Effects of Perceived Agency and Message Tone in Responding to a Virtual Personal Trainer

Gale M. Lucas
USC Institute for Creative Technologies
Los Angeles
United States
lucas@ict.usc.edu

Nicole Krämer, Clara Peters, Lisa-Sophie Taesch,
University of Duisburg-Essen
Duisburg
Germany
nicole.kraemer@uni-due.de

Johnathan Mell & Jonathan Gratch
USC Institute for Creative Technologies
Los Angeles
United States
gratch@ict.usc.edu

ABSTRACT

Research has demonstrated promising benefits of applying virtual trainers to promote physical fitness. The current study investigated the value of virtual agents in the context of personal fitness, compared to trainers with greater levels of perceived agency (avatar or live human). We also explored the possibility that the effectiveness of the virtual trainer might depend on the affective tone it uses when trying to motivate users. Accordingly, participants received either positively or negatively valenced motivational messages from a virtual human they believed to be either an agent or an avatar, or they received the messages from a human instructor via skype. Both self-report and physiological data were collected. Like in-person coaches, the live human trainer who use negatively valenced messages were well-regarded; however, when the agent or avatar used negatively valenced messages, participants responded more poorly than when they used positively valenced ones. Perceived agency also affected rapport: compared to the agent, users felt more rapport with the live human trainer or the avatar. Regardless of trainer type, they also felt more rapport - and said they put in more effort - with trainers that used positively valenced messages than those that used negatively valenced ones. However, in reality, they put in more physical effort (as measured by heart rate) when trainers employed the more negatively valenced affective tone. We discuss implications for human–computer interaction.

CCS CONCEPTS

• **Human-centered computing → Human computer interaction (HCI) →** Empirical studies in HCI

KEYWORDS

Social agents, physical activity, affective tone, rapport, persuasion

ACM Reference format:

G. M. Lucas, N. Krämer, C. Peters, L.S. Taesch, J. Mell and J. Gratch. 2018. Effects of Perceived Agency and Message Tone in Responding to a Virtual Personal Trainer. In IVA '18: 2018 ACM International Conference on Intelligent Virtual Agents, November 5–8, 2018, Sydney, Australia. ACM, NY, NY, USA, 8 pages. https://doi.org/10.1145/3171221.3171258

1 INTRODUCTION

Globally, there has been an increase in physical inactivity due to the increasingly sedentary nature of humankind. Common health consequences of inactivity and resultant obesity are cardiovascular diseases, which were the leading cause of death in 2012, diabetes as well as some cancers, including liver, gallbladder, kidney and colon cancer [1]. In spite of these consequences, many people have trouble motivating themselves to engage in physical fitness and exercise. However, technological advances (e.g., FitBit) provide opportunities for wide-spread interventions to increase physical activity. Technologies designed to change a person's behavior through social influence have been successful in promoting health behavior [2]. Virtual fitness trainers, specifically, have shown promising benefits for users such as increased motivation, enjoyment and even effort expended and thus physiological responses [3-8]. For example, Bickmore [8] tested an exercise adoption program involving a virtual trainer agent named Laura. For four weeks, participants were able to log in as often as they liked and to spend as much time as they liked viewing educational pages related to the training program. Users were motivated by the agent: those that interacted with Laura viewed significantly more educational pages than participants in the no-agent condition.

Although research has not directly compared how virtual trainers stack up against human personal trainers, there is both a theoretical and empirical basis to expect some benefit of using virtual trainers (compared to real humans). Nass and colleagues posit that, to the extent to which its able to act like humans, people will respond fundamentally to media (e.g., fictional characters, cartoon depictions, virtual humans) as they would humans [9]. For example, when asked to rate the performance of an advice-giving agent, users tried to be as polite as with humans [9]. Yet, people do not afford such consideration to other virtual objects (without human-like features) [10].

Subsequent work has considered to what extent people treat social actors differently based on whether they *think* they are real people or artificial intelligence (perceived agency [11]). Research often studies this distinction by comparing agents to avatars [12-14]. Both agents and avatars are depicted as media, whether it be cartoon depictions or realistic virtual humans; in fact, an agent and an avatar could look identical (as they do in our study). Perceived agency is about who or what users believe is behind the depiction: avatars are operated by human users, whereas agents are controlled by artificial intelligence of some kind [15]. Using this comparison, there is some evidence that the social influence of agents is comparable to that of humans [16-18].

However, users will not *always* treat agents as if they were humans. The first proposed boundary condition that we consider here is the behavioral realism. We treat agents that behave in a realistic manner like humans, but not agents that do not. For example, if several agents are present in the virtual space and they all look and behave the same way, users will treat them less like they would a human [19]. Perceived agency, however, is predicted to have an effect here: avatars (presumed to be human-operated) would be treated like humans regardless of behavior. Although there is some support for this [20-21], other research has failed to find support: when behavioral realism was explicitly manipulated for both an agent and avatar, people treated the more realistic agent like a human regardless of whether they believed it was an agent or an avatar [22].

While support for the first boundary condition (behavioral realism) is mixed, there is also another important boundary condition: there are some contexts where virtual humans are treated differently. In contexts where they might be judged, people are more comfortable interacting with an agent than an avatar. Research has found that people respond more positively to agents than avatars in settings where they feel judged: clinical interviews about their mental health [12,23-24], when asked about their personal financial situation [13], and even interactions where they have to negotiate over resources [14]. Some of this work shows that people are only more comfortable with agents if they are concerned about being judged [23].

In the current work, we address two main research questions (R_1 and R_2). First, we consider how people respond to an agent (vs. avatar vs. human) when the system gives the user feedback during a workout session (R_1). If receiving feedback during a workout makes people feel judged, we expect to find that users will be more comfortable receiving feedback from agents than avatars or humans. Yet, as people are only more comfortable with agents if they are concerned about being judged [23], if users *do not* feel judged in this context, we would not find that people are more comfortable with the agent.

Although evidence around behavioral realism as a moderator was mixed, we also considered its impact. Specifically, in addition to agent and avatar (which are both depicted via media, as detailed above), we added a third condition: participants could instead interact with a real live human via computer-mediated interaction (i.e., skype). Some research has compared three similar conditions in the context of clinical interviews [25-26]. However, rather than only being about perceived agency, the differences were also about *actual* agency: this prior research [25-26] compared an actual agent (virtual human operated by AI and that was known to participants), an avatar-thought-to-be-agent (virtual human piloted by a *Wizard-of-Oz* human operator but users thought it was a computer), and a live human. In this context, participants felt more rapport with a virtual human they thought was operated by computer (but was actually operated by a human operator) than with the actual agent or human. So, it appears that rapport might have been influenced by behavioral realism and perceived agency: like in the agent condition, participants in the avatar-thought-to-be-agent condition thought it was an agent, but unlike the agent condition, they experienced a system that used non-verbal cues better (due to being Wizard-of-Oz operated rather than AI). Although the human interaction partner used these cues as well -if not better, rapport was probably reduced by fear of being judged in the human condition. However, research in the context of negotiation finds no differences in users' response to an avatar-thought-to-be-agent (again, operated by a human but thought to be computer) and a live human [27]; so behavioral realism needs to be tested in the specific domain – and, here, we do so for physical fitness.

In contrast to those previous studies [25-26], both the agent and avatar conditions that we use in the current study use a virtual human that is operated via Wizard-of-Oz, and instead these conditions here differ in the belief about who is operating the system (computer vs. human, respectively, as in [12-14]). Thus, in this work, only the human condition has greater behavioral realism (i.e., real facial expressions). Using our three conditions, we are able to compare the impact of behavioral realism to concerns about being judged in this domain. If users respond better to the agent than the avatar or human, it could be because concerns over being judged are prevalent in our physical fitness context, whereas if they respond better to the human than the agent or avatar, it would provide evidence for the importance of behavioral realism in this domain (R_1).

Additionally, in this work we explored the possibility that the effectiveness of the virtual personal trainer might also depend on whether it employed positively or negatively valenced motivational messages. Designers of virtual trainers have considered using motivational messages that differed in affective tone. For example, systems (e.g., [28]) have contained both positive (e.g. "good job!", "fantastic!") and negative motivational messages (e.g. "get moving!"). These previous systems used positively valenced motivational messages when users increased or maintained their effort, and negatively valenced motivational messages when they did not maintain their effort. While this was a step towards examining the role of message tone, because users received *both* types of messages, it was not possible to compare the effectiveness of positively vs negatively valenced motivational messages. The present research addresses this important gap (R_2).

Because negatively valenced messages might make people feel judged, when agents utilize such messages they may no longer have an advantage (over avatars or humans) in making users feel comfortable. Also, research has found that agents cannot use angry persuasive messages like humans can; instead such negatively valenced messages backfire for agents, reducing persuasiveness [29]. For both these reasons, we expect agents will not be as effective at using negatively valenced messages as positive ones. In contrast, human coaches can use negative affective tone to motivate their clients as long as the

motivational messages are not authoritarian or developmentally inappropriate [30-31]. Thus, humans should be able to use negatively valenced messages more effectively than agents (R_2).

2 PRESENT RESEARCH

To test these research questions, we conducted an empirical study using a Wizard of Oz paradigm. Participants received motivational messages from a virtual human they believed to be either an agent or an avatar. In a third condition, they instead received the messages from a human instructor via skype. Participants interacted with one of these three forms of the personal trainer, and the trainer provided either positively or negatively valenced motivational messages to encourage the user to exercise harder. We compared these conditions in terms of both psychological and physiological responses.

2.1 Participants

There were 216 participants recruited, and they were compensated $30 for participation. Four of them were excluded from analysis due to either incomplete data or failure to follow instructions, which left 212 participants (100 female, 112 male) remaining for analysis. 46% were under 35, 48% were between 35 and 65, and 6% were over 65.

2.2 Design and procedure

In a 3 (trainer type) x 2 (message tone) between-subjects design, participants were randomly assigned to one of three different personal trainers: a virtual agent, an avatar or a real person via skype, as well as to one of two message tone conditions: positively valenced or negatively valenced. After participants consented, the experimenter introduced the cover story. Participants were told that, due to the high price of personal trainers, the study was testing a less expensive fitness system. To manipulate trainer type, participants were told that this system makes use of an artificial intelligence (in the agent condition) or a remote worker (in the avatar and real human condition) as a fitness trainer. During the workout session, we used pre-recorded scripts to manipulate the second factor of message valence (described in detail below).

The experiment room was equipped with a computer, a stationary bike (Proform 8.0 Ex Upright,) a video camera (set up on a tripod behind the bike in an over-the-shoulder angle), a timer (to count down the 10 minute workout), and a tablet attached to the bike (which was loaded with the video-conferencing service skype). Before the participants started the workout, they were asked to complete demographic measures. A heart rate monitor was attached to the participant via a chest strap. Once the participant had been set up on the bike, participants were told to comply with the trainer's requests only if they felt able to. The trainer then called them via skype, after which the experimenter started the timer and left the room. The participant then worked out for 10 minutes while receiving communications from the trainer over the skype platform.

The appearance of the trainer on the skype screen depended on condition (Fig. 1). In the agent and avatar conditions, the appearance was identical: participants saw a virtual human. The only difference was their belief about whether the virtual human was operated by an AI or human. In the human condition, the confederate communicated with the user through skype. While, in all conditions, the trainer's non-verbal behavior was aligned to the participants' (e.g., smiled if they smiled), the trainer's verbal behavior was pre-recorded (see description of trainer's script below). The confederate in the real human condition restrained her nonverbal behavior to match that of the virtual human. Also, the skin, hair, and eye color of the virtual human were matched to the confederate.

To operate the virtual trainer, a wizard controlled the virtual human from another room. She also served as the confederate in the live human condition on skype. Due to placement of the camera, the wizard could observe the participants during the workout and to react appropriately. All pre-recorded utterances as well as the nonverbal features were implemented in a native Windows GUI, through which the confederate could actuate the virtual human's behaviors. The trainer was always streamed live via skype, the only difference was the source: in the human condition, it came from a webcam trained on the confederate, whereas in the avatar and agent conditions, the animation of the virtual human came from SmartBody [32]. To keep all verbal behavior consistent across conditions, the trainer's script was pre-recorded. The wizard served as the voice actress, recording all of the lines. Scripts for the workout were identical across the two message conditions except for valence.

Social comparison was employed as a persuasion technique. To employ the social comparison technique, the trainer used lines such as "I've pulled it off, I know you can too!" (positively valenced) and "I did it and it's so easy. Try harder!" (negatively valenced). The frequency of the trainer's utterances during the ten-minute workout was set to one utterance every 30 seconds for the first eight minutes and to 20 seconds for the last two minutes. Following this structure, a total of 24 utterances were phrased for each of the two message valence conditions. Importantly, the same pre-recorded lines were used across all trainer conditions. In the agent and avatar condition, SmartBody [32] aligned the non-verbal behavior to the audio. For the human condition, the same GUI was used to activate the pre-recorded line, and the wizard lip-synced the utterance via skype. The human trainer only used the pre-recorded messages. To prevent participants from asking questions during the workout, they were told that due to technical problems the trainer could not hear them (but could see them and their progress on the stationary bike during the workout). After the workout, the experimenter returned to supply the participant with water and removed the chest strap. Participants were then able to dismount the stationary bike to complete self-report questionnaires.

Figure 1: *Human trainer (via skype) next to the virtual trainer.*

2.3 Measures

A number of self-report measures were taken. First, participants were asked to report their motivation during the workout using a 7-point scale ranging from 1 (minimum) to 7 (maximum). To index perceived quality of the trainer, participants were asked to rate the extent to which they agreed with "I think that they were a good personal trainer" and "I think the personal trainer performed well" on a scale from 1 (strongly disagree) to 7 (strongly agree). Total scores ranged from 2 to 14. Next, participants were asked to judge the trainer's personality on 32 personality characteristics using bipolar scale from 1 to 7, with each end-point representing the extreme of that personality characteristic (e.g., friendly to unfriendly), as in [33]. A principle axis factor analysis was conducted with promax rotation with standard exclusion criteria [34], leading to exclusion of 7 items. One item (tough to soft) loaded above .7 on both factors. It was therefore treated as an own factor ("***toughness***"), and served as a manipulation check for the message tone manipulation. The rest of the remaining 24 items loaded onto two clear factors: fifteen items represented disagreeableness (α = .95), and nine items representing lack of character (α = .80). Total scores could range from 1 to 7 for toughness, from 9 and 63 for lack of character, and from 15 and 105 for disagreeableness. Next, to measure rapport with the virtual trainer, seven items from the Rapport Scale [35] were used. Participants answered items such as "I tried to create a sense of closeness between us" and "My partner created a sense of closeness between us" on a scale from 1 (strongly disagree) to 5 (strongly agree) (α = .88).

Finally, effort was measured. For a self-report, participants were asked how much effort they expended during the workout on a scale ranging from 1 (minimum) to 7 (maximum). Heart rate increase, which is associated with increased physical activity [36], served as an objective measure of effort. Participants wore a chest strap that measured their heart rate per second during the workout. Extreme values were manually excluded, and the heart rate data of 12 participants had to be excluded from analysis due to measuring errors. A heart rate baseline was calculated by taking the average heart rate of the first ten seconds for each participant. Further, the heart rate percentage change from the baseline was computed for each data point (one per second). The mean of all percent changes was taken. By including the baseline and calculating the average change, the heart rate dataset could be used as a comparable measure that is independent of differences between participants such as age, weight and sex.

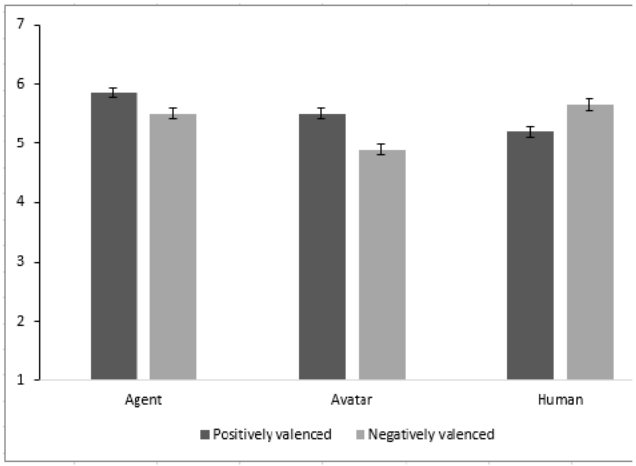

Figure 2: *Effect of agency and message tone on motivation.*

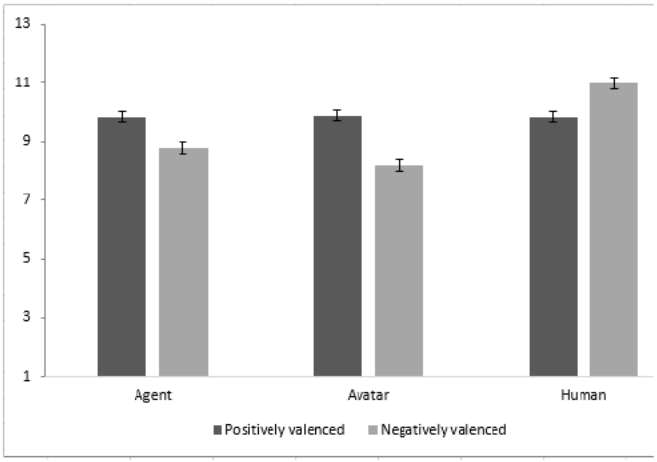

Figure 3: *Effect of agency and message tone on trainer quality.*

3 RESULTS

A series of 3 (trainer type) x 2 (message valence) ANOVAs was run, one for each of the measures. For reported motivation, there was a significant effect of trainer type ($F(2,205) = 3,79$, $p = .02$), and although there was no main effect of message tone ($F(1,205) = 0.46$, $p = .50$), there was an significant interaction with message tone ($F(2,205) = 3.35$, $p = .04$). As depicted in Fig. 2, participants were more motivated by negatively valenced messages (than positively valenced ones) when provided by a live human trainer, but less motivated by these same negative messages (compared to positively valenced ones) with agents and avatars.

Next we analyzed perceived quality of the trainer. As with reported motivation, there was a significant effect of trainer type ($F(2, 212) = 3.45$, $p = .03$), no main effect of message tone ($F(1, 212) = 1.579$, $p = .21$), and a significant interaction ($F(2, 212) =$

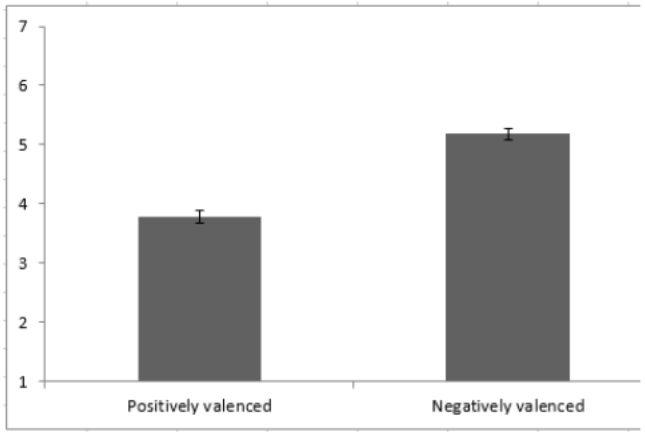

Figure 4: *Effect of message tone on perceived toughness.*

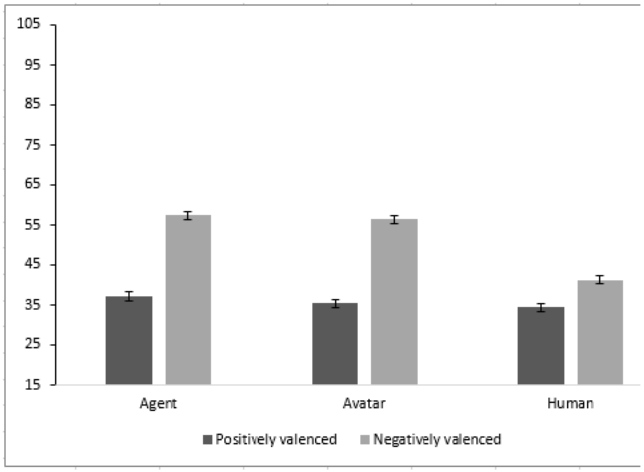

Figure 5: *Effect of agency and message tone on disagreeableness.*

3.61, $p = .03$). As depicted in Fig. 3, participants also rated the live human trainer as higher quality when using negatively valenced messages (than positively valenced ones), but rated the agent and avatar as lower in quality when they used negatively valenced messages (than positively valenced ones).

For perceptions of the personal trainer's personality, we analyzed toughness, disagreeableness, and lack of character in turn. First, there was a significant main effect of message tone on perceived toughness, $F(1,211) = 42.82$, $p < .001$; serving as a manipulation check, people rated the negatively valenced trainer as tougher than the positively valenced one (see Fig. 4). None of the other effects reached significance for perceived toughness (Fs < 0.8, *ps* > .47).

For disagreeableness, main effects of trainer type ($F(2, 210) = 6.70$, $p = .002$) and message tone ($F(1, 210) = 49.67$, $p < .001$) were qualified by an interaction ($F(2,210) = 4.16$, $p = .02$). As seen in Fig. 5, negatively valenced personal trainers were seen as more disagreeable than positively valenced ones in all trainer conditions, but there was more of a penalty for using negatively valenced messages for agents and avatars than for humans. In contrast, for lack of character, none of the effects reached significance (Fs < 2.0, *ps* > .14).

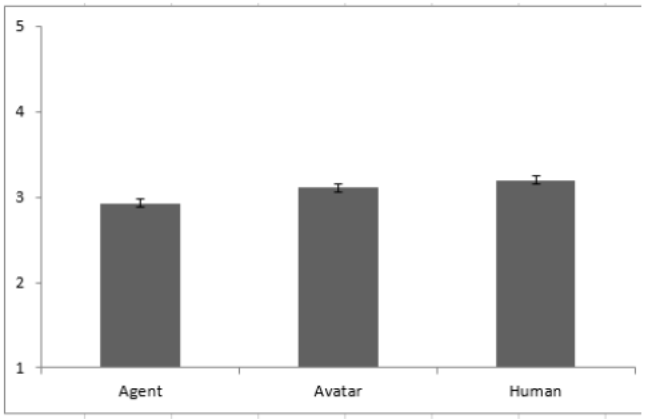

Figure 6: *Effect of agency on felt rapport.*

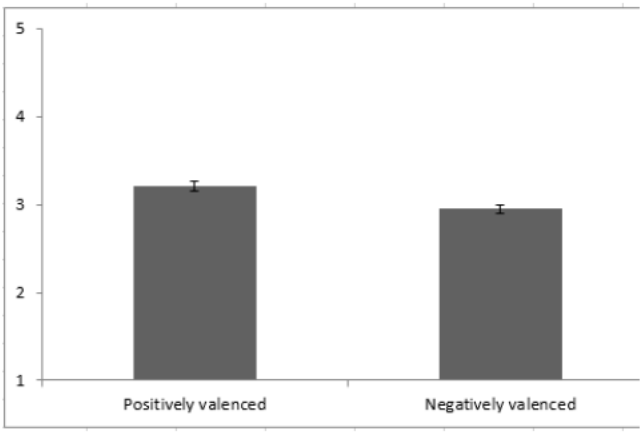

Figure 7: *Effect of message tone on felt rapport.*

Next, we analyzed rapport with the trainer. Unlike the previous outcomes, there was a marginal main effect for trainer type ($F(2,211) = 2.53$, $p = .08$) and a significant main effect for message tone ($F(1,211) = 6.65$, $p = .01$), but no interaction between the two ($F(2,211) = 1.21$, $p = .30$). Participants reported feeling marginally less rapport with the agent than with the avatar or human (see Fig. 6). Users also felt less rapport with the trainer when negatively valenced messages were used than when positively valenced messages were used (see Fig. 7).

Next we considered effort extended by the participant, first by analyzing self-reported effort and then a physiological index (increase in heart rate). For self-reported effort, there was a marginally significant effect of message valence, $F(1,206) = 3\,29$, $p = .07$. As with rapport felt with the trainer, self-reported effort was lower in the negatively valenced message condition than the positively valenced condition (Fig. 8). Although participants said they worked harder when provided motivational messages with a positive tone, none of the other effects reached significance for self-reported effort (Fs < 1.3, *ps* > .27).

There was also a significant effect of message valence on heart rate change, $F(1,194) = 6.56$, $p = .01$. However, in this case, heart rate change was ***higher*** in the negatively valenced message condition than the positively valenced condition, suggesting greater effort was actually expended in response to the

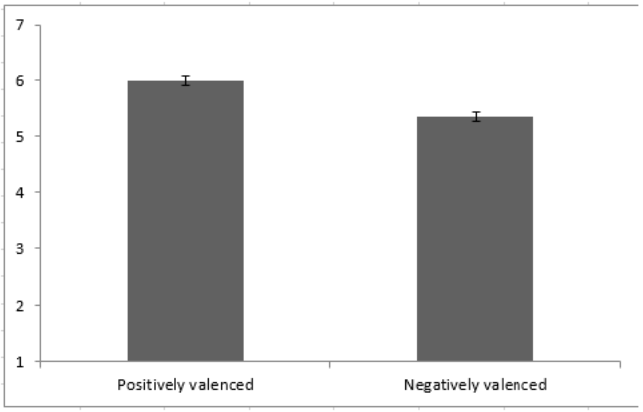

Figure 8: *Effect of message tone on self-reported effort.*

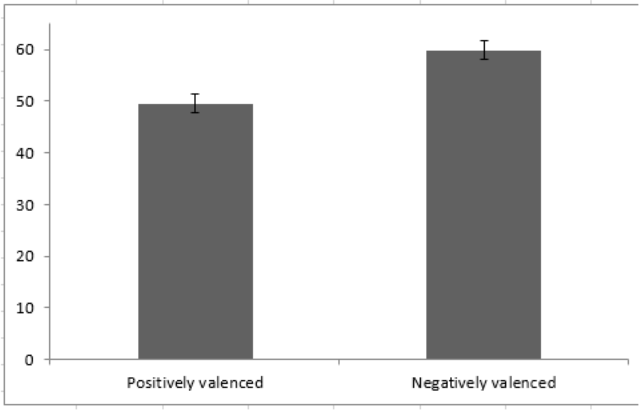

Figure 9: *Effect of message tone on heart rate increase.*

negatively valenced messages (see Fig. 9). None of the other effects reached significance for heart rate change either (Fs < 1.4, *ps* > .25).

4 DISCUSSION

4.1 Discussion of findings

The current work address two main research questions: how people respond to an agent (vs. avatar vs. human) when the system gives the user feedback during a workout session (R_1), and whether this also depends on if positively or negatively valenced motivational messages are employed (R_2). Our results suggest that, as expected, humans can use negatively valenced messages more effectively than agents. Users were less motivated and rated the quality of the trainer as lower when agents (or avatars) used negatively valenced messages (compared to when they used positively valenced messages, as expected). In contrast, they were *more* motivated and rated the quality of the trainer as *higher* when a live human used negatively valenced messages (compared to when they used positively valenced messages). The results for ratings of personality, specifically disagreeableness, also showed a similar pattern. Because the same verbal messages were used in all conditions, verbal realism could not account for this difference. There was only one difference was that the human trainer condition used a real, live human face with real facial expressions rather than virtual ones. With this greater behavioral realism, participants accepted the negatively valenced messages as well as (or even better than) the positively valenced ones. In the conditions without the real face, these negatively valenced messages were detrimental.

Based on these results, it seems that behavioral realism matters for some of these outcomes in the domain of physical fitness; in contrast, fear of being judged does not seem to be driving responses in this domain. Indeed, unlike prior research [12-14,23-24], we did not find that people were more comfortable with agents than avatars or humans. Some of this prior work suggests that people are only more comfortable with agents when they are afraid of being judged [23], so participants were probably not particularly concerned about being judged in our study. Furthermore, while they did not find the negatively

Outcome	Test	P value	Fig.
Motivation	**Effect of Agency**	**.02**	
	Effect of Message	**.50**	
	Interaction Effect	**.04**	**2**
Trainer quality	**Effect of Agency**	**.03**	
	Effect of Message	**.21**	
	Interaction Effect	**.03**	**3**
Toughness	**Effect of Agency**	**.47**	
	Effect of Message	**<.001**	**4**
	Interaction Effect	**.99**	
Disagreeableness	**Effect of Agency**	**.002**	
	Effect of Message	**<.001**	
	Interaction Effect	**.02**	**5**
Lack of character	**Effect of Agency**	**.21**	
	Effect of Message	**.73**	
	Interaction Effect	**.14**	
Rapport	**Effect of Agency**	**.08**	**6**
	Effect of Message	**.01**	**7**
	Interaction Effect	**.30**	
Reported effort	**Effect of Agency**	**.45**	
	Effect of Message	**.07**	**8**
	Interaction Effect	**.27**	
Heart rate increase	**Effect of Agency**	**.25**	
	Effect of Message	**.01**	**9**
	Interaction Effect	**.85**	

Figure 10: *Summary of results.*

valenced messages particularly pleasant, those messages did not evoke a sense of judgement. In fact, they responded *better* to this negative tone when it was used by the live human trainer, who they would have theoretically felt the most judged by.

There also seem to be implications for rapport with virtual trainers. During clinical interviews, rapport with the live human might be reduced by fear of being judged; participants felt less rapport with a live human than with a virtual human they thought was run by a computer [25-26]. In our research, where the participants appear to be relatively unconcerned about being judged, we found the opposite effect: people felt *more* rapport with the live human (or avatar) than with a virtual human they thought was run by a computer. This may help to explain other conflicting results as well. For example, some prior work in the context of negotiation found no difference in rapport between live humans and virtual humans thought to be run by a computer [27]; perhaps that was because, there, participants were *more* concerned about being judged than in our physical fitness context, but *less* concerned than during clinical interviews.

There are, however, alternatives explanations for our findings. For example, it is possible that people treated the human trainer differently in our study because of the specific kind of persuasive technique employed - *social comparison*. Perhaps seeing a physical body on the screen, rather than a virtual one, made social comparison seem more compelling. While they are necessary when using the social comparison technique, phrases like "I did it, you can do it too!" might feel more appropriate coming from a live human than a virtual human. This might be true even if you know it is being piloted by a real person: after all, the person operating that virtual representation could themselves be physically unfit, even though

the virtual human looked fit (like the confederate did in the live human condition). Similarly, in the agent condition, participants might be bothered by comments such as "I did it" coming from an unembodied artificial intelligence, which lacks the ability to "do" anything physical at all. While work on agent backstories suggests that users are not at all bothered by agents taking the role of a real human [33], that may not be the case in the context of physical fitness. While some persuasion techniques have already been shown to be effective when employed by agents (e.g., reciprocity [37-38]), to our knowledge, social comparison technique had not been examined previously. The current work suggests it may be less effective when used by agents in certain circumstances.

Indeed, there are some persuasion techniques that agents cannot use as well as humans can, and the current work reinforces prior findings in this area. Just as we found that virtual trainers cannot use negative motivation messages as well as live human personal trainers can, work in decision-making contexts has found that agents cannot use anger to persuade users like humans can [29]. Furthermore, this prior research found that agents were unable to use these strategies *because* they were seen -by default- as low in power, and thus anger was inappropriate [see also [39]). Thus, in our current work, it is possible that agents (and avatars) were seen as relatively lower in power, and thus their use of this persuasion technique was less effective.

We also found that the tone of motivational messages can have an impact regardless of whether trainers are virtual or real. Participants felt more rapport with trainers who used positively valenced messages compared to negatively valenced ones. This, however, had different consequences for effort depending on whether it was self-reported or indexed in terms of increase in cardio output (i.e., heart rate). Although users feel better when virtual trainers use positively valenced messages, they could put forth greater effort with more negatively valenced messages.

4.2 Implications for design, limitations, and future work

This finding -that positively valenced messages foster more favorable perceptions of the trainer but negatively valenced messages might result in greater effort- has implications for agent design. Indeed, this tradeoff between perception and effort should be considered when designing applications (as in [15]): positively valenced messages could reduce efficacy, whereas negatively valenced ones could reduce use when coming from a virtual trainer. Perhaps future applications could monitor both perception and physical effort online, and program virtual agents to employ positively and negatively valenced messages as needed, respectively. However, given the more negative perceptions felt by users in the negatively valanced message condition, it is possible that feelings of agitation contributed to the heightened heart rate increase. Because the difference in heart rate increase observed in this study is likely too large to be completely attributed to such agitation, such reactions would -if anything- only partially account for increase in heart rate.

One aspect that should be considered in future studies is personal preferences and personality traits. Previous studies have shown that the effectiveness of motivational strategies might differ depending on personality [40-41]. For example, researchers [41] found significant correlations between the effectiveness of motivating technology and the "big 5" personality traits. In future technologies that promote exercise, this could be realized using personality screening.

The participants in this study also commented on possible improvements of the virtual fitness trainer system. These comments have implications for features that could improve future fitness technologies. First, many participants noted that they would prefer it if the fitness trainer simultaneously exercised with the user rather than being static. While the constraints of experimental control limited us to using a script (described above), participants' comments suggested that a more interactive experience could be more motivating. For example, a more interactive and personalized feel could be achieved by providing individualized feedback on heart rate values, burnt calories and covered distance. Using text-to-speech could allow for a script to be customized with such feedback, while still maintaining internal validity. Finally, participants also requested to choose the look and the gender of the trainer.

Along these lines, it is also possible that the virtual trainer's gender could account for the reduced self-reported effort and felt connection from the more negatively valenced motivational messages. The negative stereotypes associated with women who are perceived to be "bossy" or "pushy" could have led participants to perceive the negatively valenced messages more unfavorably than if the virtual trainer were male. Future studies could explore the effects of trainer gender and thereby provide insight regarding the impact of male and female stereotypes with virtual humans [42] in the context of fitness. Instead of revealing stereotypes, we may find that people are more motivated by agents of the opposite gender as long as they are positively engaging, as in prior research with an academic context [43]. Ultimately, research should compare different combinations of positively and negatively valenced messages used by female (and male) virtual trainers to optimize positive perceptions and actual effort exerted. While such future research could further bolster our understanding in technology for physical fitness, this work takes an important step in this direction.

4.3 Conclusion

Overall, we find interesting differences in computer-mediated interactions between humans and media depictions (agents and avatars). Behavioral realism (in terms of non-verbal behaviors of the live human trainer) impacted regard for the trainer (motivation, and ratings of quality, disagreeableness, and rapport). Like in-person coaches, the live human trainer who use negatively valenced messages were well-regarded; however, when the agent or avatar used negatively valenced, participants responded more poorly than when they used positively valenced ones. In contrast, with rapport, perceived agency appears to take precedence: compared to the agent, users felt more rapport with the live human trainer or the avatar. Regardless of trainer type, they also felt more rapport - and said they put in more effort - with trainers that used positively valenced messages than those that used negatively valenced ones. However, in reality, they put in more physical effort (as measured by heart rate) when trainers employed the more negatively valenced affective tone. It seems that "what doesn't destroy you makes you stronger."

ACKNOWLEDGMENTS

This work was supported in part by the U.S. Army. Any opinion, content or information presented does not necessarily reflect the position or the policy of the United States Government, and no official endorsement should be inferred.

REFERENCES

[1] WHO. (2016). Obesity and overweight: Fact Sheet. Retrieved from http://www.who.int/mediacentre/factsheets/fs311/en/

[2] Orji, R., & Moffatt, K. (2016). Persuasive technology for health and wellness: State-of-the-art and emerging trends. *Health informatics journal.*

[3] Albaina, I. M., Visser, T., van der Mast, C. A., & Vastenburg, M. H. (2009). Flowie: A persuasive virtual coach to motivate elderly individuals to walk. In G. Troester, K. Connelly, & B. Arnrich (Eds.), *Proceedings of the 3d International ICST Conference on Pervasive Computing Technologies for Healthcare* ICST.

[4] Fox, J., & Bailenson, J. N. (2009). Virtual Self-Modeling: The Effects of Vicarious Reinforcement and Identification on Exercise Behaviors. *Media Psychology, 12(1),* 1–25.

[5] IJsselsteijn, W. A., Kort, Y. de, Bonants, R., Jager, M. de, & Westerink, J. (2004). Virtual Cycling: Effects of immersion and a virtual coach on motivation and presence in a home fitness application. *In Proceedings Virtual Reality Design and Evaluation Workshop* (pp. 22–29).

[6] King, A. C., Bickmore, T. W., Campero, M. I., Pruitt, L. A., & Yin, J. L. (2013). Employing virtual advisors in preventive care for underserved communities: Results from the COMPASS study. *Journal of health communication, 18(12),* 1449–1464.

[7] Buttussi, F., Chittaro, L., & Nadalutti, D. (2006). Bringing mobile guides and fitness activities together: a solution based on an embodied virtual trainer. In *Proceedings of the 8th conference on Human-computer interaction with mobile devices and services* (pp. 29-36). ACM.

[8] Bickmore, T. W. (2003). Relational agents. Effecting change through human-computer relationships (Doctoral dissertation).

[9] Reeves, B., & Nass, C. (1996). *The Media Equation: How people treat computers, television, and new media like real people and places* New York: Cambridge University Press.

[10] Brave, S., & Nass, C. (2003). Emotion in human-computer interaction. *Human-Computer Interaction*, 53–68.

[11] Nowak, K. L., & Biocca, F. (2003). The effect of the agency and anthropomorphism on users' sense of telepresence, copresence, and social presence in virtual environments. In *Presence Teleoperators and Virtual Environments* (pp. 481–494).

[12] Lucas, G. M., Gratch, J., King, A., and Morency, L. P. (2014). It's only a computer: Virtual humans increase willingness to disclose. *Computers in Human Behavior, 37, 94*-100.

[13] Mell, J., Lucas, G. M., & Gratch, J. (2017). Prestige questions, online agents, and gender-driven differences in disclosure. *Proceedings of the 17th International Conference on Intelligent Virtual Agents*, 273-282.

[14] Gratch, J., DeVault, D., & Lucas, G. M. (2016). The benefits of virtual humans for teaching negotiation. *Proceedings of the 16th International Conference on Intelligent Virtual Agent*, 276-287.

[15] Bailenson, J. N., & Blascovich, J. (2004). Avatars. *Encyclopedia of Human-Computer Interaction* Berkshire Publishing Group.

[16] Baylor, A. L. (2007). Pedagogical agents as a social interface. *Educational Technology, 47(1),* 11–14.

[17] Kim, Y., & Baylor, A. L. (2007). Pedagogical agents as social models to influence learner attitudes. *Educational Technology, 47(1),* 23–28.

[18] Lee, J. E., Nass, C., Brave, S., Morishima, Y., Nakajima, H., & Yamada, R. (2007). The Case for Caring Colearners: The Effects of a Computer-Mediated Colearner Agent on Trust and Learning. *Journal of Communication, 57(2),* 183–204.

[19] Sheehan, P. W., Statham, D., & Jamieson, G. A. (1991). Pseudomemory effects and their relationship to level of susceptability to hypnosis and state instruction. *Journal of Personality and Social Psychology, (60),* 130–137.

[20] Blascovich, J. (2002). A theoretical model of social influence for increasing the utility of collaborative virtual environments. *Proceedings of the 4th international Conference on Collaborative Virtual Environments,* 25–30.

[21] Bailenson, J.N., Yee, N., Merget, D., & Schroeder, R. (2006). The effect of behavioral realism and form realism of real-time avatar faces on verbal disclosure, nonverbal disclosure, emotion recognition, and copresence in dyadic interaction. *PRESENCE: Teleoperators and Virtual Environments, 15,* 359-372.

[22] Von der Pütten, A. M., Krämer, N. C., Gratch, J., & Kang, S.-H. (2010). "It doesn't matter what you are!: " Explaining social effects of agents and avatars. *Computers in Human Behavior, 26(6),* 1641–1650.

[23] Pickard, M. D., Roster, C. A., and Chen, Y. (2016). Revealing sensitive information in personal interviews: Is self-disclosure easier with humans or avatars and under what conditions? *Computers in Human Behavior, 65,* 23-30.

[24] Slack, W. V., and Van Cura, L. J. (1968). Patient reaction to computer-based medical interviewing. *Computational Biomedical Research, 1(5),* 527-531.

[25] Rizzo, A., Shilling, R., Forbell, E., Scherer, S., Gratch, J., & Morency, L. (2016). Autonomous Virtual Human Agents for Healthcare Information Support and Clinical Interviewing. *Artificial Intelligence in Behavioral and Mental Health Care,* 53–79.

[26] DeVault, D., Artstein, R., Benn, G., Dey, T., Egan, A., Fast, E., Georgila, K., Gratch, J., Hartholt, A., Lhommet, M., Lucas, G. M., ... & Morency, L. P. (2014). SimInterviewer: A virtual human interviewer for healthcare decision support. *Proceedings of the 13th conference on Autonomous Agents and Multi-Agent Systems,* 1061-1068.

[27] Gratch, J., DeVault, D., Lucas, G. M., & Marsella, S. (2015). Negotiation as a challenge problem for virtual humans. *Proceedings of the 15th International Conference on Intell gent Virtual Agents,* 201-215.

[28] Davis, J. W., & Bobick, A. F. (1998). Virtual PAT: A virtual personal aerobics trainer. In *Workshop on Perceptual User Interfaces* (pp. 13–18).

[29] Wang, Y., Lucas, G. M., Khooshabeh, P., de Melo, C., & Gratch, J. (2015). Effects of emotional expressions on persuasion. *Social Influence, 10,* 236-249.

[30] Flett, M. R., Gould, D., Griffes, K. R., & Lauer, L. (2013). Tough love for underserved youth: A comparison of more and less effective coaching. *The Sport Psychologist, 27(4),* 325-337.

[31] Erickson, K., & Côté, J. (2015). The intervention tone of coaches' behaviour: Development of the Assessment of Coaching Tone (ACT) observational coding system. *International Journal of Sports Science & Coaching, 10(4),* 699-716.

[32] Thiebaux, M., Marsella, S., Marshall, A. N., & Kallmann, M. (2008, May). Smartbody: Behavior realization for embodied conversational agents. In *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems-Volume* 1 (pp. 151-158).

[33] Gilani, S., Sheetz, K., Lucas, G. M., & Traum, D. (2016). What kind of stories should a virtual human swap? *Proceedings of the 16th International Conference on Intell gent Virtual Agents,* 124-136.

[34] Ferguson, E., & Cox, T. (1993). Exploratory Factor Analysis: A User's Guide. *International Journal of Selection and Assessment, 1(2),* 84–94.

[35] Kang, S.-H., & Gratch, J. (2012). Socially anxious people reveal more personal information with virtual counselors that talk about themselves using intimate human backstories. In *B K Wiederhold & G Riva (Eds), Studies in health technology and informatics: Vol 181* Amsterdam: IOS Press.

[36] American Heart Association. *American Heart Association Guidelines: Physical Activity*

[37] Khashe, S., Lucas, G. M., Becerik-Gerber, B., & Gratch, J. (2017). Buildings with persona: Towards effective building-occupant communication. *Computers in Human Behavior, 75,* 607-618.

[38] Khashe, S., Heydarian, A., Becerik-Gerber, B., & Wood, W. (2016). Exploring the effectiveness of social messages on promoting energy conservation behavior in buildings. *Building and Environment, 102,* 83-94.

[39] Khooshabeh, P., & Lucas, G. M. (2018). Virtual human role players for studying social factors in organizational decision making. *Frontiers in Psychology, 9*:194.

[40] Kaptein, M., Lacroix, J., & Saini, P. (2010). Individual Differences in Persuadability in the Health Promotion Domain. In T. Ploug, P. F. V. Hasle, & H. Oinas-Kukkonen (Eds.): *Vol 6137 LNCS sublibrary SL 3, Information systems and application, Incl Internet/Web and HCI, Persuasive technology 5th international conference, PERSUASIVE 2010,* Copenhagen, Denmark, June 7-10, 2010 ; proceedings (pp. 94–105). Berlin: Springer.

[41] Halko, S., & Kientz, J. A. (2010). Personality and Persuasive Technology: An Exploratory Study on Health-Promoting Mobile Applications. In D. Hutchison, T. Kanade, J. Kittler, J. M. Kleinberg, F. Mattern, J. C. Mitchell, H. Oinas-Kukkonen (Eds.), *Lecture Notes in Computer Science Persuasive Technology* (pp. 150–161). Berlin, Heidelberg: Springer.

[42] Nass, C., Moon, Y., & Green, N. (1997). Are Machines Gender Neutral? Gender-Stereotypic Responses to Computers With Voices. *Journal of Applied Social Psychology, 27(10),* 864–876.

[43] Krämer, N., Karacora, B., Lucas, G. M., Dehghani, M., Rüther, G., & Gratch, J. (2016). Closing the gender gap in STEM with male instructors? On the effects of rapport behavior and gender of a virtual agent in an instructional interaction. *Computers and Education, 99,* 1-13.