Dancing with Physio: A Mobile Game with Physiologically Aware Virtual Humans

Jorge Arroyo-Palacios and Mel Slater

Abstract—This study presents an evaluation of a mobile game with physiologically aware virtual humans as an approach to modulate the participant's affective and physiological state. We developed a mobile version of a virtual reality scenario where the participants were able to interact with virtual human characters through their psychophysiological activity. Music was played in the background of the scenario and, depending on the experimental condition, the virtual humans were initially either barely dancing or dancing very euphorically. The task of the participants was to encourage the apathetic virtual humans to dance or to calm down the frenetically dancing characters, through the modulation of their own mood and physiological activity. Results from our study show that by using this mobile game with the physiologically aware and affective virtual humans the participants were able to emotionally arouse themselves in the *Activation* condition and were able to relax themselves in the *Relaxation* condition, during the same session with only a brief break between conditions. The self-reported affective data was also corroborated by the physiological data (heart rate, respiration and skin conductance) which significantly differed between the *Activation* and *Relaxation* conditions.

Index Terms—Biofeedback, physiological computing, virtual human characters, mood modulation, mobile game

✦

1 INTRODUCTION

DEPRESSION and anxiety are common mental disorders. According to the World Health Organization (WHO), around 14 percent of the world population suffers from at least one severe mood disorder during the course of their life. In Europe, depression and anxiety, affects around 25 percent of the population, cause up to 50 percent of the total chronic sick leaves, and cost about €170 billion per year [1]. One paradox of this global problem is that these disorders can be treated; however, as reported by WHO, more than 50 percent of the people suffering from these disorders do not receive treatment. Four factors contributing to this incongruity have already been identified in [2]: cost of medication or therapy, insufficient mental health professionals, difficulty to access mental health services (due to distance or time), and reluctance to look for professional help. Currently, research organizations, like the National Institute of Mental Health in the United States of America see smartphones, wearable sensors and video games as behavioral intervention technologies that can support and facilitate the access to treatment. Nonetheless, as they acknowledge in their strategic plan for research [2], further investigation is needed to understand the capabilities of these technologies and how they can be used more effectively.

In this study we contribute to these research efforts by investigating the feasibility of interaction with physiologically aware virtual human characters in a mobile game context as one approach to mood and physiological modulation. With increased exposure to computer systems, interaction with such virtual characters has become more common. Virtual characters are a tool already in use in a number of applications in different areas such as education, medicine, training, psychology, entertainment, customer services, etc., for comprehensive reviews see [3], [4]. A number of studies have investigated the various effects of interaction with virtual characters on learning, e.g., [5], task performance, e.g., [6], behavior, e.g., [7], decision making, e.g., [8], perception of task difficulty, e.g., [9] and the "realistic" responses of humans to the behavior of virtual humans, e.g., [10]. Despite the increase in the number of studies that propose new applications for virtual characters, even in complex social interactions with humans for instance for psychotherapy [11], there has been less research into the affective states or moods portrayed by the virtual characters and the affective and physiological effects on the participants.

The work presented in this paper combines two areas of research: interaction with affective virtual humans and biofeedback. We developed an interactive game system that simulates a dancing scenario for mobile platforms (iOS and Android). The interaction method used between the participants and the virtual humans consisted of a bidirectional physiological and mood contagion (Fig. 1). The participants were told that they would have to change their own psychophysiological state to influence the mood of the virtual humans, and at the same time the observable changes in the affective state and behavior of the virtual characters were aimed to reinforce the psychophysiological changes of the participants. One of the main motivations to carry out the evaluation of the system on a portable device was to

- *J. Arroyo-Palacios is with the Psychology, EVENTLAB, University of Barcelona, Barcelona 08011, Spain. E-mail: jorgearroyo@ub.edu.*
- *M. Slater is with the EVENTLAB, University of Barcelona, Barcelona, Spain, the Department of Computer Science, University College London, United Kingdom, and Institució Catalana de Recerca I Estudis Avançats, Barcelona, Spain. E-mail: melslater@ub.edu.*
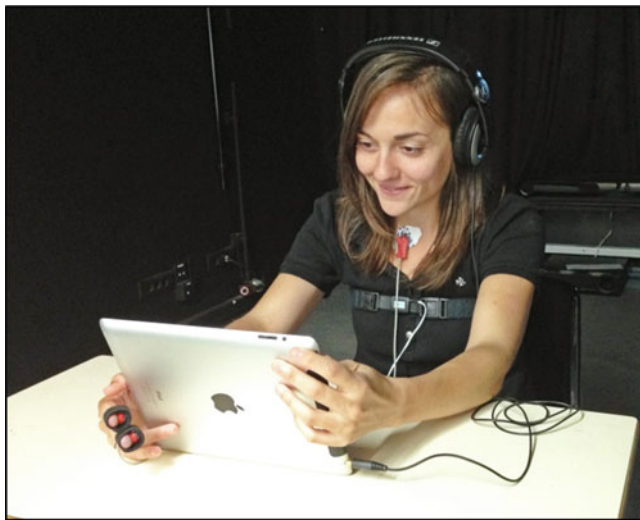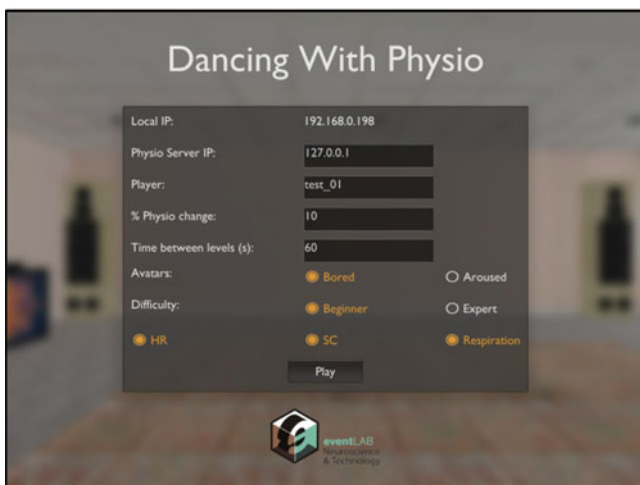
Fig. 1. a) Participants interacting with the virtual humans of the game through the use of their own psychophysiological activity, b) Initial configuration screen of the game.

evaluate the mood and physiological effects of the application without the influence of physical activity or overt movements from the participants. A second motivation was the expanding use of home health technologies: from 14.3 million consumers in 2014 to an estimate of 78.5 million by 2020 [12]. The current rise of wearable computing makes this a propitious moment to exploit this proposed "natural" and intuitive human-computer psychophysiological interaction method.

The intrinsic characteristics of video games (fun, engaging, interactive) offer favorable characteristics for healthcare applications. The increasing research and interest in their use has motivated the creation of dedicated conferences and journals such as Games for Health, as part of the Serious Games initiative. In particular, recent studies provide clinically useful information about the use of videogames for the treatment of adolescent depression [13] and for decreasing stress [14].

The research carried out in this study provides two main contributions. Firstly, we propose a novel application for relaxation and activation training. Secondly, we provide empirical evidence for the effectiveness of the use of a game with physiologically aware and affective virtual humans as a mood and physiological modulation tool, even in a portable device with non-highly immersive virtual reality capabilities. We compared the affective and physiological changes in two versions of the game, one that aims to arouse and another that aims to relax. In addition to self-reports on the affective state we also provide objective data from three physiological measures: heart rate (HR), respiration and skin conductance (SC).

Given the engaging capabilities of videogames, the affective capabilities of virtual humans, and advances in wearable devices, the findings from this study could be useful and encouraging to have more research and future development of home or on-the-go applications to support and complement the treatment of mood disorders such as depression and anxiety.

2 RELATED WORK

2.1 Emotional Contagion and Virtual Characters

Emotional contagion can be defined as the transfer of mood from person to person in a group [15]; the tendency to catch or adopt the emotions of other people by unconscious mimicry and synchronization of expressions, vocalizations, postures and movements [16]. Research has demonstrated that emotional contagion is not solely limited to direct interaction with real humans but also with the expression of the emotions delivered by different media and technology, such as static pictures of human facial expressions [17], video clips [18] and robots [19]. A recent study conducted by Facebook, with substantial media coverage due to ethical concerns revealed that emotional contagion is also possible through social networks [20].

A growing number of studies are investigating the potential of virtual characters or agents with emotional expression to enhance interaction with computers, user experience and the effectivity of the applications, for a review see [21], [22], [23], [24], [25], [26], [27]. This research trend on affective virtual characters can be observed in dedicated international research projects (e.g., HUMAINE,[1] SEMAINE[2] and EMOTE[3]), conferences, workshops and journals. For instance, in the area of education and learning, Maldonado et al. [28] found that an emotional agent that acted as a co-learner had a more positive influence on the user performance and experience than an unemotional agent in a web-based English language learning environment. Another study by Burleson and Picard [29] found gender-specific differences in the perception of learning companions that were offering either affective or task support.

In gaming applications, emotional agents have been used to make the user experience more engaging and entertaining. Agents with empathic emotions, where virtual characters would express positive emotions when the user wins a game and negative emotions when the user loses, have been perceived as more caring, likeable, trustworthy and supportive [30]. This influence on the user of empathic agents

---

1. http://emotion-research.net/projects/humaine
2. http://www.semaine-project.eu/
3. http://www.emote-project.eu/

has also been corroborated with physiological data in [31] where the skin conductance of subjects interacting with an empathic agent was significantly lower than the skin conductance of subjects interacting with a non-empathic agent. Klein et al. [32] and Hone [33] have found that empathic agents reduce feelings of frustration more effectively than agents that are not empathic.

Researchers have also explored affective agents for behavior change, for instance, for completing required levels of exercise [34], for therapeutic interventions for fear of public speaking [35] and social anxiety [10].

Despite increasing interest in virtual characters, there is limited research into the contagion effect of the emotion or mood portrayed by the characters. In particular, there is little research into the actual affective and physiological responses of the participants, and how these affective states portrayed by the characters can be used in a biofeedback protocol to assist and reinforce the physiological and mood modulation of participants.

Some of the few studies that specifically investigate emotional contagion with virtual characters include the work of Tsai et al. [36] where participants reported their emotional experience when viewing static images of emotional characters. Nonetheless, there is no physiological data that corroborates that the emotional self-reported responses corresponded to emotions actually being experienced by the users, rather than corresponding to the affective attributions of the pictures (a commonly discussed issue on the study of emotions, see [37].

Another relevant example is the study of Wu et al. [38] where a virtual human animation in a medical virtual reality system for education was used to investigate how it affected the emotions of users. Self-reported affective and physiological responses of participants were examined, however the study was focused more on the emotional responses rather than on emotional contagion. Other research involving emotional contagion includes the studies of laugh-aware virtual agents [39].

2.2 Biofeedback

The majority of currently available biofeedback systems employ the use of auditory (sonification or musical), visual or vibrotactile cues to provide information about the ongoing psychophysiological activity of the user (see [40], [41], [42] for a review). However, advances in computing technology have enabled researchers to use state-of-the-art systems with 3D graphics and immersive virtual reality to provide physiological feedback (e.g., in [43]). As compared to traditional biofeedback systems, these newer approaches provide a higher effectivity, while also making the experience much more engaging and enjoyable for the participant [44].

Although there are many commercial and research applications using different approaches for biofeedback, there is little research that has investigated the combined use of moods in virtual human characters and biofeedback as a tool for psychophysiological modulation. One example is the recent study by Chittaro and Sioni [45] where a relaxation training application provides the participant with explicit feedback about their detected affective state. The

participant is presented with a virtual character that has to complete an office task. The more relaxed the participant is, the more relaxed and happy this makes the avatar which leads to more progress in the task. As the participant's stress level increases (e.g., with distracting stressors such as a phone ringing constantly), the avatar becomes more and more unsettled and less focused on the work.

A few mobile applications that employ the traditional visual or auditory biofeedback cues can also be found in the AppStore and Google Play. There are even less mobile applications that support mood disorders, for example: iMoodJournal, eMoods, T2 Mood Tracker, Optimism, Moody Me, etc. (see [46]). However, these apps function more like personal journals to help the user to keep track of moods, emotions, thoughts, along with other symptoms.

In this study we exploit the favorable characteristics of different areas: the engaging capabilities of videogames, the affective capabilities of virtual humans, and the advances in biofeedback and wearable computing. The question that we aim to answer is: how feasible and effective is the combined use of affective virtual humans in a portable biofeedback game protocol to assist and reinforce the mood and physiological modulation?

Our work contributes to the research in interaction with affective virtual humans, extending the work in various directions such as by taking forward the work reported by Tsai et al. [36] and providing physiological data that corroborates the change in the affective state of participants. Also, this study broadens the work carried out by Chittaro and Sioni [45] as it reports the results not only of relaxation but also activation, and employs different physiological measures. Lastly, our study differs from the work of Wu et al. [38] by focusing on the reinforcing contagion effects of the same affective states portrayed by the virtual characters in a bio-feedback protocol, rather than on emotional responses that differ from the affective animations of the characters.

3 MATERIALS AND METHODS

3.1 Experimental Design

A within-subjects experimental design explored the effects of two different conditions: *Activation* and *Relaxation*. Each condition had three different levels, and in order to move from one level to the next, participants had to reach predefined physiological state targets. A total of 21 subjects participated in this experiment, 13 females and eight males. The mean age was 22 years (S.D. four years). They were recruited from the University Campus and by email using the database of participants from our research group.

The central hypothesis in our experiment was that the biofeedback game with the physiologically aware and affective virtual human characters would be effective to significantly influence the mood and physiological activity of the participants, even in a portable and non-immersive virtual reality setting. We formulated our hypothesis based on the literature about the affective capabilities of virtual human avatars and our preliminary data gathered in our pilots. With regards to the specific experimental conditions we hypothesized that by using the game participants would be able to achieve opposite affective states in the arousal dimension (measured by SAM questionnaire) and opposite

Fig. 2. a) Apathetic virtual humans barely dancing, b) euphoric virtual humans dancing frenetically.

physiological states (measured by the heart rate, respiration and skin conductance), during the same session with only a brief break between conditions.

3.2 Virtual Scenario and Game Mechanics

An iPad3 was used for the presentation of the virtual scenario. The virtual scenario used for the experimental study was a large room with a sound system, a set of large speakers and a DJ desk on one side of the virtual room. Music was played in the background and two virtual human characters were dancing in the center of the room, in front of the participant.

The experimental task was designed with videogame-like features in order to make it more fun and engaging: it therefore had challenges, achievements, time limits and progression through levels. The goal and the progression of the game varied depending on the experimental condition. In the *Activation* condition participants were encouraged to increase their physiological activity to animate the apathetic virtual humans that are seen not to be dancing (Fig. 2a). The game starts with a reference period of 60 seconds during which time, slow tempo music is played and the virtual humans are idle. After these 60 seconds, level 0

starts with a change in the music and in the animation of the virtual characters. Participants have 60 seconds to increase their physiological activity and reach the targeted physiological change to move to the next level. If the target is achieved at any point during the period of time, the game moves to the next level, indicated with a change in the music and in the dance style of the virtual humans. At the new level, participants have another 60 seconds to reach the next target and move to the following level. If the participant does not reach the targeted physiological change in the period of time for the level, the application ends and displays a "Game Over" message. On the other hand, if the participant gets to level 3, the characters perform a choreographed dance of "Gangnam Style" followed by a "Game Completed" message.

In the *Relaxation* condition participants were instructed to relax and slow down their physiological activity to calm the overexcited virtual humans (Fig. 2b). The reference period consists of an energetic piece of music with fast tempo and animations of excited characters dancing euphorically. After the reference period, level 0 starts with the virtual humans dancing less energetically to slightly calmer music. Participants then have 60 seconds to decrease their physiological activity and reach the targeted physiological change. As the participant moves forward in the levels of the game, the music and the characters' dancing slow down. The game ends at level 3 with the virtual humans dancing the Waltz while slow tempo pop music is played in the background.

The application can be easily set with different types of initial configurations (e.g., percentage of physiological change required to move from level to level, physiological measures to monitor and its weights, game difficulty, etc., see Fig. 1b).

In this first evaluation of this type of game with physiologically aware virtual humans three physiological measures were used with equal weight: SC, HR and Respiration. The game difficulty was set to an easy mode where if any of the physiological signals passed the threshold at least once during the level duration, the game moved forward to the next level. The idea behind this decision was to allow users to realize that they can control their physiological activity and at the same time reinforce their psychophysiological changes with the empathic and evocative affective response of the virtual human characters.

The physiological percentage changes (*ppCh*) required to move through the levels of the game were set to: 10 percent for level 1, 20 percent for level 2 and 30 percent for level 3, with respect to the reference period (baseline). These physiological percentage changes were based on results from some pilot tests, paying particular attention to the trade-off of keeping the game effective and fun, without making it too difficult to avoid frustration. For each participant, the specific heart rate, respiration rate and skin conductance threshold values (*tVal*), at a particular game level (*i*), were calculated as shown in Equation (1):

$$tVal_i = Baseline \pm (Baseline * ppCh_i). \qquad (1)$$

The experimental condition determined the direction of the physiological change required to move forward in the levels of the game (i.e. if the participant should increase or

TABLE 1
Music for the Different Game Levels

	Reference	Level 0	Level 1	Level 2	Level 3
	60 seconds	60 seconds	60 seconds	60 seconds	60 seconds
	-	10%	20%	30%	N/A
Activation	Life_Sleep (Iamus) 40 bpm.	PD1_Slight (Iamus). 60-80 bpm	E5_Decent (Iamus). 100-130 bpm.	PD3_Exciting (Iamus). 150-170 bpm.	Gangnam Style (Psy). 132 bpm.
Relaxation	Misirlou (Dick Dale). 173 bpm.	E4_Slight (Iamus). 60-80 bpm	Life_Nap (Iamus). 60 bpm	Life_Relaxing (Iamus). 40 bpm	Just give me a reason (Pink). 95 bpm

Music, tempo in beats per minute (bpm), duration, and physiological percentage change targets used for the different stages of the game for each of the experimental condition.

decrease their physiological activity as compared with the activity of the reference period).

Given this configuration of the game, the fact that participants are able to move from level to level doesn't guarantee that there was an important physiological change. A very brief response or effort could cause the game to move forward on the levels. Therefore it is important to assess and quantify the true impact on the physiological activity of the participants while playing the game with the affective virtual humans.

3.3 Music

The music plays an important role in the biofeedback protocol of the experiment. It has a dual purpose: firstly to complement the stimulation and help the participants to modulate their own psychophysiological activity, and secondly to provide feedback about the change of the level in the game.

Most of the music was provided by the Melomics computational system.[4] This music is composed automatically without human intervention. However, three of the songs are popular commercial ones. Table 1 presents a summary showing the music, tempo, length and targeted percentage of physiological change of each level of the game.

3.4 Virtual Humans' Animations

The dance animations of the virtual humans were recorded from volunteers who were asked to dance to the different pieces of music previously selected for each of the levels of the game. OptiTrack ARENA from NaturalPoint was used for the motion tracking.

The volunteers wore the motion tracking suit with markers that were tracked by 12 infrared cameras (see Fig. 3). The volunteers' dance moves were recorded and the animations were refined in Motion Builder to be applied to the virtual humans. Only one animation (for the Gangnam Style song) was downloaded from the Mixamo online platform.[5]

3.5 Measures

3.5.1 Physiological Measures

Heart rate, skin conductance and respiration measures from the participants were recorded. These measures were selected due to their relevance in the discrimination of

arousal and relaxation, their easy acquisition and their practicality for the current and future applications.

ECG, GSR and respiration sensors from g.tec[6] in combination with Matlab code was used for real-time recording and processing of the physiological signals (physiological recording system described in [47]). A bipolar ECG recording with an abdominal placement setup was employed: the positive electrode was placed below the ribs on the left, the ground electrode at the same level on the right, and the negative electrode was placed on the upper sternum area. The GSR electrodes were placed on the index and middle fingers without electrode gel. The respiration sensor consisted of a piezo-electric crystal sensor embedded on a belt system which was placed around the chest of the participant.

3.5.2 Mood and Affective Measures

The Emotional Contagion Scale [48] was used to measure the participants' susceptibility to being influenced by the emotions of others. It consists of a 15-item index to examine a person's tendency to be influenced by surrounding people in five basic emotions: sadness, fear, anger, happiness and love. Each item is scored using a four-point scale with the following key: 4—always; 3—Often; 2—Rarely; 1—Never. The higher the total score, the more susceptible to emotional contagion a person would be said to be.

The mood of the participants was evaluated before and after interaction with the virtual humans. It was assessed using the Profile of Mood States (POMS) questionnaire [49], which provides a method to assess transient and fluctuating mood states. The POMS version used in this study contains 58 items which are rated on a 5-point scale ranging from "Not at all" to "Extremely". The POMS measures six identified mood factors: *tension, depression, anger, vigor, fatigue and confusion.*

A 5-point scale Self-Assessment Manikin (SAM) questionnaire [50] was also used to assess the affective state of the participants (in particular the valence and arousal) before and after experiencing the portable game with the physiologically aware virtual human characters. The SAM questionnaire uses graphic scales depicting cartoon characters expressing different levels of affective states in three dimensions: valence, arousal and dominance. In this study we were interested in measuring affective states only in the valence and arousal dimensions.

4. http://melomics.com/
5. https://www.mixamo.com/

6. http://www.gtec.at/content/view/full/2

Fig. 3. Motion capture setting of the dance animations.

Finally, a post-experiment questionnaire was used to evaluate the feeling of presence in the virtual scenario and the general experience of the game while interacting with the virtual human characters.

3.6 Experimental Procedure

First, the participants read an information sheet including an explanation about the study. After that, they were given a consent form and completed a demographics questionnaire, the Emotional Contagion Scale and the POMS questionnaire.

It was explained to the participants that the application on the iPad was similar to a videogame with three levels where the objective was to cheer up and encourage dancing or relax and calm down the virtual humans characters (depending on the experimental condition). They were also informed that they would directly influence the behavior of the characters and move to the next level with changes in their own physiological activity and affective state. At the end of the explanation, participants were taken to one lab, seated on an office chair with a desk placed in front of them. They were instructed to rest their arms over the desk, hold the iPad with both hands and avoid overt movements to reduce the number of artifacts. Physiological sensors and headphones were then placed and the experiment started.

The conditions (*Activation* or *Relaxation*) were presented in counter balanced order to the participants. After each condition the POMS and SAM questionnaires were given to the participant. Each session lasted 45 minutes approximately and the participant played the game on the iPad for around 12 minutes. At the end of the session the participants filled out a post-session questionnaire, had a brief semi-structured interview and were paid 5 euros for their participation.
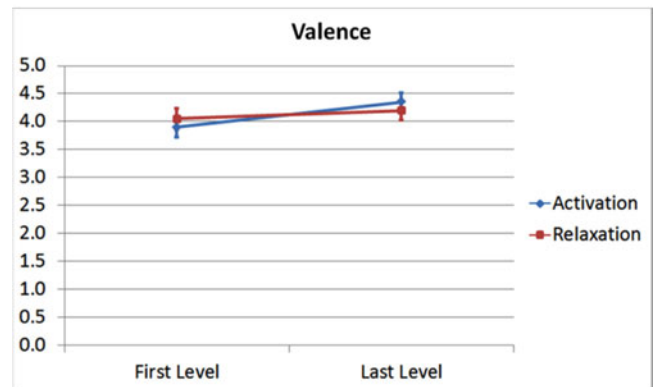
4 RESULTS

4.1 Game Completed

The data of one out of the 21 participants had to be discarded due to technical problems with the sensors. The
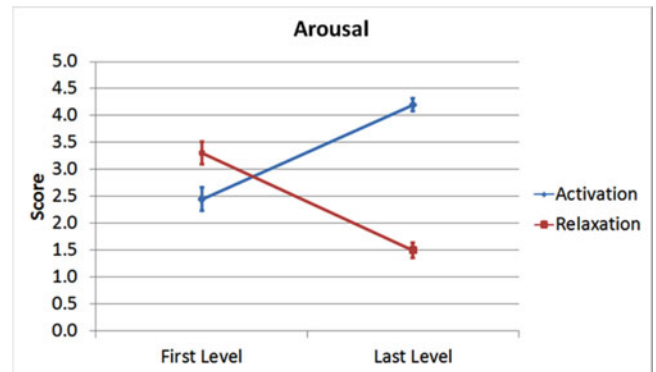
Fig. 4. Results from the five-point scale SAM questionnaire: means and standard errors in the a) arousal and b) valence dimensions.

participants played each of the two conditions of the game once only. In total 20 participants successfully finished the three levels of the application on the *Relaxation* condition and 19 participants successfully finished the three levels on the *Activation* condition. Only one participant did not finish the three levels of the game, reaching only level 2 in the *Activation* condition.

4.2 Affective State and Mood

Participants had a mean total score of 44.09 (SD 6.06) on the susceptibility to being influenced by the emotions of others given by the Emotional Contagion Scale (the maximum possible score value is 60). The following analyses of the SAM and POMS questionnaires show the changes in the mood and affective state after experiencing the conditions on the iPad application.

4.2.1 SAM Questionnaire

Fig. 4 shows the affective state of the participants on the arousal and valence dimensions obtained from the SAM questionnaire. Fig. 4a shows how the participants on the *Activation* condition started with an average arousal score of 2.45 (SD 0.94) and experienced an increase in their arousal level at the end of the last level of the game by reaching an average arousal of 4.2 (SD 0.52), on a 5-point scale. On the other hand, the *Relaxation* condition produced the opposite effects: participants started with an average arousal score of 3.3 (SD 0.92) and at the end of the last level their average arousal score decreased to 1.5 (SD 0.60).
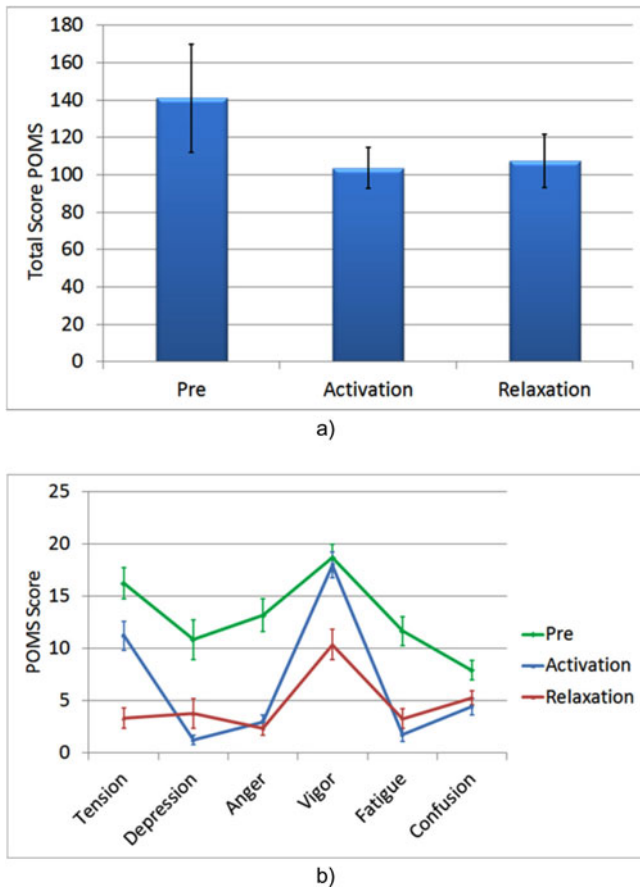
Fig. 5. Results from POMS questionnaire: mean and standard errors in a) total scores, b) scores by factors.

Fig. 4b presents the scores on the valence affective dimension, which reflect how pleasant or unpleasant the participant's experience was at the start and end of each of the conditions on the iPad application. Participants reported a pleasant experience in both conditions with

a small positive change between the first and last level of the game.

Wilcoxon signed-rank tests were used to compare the scores between the first and last level of both *Activation* and *Relaxation* conditions. The results from these tests showed that during the *Activation* condition, participants significantly increased their levels of arousal ($Z = 3.89, p < 0.001$, $r = 0.61$), however the valence scores remained positive, without reaching significant changes ($Z = -1.81, p = 0.07$, $r = 0.29$). Results from the *Relaxation* condition showed that participants significantly reduced their arousal levels after the last level of the game ($Z = -3.77, p < 0.001, r = 0.60$) and maintained a pleasant experience without significant changes on the valence scores from the first level ($Z = -0.79, p = 0.42, r = 0.13$).

4.2.2 POMS Questionnaire

By looking at Fig. 5a it can be observed that there is a change in the mood of participants when comparing the total POMS scores (indicator of total mood disturbance): i) before and after the *Activation* condition and, ii) before and after the *Relaxation* condition. An ANOVA revealed that the difference between the POMS scores before the experiment (*PRE*), after *Activation* and after *Relaxation* is significant ($F(2, 38) = 30.24, p < 0.001, \eta^2 = 0.61$). Post-hoc analysis also corroborated that there were significant differences when comparing the POMS scores between *PRE* and *Activation* ($p < 0.001$) and between *PRE* and *Relaxation* ($p < 0.001$).

The changes in the individual factors evaluated in the POMS questionnaires can be appreciated in Fig. 5b. Tension decreased after the *Activation* condition but the decrease is even greater after the *Relaxation* condition. Other factors such as depression, anger, fatigue and confusion also decreased in both conditions. As it would be expected, the factor vigor only decreased significantly on the *Relaxation* condition. Table 2 presents a summary of the results from

TABLE 2
Statistical Results in the Different POMS' Factors

Factor	Comparisons			Z	p	Adjusted significance level $p_i = \frac{Q}{n} * i$
Tension	PRE	vs.	ACT	-2.56	*0.010	$(0.05/18)^*13 = 0.036$
	PRE	vs.	REL	-3.926	*0.000	$(0.05/18)^*1 = 0.003$
	ACT	vs.	REL	-3.926	*0.000	$(0.05/18)^*2 = 0.006$
Depression	PRE	vs.	ACT	-3.87	*0.000	$(0.05/18)^*3 = 0.008$
	PRE	vs.	REL	-3.83	*0.000	$(0.05/18)^*4 = 0.011$
	ACT	vs.	REL	-2.06	*0.040	$(0.05/18)^*15 = 0.042$
Anger	PRE	vs.	ACT	-3.92	*0.000	$(0.05/18)^*5 = 0.014$
	PRE	vs.	REL	-3.87	*0.000	$(0.05/18)^*6 = 0.017$
	ACT	vs.	REL	-1.90	0.058	$(0.05/18)^*16 = 0.044$
Vigor	PRE	vs.	ACT	-0.57	0.571	$(0.05/18)^*18 = 0.050$
	PRE	vs.	REL	-3.50	*0.000	$(0.05/18)^*7 = 0.019$
	ACT	vs.	REL	-3.18	*0.001	$(0.05/18)^*10 = 0.028$
Fatigue	PRE	vs.	ACT	-3.73	*0.000	$(0.05/18)^*8 = 0.022$
	PRE	vs.	REL	-3.62	*0.000	$(0.05/18)^*9 = 0.025$
	ACT	vs.	REL	-1.44	0.150	$(0.05/18)^*17 = 0.047$
Confusion	PRE	vs.	ACT	-3.29	*0.001	$(0.05/18)^*11 = 0.031$
	PRE	vs.	REL	-2.75	*0.006	$(0.05/18)^*12 = 0.033$
	ACT	vs.	REL	-2.30	*0.022	$(0.05/18)^*14 = 0.039$

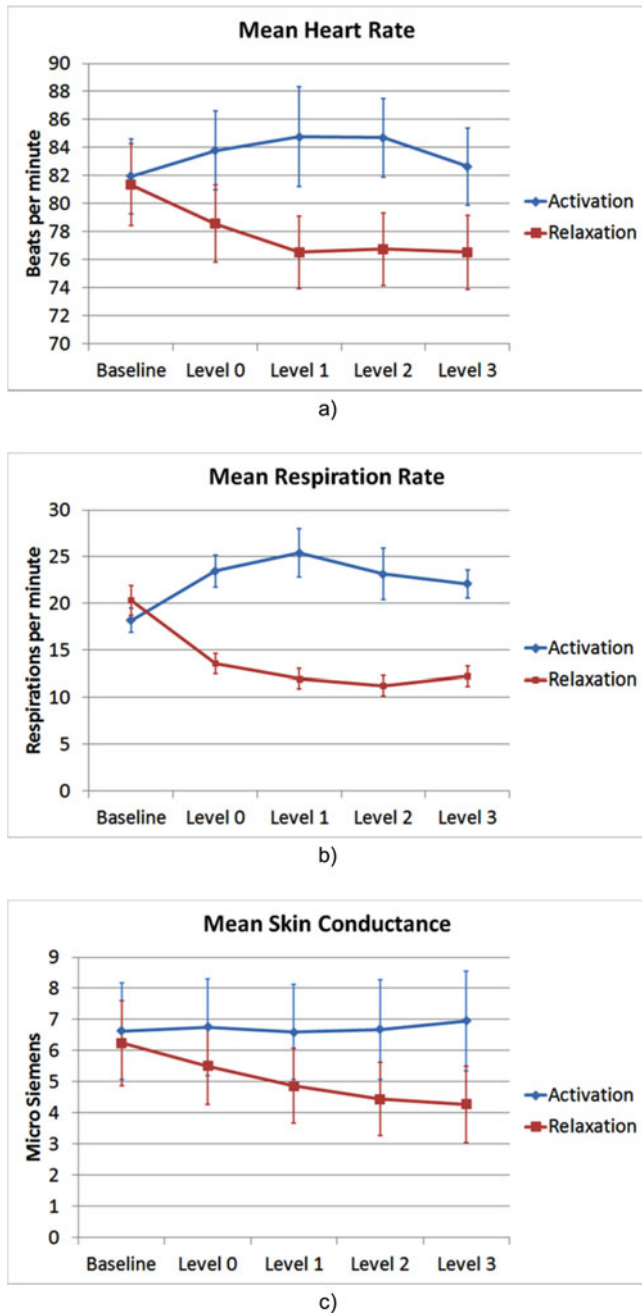*ACT = Activation, REL = Relaxation, * = significant p-value (< adjusted significance level).*

Fig. 6. Means and standard errors of the physiological variables in the two experimental conditions: a) heart rate, b) respiration rate, c) skin conductance.

the Wilcoxon signed rank tests for each of the factors evaluated in the POMS questionnaire.

4.3 Physiological Measures

The analysis of the physiological data was carried out with a 2×5 repeated measures ANOVA: condition (*Activation* and *Relaxation*) and *level* of the game (baseline, level 0, level 1, level 2, level 3). For the post-hoc analyses the significance level was adjusted for multiple comparisons using the method by Benjamini & Hochberg [51].

A visual analysis of the physiological variables (see Fig. 6) shows how the participants started with a similar baseline in both experimental conditions. However, the

TABLE 3
Comparisons of Heart Rate

Comparison Activation versus Relaxation	p	Adjusted significance level $p_i = \frac{Q}{n} * i$
Level 3	$P < 0.000^*$	$(0.05/5)^*1 = 0.01$
Level 2	$P = 0.002^*$	$(0.05/5)^*2 = 0.02$
Level 1	$P = 0.015^*$	$(0.05/5)^*3 = 0.03$
Level 0	$P = 0.033^*$	$(0.05/5)^*4 = 0.04$
Baseline	$P = 0.786$	$(0.05/5)^*5 = 0.05$

*Comparisons of Heart Rate between the corresponding levels from the Activation and Relaxation conditions. * = significant p-value (< adjusted significance level).*

participants experienced greater heart rate, respiration and skin conductance activity in each of the levels of the *Activation* condition compared to each of the corresponding levels of the *Relaxation* condition.

4.3.1 Heart Rate

The 2×5 ANOVA on HR corroborated that there is a significant effect by the experimental *condition* factor $F(1, 18) = 10.92, p = 0.004, \eta^2 = 0.38$ and in the interaction *condition* x *level* of game $F(2.70, 48.55) = 3.89, p = 0.017, \eta^2 = 0.18$. As expected, there was no significant difference in the factor *level* of game $F(2.14, 38.44) = 0.758, p = 0.48, \eta^2 = 0.04$, due to the fact that the HR was increasing in the *Activation* and decreasing in the *Relaxation* condition, thereby reducing the total change between the different levels.

Post-hoc analyses corroborate that participants had a significantly higher HR in each of the levels of the game in the *Activation* condition than in the *Relaxation* condition except during the baseline period. Table 3 presents a summary of these multiple comparisons with the corresponding adjusted p-value.

4.3.2 Respiration

The 2×5 ANOVA on the respiration rate corroborated that there is a significant effect on the experimental *condition* factor $F(1, 18) = 21.75, p < 0.001, \eta^2 = 0.55$ and in the interaction *condition* x *level* of game $F(2.16, 38.93) = 11.95, p < 0.001, \eta^2 = 0.40$. However, as each condition evokes respiration changes in opposite directions, there was no significant difference in the factor *level* of the game $F(2.37, 42.73) = 2.02, p = 0.137, \eta^2 = 0.10$.

Post-hoc analysis corroborates that the respiration rate was significantly higher in each of the game levels on the *Activation* condition than in the corresponding game levels of the *Relaxation* condition. There was no significant difference between the baselines of each condition. Table 4 presents a summary of these multiple comparisons with the adjusted p-value.

4.3.3 Skin Conductance

The analysis of the skin conductance was carried with Wilcoxon signed-rank tests due to the non-normally distributed residuals. The results from these tests corroborate that participants had a significantly higher skin

<div style="text-align:center">

TABLE 4
Comparisons of Respiration Rate

</div>

Comparison Activation versus Relaxation	p	Adjusted significance level $p_i = \frac{Q}{n} * i$
Level 0	$P < 0.000^*$	$(0.05/5)^*1 = 0.01$
Level 1	$P < 0.000^*$	$(0.05/5)^*2 = 0.02$
Level 3	$P = 0.000^*$	$(0.05/5)^*3 = 0.03$
Level 2	$P = 0.001^*$	$(0.05/5)^*4 = 0.04$
Baseline	$P = 0.247$	$(0.05/5)^*5 = 0.05$

Comparisons of Respiration Rate between the corresponding levels from the Activation and Relaxation conditions. * = significant p-value (< adjusted significance level).

<div style="text-align:center">

TABLE 5
Comparisons of Skin Conductance

</div>

Comparison Activation versus Relaxation	p	Adjusted significance level $p_i = \frac{Q}{n} * i$
Level 3	$P = 0.006^*$	$(0.05/5)^*1 = 0.01$
Level 2	$P = 0.007^*$	$(0.05/5)^*2 = 0.02$
Level 1	$P = 0.048$	$(0.05/5)^*3 = 0.03$
Level 0	$P = 0.263$	$(0.05/5)^*4 = 0.04$
Baseline	$P = 0.911$	$(0.05/5)^*5 = 0.05$

Comparisons of Skin Conductance between the corresponding levels from the Activation and Relaxation conditions. * = significant p-value (< adjusted significance level).

conductance during levels 2 and 3 of the *Activation* condition than in the corresponding levels of the *Relaxation* condition. The test on the difference of level 1 from the two conditions reaches a p-value of 0.048, however the adjusted p-value, due to the multiple comparison, leaves the p-value of this level comparison as non-significant. The results also reveal no significant difference during the baseline period of each condition. Table 5 presents a summary of these multiple comparisons with the adjusted significance level.

5 DISCUSSION

The results from this study provide empirical evidence about the feasibility and effectivity of a portable game, with physiologically aware and affective virtual human characters, to help users to change their mood and control their physiology.

We had two main motivations to carry out the first evaluation of the system on a portable device. Primarily, we wanted to see if the interaction with affective virtual characters in a biofeedback game context, even using a non-immersive portable device, would be persuasive enough to influence the mood and physiological activity of participants. If effective, this would encourage development of accessible (home or on-the-go) applications to support and complement the treatment of mood disorders such as depression and anxiety. Secondly, we wanted to evaluate the mood and physiological effects of the application, without the influence of physical activity or movements from the participants.

The results from the SAM questionnaire show the significant change in the self-reported arousal between the start and end of each condition of the game. Moreover, the direction of the arousal changes corresponds to the experimental conditions, increase on the *Activation* and decrease on the *Relaxation*. The responses on the valence dimension of the SAM questionnaire are also positive. They corroborate that regardless of the intention of the game (relaxation or activation) the participants had a positive experience after playing it.

The results on the total score of the POMS questionnaire indicate that there is indeed a change in the mood of the participants before and after playing the game in each of the experimental conditions. Analyses into the individual factors that compose the POMS questionnaire suggest that the application was effective in reducing the levels of depression, tension, anger, fatigue and confusion. As it would be expected, due to the intrinsic nature of the *Activation* condition, the factor *vigor* was only reduced in the *Relaxation* condition.

The physiological data corroborated the mood and arousal changes evoked by the two experimental conditions. The three physiological measures followed the expected direction with the changes, increasing in the *Activation* condition and decreasing in the *Relaxation* condition. Moreover, post-hoc comparisons revealed that the heart rate and respiration rate in each of the game levels of the *Activation* condition were significantly higher than the corresponding level in the *Relaxation* condition; the skin conductance was also higher in the *Activation* condition on levels 1, 2 and 3 but not on level 0.

There are two aspects that strengthen our results, and are important to take into consideration when assessing the effectiveness of this application for mood and physiological modulation. Firstly, while many studies on emotion analyze a short period of time, which usually correspond to the initial response to a stimulus, we used a more conservative approach by analyzing the mean of physiological responses during the whole duration of each game level. Our results showed statistical significance with good effect sizes, even though the mean of a relatively long period of time cancels out short responses, and may include the natural decay that occurs due to habitation effects.

Secondly, we carried out the evaluation of the application in a within-subjects design. This implied that the subjects had to play both conditions of the game in the same session, even though each condition was oriented to evoke completely opposite affective states. Thus, the application encouraged the participants to reach a relaxation state with a 30 percent decrease in the physiological activity, and then after a brief break, the application motivated subjects to arouse and activate themselves by increasing their physiological activity by 30 percent. Although carry over effects from the first random experimental condition could have reduced the effects of the second condition, the game was still effective to evoke both opposite affective states in the participants.

The current focus of companies like Apple, Google, Microsoft and Samsung on wearable devices (e.g., iWatch, Android Wear, Microsoft Band, Simband) physiological sensors, and software platforms (e.g., Apple HealthKit, Google Fit, Microsoft Health, Samsung SAMI), offers a promising

panorama for the development of accessible tools to support the treatment of people with mood disorders.

Although the participants in this study were University students, the information gathered is useful as a first evaluation of this type of application. We hope the results from this study encourage further research into wider types of populations.

The evidence from this study is also relevant given the current and future healthcare trends, and supports the feasibility of the use of affective and physiologically aware virtual humans through portable (home or on-the-go) applications to assist in the modulation of mood and physiological activity.

6 FUTURE WORK

In future studies we will investigate the human-avatar mood contagion within an immersive virtual reality setting. During some pilots carried out in a CAVE system we noticed that the participants were so influenced by the virtual humans that they were actually dancing with them. Given how the virtual humans were able to motivate the participants to dance in the CAVE version, we are planning to evaluate some behavioral responses of the participants, as this could provide us with a higher level understanding of how the affective state displayed by virtual characters can influence people interacting with them. Finally, we also aim to compare the effectivity of affective virtual human avatars for mood and physiological modulation between a portable and an immersive virtual reality system.

REFERENCES

[1] WHO. (2014). Depression in Europe: Facts and figures [Online]. Available: http://www.euro.who.int/en/health-topics/non-communicable-diseases/pages/news/news/2012/10/depression-in-europe/depression-in-europe-facts-and-figures

[2] NIMH, "National Institute of Mental Health strategic plan for research," vol. 1, 2015.

[3] E. Van Lun, X. Verduijn, K. Kuligowska, A. De Wolf, and D. Morton. (2014). Chatbots.org [Online]. Available: http://www.chatbots.org/

[4] J. Cassell, J. Sullivan, S. Prevost, and E. Churchill, *Embodied Conversational Agents*. Cambridge, MA, USA: MIT Press, 2000.

[5] W. Johnson, J. W. Rickel, and J. C. Lester, "Animated pedagogical agents: Face-to-face interaction in interactive learning environments," *Int. J. Artif. Intell. Educ.*, vol. 11, pp. 47–78, 2000.

[6] S. Babu, E. Suma, T. Barnes, and L. F. Hodges, "Can immersive virtual humans teach social conversational protocols?" in *Proc. IEEE Virtual Reality Conf.*, 2007, pp. 215–218.

[7] N. Yee, J. N. Bailenson, and N. Ducheneaut, "The proteus effect: Implications of transformed digital self-representation on online and offline behavior," *Commun. Res.*, vol. 36, pp. 285–312, 2009.

[8] C. M. De Melo, J. Gratch, and P. J. Carnevale, "The effect of agency on the impact of emotion expressions on people's decision making," in *Proc. Humaine Assoc. Conf. Affective Comput. Intell. Interaction*, 2013, pp. 546–551.

[9] H. Prendinger, J. Mori, and M. Ishizuka, "Using human physiology to evaluate subtle expressivity of a virtual quizmaster in a mathematical game," *Int. J. Hum. Comput. Stud.*, vol. 62, no. 2, pp. 231–245, Feb. 2005.

[10] X. Pan, M. Gillies, C. Barker, D. M. Clark, and M. Slater, "Socially anxious and confident men interact with a forward virtual woman: An experimental study," *PLoS One*, vol. 7, no. 4, p. e32931, Jan. 2012.

[11] C. L. Lisetti, "Embodied conversational agents for psychotherapy," in *Proc. Conf. Workshop Technol. Mental Health*, 2008, pp. 1–12.

[12] Tractica, "Home health technologies," 2015.

[13] S. N. Merry, K. Stasiak, M. Shepherd, C. Frampton, T. Fleming, and M. F. G. Lucassen, "The effectiveness of SPARX, a computerised self help intervention for adolescents seeking help for depression: Randomised controlled non-inferiority trial," *BMJ*, vol. 344, no. apr18 3, pp. e2598–e2598, 2012.

[14] C. V. Russoniello, K. O'Brien, and J. M. Parks, "The effectiveness of casual video games in improving mood and decreasing stress," *J. Cyber Ther. Rehabil.*, vol. 2, no. 1, pp. 53–66, 2009.

[15] S. G. Barsade, "The ripple effect: Emotional contagion and Its influence on group behavior," *Adm. Sci. Q.*, vol. 47, no. 4, pp. 644–675, Dec. 2002.

[16] E. Hatfield, J. T. Cacioppo, and R. L. Rapson, "Emotional contagion," *Curr. Dir. Psychol. Sci.*, vol. 2, pp. 96–99, 1993.

[17] B. Wild, M. Erb, and M. Bartels, "Are emotions contagious? Evoked emotions while viewing emotionally expressive faces: Quality, quantity, time course and gender differences," *Psychiatry Res.*, vol. 102, no. 2, pp. 109–124, Jun. 2001.

[18] U. Hess and S. Blairy, "Facial mimicry and emotional contagion to dynamic emotional facial expressions and their influence on decoding accuracy," *Int. J. Psychophysiol.*, vol. 40, no. 2, pp. 129–141, Mar. 2001.

[19] J. Xu, J. Broekens, K. Hindriks, and M. A. Neerincx, "Robot mood is contagious: Effects of robot body language in the imitation game," in *Proc. Int. Conf. Auton. Agents Multi-Agent Syst.*, 2014, pp. 973–980.

[20] A. D. I. Kramer, J. E. Guillory, and J. T. Hancock, "Experimental evidence of massive scale emotional contagion through social networks," *Proc. Nat. Acad. Sci. U.S.A.*, vol. 111, no. 24, pp. 8788–8790, Jul. 2014.

[21] R. Beale and C. Creed, "Affective interaction: How emotional agents affect users," *Int. J. Hum. Comput. Stud.*, vol. 67, pp. 755–776, 2009.

[22] N. Yee, J. N. Bailenson, and K. Rickertsen, "A meta-analysis of the impact of the inclusion and realism of human-like faces on user experiences in interfaces," in *Proc. SIGCHI Conf Human Factors Comput. Syst.*, 2007, p. 1.

[23] F. De Rosis, C. Pelachaud, I. Poggi, V. Carofiglio, and B. De Carolis, "From Greta's mind to her face: Modelling the dynamics of affective states in a conversational embodied agent," *Int. J. Hum. Comput. Stud.*, vol. 59, no. 1–2, pp. 81–118, 2003.

[24] C. Pelachaud, "Modelling multimodal expression of emotion in a virtual agent," *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, vol. 364, no. 1535, pp. 3539–3548, 2009.

[25] J. Gratch and S. Marsella, "Tears and fears: Modeling emotions and emotional behaviors in synthetic agents," in *Proc. Int. Conf. Auton. Agents*, 2001, pp. 278–285.

[26] J. Rickel, S. Marsella, J. Gratch, R. Hill, D. Traum, and W. Swartout, "Toward a new generation of virtual humans for interactive experiences," *IEEE Intell. Syst. Their Appl.*, vol. 17, no. 4, pp. 32–38, 2002.

[27] G. McKeown, M. Valstar, R. Cowie, M. Pantic, and M. Schröder, "The SEMAINE database: Annotated multimodal records of emotionally colored conversations between a person and a limited agent," *IEEE Trans. Affect. Comput.*, vol. 3, no. 1, pp. 5–17, Jan.–Mar. 2012.

[28] H. Maldonado, J.-E. R. Lee, S. Brave, C. Nass, H. Nakajima, R. Yamada, K. Iwamura, and Y. Morishima, "We learn better together: Enhancing eLearning with emotional characters," in *Proc. Int. Conf. Comput. Supported Collaborative Learning: The Next 10 Years* 2005, pp. 408–417.

[29] W. Burleson and R. W. Picard, "Gender-specific approaches to developing emotionally intelligent learning companions," *IEEE Intell. Syst.*, vol. 22, no. 4, pp. 62–69, Jul./Aug. 2007.

[30] S. Brave, C. Nass, and K. Hutchinson, "Computers that care: Investigating the effects of orientation of emotion exhibited by an embodied computer agent," *Int. J. Human Comput. Stud.*, vol. 62, pp. 161–178, 2005.

[31] H. Prendinger, S. Mayer, J. Mori, and M. Ishizuka, "Persona effect revisited—Using bio-signals to measure and reflect the impact of character-based interfaces," in *Proc. 4th Int. Working Conf. Intell. Virtual Agents*, 2003, pp. 283–291.

[32] J. Klein, Y. Moon, and R. W. Picard, "This computer responds to user frustration:," *Interacting with Comput.*, vol. 14, pp. 119–140, 2002.

[33] K. Hone, "Empathic agents to reduce user frustration: The effects of varying agent characteristics," *Interact. Comput.*, vol. 18, pp. 227–245, 2006.

[34] T. W. Bickmore and R. W. Picard, "Establishing and maintaining long-term human-computer relationships," *ACM Trans. Comput.-Human Interaction*, vol. 12, pp. 293–327, 2005.

[35] M. Slater, D.-P. Pertaub, C. Barker, and D. M. Clark, "An experimental study on fear of public speaking using a virtual environment," *Cyberpsychol. Behav.*, vol. 9, pp. 627–633, 2006.

[36] J. Tsai, E. Bowring, S. Marsella, W. Wood, and M. Tambe, "A study of emotional contagion with virtual characters," in *Proc. 12th Int. Conf. Intell. Virtual Agents*, 2012, pp. 81–88.

[37] R. W. Levenson, "Autonomic specificity and emotion," in *Handbook of Affective Sciences*, R. J. Davidson, K. R. Scherer, and H. H. Goldsmith, Eds. New York, NY, USA: Oxford Univ. Press, 2003.

[38] Y. Wu, S. V Babu, R. Armstrong, J. W. Bertrand, J. Luo, T. Roy, S. B. Daily, L. C. Dukes, L. F. Hodges, and T. Fasolino, "Effects of virtual human animation on emotion contagion in simulated interpersonal experiences," *IEEE Trans. Vis. Comput. Graph.*, vol. 20, no. 4, pp. 626–635, Apr. 2014.

[39] R. Niewiadomski, J. Hofmann, J. Urbain, T. Platt, J. Wagner, and P. Bilal, "Laugh-aware virtual agent and its impact on user amusement," in *Proc. 12th Int. Conf. Auton. Agents Multiagent Syst.* 2013, pp. 619–626.

[40] E. Peper, R. Harvey, and N. Akabayashi, "Biofeedback an evidence based approach in clinical practice," *Japanese J. Biofeedback Res.*, vol. 36, pp. 3–10, 2009.

[41] M. S. Schwarts and F. Andrasik, *Biofeedback: A Practitioner's Guide*, 3rd ed. New York, NY, USA: The Guilford Press, 2003.

[42] I. Bergstrom, S. Seinfeld, J. Arroyo-Palacios, M. Slater, and M. V Sanchez-Vives, "Using music as a signal for biofeedback," *Int. J. Psychophysiol.*, vol. 93, no. 1, pp. 140–149, 2014.

[43] H. Huang, T. Ingalls, L. Olson, K. Ganley, T. Rikakis, and J. He, "Interactive multimodal biofeedback for task-oriented neural rehabilitation," in *Proc. Conf. IEEE Eng. Med. Biol. Soc.*, 2005, vol. 3, pp. 2547–2550.

[44] D. Bersak, G. McDarby, N. Augenblick, P. McDarby, D. McDonnell, B. McDonald, and R. Karkun, "Intelligent biofeedback using an immersive competitive environment," in *Proc. UbiComp Workshop Des. Ubiquitous Comput. Games*, [Online]. Available: http://www.cse.chalmers.se/research/group/idc/ituniv/student/2002/reevolution/CoopGames/bersak.pdf, 2001.

[45] L. Chittaro and R. Sioni, "Affective computing vs. affective placebo: Study of a biofeedback- controlled game for relaxation training," *Int. J. Hum. Comput. Stud.*, vol. 72, pp. 663–673, 2014.

[46] A. Pietrangelo. (2014). The best bipolar disorder iphone and android apps of the year. *Healthline* [Online]. Available: http://www.healthline.com/health-slideshow/top-iphone-android-apps-bipolar-disorder#1

[47] B. Spanlang, J.-M. Normand, D. Borland, K. Kilteni, E. Giannopoulos, A. Pomes, M. Gonzalez-Franco, D. Pérez-Marcos, J. Arroyo-Palacios, X. Navarro Muncunill, and M. Slater, "How to build an embodiment lab: Achieving body representation illusions in virtual reality," *Front. Robot. AI*, vol. 1, no. 9, pp. 1–22, 2014.

[48] R. W. Doherty, "The emotional contagion scale: A measure of individual differences," *J. Nonverbal Behav.*, vol. 21, no. 2, pp. 131–154, 1997.

[49] D. M. McNair, M. Lorr, and L. F. Droppleman, *Manual for the Profile of Mood States*. San Diego, CA, USA: Educational & Ind. Testing Service, 1971.

[50] M. Bradley and P. J. Lang, "Measuring emotion: The self-assessment manikin and the semantic differential," *J. Behav. Ther. Exp. Psychiatry*, vol. 25, no. I, pp. 49–59, 1994.

[51] Y. Benjamini and Y. Hochberg, "Controlling the false discovery rate: A practical and powerful approach to multiple testing," *J. Roy. Stat. Soc.*, vol. 57, no. 1 , pp. 289–300, 1995.

Jorge Arroyo-Palacios graduated from the Instituto Tecnológico de Ciudad Victoria in 2003 with an Honours BEng degree in computer systems. He received the MSc degree in advanced computer science in 2006 and the PhD degree in computer science in 2011 from the University of Sheffield. He received a research scholarship from the Mexican Research Council of Science and Technology (CONACYT) to continue his postgraduate studies. Currently, he is a postdoctoral researcher at the EVENTLAB, University of Barcelona. His research aims to explore the use of physiological signals and machine learning methods to design and implement innovative intelligent computer systems able to: i) adapt to the users' health, physical and affective state, ii) interact with the users in a natural and intuitive way, and iii) support healthcare practitioners and patients.

Mel Slater received the BSc, MSc, DSc degrees. He is an ICREA research professor at the University of Barcelona, and leads the Event Lab (www.event-lab.org). He became a professor of virtual environments at the University College London in 1997. He was a United Kingdom Engineering and Physical Sciences Senior Research Fellow from 1999 to 2004. Thirty-two of his PhD students have obtained their PhDs since 1989. In 2005, he received the Virtual Reality Career Award by IEEE Virtual Reality "In Recognition of Seminal Achievements in Engineering Virtual Reality." He is a coordinator in the EU 7th Framework Integrated Project VERE (www.vereproject.org), and was a scientific leader in the Integrated Project BEAMING (www.beaming-eu.org). He holds a European Research Council grant TRAVERSE (www.traverserc.org). He has pioneered the scientific study of virtual reality and contributed to its application in psychology and cognitive neuroscience concentrating on the issues of body ownership and agency.

▷ **For more information on this or any other computing topic, please visit our Digital Library at** www.computer.org/publications/dlib.