealth Group, Roessingh Research and Development, Roessinghsbleekweg 33b, Enschede 7522, AH, The Netherlands*

<sup>2</sup>*Faculty of Electrical Engineering, Mathematics and Computer Science, University of Twente, Drienerlolaan 5, Enschede 7522 NB, The Netherlands*

<sup>3</sup>*Faculty of Behavioural, Management and Social Sciences, University of Twente, Drienerlolaan 5, Enschede 7522 NB, The Netherlands*

\*Corresponding author: [s.sterstal@utwente.nl](mailto:s.sterstal@utwente.nl)

**Embodied conversational agents (ECAs) could engage users in eHealth by building mutual understanding (i.e. rapport) via emotional expressions. We compared an ECA's emotions expressed in text with an ECA's emotions in facial expressions on users' perceptions of rapport. We used a 2 × 2 design, combining a happy or neutral facial expression with a happy or neutral textual expression. Sixty-three participants (mean, 48±22 years) had a dialogue with an ECA on healthy living and rated multiple rapport items. Results show that participants' perceived rapport for an ECA with a happy facial expression and neutral textual expression and an ECA with a neutral facial expression and happy textual expression was significantly higher than the neutral value of the rapport scale ( $P = 0.049$  and  $P = 0.008$ , respectively). Furthermore, results show no significant difference in overall rapport between the conditions ( $P = 0.062$ ), but a happy textual expression for an ECA with a neutral facial expression shows higher ratings of the individual rapport items helpfulness ( $P = 0.019$ ) and enjoyableness ( $P = 0.028$ ). Future research should investigate users' rapport towards an ECA with different emotions in long-term interaction and how a user's age and personality and an ECA's animations affect rapport building. Optimizing rapport building between a user and an ECA could contribute to achieving long-term interaction with eHealth.**

## RESEARCH HIGHLIGHTS

- Perceived rapport towards embodied conversational agents (ECAs) having different facial and textual emotional expressions is not significantly different after a short interaction in eHealth.
- After a short interaction with an ECA in eHealth, an ECA with a neutral face and happy text is perceived as more helpful and enjoyable than an ECA with a neutral face and neutral text.
- Future research should investigate users' rapport towards an ECA with different emotions in long-term interaction and how a user's age and personality and an ECA's animations affect rapport building.

*Keywords: embodied conversational agent; emotion; facial expressions; textual expressions; rapport; eHealth*

*Handling Editor: Helen Petrie*

*Received 8 November 2020; Revised 9 May 2021; Accepted 25 May 2021*

## 1. INTRODUCTION

eHealth applications can be used to provide long-term care and coaching, also without a health care professional being involved. Via this long-term care and coaching, people can be supported in having a healthy lifestyle, thereby supporting the onset or development of disease (progression). However, most of the applications are used without human support, which makes it difficult to maintain users' engagement. To engage users, the use of embodied conversational agents (ECAs) as health coaches is investigated. ECAs are more or less autonomous and intelligent software entities with an embodiment used to communicate with the user (Ruttkey *et al.*, 2004). ECAs can contribute to engagement by providing users with social support, which is one of most important persuasive drivers in eHealth (Kelders *et al.*, 2012). ECAs can provide this social support by actively building mutual understanding or 'rapport' with the user. Evidence has been found that such rapport between an ECA and the user can lead to companionship and higher user engagement (Gratch *et al.*, 2007a, b).

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

According to Tickle-deggen & Rosenthal (1990), one of the variables that strongly influences 'rapport' is non-verbal behaviour. More specifically, when there are expressions of positive emotions, rapport can be built (Tickle-deggen & Rosenthal, 1990). Thus, we could hypothesize that expressions of positive emotions by the ECA could increase rapport between the ECA and the user and increase the user's engagement. An ECA can express these positive emotions in several ways, such as via speech, gaze behaviour and hand and body gestures. In addition, an ECA can communicate these emotions via (i) *textual expressions* and (ii) *facial expressions*. First, humans are experts in recognizing emotions based on

facial expressions (Ekman & Friesen, 1978). Second, with respect to emotion in text, initially, it was suggested that within computer-mediated textual communication, the lack of non-verbal cues might make it more difficult for individuals to recognise expressed emotions (Short *et al.*, 1976, as cited by Hancock *et al.*, 2007). However, people might be more able to adapt to emotional expressions through computer-mediated textual communication than previously suggested (Hancock *et al.*, 2007, Walther, 1992).

Several studies have researched an ECA's emotional expressions. For example, studies investigated how to model an ECA's emotions in its facial expressions (Bevacqua *et al.*, 2007, Llorach *et al.*, 2019, Pelachaud, 2009). In addition, Schröder (2012) built a promising real-time interactive multi-modal dialogue system for ECAs with a focus on emotional and non-verbal interaction capabilities (called the sensitive artificial listener). Although these studies are great examples of research on how to technically model an ECA's emotions, they do not primarily evaluate users' perceptions of emotional expressions. Other studies do focus on evaluating users' perceptions, researching effects of positive, negative and neutral emotional expressions of an ECA on students' social judgements, interest and self-efficacy (Kim *et al.*, 2007) and effects of an ECA's emotional facial expressions and tone of voice on students' emotional states (Moridis & Economides, 2012). Although these studies show positive effects of an ECA's emotional expressions, they do not investigate effects of an ECA's emotional expressions on the quality of the relationship between the ECA and its users.

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

Some studies researched effects of an ECA's emotional expressions on the quality of the relationship between an ECA and its users in the context of eHealth specifically. Creed *et al.* (2015) describe that an emotionally expressive ECA, in the role of a nutritional coach, is perceived more likeable and caring than an unemotional ECA. In addition, Amini *et al.* (2014) report that an empathic ECA, assessing a user's

**TABLE 1.** Hypotheses to be tested in this research.

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

alcohol dependence, scores higher on perceived sociability, social presence, trust and social influence than a non-empathic version and text-only interface. However, to the best of our knowledge, only a few studies explicitly researched effects of an ECA's emotional expressions in eHealth on a user's perception of rapport towards an ECA. DeVault *et al.* (2014) developed the SimSensei Kiosk, offering interviewing with a virtual human to perform automatic assessment of distress indicators. They analysed face-to-face human interactions to identify potential distress indicators, dialogue policies and virtual human gestures and translated them into technical solutions but also evaluated users' experiences with both a Wizard-of-Oz and fully autonomous version of the system. For both versions of the systems, participants reported feelings of rapport comparable to that of a face-to-face interview. Furthermore, Ranjartabar *et al.* (2019) compared an empathic therapist with a neutral therapist. They show that users interacting with an empathic ECA will not necessarily build more rapport than those interacting with a neutral ECA. They note that rapport building towards the ECA depends upon the emotional feelings the user is initially expressing towards the problem addressed by the ECA.

Thus, although research has been performed on developing and evaluating an ECA's emotional expressions, no studies specifically disentangled effects of an ECA's emotion in facial expressions from effects of an ECA's emotion in textual expressions. Specifically, how do these aspects affect users' perceptions of rapport? Therefore, this research aims to *compare an ECA's emotions expressed in text with an ECA's emotions in facial expressions on users' perceptions of rapport* in the context of eHealth. The expression of positive emotions (e.g. happiness) is compared to the absence of emotions (e.g. 'neutral' emotions). The hypotheses to be tested in this research can be seen in Table 1.

**TABLE 2.** Participants are randomly assigned to one of the four conditions.

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

## 2. METHOD

### 2.1. Study design

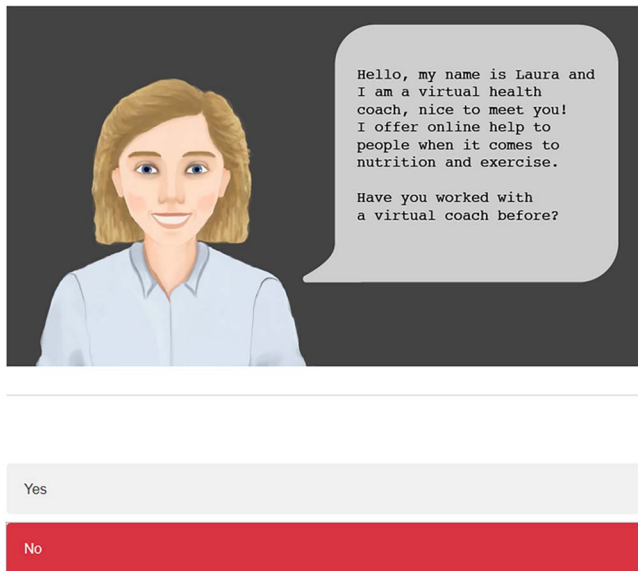
We researched the perceptions of rapport towards an ECA using a  $2 \times 2$  factorial design, consisting of the independent variables 'type of facial expression' and 'type of textual expression'. For both the facial expression and the textual expression, two types of emotions were examined: (i) neutral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

### 2.2. The dialogues

The topic of the dialogue with the ECA, Laura, was gaining a healthy lifestyle (i.e. being physically active and having a healthy nutrition). More specifically, the ECA introduced herself, explained the role of an ECA and the topic of the dialogue, to finally work towards concrete tips on physical activity and healthy nutrition. The dialogues were presented to the participants in Qualtrics, an online survey tool. Every dialogue consisted of multiple dialogue steps. As seen in Fig. 1, each dialogue step consisted of the ECA with its message (presented as one static image in a question in the survey tool) and a set of response options for the user (presented as responses in the survey tool).

### 2.3. The ECA's facial expressions

For the ECA's facial expressions, the Facial Action Coding System (FACS) developed by Ekman & Friesen (1978) was used. The FACS classifies multiple action units that, when combined, lead to facial expressions of emotions. Examples of these action units are 'Frontalis, pars medialis' (raising inner corner of eyebrow) and 'pterygoids and digastric' (mouth stretched open). The ECA's happy and neutral facial expressions were



**FIGURE 1.** Example step of the dialogue between the ECA and the participant; every dialogue step consisted of an image of the message presented by the ECA (as a question in the survey tool) and a set of response buttons for the participant (as responses in the survey tool).

developed using the action units as classified by the FACS. For a neutral expression, no muscle contractions take place that lead to an expression of emotion; therefore, no action units were active for this expression.

For the happy facial expression, action unit 6 ('Orbicularis oculi, pars orbitalis'—Cheeks raised, eyes narrowed) was combined with action unit 12 ('Zygomaticus major'—Lip corners pulled up and laterally), a common combination to express happiness (Kanade *et al.*, 2000). Table 3 shows an image of both an ECA with a happy and neutral facial expression. The ECA's facial expression was static during the complete dialogue (i.e. the ECA was displayed as an image that did not change throughout the interaction).

#### 2.4. The ECA's textual expressions

As suggested by Hancock *et al.* (2007), positive emotions as opposed to negative emotions can be added to text by implementing the following: (i) more words, (ii) fewer negative affect terms, (iii) fewer negotiations, (iv) fewer assents, (v) less agreement and (vi) more punctuation and exclamation marks. Therefore, compared to the neutral textual expressions, the happy textual expressions contained more words, less negative affect terms, more positive affect terms and more punctuation and exclamation marks. Though, it should be noted that the neutral textual expression was framed less negative than a negative textual expression would have been. Table 3 shows an example of both a neutral and happy textual expression of the ECA.

#### 2.5. Participants

Respondents to the questionnaire were at least 18 years old and fluent in the Dutch or the English language. No other inclusion or exclusion criteria were set. We recruited the respondents via our panel of adults (Dutch adults, living in Twente, the Netherlands) that indicated they were interested in participating in research on eHealth and through snowball sampling via social media and personal connections. The questionnaire was accessible via a public link of the survey program Qualtrics and available for 5 weeks, in October and November 2018. Participants gave an informed consent prior to the questionnaire.

A total of 63 respondents filled out the complete questionnaire. Of these respondents, 27 were male and 36 were female. The age of the respondents ranged from 20 to 88 years (mean,  $48 \pm 22$  years). The respondents were distributed over the conditions as follows: NFNT: 15 respondents (mean age, 54 years), HFNT: 16 respondents (mean age, 42 years), NFHT: 19 respondents (mean age, 44 years) and HFHT: 13 respondents (mean age, 57 years). An overview of all characteristics of the respondents can be seen in Table 4.

#### 2.6. Measurements

The following data were collected via the online questionnaire:

- *Respondents perceptions of rapport with the ECA*—Measured using an adaptation (Acosta & Ward, 2011) of the 'rapport scale' originally created by Gratch *et al.* (2007a). The scale covers 10 different dimension of feelings of rapport: emotional rapport, cognitive rapport, helpfulness, trustworthiness, likeability, naturalness, enjoyableness, human-likeness, persuasiveness and recommend ability (all rated on 7-point Likert scales ranging from strongly disagree to strongly agree)
- *Characteristics of respondents* (age, gender and education and technology use)

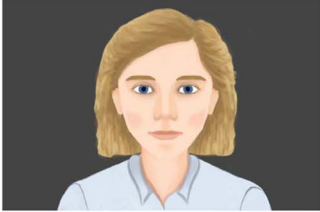

#### 2.7. Procedure

In the introduction to the questionnaire, the context of the research was explained and informed consent was obtained. Users were asked to imagine being supported by a virtual coach on healthy living. Then, the respondent was randomly assigned to one of the four conditions. The respondent engaged in a short dialogue with a ECA, after which he or she rated the ECA's rapport by statements on 7-point Likert scales. Lastly, participants filled out some questions regarding their demographics. In total, the questionnaire took around 15 minutes. Participants did not receive any compensation for taking part in the study.

#### 2.8. Data analyses

SPSS 25 was used to perform statistical analyses. The respondent's overall rapport towards the ECA was calculated by

**TABLE 3.** The differed facial and textual expressions used in this study.

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

**TABLE 4.** Characteristics of the respondents to the questionnaire.

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

Numbers represent number of respondents per condition.

taking the mean of all 10 rapport items. Data of the overall rapport and individual rapport items were normally distributed. Hence, one-sample *t*-tests were used to analyse whether the mean rapport ratings of the four conditions differed significantly from the neutral point of the rapport scale (i.e. 4). To test for a difference in overall rapport mean value between the four

conditions, one analysis of variance (ANOVA) between subjects test (with the four mean rapport values of the four conditions input) was conducted. Differences were further analysed by performing Bonferroni *post hoc* tests. Similar to previous research by Acosta & Ward (2011), of which we used the 'Rapport scale', we analysed whether differences in the indi-

**TABLE 5.** Results of the one-sample *t*-tests, testing for a difference between the mean rapport values of the four conditions and the neutral point of the rapport scale (i.e. 4).

	Neutral face, neutral text	Happy face, neutral text	Neutral face, happy text	Happy face, happy text
<i>Overall rapport, mean(SD)</i>	3.69(1.09)	4.39(0.73)	4.57(0.84)	4.09(1.25)
<i>P</i>	0.238	*0.049	*0.008	0.794
<i>Cohen's d</i>	0.288	0.537	0.681	0.074

\*  $P < 0.05$ **TABLE 6.** Results of the ANOVA comparing the mean values for overall rapport of the four conditions.

	Neutral face, neutral text	Happy face, neutral text	Neutral face, happy text	Happy face, happy text	<i>F</i> -value	<i>P</i>
	<i>mean(SD)</i>	<i>mean(SD)</i>	<i>mean(SD)</i>	<i>mean(SD)</i>		
<i>Overall rapport</i>	3.69(1.09)	4.39(0.73)	4.57(0.84)	4.09(1.25)	2.587	0.062

vidual rapport items could be found between the conditions. A multivariate analysis of variance (MANOVA) between subjects test was conducted to test for a difference in mean value for every individual rapport item across the four conditions. Again, the differences were further analysed by performing Bonferroni *post hoc* tests. All tests used a 95% confidence interval.

### 3. RESULTS

#### 3.1. Overall rapport

First, we analysed whether participants did establish rapport in general. We compared the mean rapport values of all conditions individually with that of the neutral point of the rapport scale (i.e. a value of 4). Table 5 shows that participants developed significantly more rapport than the neutral value of the scale in the conditions HFNT and NFHT ( $P = 0.049$  and  $P = 0.008$ , respectively).

Second, Table 6 shows the results of the ANOVA comparing the mean values for overall rapport of the four conditions. Results show no significant difference ( $P = 0.062$ ). To validate that there is not a significant difference in means between any of the conditions, a Bonferroni *post hoc* test was performed; again it appears that there is no significant difference in means.

#### 3.2. Individual rapport items

Table 7 shows the mean values of the individual rapport items for the four conditions.

The means of the individual items have been compared using a MANOVA between-subject design test. With  $F(64, 3) = 3.441$ ,  $P = 0.022$ , this test shows that there is at least a significant difference within the helpfulness, enjoyableness and persuasiveness items ( $P = 0.024$ ,  $P = 0.032$  and  $P = 0.021$ , respectively, as can be seen in Table 8).

These differences were further analysed by performing Bonferroni *post hoc* tests, as can be seen in Fig. 2. These tests show that

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

Table 9 shows the findings for the hypotheses of this research. This research only partly confirms hypothesis H1 and does not support hypotheses H2 and H3.

### 4. DISCUSSION

This research compared an ECA's emotions expressed in text with an ECA's emotions in facial expressions on users' perceptions of rapport in the context of eHealth. We compared expressions of positive emotions (e.g. happiness) with absence of emotions (e.g. 'neutral' emotions). Our results show that *users perceive an ECA with a neutral face and happy text as more helpful and enjoyable than an ECA with a neutral face and neutral text*. This partly confirms our first hypothesis

**TABLE 7.** Results of the individual rapport items for the four conditions.

	Neutral face, neutral text <i>mean(SD)</i>	Happy face, neutral text <i>mean(SD)</i>	Neutral face, happy text <i>mean(SD)</i>	Happy face, happy text <i>mean(SD)</i>
<i>Emotional rapport</i>	3.13(1.55)	3.69(1.20)	3.74(1.15)	3.69(1.55)
<i>Cognitive rapport</i>	3.53(1.60)	4.00(1.21)	4.11(1.24)	4.23(1.64)
<i>Helpfulness</i>	4.53(1.30)	5.06(1.00)	5.79(0.92)	4.92(1.55)
<i>Trustworthiness</i>	4.27(1.22)	4.75(1.07)	5.42(1.22)	4.62(1.45)
<i>Likability</i>	4.07(1.62)	5.12(1.20)	5.11(0.99)	5.08(1.32)
<i>Naturalness</i>	3.00(1.60)	3.75(1.44)	3.63(1.26)	3.23(1.48)
<i>Enjoyableness</i>	3.53(1.30)	4.50(1.03)	4.74(1.28)	4.15(1.07)
<i>Human-likeness</i>	3.67(1.18)	4.31(1.01)	4.21(1.48)	3.85(1.63)
<i>Persuasiveness</i>	3.60(1.35)	4.63(0.96)	4.74(1.37)	3.69(1.38)
<i>Recommendability</i>	3.53(1.41)	4.13(1.41)	4.26(1.24)	3.46(1.61)

**TABLE 8.** Test of between-subject design effects.

Source	<i>F</i> -value	<i>P</i>
<i>Emotional rapport</i>	0.695	0.559
<i>Cognitive rapport</i>	0.685	0.565
<i>Helpfulness</i>	3.382	0.024*
<i>Trustworthiness</i>	2.649	0.057
<i>Likability</i>	2.487	0.069
<i>Naturalness</i>	0.921	0.436
<i>Enjoyableness</i>	3.148	0.032*
<i>Human-likeness</i>	0.804	0.497
<i>Persuasiveness</i>	3.507	0.021*
<i>Recommendability</i>	1.322	0.276

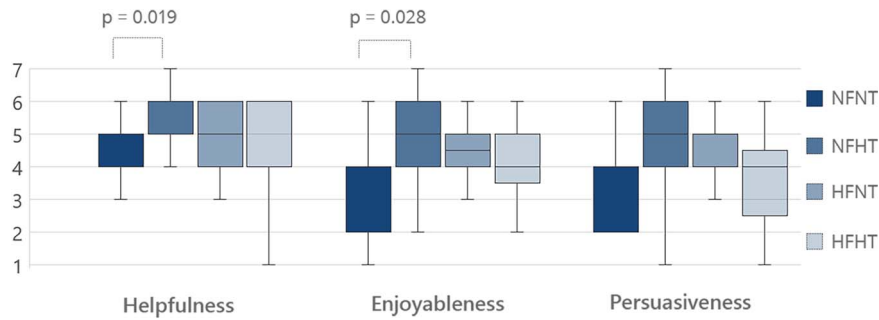
\*  $P < 0.05$ 

stating that users' perceptions of rapport towards an ECA with happy textual expressions are higher than towards an ECA with neutral textual expressions. Our second hypothesis, stating that users' perceptions of rapport towards an ECA with happy facial expressions are higher than towards an ECA with neutral facial expressions, and third hypothesis, stating that users' perceptions of rapport towards an ECA with both happy textual and happy facial expressions are higher than towards an ECA with only happy facial or happy textual expressions, were both rejected. In the rest of this section, we discuss our findings in more detail.

First, although in the HFNT and NFHT condition participants developed significantly more rapport than the neutral value of the scale, *overall rapport values found were limited*. As explained by Tickle-deggen & Rosenthal (1990), regardless of the time at which rapport is assessed, it consists of three components—*mutual attention, positivity and coordination*—, but the relative importance of these components changes over time. At initial encounters, development of rapport is strongly connected to the presence of *positivity* (warmth and friendli-

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

Second, *we did not find a difference between the overall rapport values of the four different conditions*. A lack of differences between the conditions could have been explained by the study population: half of the participants in our study were 50 five years or older. As shown by Beer *et al.* (2015), older adults show a decrease in emotion recognition accuracy for an ECA's emotions of anger, fear and happiness. The older adults that participated in our study might have had difficulties in recognising the ECA's happy emotions, and therefore, they might have



**FIGURE 2.** Results of the Bonferroni *post hoc* test comparing the mean values for the ECA's helpfulness, enjoyableness and persuasiveness for the four conditions.

**TABLE 9.** Findings for the hypotheses tested in this research.

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

experienced little differences between the conditions. Future work could differentiate between perceptions of younger and older adults specifically. Next to the participant's age, the personality of the participants might have affected our results. As shown by [Cerekovic et al. \(2014, 2017\)](#), a user's personality could affect his or her perceptions of rapport towards an ECA. Future work could investigate how a user's personality affects building of rapport towards an ECA in the eHealth context specifically. By investigating possible differences in rapport building due to a user's age or personality, the interaction with an ECA can be optimized for different users.

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

can affect user's perceptions of aspects of rapport in an early stage.

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



study, which might have resulted in participants having difficulties with assessing the non-verbal correlates of rapport in the ECA's facial expressions. Future work could test the effect of emotion in facial expressions in animated photo-realistic ECAs instead of in static cartoon ECAs. Since movements in the face help users to assess facial expressions, the contrast between the happy and neutral facial expressions most likely increases, which might affect the levels of rapport.

#### 4.1. Strengths and limitations

The strength of this study is that we tested the effect of an ECA's emotional expressions in a realistic scenario in the context of eHealth. In addition, this study particularly researched the effect of emotional expressions in cartoon ECAs, whereas much research focuses on emotional expressions in more photo-realistic ECAs having dynamic animations (DeVault *et al.*, 2014, Pelachaud, 2009, Schröder, 2012). Testing the effect of emotions via facial expressions in static cartoon ECAs was also a limitation of this study, since the contrast between the happy and neutral facial expression might have been less than it could have been in a more photo-realistic ECA. Cartoon ECAs require less technological development, and, therefore, the threshold to implement these ECAs in eHealth is lower. Also, expressing an ECA's emotion in text remains a challenge, since emotion cannot be communicated via intonation, whereas this is an important factor for communication of emotion in speech. Therefore, rapport towards the ECA might have been little compared to an ECA that would have used speech to communicate emotion. In addition, we did not perform a power analysis. Future research might study our results in more detail by performing a power analysis upfront and including the required number of participants in the experiment. Furthermore, participants only interacted with the ECA shortly. The interaction time might have been too short to actually build rapport within all components. Lastly, we built our study upon previous research by using an adaptation (Acosta & Ward, 2011) of the 'Rapport scale' originally created by Gratch *et al.* (2007a). To the best of our knowledge, the researchers do not describe the psycho-metrical properties of the scale. Therefore, the creation of a more robust, psycho-metrically developed scale of rapport of which validity and reliability are tested is recommendable. For future studies, we would recommend to create animated, more photo-realistic ECAs and to evaluate them in long-term interaction.

#### 5. CONCLUSION

We did not find any difference between an ECA's emotions expressed in text and an ECA's emotions in facial expressions on users' overall perceptions of rapport in the context of eHealth. However, our ECA was perceived as more helpful and enjoyable when happy textual expressions were used for an ECA having a neutral facial expression. Therefore,

we suggest to carefully design an ECA's textual expressions, since their framing could influence users' perceptions of the ECA's helpfulness and enjoyableness in an early stage. Future research should investigate users' rapport towards ECAs with different emotions in long-term interaction, both subjectively and objectively, and how a user's age and personality and an ECA's animations affect rapport building. Optimizing users' interaction with ECAs could contribute to achieving long-term interaction with eHealth.

#### REFERENCES

- [1] Acosta, J. C. and Ward, N. G. (2011) Achieving rapport with turn-by-turn, user-responsive emotional coloring. *Speech Commun.*, 53, 1137–1148.
- [2] Amini, R., Lisetti, C. and Yasavur, U. (2014) Emotionally responsive virtual counselor for behavior-change health interventions. In *Int. conf. design science research in information systems*, pp. 433–437. Springer, Cham.
- [3] Beale, R. and Creed, C. (2009) Affective interaction: how emotional agents affect users. *Int. J. Hum. Comput. Stud.*, 67, 755–776.
- [4] Beer, J. M., Smarr, C. A., Fisk, A. D. and Rogers, W. A. (2015) Younger and older users' recognition of virtual agent facial expressions. *Int. J. Hum. Comput. Stud.*, 75, 1–20.
- [5] Bevacqua, E., Mancini, M., Niewiadomski, R. and Pelachaud, C. (2007) An expressive ECA showing complex emotions. In *AISB'07: artificial and ambient intelligence*, 208–215. IEEE Computer Society.
- [6] 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 *HBU 2014: Human Behavior Understanding. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Vol. 8749, pp. 1–15. Springer.
- [7] Cerekovic, A., Aran, O. and Gatica-Perez, D. (2017) Rapport with virtual agents: what do human social cues and personality explain? *IEEE Trans. Affective Comput.*, 8, 382–395.
- [8] Creed, C., Beale, R. and Cowan, B. (2015) The impact of an embodied agent's emotional expressions over multiple interactions. *Interact. Comput.*, 27, 172–188.
- [9] DeVault, D. *et al.* (2014) SimSensei kiosk: a virtual human interviewer for healthcare decision support. In *13th int. conf. autonomous agents and multiagent systems, AAMAS 2014*, Vol. 2, pp. 1061–1068. ACM.
- [10] Ekman, P. and Friesen, W. (1978) *Facial Action Coding System: A Technique for the Measurement of Facial Movement*. Consulting Psychologists Press, Palo Alto, California.
- [11] Gratch, J., Wang, N., Gerten, J., Fast, E. and Duffy, R. (2007a) Creating rapport with virtual agents. In *Int. conf. intelligent virtual agents*, pp. 125–138. Springer.
- [12] Gratch, J., Wang, N., Okhmatovskaia, A., Lamothe, F., Morales, M., van der Werf, R. and Morency, L.-P. (2007b) Can virtual humans be more engaging than real ones? In *12th int. conf. human-computer interaction*, pp. 1–10. Springer.

- [13] Hancock, J. T., Landrigan, C. and Silver, C. (2007) Expressing emotion in text-based communication. In *Proc. SIGCHI conf. human factors in computing systems—CHI '07*, pp. 929. ACM.
- [14] Kanade, T., Cohn, J. F. and Tian, Y. (2000) Comprehensive database for facial expression analysis. In *Proc. fourth IEEE int. conf. automatic face and gesture recognition (FG'00)*, pp. 484–490. IEEE Computer Society.
- [15] Kelders, S. M., Kok, R. N., Ossebaard, H. C. and Van Gemert-Pijnen, J. E. W. C. (2012) Persuasive system design does matter: a systematic review of adherence to web-based interventions. *J. Med. Internet Res.*, 14, 1–24.
- [16] Kim, Y., Baylor, A. L. and Shen, E. (2007) Pedagogical agents as learning companions: the impact of agent emotion and gender. *J. Comput. Assist. Learn.*, 23, 220–234.
- [17] Llorach, G., Agenjo, J., Blat, J. and Sayago, S. (2019) *Web-Based Embodied Conversational Agents and Older People*. Springer International Publishing.
- [18] Loveys, K., Sebaratnam, G., Sagar, M. and Broadbent, E. (2020) The effect of design features on relationship quality with embodied conversational agents: a systematic review. *Int. J. Social Rob.*, 12, 1293–1312.
- [19] Moridis, C. N. and Economides, A. A. (2012) Affective learning: empathetic agents with emotional facial and tone of voice expressions. *IEEE Trans. Affective Comput.*, 3, 260–272.
- [20] Novick, D. and Gris, I. (2014) Building rapport between human and ECA: a pilot study. In *Int. conf. Human-Computer Interaction*. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), Vol. 8511 LNCS(PART 2), pp. 472–480. Springer.
- [21] Pelachaud, C. (2009) Studies on gesture expressivity for a virtual agent. *Speech Commun.*, 51, 630–639.
- [22] Ranjbartabar, H., Richards, D., Bilgin, A. and Kutay, C. (2019) First impressions count! The role of the human's emotional state on rapport established with an empathic versus neutral virtual therapist. *IEEE Trans. Affective Comput.*
- [23] Ruttkey, Z., Dormann, C. and Noot, H. (2004) Embodied Conversational Agents on a Common Ground: A Framework for Design and Evaluation. In *From Brows to Trust*, pp. 27–66. Kluwer.
- [24] Schröder, M. *et al.* (2012) Building autonomous sensitive artificial listeners. *IEEE Trans. Affective Comput.*, 3, 165–183.
- [25] Tickle-deggen, L. and Rosenthal, R. (1990) The nature of rapport and its nonverbal correlates. *Psychol. Inq.*, 1, 285–293.
- [26] Walther, J. B. (1992) Interpersonal effects in computer-mediated interaction: a relational perspective. *Commun. Res.*, 19, 52–90.