A personality-based emotional model for embodied conversational agents: Effects on perceived social presence and game experience of users

Pejman Sajjadi\textsuperscript{a,b,⁎}, Laura Hoffmann\textsuperscript{a}, Philipp Cimiano\textsuperscript{a}, Stefan Kopp\textsuperscript{a}

\textsuperscript{a} CITEC, Bielefeld University, 33619 Bielefeld, Germany
\textsuperscript{b} Pennsylvania State University, 16803, State College, USA

Keywords:
Embodied conversational agent
Virtual reality
Game experience
Social presence
Personality
Behavioral model

ARTICLE INFO

This paper reports on an experiment that investigates the effect of interacting with a personality-driven embodied conversational agent (ECA) on the perceived social presence and game experience of people in a VR social simulator. Furthermore, the dynamics between the different metrics of game experience and social presence of people are explored to determine which game experience metrics are the strongest predictors of perceived social presence in this context. A personality-based emotional model is used for personifying the employed ECA, which governs the manifestation of its non-verbal behaviors. Three experimental conditions manipulating the existence and intensity of non-verbal behaviors exhibited by the ECA were used to investigate the effect of this proposed approach. The results of the experiment with 41 participants indicate that people who were exposed to an extrovert ECA experienced significantly higher levels of behavioral involvement as part of their social presence compared to the other conditions. These results suggest that incorporating personality by means of non-verbal behavior in the emotional model of an ECA influences users perceived feeling of social presence. Furthermore, our results reveal that there is a bidirectional relationship between game experience metrics and perceived social presence of people as each predict the other.

1. Introduction

Over the past years we have witnessed an increasing interest in the use of interactive and immersive applications such as games \cite{1} and simulations for serious purposes (e.g. social skill training \cite{2}, physical rehabilitation \cite{3}, logical problem solving \cite{4}). In contexts where social interactions are considered as an important essence of learning, social entities generally referred to as embodied conversational agents (ECA) \cite{5} capable of imitating the behavior of real people are crafted and utilized by these systems. ECAs facilitate social interactions to take place with users through verbal, non- and para-verbal behavioral cues (e.g. \cite{6}). As learning in these systems occur through socialization with an ECA, perceiving the social presence (a feeling of sensing another entity being present) of the interlocutor (i.e. the ECA) is of key importance. Research has shown that there is a correlation between the perceived social presence and learning where higher levels of perceived social presence in online education lead to higher levels of perceived learning and satisfaction with the instructor \cite{7}. Social presence can be studied in light of constructs such as immediacy (psychological distance) and intimacy (interpersonal closeness) \cite{8}. These experiences are highly dependent on the degree and quality of verbal, non-verbal, and paraverbal behaviors demonstrated by the ECA, as well as the game/simulation experiences of users (e.g. immersion, flow, positive and negative affects). These ECA behaviors provide unobtrusive, and yet extremely expressive feedback to the users about the effect of their actions \cite{9}. Such feedback could evoke a sense of empathy towards the ECA, elicit a sense of behavioral dependency, and drive the learners to adjust their actions. In educational contexts, with an objective of behavior adjustment or improvement, this is highly desired and at the core of the learning strategy.

Extensive argumentation can be found in the literature highlighting the importance of good game experience, particularly in light of metrics such as immersion and flow for effective learning (e.g. \cite{10–12}). Furthermore, it has been suggested that the immersiveness of an experience actually provides the boundaries within which presence (the feeling of being there) can occur \cite{13}. As such, in serious games and social simulations where learners have to actively engage in social interaction with ECAs, it is highly desirable to evoke both: good game experience and high levels of perceived social presence to facilitate effective learning. In light of this, we have opted to investigate whether...
the perceived social presence and game experience of people can be positively affected in a social simulation environment perceived through virtual reality (VR), by employing an ECA that expresses its emotional states through non-verbal behavioral cues governed by its personality. Similar research works have investigated the effect of personified ECAs with different characteristics on the perceived social presence of people. However, investigating the effect of manipulating the behavioral model of an ECA based on personality, and its consecutive influence on social presence and game experience has not been sufficiently addressed. Furthermore, although the literature suggests that immersion is an influential metric that informs the elicitation of presence, it is not clear whether the same principle applies for social presence. As such, it is important to understand what game experience metrics are the strongest contributing factors to the feeling of perceived social presence. Such insight can inform and improve the design of future virtual social experiences.

On these grounds, this paper introduces a multi-modal prototype that utilizes a personality-driven ECA equipped with verbal and non-verbal behaviors. Moreover, it presents the results of a pilot study investigating its effects on the perceived social presence and game experience of users, as well as the dynamics between them. In a pilot study with 41 participants, the existence and intensity of the non-verbal behaviors of the used ECA was manipulated in three conditions to assess its effect on the reported game experience and perceived social presence of users: (1) no apparent non-verbal behaviors exhibited by the ECA (2) non-verbal behaviors governed by an extrovert-based emotional model, in which the exhibitions by the ECA are more assertive and pronounced (3) non-verbal behaviors governed by an introvert-based emotional model, in which the exhibitions are more submissive and minimal. Our results suggest that the projection of non-verbal behaviors as a result of incorporating personality as part of the emotional model of an ECA, influence the elicitation of the feeling of social presence from the users. More precisely, the results suggest that more assertive and pronounced non-verbal behaviors produced by an ECA are more noticeable by users and seem to have higher impacts on their feeling of social presence, compared to more submissive and minimal non-verbal behaviors. Furthermore, our results show that among the measured game experience metrics, the experienced levels of flow, immersion, and tension are the strongest predictors for the perceived feeling of social presence with flow as the strongest predictor. These results suggest that the aesthetics and richness of the experience, in addition to minimal experienced levels of tension (although significantly influential), are surpassed by a feeling of losing track of time and being occupied and engaged with activities involving interaction with a human-like entity in relation to experiencing social presence.

The rest of this paper is organized as follows: Section 2 discusses related works on the effect of computer-mediated artifacts (mostly with a focus on ECAs) on perceived social presence, as well as the role of game experience in effective learning. Section 3 discusses the underlying model of the used ECA, and Section 4 elaborates on the used VR simulation environment. Section 5 discusses the experiment with the prototype, and Section 6 concludes the paper while addressing limitations and future work.

2. Background & related work

ECAs have been a topic of interest in the human-agent interaction research community for decades. Extensive works can be found in the literature on the use of ECAs in social interaction scenarios for serious purposes. For instance, in the work of Hoque et al. social agents are used as interlocutors to train individuals for job interviews [14], and in the work of Linssen et al. they are used as crime suspects for training police officers on how to perform effective interrogation [2]. Depending on the objective and context of use, ECAs are personified on various human-like characteristics, such as personality, mood, emotion, moodiness (e.g. temperamental vs. lethargic), and motivation (e.g. [15-18]), and manifested as an interlocutor on different levels of realism. Numerous research projects have focused on designing behavioral models that exert personified ECAs (e.g. [17,15]). The term personification is generally used to refer to the process of giving human-like qualities to an agent, and can be studied from the facets: physical, expres-sional, logical, and emotional [17].

One of the frequently researched aspects of human-agent/robot interaction is the level of perceived social presence by humans when interacting with these entities. This special form of presence (“the feeling of being located in a perceptible external world around the self”[19], page 39), the so called “social presence” has been defined differently by various researchers (e.g. [20,21]). One of the most frequently cited definitions for this is by Biocca and Harms as “having some level of access or insight into someone else’s affective, cognitive, and inten-tional states” [22]. In this paper, we subscribe to this definition for the notion of social presence. Furthermore, extensive argumentation can be found in the literature highlighting the importance of virtual reality in eliciting social presence (e.g. [23]). Being exposed to an inherently immersive experience in which one experiences the non- and para-verbal behaviors of a virtual counterpart from a natural viewpoint e.g., though VR glasses, is one reason to sustain this claim.

Different factors (e.g. human controlled avatars vs. autonomous agents, degree of personification, and realism of non-verbal and para-verbal signals) may contribute to the perceived social presence of a person when interacting with an ECA. For instance, in the threshold model of social influence proposed by Blasovich et al. [23], it is argued that if users are aware of the fact that they are interacting with a human controlled avatar, the behavioral realism of the avatar is of less importance to cause social influence on the user. This claim is based on the assumption that this awareness already leads to a high level of social presence. Conversely, an autonomous agent is low on the agency axis, and thus more behavioral realism in the appearance is needed to evoke the same social influence than an avatar with low behavioral realism. Notwithstanding, von der Pütten et al. [24] showed that such awareness of interacting with an autonomous agent versus a human-controlled avatar does not affect the perceived social presence of users. When participants were led to believe that they were interacting with either an autonomous agent or a human controlled avatar in a social situation (although both conditions were using the same agent and no avatar), no significant differences between the self-reported measures of perceived social presence between the two conditions were observed.

With respect to the effect of personification level exhibited by an agent, Nowak and Biocca [25] have demonstrated that higher levels of anthropomorphism does not by default lead to a higher sense of perceived social presence. It was reported that participants who were exposed to a less anthropomorphic artifact had experienced a higher level of social presence than those who were exposed to either no artifact, or a highly anthropomorphic one. This finding was interpreted as high levels of anthropomorphism increasing the expectation of the people, and reducing their social presence when the expectations are not met. Von der Pütten and colleagues, [24] however, offer an alternative perspective. An experiment on the effect of behavioral realism fidelity exhibited by an ECA on the social presence of the participants showed to be significant. When participants were exposed to two versions of an ECA, one with rich non-verbal behavior signals and one with less richer signals, the self-reported social presence was significantly higher in the rich condition. On similar grounds, Chuah et al. [26] state that increasing the physicality level of the ECA increases the perception of social presence by users. Furthermore, Daher et al. [27] reported that observing two intelligent virtual agents (IVAs) interact with each other (priming) has a significant positive effect on the perceived social presence of a user in the consecutive interaction with one of the IVAs. Therefore, it can be used as a mechanism for enhancing the social interaction experience with IVAs (ECAs).

Furthermore, Lee et al. [28] have shown that the para-verbal characteristics of a voice-mediated interaction between users and
In the context of human-robot interaction, Lee and colleagues [29] report that when the personality of the user is not similar to the one of the robot, they perceived the robot to be more intelligent and socially present. The work of Gajadhar et al. [30], although focusing on interactions among humans, shows that rich social interactions occur when the involved parties are co-located in the same environment where verbal and non-verbal behaviors can be continuously monitored. This observation relates to social processes such as immediacy, mirroring, and emotional contagion that characterize the feeling of social presence, which according to De Kort [31] affect game experience. Gajadhar et al. have shown that being co-located with a human opponent has a significant effect on the game experience in comparison to playing against a computer (with no embodied agents involved), or a mediated co-player. They conclude that being co-located with a human co-player (immediacy) significantly improves positive affect, decreases tension, and increases a feeling of competence. These findings are in line with results reported by Mandryk and Inkpen, and Ravaja et al. [32,33] indicating that immediacy plays a determining role in experiencing enjoyment. Furthermore, De Kort et al. [31] have stated that "higher levels of social presence may be attained between remote players who are continuously and mutually engaged in a collaborative game, than between co-located players who are each concentrated on attaining their individual goals without the need to interact or share" ([31, ], page 7).

Moreover, extensive argumentation can be found in the literature highlighting the importance of good game experience for effective learning. The degree to which factors such as involvement (flow) and immersion are experienced in an environment designed for educational purposes can be predictors for its success [10]. Additionally, According to Witmer and Singer [34] both immersion and involvement are prerequisites for experiencing presence in a virtual environment. They have defined immersion as "a psychological state characterized by perceiving oneself to be enveloped by, included in, and interacting with an environment that provides a continuous stream of stimuli and experiences" ([34], page 227); and involvement as "a psychological state experienced as a consequence of focusing one’s energy and attention on a coherent set of stimuli or meaningfully related activities and events" ([34], page 227). Involvement is a prerequisite for experiencing the flow state introduced by Csikszentmihalyi [35] as a state of absolute absorption to a task to a point of losing self-consciousness where the activity itself becomes rewarding in its own, and this enables an individual to function at his/her fullest capacity. Moreno and Mayer [36] argue that a massive amount of information imposed by an activity (e.g. game) can overload the working memory capacity of the player, leading to weak or incorrect learning. Conversely, if the experience is under-loading, it can lead to boredom and disengagement, and consequently deterioration of performance. Similar arguments are made for the necessity to maintain a balance between challenge and competence based on the flow theory of Csikszentmihalyi. On similar grounds Dedé [12] has stated that apart from its rich entertainment value, a high level of immersion can enhance learning.

In summary, the state of the art highlights the importance of studying the effect of an artifact (particularly ECAs) on the elicited feeling of social presence from users. Furthermore, it demonstrates how manipulating different features of an ECA can influence this feeling. Moreover, there are indications to the existence of a relationship between game experience, particularly immersion (feeling of co-location and immediacy), involvement (mutual engagement in activities) and presence. Also it is known that these experiences are important for the realization of effective learning. What remains open is the question to what extent social presence and game experience can be enhanced by utilizing an ECA that expresses its emotional states through non-verbal behavioral cues, governed by its personality. Moreover, it is not clearly known which game experience metrics are the strongest predictors for the extent to which one experiences a special form of presence most relevant in social contexts: social presence. In light of this, the research presented in the paper aims at investigating the effect that incorporating personality as part of the emotional model of an ECA would have on the perceived social presence and game experience of users. Furthermore, we will investigate which metrics of game experience are the strongest predictors for the feeling of perceived social presence of users in a social interaction simulation that employs a personality-driven ECA equipped with non-verbal behaviors.

As such, we hypothesize: H1: A highly emotionally-personified agent with apparent non-verbal behavior, governed by a personality-driven behavioral model would elicit (a) a higher level of social presence, and (b) a better game experience, compared to a less emotionally-personified agent with no apparent non-verbal behavior and H2: Immersion and involvement are prerequisites for social presence and thus predict the perceived sense of social presence. To test these hypotheses we built a prototype consisting of an ECA with a personality-driven model that governs the expression of her emotional state using a variety of non-verbal cues (see Sections 3 and 4 for details), and evaluated its effect in a pilot study (see Section 5).

3. The embodied conversational agent: Linda

An embodied conversational agent named Linda was designed in the context of this research. Linda is personified on the three facets: physical, expressional, and emotional. Her corporeal representation which embodies her facial and bodily features in form of a realistic female (can be also changed to a male character) in mid thirties provides her physical personification. Furthermore, a series of facial expressions, gestures, and postures in form of synthesized animations constitute her expressional personification. The emotional personification facet however is more complex. This facet is considered as the “mind” of the ECA that guides its emotional responses [17], and is therefore a composition of multiple factors. Inspired by the emotional model proposed by Becker et al. [16,15], Linda was emotionally personified based on the following human-like characteristics: personality, emotion and mood valences, moodiness, and boredom. Personality is defined as “characteristics of a virtual human that distinguishes him from the others”, while emotion is defined as “a state of mind that is only momentary”, and mood as “a prolonged state of mind, resulting from a cumulative effect of emotions” [17].

Our model calculates a concrete emotional state to be exhibited by Linda, by taking into account the interrelations between the aforementioned characteristics during each interaction instance with the user. Briefly, the presence of a stimulus (i.e. what the user says to Linda) will have either a fortifying or an alleviating affect (depending on the negative or positive nature of the stimulus) on the interrelations between emotion and mood valences, moodiness, and boredom characteristics of Linda. Moreover, if no stimuli is produced by the user, the boredom characteristic of Linda will gradually increase after a specific period of time, which in turn affects the interrelations between the aforementioned characteristics. These newly derived characteristics are then translated to a point of reference (e.g. point P in Fig. 1) in a three independent bipolar dimensional space represented by pleasure (displeasure), arousal (sleepiness), and dominance (submissiveness) (PAD) introduced by Russel and Mehrabian [37] (for a more elaborate explanation regarding the underlying model please refer to [15,16]).

According to the creators of PAD, the three dimensions of pleasure,
arousal, and dominance are sufficient to adequately define and represent any emotional state. In light of this, Russel and Mehrabian have suggested 181 concrete emotional states represented as triples in the PAD space. Similar to Becker, other research works (e.g. [18]) have either directly incorporated a subset of these emotional states as part of their model, or have slightly adapted them to their context before doing so. Based on the interaction scenario of the social simulation used in this research, we have incorporated 18 emotional states (Es) into our behavioral model, inspired by the triples listed by Becker et al., and Boulic et al. [15, 18]. These states with their corresponding PAD values can be seen in Fig. 1.

In order to guide Linda on which emotional state to exhibit, an emotion awareness likelihood function [15,16] is used to calculate the proximity of the translated reference point P to either of the emotional states Es defined in the PAD space. This proximity is calculated on two levels. Each Es is assigned two zones called activation (Δ) and saturation (ϕ), and a base intensity level (i) as its attributes. The activation zone has a higher radius than the saturation zone. The emotion awareness likelihood for each Es would increase, the closer the point of reference gets to it. An example for these zones can be seen for the Es “Sad” in Fig. 1.

At every interaction instance (i.e. presence of a stimuli or lack thereof), and for every Es point, if the distance between the point P and that Es is less than the activation zone of that Es (i.e. ΔEs), the calculation of its awareness likelihood ωEs (Eq. (1)) will commence. However, it is possible for the reference point P to fall into the activation zone of more than one Es. In such situations, the Es with the highest likelihood will be selected as the most appropriate candidate. Moreover, if the distance between the reference point and each of the Es points is less than the saturation zone of that Es (i.e. ΔEs), the awareness likelihood for those Es(s) will be calculated. If the reference point falls into the saturation zone of more than one Es, the Es with the highest likelihood will be selected as the most appropriate candidate. Using this mechanism, Linda’s emotional model decides which emotional state should be exhibited by her. It is important to note that although the point of reference P will move within the PAD through the course of interaction with Linda, there is a possibility that P is not within the saturation or activation zone of any Es. In this situation, Linda will maintain her previous emotional state.

\[
\omega_{Es} = 1 - \frac{\text{Distance} - \Delta_{Es}}{\Phi_{Es} - \Delta_{Es}}
\]  

(1) \[\Phi_{Es} > \Delta_{Es} \forall Es \in \{E_{S1}, E_{S2}\}\]

Furthermore, Linda is assigned a personality profile limited to the extrovert-introvert dimension of the Five Factor Model (FFM) [38]. Depending on whether Linda is extrovert or introvert, a personality profile is synthesized and expressed in the PAD space (PPAD in Fig. 1) based on the characteristics of this personality dimension. As such the PPAD point can be looked at as a point where the ECA is most comfortable since it corresponds to its personality. The PPAD point serves two purposes: First, it manipulates the activation and saturation zone radius, as well as the base intensity of certain Es(s). This is performed based on the relationship between personality and the likelihood of exhibiting (or not) certain emotional states. This mechanism will increase (or decrease) the probability of certain Es(s) to be exhibited by the ECA.

Second, the PPAD acts as a center of gravity when moving P into the PAD space after each interaction. If through the course of interaction P gets farther from (i.e. away from ECA’s comfort zone) PPAD (e.g. P’ in Fig. 1), its movement is manipulated in such a way that it travels less distance. Conversely, if the P’ is getting closer to (i.e. towards the ECA’s comfort zone) the PPAD, it will travel more distance. This mechanism governs the emotional state exhibition by the ECA based on its personality.

4. The simulation

To produce an immersive experience for people when interacting with the ECA, a social simulation in virtual reality was created. This experience takes place in a virtual office environment built using Unity3D, where the characters are sitting across a table from each other. The user assumes the role of a company manager, and is required to review the performance of one of his/her employees (i.e. our ECA Linda). The simulation is perceived through a virtual reality headset (using Oculus Rift1), and the user has a 360 degree view of the room.

This mechanism promotes a feeling of co-location with the ECA, and therefore immediacy. As such, the user experiences the non-verbal behaviors of the ECA from a point of view natural to them. This design choice was made based on the assumption that it will positively affect the game experience of users, and consequently their level of perceived social presence as proposed by the research presented in [30]. Additionally, a Leap Motion\textsuperscript{2} device is attached on top of the VR headset that facilitates the virtual representation of user’s hands if s/he brings them within its field of vision. This mechanism provides partial virtual body ownership and was added under the assumption that it will motivate the users to employ body language (in terms of hand gestures) when interacting with the ECA. The primary mode of interaction with the ECA however is by voice (using Microsoft speech API) in a turn-taking system. At each turn, multiple conversation options are available for both the user and ECA. The user would need to say out loud the content of the option they want to utter to the ECA. As a consequence, that option will be highlighted and then selected (Fig. 2). Based on the option chosen by the user, the ECA will select the appropriate response and communicate it to the user by means of speech synthesis, augmented with lip-syncing, eye gazing, facial expressions, postures, and gestures. The state of the art on open-ended, but meaningful communication between a human and an intelligent system is not advanced enough to allow for the implementation of such mechanism. Therefore, a pre-scripted conversation graph was used in this social simulation, in which the conversation options for both characters are synthesized beforehand.

The conversation that takes place between the user and the ECA is focused on reviewing the performance of the ECA (i.e. the virtual employee) by the user, and the responses of the ECA to these revisions. It is worth mentioning that this topic was chosen since it directly relates to the educational aspect of this social simulation. However, the evaluation of our simulator’s success in teaching users how to effectively evaluate the performance of an employee is out of the scope of this article and is not further discussed. Throughout the course of interaction, each conversational choice selected by the user will introduce an effect (i.e. the stimuli) on the human-like characteristics of the ECA (Section 3). These effects would be of either positive or negative nature and feature different levels of intensity. However, for a preliminary evaluation of our proposed model the conversation options available to the user were restricted to only the ones that would introduce a negative effect (with the same level of intensity) on the emotional state of the ECA. This choice was made to ensure that the experience of users in relation to interaction with the ECA remain consistent regardless of which option they choose.

As such, the dynamic between Linda’s characteristics will constantly grow towards the negative end of the spectrum, and therefore their translation into a point of reference in the PAD space would eventually result in a negative value for the Pleasure dimension. In light of this, the scenario of the simulation can be looked at as the user constantly giving harsh comments to Linda, thus violating the generally accepted rules for effectively reviewing an employee’s performance (e.g. be specific and complete, evaluate performance but not personality, stick to facts not assumptions).

5. Experiment

To be able to adequately evaluate the effect of our proposed personality-driven ECA on perceived social presence and game experience of users, three experimental conditions were devised: neutral, extrovert and introvert. Each condition uses a different variation of the personality-driven behavioral model for governing the exhibition of emotional states by Linda, as well as their manifestation into non-verbal behaviors.

\footnote{\url{https://www.leapmotion.com}.}
behavioral cues in the extrovert condition are more pronounced and assertive, and in the introvert condition they are more subtle and submissive. For instance, Fig. 3 demonstrates the expression of two emotional states Angry (−70, 80, 100) and Afraid (−70, 80, −100) that are identical to each other except for their dominance value. Furthermore, in the extrovert condition, the ECA can exhibit the Angry emotion manifested by mild frowned eyebrows, direct eye contact, shoulders up, and leaning to the side, and in the introvert condition the ECA can exhibit the Afraid emotion manifested by slightly raised eyebrows, avoiding eye contact, dropped shoulders, and hands on legs.

5.1. Methodology

The perceived social presence and the game experience of users, as well as the dynamics between these constructs were evaluated in a laboratory experiment with N = 41 participants. We used a between-subjects design in which participants were exposed to either the neutral, introvert or extrovert condition. In addition, the effect of participants’ sex (gender) on the reported experiences was investigated.

5.1.1. Measures

Participant’s perception and experiences were collected via questionnaires. A general questionnaire was administered before interaction with the prototype that asked for demographic information (age, gender, frequency of gaming, past experiences with creating ECAs, and with virtual reality) from participants, in addition to measuring their empathy level. The Toronto empathy questionnaire [40] was used to measure the empathy level of participants by summing their scores on 16 statements (with lowest overall possible score 16, and highest 80).

After interaction with the prototype we administered a second questionnaire that asked for participants’ perceived social presence, game experience, and empathy towards the agent. Several instruments were used, and their scores combined to measure participants’ levels of perceived social presence: five items by Bailenson et al. [41], the social presence module of GEQ [42] (in terms of empathy, and behavioral involvement), and the attentional allocation scale of Biocca et al. [43]. Using a combination of these instruments, we were confident that we could adequately measure social presence. Furthermore, a combined value of social presence was calculated by averaging the scores on the aforementioned four social presence metrics. This combined value presents a single score for users’ levels of perceived social presence.

The game experiences of participants were assessed using the core module of GEQ [42] (based on immersion, flow, tension/annoyance, and negative and positive affects). Furthermore, participants’ perception of and attitude towards the ECA were measured. Given the negative nature of the interaction scenario, it was of interest to assess how well the participants could estimate the ECA’s affective states. Therefore, we measured their perception of Linda’s negative feelings using the PANAS instrument of Watson et al. [44]. Additionally, we measured their level of empathy towards Linda using the instrument of von der Putten et al. [45]. Furthermore, using the likeability scale of Bailenson et al. [41], it was measured to what extent the participants liked Linda. Lastly, using six items, the perception of Linda’s features were assessed in terms of the degree to which the users noticed changes in her facial expressions, gestures and postures, lip-syncing, blinking, gazing, and following the head movements of the user. All items were rated on a Likert-type scale from 1 (“strongly disagree”) to 5 (“strongly agree”). For data analyses, the items were collapsed into average scores per subscale.

5.1.2. Procedure

Before interacting with the prototype, the participants were given a brief presentation which introduced them to the scenario and mechanics of the interaction. After the presentation, the first questionnaire was administered (i.e. five demographic questions and the Toronto empathy questionnaire). Afterwards, the participants were given ample opportunity to tryout the experience by exchanging some generic greetings statements in a sample interaction. The tryout was administered to familiarize the participants with the mechanics of the prototype, while providing them with the opportunity to ask any questions they might have. Afterwards, the participants were randomly assigned to either neutral, introvert, or extrovert conditions, and interacted with the prototype. The duration of interaction with the prototype was on average 15 min. During this phase, the participants were left alone in the room so they would feel comfortable when interacting with Linda verbally and possibly non-verbally (e.g. by using hand gestures). Once the interaction was over, the participants had to fill in the second questionnaire for the evaluation of their perceptions and experiences (see Section 5.1.1). Furthermore, an informal interview was conducted by the experimenter to gather further insights on individual impressions. Finally, participants were fully debriefed at the end of the session.

5.1.3. Participants

41 participants (20 male, 19 females, 1 who preferred not to answer, and 1 indicating “other”) took part in the preliminary evaluation of this prototype. Out of them, 13 were assigned to the neutral condition and 14 to each the extrovert and introvert conditions, respectively. The mean age of the participants was 25.73 (SD = 4.46), ranging from 20 to 40 years old. The participants were recruited using a university experiment participation platform and flyers, and each was compensated €5 for 20–30 min (includes all the steps in the procedure) of their participation time. To explore the effect of gender on how the prototype is perceived, we compared male and female participants using an independent t-test, which revealed that females were in general more empathetic (t = −2.36, p < .05), experienced less tension/annoyance (t = 2.32, p < .05), less negative affect (t = 3.5, p < .001), and were more empathetic towards Linda (t = −2.310, p < .05). However, gender did not affect our main findings reported in this paper.

5.2. Analysis

To investigate the formulated hypotheses, at first a series of ANOVAs were calculated to unveil the differences between the conditions. We controlled for the influence of participants’ empathy state as a covariate in the analysis, but since it did not affect the results, it is not further discussed. In the case of ANOVA indicating a significant difference, the LSD post hoc test was used to unveil which conditions are significantly different from each other.

To explore the relationship between social presence and game experiences, a series of regression analyses were performed. To investigate whether game experience can predict social presence, a stepwise multiple regression was calculated with each of the measured metrics of social presence as the dependent variable, and each of the measured game experience factors as an independent variable. Conversely, to investigate whether social presence can predict game
experience, a number of linear regressions were calculated. In each calculated linear regression, one of the measured game experience metrics was the dependent variable and the social presence metric the independent variable. All analyses were performed using IBM SPSS Statistics 22.

5.3. Results & discussion

The overall scores (all participants across all conditions) for the measured metrics are represented in the overall column of Table 1. As is evident from the these scores, interaction with an ECA inside virtual reality, and using voice as a natural communicative mean, prompts a rather rich experience. This is corroborated by the above average scores for the mean overall social presence (3.36) of Bailenson et al., the mean overall empathy (3.34) and behavioral involvement (3.67) as part of the social presence module of GEQ, the mean overall attention allocation (3.76) of Biocca et al., and the mean overall combined score for social presence (3.53). These scores indicate a positive evaluation of the prototype in successful elicitation of a feeling of perceived social presence as a result of interaction with a personality-driven ECA. Furthermore, the results indicate that participants had a positive experience interacting with the prototype in light of the game experience metrics. This is corroborated by their above average mean scores on immersion (3.62), flow (3.72), and positive affect (3.66), and low mean scores on tension/annoyance (1.94) and negative affect (1.98) of GEQ. In summary, the mean scores indicate a promising positive evaluation of the prototype (note that the lowest possible score is 1 and the highest 5).

Furthermore, we observed that mean score for the perception of ECA's negative feeling is slightly below average (2.73). Given the negative nature of the interaction scenario, this was expected and can be interpreted as the ECA being successful in expressing its negative emotional states in a comprehensible fashion. Moreover, we observe that the expression of these negative feelings has resulted in a high mean score for empathy towards the ECA (3.16). Interestingly, however, feeling empathetic towards the ECA did not result in participants liking the ECA, as the mean score for this factor is only slightly below average (2.89). We attribute this result to the negative nature of the experience, where the ECA had to respond negatively to the actions of the user. However, empathy towards the ECA and its likableness were shown to be positively and significantly (p < .05) correlated to one another.

A post hoc test (LSD) indicated that the behavioral involvement score in the extrovert condition was significantly higher than the neutral (p < .05, Mean difference = .366), and the introvert (p < .05, Mean difference = .339) conditions, but the introvert condition was not significantly different from the neutral (see Table 1 for means). The behavioral involvement factor as part of the social presence module of GEQ was measured using items that asked participants to what extent they felt their actions were dependent on the actions of the ECA and vice versa, to what extent they paid close attention to the ECA and vice versa, and to what extent the ECA affected what they did and vice versa. The results indicate that an emotionally-personified extrovert ECA with more pronounced and assertive non-verbal behaviors elicits a higher feeling of behavioral involvement in the interaction with the ECA, compared to both a less emotionally-personified ECA with no apparent non-verbal behavior, and an emotionally-personified introvert ECA with more subtle and submissive non-verbal behaviors. As such, our results suggest that the manifestation of a personality-driven emotional model in which the non-verbal behavior of the ECA is more elaborate and profound (i.e. more assertive and pronounced gestures, and higher rate of making direct eye contact with the user) evokes a higher feeling of behavioral involvement.

The results therefore only partially support our hypothesis H1a for the extrovert version evoking a higher sense of perceived social presence than the neutral condition, whereas no anticipated difference for the introvert condition was observable. To investigate why the introvert condition did not differ from the neutral, we re-considered the control questions that addressed in how far participants recognized changes in the ECA's behavior. ANOVA revealed a significant effect (F(2,41) = 14.398, p < .005, part.eta$^2$ = .431) for changes in the ECA's posture and gestures. Compared to the neutral condition, in both extrovert (P = .000, Mean difference = 1.96) and introvert (P = .016, Mean difference = .964) conditions, participants noticed that Linda changed her posture and gesture a few times. This was expected since in the neutral condition, there was no apparent non-verbal behavior from Linda. However, there was also a significant difference between the extrovert and introvert conditions for this feature (p < .01, Mean difference = 1) indicating that people noticed changes in the non-verbal behavior of Linda in form of gestures and postures significantly more in the extrovert version. The results highlight the importance of how a personality-driven emotional model should be manifested to ensue the highest effect, since a significant difference between the extrovert and introvert conditions was surprisingly observable. Furthermore, similar to what was reported by von der Pütten et al. [24], our results do not support the claim of Nowak and Biocca [25] that high levels of anthropomorphism raises expectations in users (i.e. expecting the ECA behavior to be identical to a real human) and this will negatively affect

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Comparison of means and standard deviations of the overall population &amp; three condition groups.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factors</td>
<td>Overall</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
</tr>
<tr>
<td>Social Presence (Bailenson et al.)</td>
<td>3.36</td>
</tr>
<tr>
<td>Social presence - empathy (GEQ)</td>
<td>3.34</td>
</tr>
<tr>
<td>Social presence - behavioral involvement (GEQ) *</td>
<td>3.67</td>
</tr>
<tr>
<td>Social presence - attention allocation (Biocca et al.)</td>
<td>3.76</td>
</tr>
<tr>
<td>Social presence (combined)</td>
<td>3.53</td>
</tr>
<tr>
<td>Game experience - immersion (GEQ)</td>
<td>3.62</td>
</tr>
<tr>
<td>Game experience - flow (GEQ)</td>
<td>3.72</td>
</tr>
<tr>
<td>Game experience - tension/annoyance (GEQ)</td>
<td>1.94</td>
</tr>
<tr>
<td>Game experience - negative affect (GEQ)</td>
<td>1.98</td>
</tr>
<tr>
<td>Game experience - positive affect (GEQ)</td>
<td>3.66</td>
</tr>
<tr>
<td>Perceptions of agent's negative feeling (PANAS)</td>
<td>2.73</td>
</tr>
<tr>
<td>Empathy towards the agent</td>
<td>3.16</td>
</tr>
<tr>
<td>Likableness of the agent</td>
<td>2.89</td>
</tr>
</tbody>
</table>
their perceived social presence when the expectations are not met. In fact, we have witnessed that being exposed to a highly personified version of Linda equipped with pronounced non-verbal behaviors, result in a higher sense of perceived social presence compared to a version with minimal non-verbal behaviors or one with non.

With regard to H1b, ANOVA with the game experience measures as dependent variables did not reveal any significant differences between the conditions and, hence H1b is rejected. Considering that the overall game experience across all conditions was quite positive, we can therefore conclude that the combination of VR, voice input, speech synthesis, lip-syncing, and eye gazing evokes an adequate level of game experience in users, regardless of the inclusion of personality-driven non-verbal behavior.

In light of the reported results, we opted to further investigate H2 by exploring the dynamics between the different game experience metrics and the perceived social presence of all participants, using the combined measure of perceived social presence. Since our prototype elicits a rather consistent game experience and perceived social presence across all conditions (with the exception of behavioral involvement as part of social presence and only for the extrovert condition), we decided to investigate the aforementioned dynamics over all participants and using a single value that represents social presence.

The result of the step-wise multiple regression analysis indicate that indeed (some of) the game experience metrics are significant predictors for the combined perceived social presence measure of users (F (3,37) = 12.113, \( p < .001 \)), with an adjusted \( R^2 \) of .455. The calculated model indicated immersion, flow, and tension/annoyance as the predictors explaining 45.5% of variance in social presence, and excluded negative and positive affect metrics. The model indicates that the participants’ predicted perceived social presence is equal to 2.284 + .179 (immersion) + .240 (flow) − .148 (tension/annoyance), where immersion is measured as how rich and impressive the experience was; how pleasant the aesthetics were; how imaginative and explorative the users felt; and how interesting the story of the simulation was. Flow is measured as how occupied and engaged with an activity that involves interaction with a human-like entity. As such, the results suggest that similar to presence, both immersion and flow (involvement) are necessary requirements for experiencing social presence. In light of these results we accept H2.

In addition, it was also revealed that the participants’ levels of perceived social presence is a significant predictor for their game experience. The results of these linear regression analyses are shown in Table 2. As such, the results of this experiment suggest that there is indeed a bidirectional relationship between the perceived social presence and game experience in the context of human-ECA interaction in VR. In light of these results, we argue that more emphasis should be placed on the design of engaging activities to be performed in relation to an ECA in human-agent interaction contexts in order to elicit a high level of perceived social presence.

6. Conclusions, limitations, & future work

The results of a pilot study with our prototype indicate an overall positive perception of it in terms of perceived social presence, game experience, and attitude towards the incorporated ECA. Based on the results of our analysis, H1a was partially accepted. It was observed that an emotionally-personified ECA with an extrovert-based personality elicits a higher sense of behavioral involvement in users, compared to a less emotion-personified agent with no non-verbal behavior, or one with an introvert-based personality. In light of these results, and limited to the context of this experiment, we conclude that (1) higher levels of personification based on incorporating personality as part of the behavioral model of an ECA seem to induce an increase in the behavioral involvement of users with an ECA as part of their social presence, (2) the way a personality-driven behavioral model is manifested as an ECA equipped with non-verbal behaviors seem to have a substantial influence on the first conclusion. These conclusions are corroborated by the observations that an extrovert ECA with more pronounced, assertive, and noticeable non-verbal behaviors evokes a higher sense of behavioral involvement in users. Furthermore, based on the results of our analysis H1b was rejected as no significant differences between the conditions were observed with respect to the measured game experience metrics.

Furthermore, our results indicate that similar to presence, immersion and flow (engagement) are bidirectionally related to social presence. Although the results do not enable us to generalize that flow and immersion are mandatory requirements for experiencing social presence particularly when interacting with an ECA in VR, we have shown that they are strong predictors for the perceived feeling of social presence of users in the context of this research. Based on this result H2 was accepted. Moreover, we have witnessed that engagement with the activity performed in the virtual environment and in relation to the ECA (actual interaction with the ECA) is a more significant predictor for perceived social presence, compared to the immersiveness of the experience itself. Furthermore, the role of tension/annoyance as part of the game experience of users that was largely neglected by previous studies when examining the relation between game experience and

<table>
<thead>
<tr>
<th>Game experience predicting social presence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flow</td>
</tr>
<tr>
<td>Immersion</td>
</tr>
<tr>
<td>Tension/annoyance</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Social presence (combined)</th>
</tr>
</thead>
<tbody>
<tr>
<td>B = .240 p &lt; .01</td>
</tr>
<tr>
<td>B = .179 p &lt; .05</td>
</tr>
<tr>
<td>B = −.148 p &lt; .05</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Social presence predicting game experience</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social presence (combined)</td>
</tr>
<tr>
<td>Flow</td>
</tr>
<tr>
<td>Immersion</td>
</tr>
<tr>
<td>Tension/annoyance</td>
</tr>
<tr>
<td>Negative affect</td>
</tr>
<tr>
<td>Positive affect</td>
</tr>
<tr>
<td>B = .856 p &lt; .01</td>
</tr>
<tr>
<td>B = .926 p &lt; .001</td>
</tr>
<tr>
<td>B = −.841 p &lt; .01</td>
</tr>
<tr>
<td>B = −.629 p &lt; .01</td>
</tr>
<tr>
<td>B = .999 p &lt; .000</td>
</tr>
</tbody>
</table>

Table 2
Results of regression analyses.
presence, was shown to be significant. Our results suggest that the less tension the users experience, the more likely they are to perceive the social presence of the ECA with respect to our experimental setup. Based on these observations, we conclude that (3) the game experience metrics defined as flow, immersion, and tension seem to be significant predictors for the perceived social presence of people in social human–ECA interaction in VR, with flow as the strongest predictor.

Notwithstanding, it is important to keep in mind that this research was carried out using a prototype and only a sample size of 41 participants. As such, it unavoidably has some limitations and presents opportunities for improvement. One of the most frequently received comments from the participants during the informal interview after the experiment sessions was related to the way conversation options were visualized and presented to them. Given the fact that participants perceived the virtual world through VR glasses, each option was presented to them as a standalone 3D object in the environment visualized on the top left, center, or top right side of the ECA (depending on the number of conversation options available at each turn). As is the case in normal conversation, people often utter more than a few words when expressing their opinion and emotions. Therefore, in some cases the content of the options spanned over two sentences and required the participants to invest the necessary time for studying them before deciding which one to choose and utter to the ECA. This forced the participants to frequently divide their attention between the conversation options and the ECA for a relatively long duration. Therefore, they could not dedicate their undivided attention to observing the ECA’s non-verbal behavioral cues all the time, which might have influenced their perceived social presence as well as game experience. As part of future work, we are working on a mechanism that would display the options one at a time, placed right under the ECA. The users can then flip through the list of options (e.g. using the controllers of the VR headset), while having the ECA present in their area of attention at all times.

Furthermore, although all participants were introduced to the functionality of Leap Motion and were told that they had the freedom to use body language in form of hand gestures when interacting with the ECA, almost no participant used this feature. This could be due to the reason that people do not tend to use no-verbal language when interacting with artificial entities, and/or the lack of participants’ full embodiment in VR (lack of a body ownership feeling). Therefore, as part of the future work, we are planning to incorporate a full embodiment of the users in this environment and perform real-time body tracking using inverse kinematic techniques. This is of interest to the continuation of this research, as it will aid us in monitoring, analyzing, and interpreting the signals produced by the users in terms of gestures and postures when interacting with the ECA. Such addition would open the door for investigating rather interesting research questions, such as the evaluation of the effect of full embodiment in VR on perceived social presence.

Moreover, with respect to the dynamics between game experience metrics and the perceived social presence of participants, our results do not enable us to draw any conclusions on what exactly affects the game experiences that predict the perceived social presence of users. In light of this, as part of the future direction of this research, different aspects of this virtual experience (e.g. interaction modality, the use of VR, content of the interaction, quality of the environment and the ECA, etc.) will have to be manipulated to assess their effects on different game experience metrics, and consequently on the perceived social presence of users.

Lastly, based on the gained insights on incorporating personality as part of the emotional model of an ECA and the effect of its manifestation on the perceived social presence and the overall game experience scores of users, we are confident in having a good enough system to start evaluating its learning effect on social behavior competency of users. Therefore, as part of future work, we will be focusing on training individuals for adequate social behavior competencies in the context of employee evaluation using our prototype.

References


