Improving Social Presence in Human-Agent Interaction

André Pereira
INESC-ID and Instituto Superior Técnico
University of Lisbon, Portugal
andre.a.pereira@ist.utl.pt

Rui Prada
INESC-ID and Instituto Superior Técnico
University of Lisbon, Portugal
rui.prada@inesc-id.pt

Ana Paiva
INESC-ID and Instituto Superior Técnico
University of Lisbon, Portugal
ana.paiva@inesc-id.pt

ABSTRACT
Humans have a tendency to consider media devices as social beings. Social agents and artificial opponents can be examined as one instance of this effect. With today's technology it is already possible to create artificial agents that are perceived as socially present. In this paper, we start by identifying the factors that influence perceptions of social presence in human-agent interactions. By taking these factors into account and by following previously defined guidelines for building socially present artificial opponents, a case study was created in which a social robot plays the Risk board game against three human players. An experiment was performed to ascertain whether the agent created in this case study is perceived as socially present. The experiment suggested that by following the guidelines for creating socially present artificial board game opponents, the perceived social presence of users towards the artificial agent improves.

Author Keywords
Human-Robot Interaction (HRI); social presence; board games; artificial opponents

ACM Classification Keywords
I.2.1. Artificial Intelligence: Applications and Expert Systems - Games; I.2.9. Artificial Intelligence: Robotics - Commercial robots and applications; H.1.2. Information Systems: User/Machine Systems - Human factors

INTRODUCTION
Techniques for representing others to evoke social presence have an ancient history that dates back to the first stone sculptures [3]. Science-fiction films or books have long included characters such as intelligent computers, robots and androids that evoke the same type of social responses from the audience or the reader [20]. With the evolution of technology, these science-fiction visions are now becoming a reality, and new interactive techniques and devices are designed to evoke social responses from users. Quasi-social relationships are beginning to be established with computers, intelligent virtual agents and robots [19]. People attribute social presence to interactive devices and treat them in social ways, even while knowing that these devices do not have real emotions or ideas [26]. Social presence can be briefly defined as the "sense of being together with another" [2]. Studying social presence can contribute to the understanding of human social behaviour while using these types of technologies. Achieving a sense of social presence is the design goal of many types of hardware and software engineering.

Creating artificial agents and robots that can interact with multiple humans is starting to become the focus of some computer entertainment and educational scenarios [17, 5]. This is a challenging task, as in multi-party interactions it is often more difficult to predict others' intentions and behaviours. The main focus of this paper is to guide researchers or engineers regarding how to design socially present agents. More specifically, we focus on intelligent agents that play board games against multiple players while still being perceived as socially present. These agents, rather than focusing on beating human players by performing millions of operations per minute, are socially aware, and their behaviours are influenced by their history with different human opponents. Board game artificial opponents are already being created both in Artificial Intelligence (AI) research and commercial applications. However, current artificial opponents still have performance and social deficits, especially in games in which they play against multiple human players. Johansson [13] stated that "bots are blind and objective, while humans may decide to eliminate the bots first, just because they are bots". This sentence shows that humans attribute a very low sense of social presence to artificial opponents. If human players do not perceive artificial opponents as socially present, their enjoyment and willingness to interact with them decreases [10].

Definitions of social presence and the factors that influence social presence in human-agent interaction are presented in the related work section of this paper. By taking these factors into account and following guidelines for creating socially present board game opponents [25], a case study was developed. This case study explores the use of today's technology to increase users' perceived social presence towards an artificial opponent. It features two main technology devices that provide an artificial opponent with the capability of playing a board game against multiple human players. First, a custom digital table was used as the interface between the human players, a robot and the game. Second, the artificial opponent is physically embodied by a robotic head with emotion rendering capabilities. We have performed an experiment in which participants played the Risk board game either against the artificial opponent described in our case study or against a control condition opponent based on a previous study [18].

CHI 2014, April 26–May 1, 2014, Toronto, Ontario, Canada.

http://dx.doi.org/10.1145/2556288.2557180

This experiment suggested that by following the design considerations described in our case study, users' perception of the social presence of the artificial opponent improves.

RELATED WORK

Social presence was initially proposed by Short, Williams and Christie [30] as "the degree of salience of the other person in the interaction and the consequent salience of the interpersonal relationships". The social presence theory allows researchers to guide the design, and to anticipate and measure differences of new types of social technology. Instead of using trial and error exploration, a better understanding of what social presence is can save valuable time and money, and can improve the end product in the design of new media technologies [20]. Many studies regarding social presence are found in new techniques of human-human communication such as computer conferencing. However, social presence is also used to measure an individual's perception of a particular interactive medium, be it a virtual reality environment [31] or the interaction with a social robot [18].

Several authors have discussed factors that influence social presence, but none of the authors have previously focused on artificial opponents. In the remainder of this section, we collect the factors that are deemed important for creating socially present artificial agents and opponents.

Contributing Factors for Social Presence

There are different modes of interacting with a virtual agent, but in terms of social presence, *face-to-face interaction* is still considered the gold standard in communication against which all platforms are compared [1]. Social presence is assumed to be highest when two people are within reach of each other and are interacting on a task [4]. Therefore, virtual agents that do not use the rich set of social behaviours and cues involved in face-to-face interaction may be considered less socially present. Face-to-face interaction is generally accompanied by *verbal communication*. While interacting with virtual characters or robots, verbal communication offers the most attractive input and output alternatives. We are familiar with verbal communication, it requires minimal effort from users and it leaves their hands and eyes free [34]. Interactions should also feel natural and quick. Systems should have *quick feedback* for the user to feel immediacy of control, as delays between actions and reactions can diminish the sense of social presence [20]. In robots or on-screen characters, having a responsive real time gaze system can produce a high sense of agency and increase the agent's perceived social presence [35].

The *number of interacting entities* (be they virtual agents or humans) can also positively influence the perception of social presence in an interactive system. Having more than one entity in media interactions can be an easy way to induce a sense of social presence, regardless of the other perceptual features of the world [11]. *Knowledge and prior experience* with the medium also influences the sense of social presence. Social presence varies across individuals and across time for the same individual. When we have been exposed to artefacts of a particular medium over time, we have a higher knowledge of interacting with it, and it is possible to have an increased feeling of social presence. However, very often, continued experiences may cause the well-known habituation or novelty effect [14]. This effect causes an initially higher sense of social presence that fades away as users become more experienced with novel technology [18]. This novelty effect is present in almost all types of media, including artificial agents or robots [8]. The ability to attribute mental states to oneself and to others is fundamental to human cognition and social behaviour [32]. Biocca [4] stressed the importance that the *theory of mind* has in social presence. He defines social presence as the sense of "being together with another" and attributes this sense to the ability to relate to or to construct mental models of another's intelligence. These models can simulate minds of people, animals, agents, aliens, gods and so on. If we interact with an agent and create a mental model of it, we can anticipate the agent's behaviour and judge its consistency. The number and quality of sensory channels are important for generating a sense of social presence. More importantly, *consistency* between all of the different modalities is one of the most relevant keys for achieving social presence: "the information received through all channels should describe the same objective world" [20]. If we do not meet this criterion, we emphasise the artificial and lessen the feeling of social presence. Correlations between actions and reactions should be credible when compared to events that would be expected in reality under similar circumstances. Slater [31] notes that another important factor for social presence is the occurence of some events not directly related to the users' actions. These events show *autonomy* in the environment or character. In a study conducted in a cave-like environment, participants spent approximately five minutes in a virtual bar interacting with five virtual characters [7]. Participants were reported to automatically respond to the virtual characters present in the bar in social ways. Though these virtual characters had limited social behaviours, mutual gaze combined with lucky randomness, was perceived by participants as the characters watching and mimicking them. *Embodiment* is also important for designing a computer to achieve a higher sense of social presence. It has been reported in virtual poker environments that the simple addition of a picture personifying players made the game more likeable, engaging and comfortable [16]. We can also find examples where physically embodied agents (or robots) are used to simulate opponents. It was shown that by using a robotic embodiment instead of an on-screen character, artificial opponents have improved feedback, immersion and social interaction [23]. We use our emotions in our social world almost constantly.

Emotional responses can elicit adaptive social responses from others. For example, someone with an angry temperament can elicit a fear response from someone else, while someone distressed may elicit an empathic response from others. We also use emotional expressions for communication, signalling and for social co-ordination. These types of natural social primitives can be interpreted by humans without the need to learn something new; a human-like computer using these cues can cause social facilitation in users. Endowing agents with *emotional behaviour* can contribute to the believability of a character and thus to its perceived social presence.

Figure 1. Risk case study

CASE STUDY

Five guidelines for designing socially present board game opponents were identified in our previous work [25]. These guidelines argue that to improve social presence, an artificial board game opponent should (1) have a physical embodiment and be able to engage in face-to-face interaction with one or multiple participants; (2) be believable and obtain players' attention by using both verbal and non-verbal behaviours; (3) have an emotion system to make better judgements and to simulate human emotions; (4) have social memory, i.e., recognise each user individually and remember its past interactions with him/her; and (5) be able to simulate social roles common in board games. We developed a case study that implements these guidelines in which an artificial opponent plays the Risk board game against three human players. The human players use a digital table as the game interface. Risk was chosen because it is a game where face-to-face interactions, social actions and strategic social reasoning are important components of the game. In the case study, the social robot is positioned on one side of a digital multi-touch table and interacts with three other players on the other sides of the table (see Figure 1). In the remainder of this section, we first present the design considerations taken in this case study. Then, we briefly describe the empirically inspired relationship variables that influence the decision process, dialogue and emotional synthesis of the robot, and finally, we describe the robot's multi-player gaze system.

Design Considerations

By using robotic embodiments, board game computer opponents are more successful at engaging in face-to-face interaction with multiple participants [25]. For embodying the social Risk opponent we used the EMYS (EMotive headY System) robotic head [15]. This robot was chosen for three different reasons. First, it can socially engage users face-to-face by quickly moving its head, eyes and eyelids towards the user, which makes it appropriate for our gaze system. Second, it can display idle behaviours and a wide range of emotions with different intensities. The idle behaviours and facial expressions used in this case study were developed by Ribeiro et al. [27]. These authors took inspiration from the principles and practices of animation and applied them to the development of idle behaviours and emotional expressions for the EMYS robot. Facial expressions are used in the case study for establishing turn-taking and for revealing emotional states such as sadness and pleasure. Finally, we chose this embodiment because of its fun and cartoonish appearance. While interacting with anthropomorphic robots, users have certain expectations of them, and not meeting such expectations can deteriorate the experience [22]. This consideration should also hold true when choosing an embodiment for an artificial opponent. The robot uses high quality text to speech (Brian from IVONA[1]) to vocalise its utterances. While speaking, the robot is able to accurately lip-sync with the text-to-speech system. We also upgraded the EMYS robotic head with a Kinect sensor that is used for detection of speech direction.

We designed and built a custom digital table for this case study. By using a digital table, players are able to freely communicate and still be aware of the game state. This means that they can more easily be engaged with both the game and each other [28]. In the table's design we took into account the ergonomic considerations described in [29]. The first design decision to improve ergonomics was to have edges considerable wider than those of commercial digital tables available at this writing. If the edges are not wide enough, users tend to lean their elbows or arms on the table surface, which interferes with the application. By allowing users to rest their arms, they can be more relaxed and focused on social interactions. The second ergonomic decision was interface related. When using digital tabletop hardware, users should have their interface area in a comfortable personal space [33]. We designed the game interface using this principle, and therefore, we took users' comfort into account so that they could focus solely on social interactions with their peers and the social robot. Finally, the height of the table was considered. When designing a table for casual interaction, short coffee tables are appropriate, desk-height tables are more suited for productivity or longer tasks such as long gaming sessions. Therefore, we chose the height of a normal desk table.

For implementing social memory in the artificial opponent, we simplified the user recognition process by requiring each user to login with their own private interface on the digital table. At login time, the robot acknowledges the presence and position of a user, greets that particular user, and updates its history with him/her. The artificial opponent stores basic data in memory about the results and dates of previous matches. This information is often mentioned in the initial interaction, where the robot can say for example: "One week ago you won, this time I am going to win!". More complex social variables are also used by the robot to map relationships established with particular users. These variables evolve during the game and are stored in the robot's memory for future interactions.

Social Relationship Variables

Dialogue recordings of humans playing the traditional Risk game have been previously categorised [24]. The most relevant categories to model in a Risk artificial opponent were

[1] http://www.ivona.com/en/

extracted and that categorisation helps pinpoint the most important social behaviours to simulate in a Risk game. Additionally, in that study, a database of possible utterances was extracted from real human social behaviour. However, to understand the humans' thought process and to know when an artificial opponent should select a particular move or utterance, a protocol analysis was performed. In this analysis, participants were asked to think aloud while playing a traditional Risk game, and the most relevant variables to simulate in a social Risk opponent were extracted. In this sub-section we briefly describe the empirically extracted variables that enable the artificial opponent to establish different social relationships with different users: *Familiarity*, *Like/Dislike* and *Luck Perception*. These variables are used for generating dialogue or for choosing the next move.

Familiarity can be an important variable to model in a social agent because the number of utterances that an artificial agent can speak is often limited, and long term studies have shown that repetitive behaviours decrease social presence and believability [18]. Therefore, it is advantageous for an artificial agent to be shyer (less talkative) towards players with whom it has interacted for only a limited time and to become more familiar (talkative) with them over time. By implementing this behaviour, we are also following empirical results showing that when players already know each other outside the game or when they have played previous games before, they are more communicative and more willing to establish alliances between themselves. In our implementation, familiarity starts at a minimal value, never decreases, and increases slightly every time the robot interacts with players or every time a player interacts with the robot.

Influenced by attacks in the current and previous games, another relationship variable was simulated. This variable can be either positive (like), zero (neutral), or negative (dislike) for each of the agent's opponents. The agent can act socially according to this variable and establish different social relationships with different users. The variable changes after relevant game events; a change can be positive or negative. When players are not attacking each other, they are nicer to each other, and the opposite occurs when they attack. Thus, when a player attacks the agent, the relationship variable towards him/her decreases. The variable increases slightly when players have the opportunity to attack the robot but do not. The variable also changes when an opponent attacks a player the agent likes or dislikes. We took inspiration for this behaviour from Heider's balance theory [12]. Following the balance theory rationale, when a player is attacking one of the robot's "most hated" opponents, the Like/Dislike variable towards him/her increases. Conversely, if a player attacks one of the agent's "friends," the relationship variable will decrease. As previously reported, the Like/Dislike variable is stored, so the robot can disclose, for example, that it holds a "grudge" against a particular player because of previous games.

Luck perception is also stored in memory, so the agent can assess and comment if a player is lucky or if he/she was lucky in previous games. Risk is a game that involves lucky dice throws, and players are constantly "storing" in memory the luck that they attribute to other players. When players are lucky and constantly winning at dice throws, other players expect them to continue winning and usually comment on that fact. When the unexpected occurs and players lose after a winning or losing streak, stronger verbal and nonverbal reactions are usually elicited. Due to the frequency of verbal and emotional content commonly found in this game event [24], such behaviours seem to be important to implement in artificial agents that play games involving luck. Simple rules and statistics were used to monitor players' luck in the game. Luck events are generated by using an anticipatory mechanism [21] that assesses the mismatch between the agent's expectations of a dice throw and the actual result.

Risk is a highly social game that supports various social roles in its game play. Players can change roles throughout a game. Social roles in board games were identified by Eriksson et al. [6]. In our implementation, these roles arise because the robot's AI uses the relationship variables to influence its social behaviour. For example, when the agent "likes" another player, it often demonstrates the social role of Helper by making encouraging comments such as "It went well this turn!". Conversely, when the agent has a negative relationship with another player, it is more likely to adopt the Violator role, for example by attacking him/her without seeking any in-game benefit.

Multiplayer Gaze System

To achieve believable face-to-face interactions, we developed a gaze system that equips our robot with the ability to simultaneously interact with multiple players in our gaming context. The gaze system uses speech direction detection, face detection, and the context of the game; it is based upon studying how humans behave in such context. We have extracted gaze patterns and created a gaze system for multi-user interaction with a social robot. The gaze system is influenced not only by the robot's own variables but also by the other players' voices and game actions. The robot also uses a camera in its "nose" for face detection and uses sound direction detection sensed from the Kinect's microphone array.

In most board games, players shift their focus between looking at different parts of the board and looking at other players. Players look at the game board when they are thinking, during their own turns and their opponents' turns or when other players make their moves on the board. When it is not a players turn, they tend to look at the active player more than any other. We noted that the factor that most influences the amount of time that participants look at the board or at other players is their concentration on the game. Inspired by this observation, we modelled these behaviours in our robot by using a *concentration* variable. This variable enables the robot to look more focused on the game during its own turn and more focused on the other players during their turns. The variable tends to be higher during the robot's turn and lower during an opponent's turn. It also decreases when other players take too long to play or when all of the events in the game are not related to the robot's game.

Once a gaze command is over or is interrupted by a relevant event for the gaze system, another gaze command is issued. When it is not the robot's turn, it performs a Concentration Test (CTest). The CTest starts by generating a random value. If that value is less than the robot's current concentration variable, the robot issues a focused gaze action. If the value is higher, it issues an unfocused gaze action. This means that if the concentration variable is low, the robot tends to be unfocused, and when the variable is high, it is more often focused. When unfocused, the robot looks at other players randomly but with a higher probability of looking at the active player. When the robot issues a focused gaze action, it looks randomly at a point on the game board, simulating that it is thinking and looking at the game action attentively.

When a user touches the interface, the robot is informed of the location of the touch, and the robot is able to gaze at it. The robot can look at points on the interface precisely. To implement this, the robot is always fixed on the same predetermined position on the digital table; parametrised gaze values were calibrated for each point on the interface. If after a touch event the robot passes a CTest, it gazes at the interface. If the CTest fails, the event is ignored. This simulates the observed behaviour that when players are thinking they are more focused on the game board.

To further improve the robot's gaze system, we found the need to implement a speech direction detection module in the robot to make it look at players when they are talking. This module was implemented by using Kinect and its beamforming algorithms from the Microsoft SDK. The Kinect microphone array reports the position of the speaking player, and we use that angle to look in the direction of the sound. However, the robot only chooses to look in the direction of speech if it fails a CTest, simulating that it only "listens" to other players when unfocused.

Finally, when the robot decides to speak to a player, the gaze system also causes the robot to look at the direction of the intended player. All gaze actions belonging to the "look at a player" group work in the following manner. First, the robot looks at a position where the player most likely is. This initial probable position was calculated by several user tests in which we fixed gaze values for looking at the three human players. Second, once the robot fixes its gaze on that position, it tries to detect a face using EmguCV[2]. If successful, it tracks the user's face and follows it for the remaining time of the gaze action. Finally, to increase accuracy, the robot stores the last position of that opponent, remembering the probable position for that user for the next time.

Every gaze action stays focused on the target, be it a player or a point on the interface, for a determined amount of time. The only exception is when speaking to a player, the robot directs its gaze to that player until it stops speaking. For every other type of gaze, we have defined a minimum and a maximum length of time, and a minimum and a maximum of gaze speed. When the robot initiates a new gaze order, it uses random values both for the speed and gaze duration between those minimum and maximum values. We fixed high speeds and shorter time spans for events in which the interface is touched or a speech direction is detected. Slower gaze speeds and longer time spans are used when the robot is inactive.

EVALUATION

In this paper, we hypothesise that an agent that follows guidelines for socially present board game opponents will be perceived as more socially present than an agent that does not implement such guidelines. For testing this hypothesis, we report a study between subjects where participants played a Risk game against a robotic opponent.

Participants

Forty five participants (12 males, 33 females) with ages ranging from 18 to 40 years old (M = 24.0, SD = 5.0) took part in the experiment. Participants were undergraduate and graduate students recruited via a program in which they received extra curriculum credits for their participation. The experiment was conducted at a Psychology University. Each session included three participants.

Procedure

At the beginning of each session, participants were told that they were going to play the Risk board game against a physical robot and each other. We used a Powerpoint presentation alongside the interface of the game to explain the Risk rules and the game interface. A sequenced presentation was used so that every participant learned the rules and how to interact with the game interface in the same manner. The three participants then sat around the digital table with EMYS, our social robot, on the remaining side. EMYS acted as an artificial opponent for thirty minutes or until it was eliminated from the game. When EMYS was eliminated or the thirty minutes were over, EMYS warned participants that the interaction had ended. Finally, participants filled in a questionnaire, and the experiment was over. Each experiment lasted approximately 1 hour.

Manipulation

The experiment had two main conditions: a socially present condition (SP) in which we used the full implementation of our social opponent, and a less social condition ($\neg$SP) in which the behaviour of the robot was not inspired by the guidelines for creating socially present opponents. We assigned 27 participants (9 groups of 3) to the SP condition and 18 participants (6 groups of 3) to the $\neg$SP condition. When recruiting participants for this study, we requested good English skills as a prerequisite. However, two participants in the SP condition were removed from the study (1 male and 1 female) because they reported in the questionnaire and to the experimenters that they did not have the English skills necessary to understand EMYS.

Socially Present Condition (SP)

Our experiment had the limitation of a single 30 minute interaction with the robot. Given this time constraint, we decided to fix some of the agent's variables to have a more controlled experiment and for users to experience the same diversified

[2] http://www.emgu.com/

behaviour that would occur in longer exposures to our artificial opponent.

We designed the familiarity variable for long term interactions with the same participants. In a single interaction, we wanted the social robot to be as interactive as possible from the beginning of the game. Therefore, we fixed this variable to its highest value, 1, for all participants.

Additionally, we randomly assigned fixed values of the Like/Dislike variable to each user. When a game started, the social robot had a *positive relationship* with one of the participants, a *negative relationship* with one of them and a *neutral relationship* with the third. By manipulating the Like/Dislike variable in this manner, we ensured that the robot displayed different social roles in each interaction and had a more diversified and constant behaviour.

Not as Socially Present Condition (¬SP)

Comparing an artificial opponent with another that has no gaze at all and no speech or non-verbal behaviour, would most certainly result in a decrease in the sense of social presence. In [18], a study was presented in which, in a similar scenario, the behaviour of a social robot was not sufficient to maintain a sense of social presence in users. This evaluation and the guidelines presented in [25] were informed by that study. We compared our SP condition with the baseline behaviour of that robot and identified the main differences between the two. As a result, in the ¬SP condition, when compared to SP, the following changes were implemented: (1) the same physical embodiment was used, but a more randomised gaze system provided less believable face to face interactions, (2) the verbal and non-verbal behaviours of our robot were limited to only non-social game events, (3) the robot exhibited fewer emotional expressions and its emotional system did not influence the moves it chose, (4) it did not store any type of social data in memory and (5) it did not display the common social roles present in board games.

To implement a randomised gaze system, the robot was always looking at random points in the interface during its turn, and during the opponent's turn, it was randomly shifting its gaze towards one of the other players. The main differences between the simplified gaze system and the complete one are the absences of speech direction detection, a face detection algorithm to more convincingly look a player in the face, and a concentration variable that informs the "randomness" of the gaze target.

Regarding the verbal and non-verbal behaviour of the robot, in this condition, our artificial opponent only commented and reacted to non-social events. To generate social events in the SP condition, we used our empirically inspired social variables such as Like/Dislike and Luck perception. These variables were set to 0 and did not change. This meant that the robot was neutral towards all users and did not comment, for example, on lucky dice throws. Familiarity was the only empirically inspired variable that we maintained in this condition. It was fixed to 1 so that there would be no differences in the speech rate across conditions. As the robot's variables did not change in this condition, the agent had no social memory and could not display social roles.

Measures

Social Presence

The measure for social presence that we considered grounded and suited to measure social presence in human-agent interaction is the Networked Minds Questionnaire [9]. This measure was originally designed to distinguish between levels of social presence between face-to-face interactions and different levels of mediated interactions. However, we have successfully used it for measuring social presence in a human-robot interaction scenario [18]. It conceptualises social presence in the following six dimensions: (1) *co-presence*, the degree to which the observer believes s/he is not alone; (2) *attentional allocation*, the amount of attention the user allocates to and receives from an interactant; (3) *perceived message understanding*, the ability of the user to understand the message from the interactant; (4) *perceived affective understanding*, the user's ability to understand the interactant's emotional and attitudinal states; (5) *perceived affective interdependence*, the extent to which the user's emotional and attitudinal state affects and is affected by the interactant's emotional and attitudinal states; and (6) *Perceived behavioural interdependence*, the extent to which the user's behaviour affects and is affected by the interactant's behaviour.

This questionnaire uses six items for each of these dimensions; subjects were asked to express their agreement or disagreement regarding each item on a seven-point Likert scale (one means "totally disagree" and seven means "totally agree"). Two experimenters separately translated all items of the questionnaire into Portuguese. They then agreed upon the most suitable and correct translation.

General Questions

To assess whether users would like to continue interacting with the robot on a regular basis and to learn what the users liked most and least about EMYS, we asked the following general questions after the social presence questionnaire: Did you enjoy playing against EMYS? Would you like to play another game of Risk with EMYS? What would you change about EMYS? Do you think EMYS has social skills? For each of these questions we also asked why?

Results

In this section, we first validate our translated social presence questionnaire by checking the normality and reliability of the scales. We then present a comparison between the SP and ¬SP conditions using the measures above.

Reliability Analysis and Normality Tests

For measuring social presence, we used a questionnaire that was already validated and had been used previously by the research community. However, because we had translated all of the items from English to Portuguese we started by performing reliability tests on our questionnaire. Additionally, we ran normality tests and examined the normality plots of each dimension to assess the normality of each measure. Of the six

Social Presence Questionnaire	Mean	t(41)	p-value	r
Attention Allocation	SP = 4.18 ¬SP = 3.78	0.99	>0.05	0.15
Perceived Aff. Understanding	SP = 4.34 ¬SP = 2.57	5.49	<0.01	0.65
Emotional Interdependence	SP = 4.21 ¬SP = 3.15	2.64	<0.01	0.38
Behaviour Interdependence	SP = 4.15 ¬SP = 2.78	3.42	<0.01	0.38
Total	SP = 4.70 ¬SP = 3.51	4.55	<0.01	0.47

Table 1. Mean, t-value, p-value and effect size for the parametric Social Presence questionnaire items.

different dimensions of the presence questionnaire, three dimensions had unacceptable Cronbach's α levels. After further analysis, we found that the reliability of those scales could be greatly increased by removing two items from each of these 3 dimensions. The other items had very high reliability values and all items from the original questionnaire were used. Regarding the normality of the distribution, most measures were parametric with the exception of Co-presence and Perceived Message Understanding, which are not normally distributed. For non-parametric items, we applied the Mann-Whitney U test and report the median value between conditions. For the parametric items, we applied independent t-tests, and the mean is reported.

Social Presence

Overall, participants in SP (Mdn = 6.17) experienced a significantly greater sense of co-presence than participants in ¬SP (Mdn = 5.00), U = 96.50, z = 3.17, $p < 0.001$, r = 0.48. In terms of Perceived Message Understanding there was also a significant difference between the SP condition (Mdn = 5.00) and the ¬SP condition (Mdn = 3.25), U = 102.50, z = 3.025, $p < 0.01$, r = 0.46.

The other dimensions were parametric and, with the exception of attention allocation, were similarly significantly classified as higher in SP, as we can observe in Table 1.

General Questions

All participants in the SP condition claimed that they enjoyed playing against our artificial opponent, and 92% stated that they would like to play against EMYS again (see Figure 2). Enjoyment (89%) and intention to continue using the system (78%) were lower but also quite high for the ¬SP condition. When we asked participants if they would change something about EMYS, only 28% would change something in our main condition in contrast to 56% in the ¬SP condition. The Chi-Square tests did not reveal significant differences between these questions. However, when users were asked if they thought that EMYS had social skills, the difference was significant ($\chi^2(1) = 11.495$, $p < 0.001$). In the SP condition, 88% claimed that EMYS had social skills, while in the ¬SP only 39% answered affirmatively. The effect size for this finding was large ($\phi = 0.517$).

These questions had a qualitative second part asking participants to further develop their answer. Next, we report the patterns identified in the qualitative data.

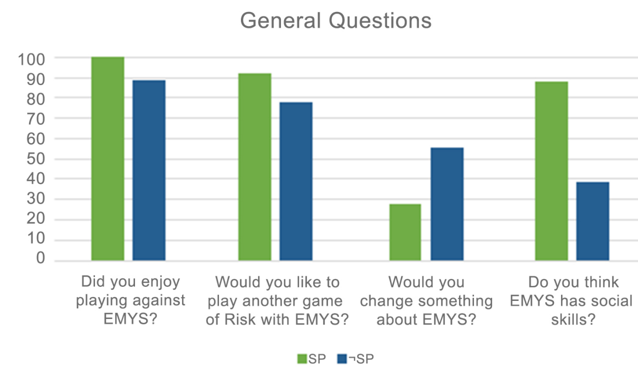

Figure 2. Percentage of positive ("yes") answers reported.

Did you enjoy playing against EMYS?

Many participants in the SP condition claimed that they enjoyed playing against EMYS because it was fun, interesting, challenging and the same as playing against a human player. Many participants also focused on social interactions with and the characteristics of the robot to explain their enjoyment. Empathy for, and feelings of rivalry or alliances towards the robot were the most reported social claims. One participant replied: "I felt like he was a real companion that socialised with all of us". In the ¬SP condition, participants focused more on the game being a new type of experience: "I enjoyed it because it was an experience in which I played against a robot". Enjoyment was high in both conditions, and we could not extract any relevant trend in the development of the negative answers.

Would you like to play another game of Risk with EMYS?

More than half of the participants in SP responded that they would like to play another game against EMYS because it was fun. Interesting and exciting were also common words used to characterise the interaction. Some participants also said that the robot had a good sense of humour and that they enjoyed interacting socially with a robot. A common answer was: "I would like to have the opportunity of getting to know EMYS a little bit more." In the ¬SP condition, participants also mentioned the robot's sense of humour and the fact that it was a fun, novel, and interesting experience. However, they did not dwell as much on the social aspects of the robot as on the challenge that the experience provided and how well the robot played. Some participants replied that they would like to play a game until the end and not only thirty minutes. One participant said: "I would like to play until the end to determine a winner". For this question, we could also find some common negative answers. In both conditions, most participants would not like to play against EMYS again because they do not like board games. However, in ¬SP, participants also replied that it was the same as playing against a computer and that the robot should be more interactive.

Would you change something about EMYS?

In the SP condition we had multiple types of suggestions to improve our artificial opponent. Some suggested changes in the embodiment, such as reducing the robot's noise or adding hands or a body to our robot. Others argued that it should recognise and speak the participants' native language, be less competitive and not so critical. Curiously, most changes suggested by participants in the ¬SP condition were already addressed in the SP condition. As expected, most participants that would not change anything about EMYS did not elaborate on their response. Typically, in both conditions we had some users saying that the "robot was already highly advanced" or that they "would not change a thing".

Do you think EMYS has social skills?
There were large differences in both conditions in responses to this question. In the SP condition, the most frequent positive answer was that the robot acted appropriately and according to the participants moves. Many participants also recognised the robot liking or disliking one of the players as a social act. Although the robot started with a predetermined value for the Like/Dislike variable and did not ally with other players, participants said that the robot did not like players who attacked him and made alliances with players who did not. One participant replied that "it can socialise, does not like it when we attack him and liked one of the players more, he even showed unhappiness when we attacked that player". In the ¬SP condition, players gave diverse reasons for their answers: some replied that it was able to interact socially but only in a superficial manner. One participant replied that it does interact socially but is more suited for younger people. Pressuring users to play was also regarded by two players as a social behaviour. Irony, competitive behaviour and emotional capabilities were also reported in this condition. One participant replied that "it can be a little social as it puts pressure on us when we take too long". In both conditions, negative responses were justified by the robot being programmed or simply because it is "just a robot".

Discussion

When comparing SP versus ¬SP, only one of the social presence dimensions, attention allocation, was not significantly different across conditions. However, even in this dimension, we had a slight positive difference benefiting the socially present condition. This, alongside with the total of the social presence questionnaire being significantly higher in the SP condition, indicates that by following the guidelines for creating socially present opponents in the design of our artificial opponent, we increased participants perceived social presence. Co-presence, one of the most important factors for achieving social presence was the highest rated dimension in both conditions. Regarding attentional allocation, the difference was not significant because it was one of the highest rated dimensions in the ¬SP condition. In both conditions, participants allocated attention to the robot. We have two reasons to explain the similarity in the results between the conditions. First, the random gaze that we implemented always looked at a random player or the interface. Although it was doing it randomly, when the robot looked at participants, most thought that the robot was paying attention to them. Additionally, participants allocate their attention towards the robot because none of them had any prior experience with a robot and may have been influenced by the novelty effect. The other four dimensions, perceived message understanding, message interdependence, affective understanding and affective interdependence, were clearly more successfully achieved in our socially present condition. By implementing a gaze system that "listened" to players speaking, followed their faces, looked at where participants touched and was influenced by the game state, we successfully made users think that the robot's gaze was dependent on their behaviour. In terms of affect interdependency and understanding, the introduction of the Like/Dislike and Luck variables, which in part resulted in different social roles, seems to have contributed to this result. For example, in the qualitative questions many users noted that they thought that the robot would become angry with them if they attacked it. This means that they associated the robot's facial expressions and future behaviours with their past actions.

We also investigated differences in social presence when comparing the three different variations of the Like/Dislike variable in the SP condition. No significant differences were found. The social presence reported by participants was not significantly affected by the robot's relationship with the user. We can explain this result because within SP every participant was exposed to each variation. When replying to the qualitative questions, participants in the neutral variation also noticed the robot's social behaviour towards the other players. These players also replied that "the robot was angry with another player because that player attacked him". Some of our highest ratings for social presence were even from participants in the Negative relationship variation, which indicates that the perception of social presence does not seem to be affected by whether the relationships with the robot were positive or negative. The social presence that participants attributed to our artificial opponent seemed to be more influenced by the believability of the robot's behaviour towards the whole group.

We did not include questions in our questionnaire to directly assess if alliances were formed during the game. However, at the end of the game, we directly asked if any alliance was established. In the SP condition, only one of the 9 user groups allied against the robot, and the reported reason for doing so was that the robot was getting too strong. In the ¬SP condition, two out of six groups allied against the robot. One of the groups was even successful in eliminating our artificial opponent before the thirty minutes were over, ending the experiment at minute 27. The reasons for the alliance argued by participants were that they wanted to know what would happen if the robot lost or simply because "it was just a robot". In the introduction section, we noted that we had already predicted that users would choose to eliminate the artificial agent first. Preventing this behaviour was one of our main motivations. It seems that we had reasonable success in fighting this effect, although we would need to have a larger number of interactions to better establish it.

CONCLUSIONS

We attempted to answer the question of how to create socially present agents that can play board games against multiple human players. Five design guidelines for increasing social presence in artificial opponents were outlined in a previous theoretical finding. To demonstrate how these guidelines can be used to create new forms of computer entertainment and to evaluate whether they indeed improve an artificial agent's social presence, we developed and evaluated a case study. An experiment was performed in which 45 participants played the Risk board game against an artificial opponent. The evaluation suggested that by following guidelines for creating socially present opponents, the users' perceived social presence towards our artificial opponent improved. We have no measure to exactly determine the success of each individual guideline. However, we can relate our results to each guideline.

Be physically embodied and be able to engage users in face-to-face interactions. Previous studies have shown the importance of a physical embodiment in an artificial opponent. The artificial opponent was therefore physically embodied in both of our conditions. However, when comparing both social conditions, our results suggest that the face-to-face interactions provided by the use of our gaze system contributed positively to the sense of presence. To further investigate differences in face-to-face interactions between conditions, facial videos of all the participants were retrieved during the experiment. For future work on this subject, a more detailed analysis using video annotations will be performed.

Exhibit believable verbal and non-verbal behaviours. By examining initial empirical studies on the target game and by carefully using those data in the implementation of the case study, we could create believable behaviour in terms of dialogue, game choices and emotional reactions. In the qualitative responses of the final user study, participants often argued that the robot's behaviour was adequate to most situations and that the behaviours were displayed in a timely manner. Additionally, there were significant differences in the dimensions of perceived message understanding and behaviour interdependence, these showed that applying data extracted from a verbal communication study and protocol analysis helped us create believable verbal and non-verbal behaviour that was displayed in a timely manner.

Comprise an emotion system. Our emotion system is composed of several variables that influence the artificial opponent's move selection, utterances and emotional behaviours. These variables, inspired by psychology models of appraisal, the human thought process while playing Risk and previous work on socially intelligent agents, seem to be sufficient for users to perceive the robot as an emotional/social being. Some participants claimed that the robot would become angry if they attacked him. Others said that he became sad when losing ground. These answers, along with significant differences in perceived affective understanding and emotional interdependence on the social presence questionnaire, indicate that participants believe that the robot played according to his emotional state and displayed coherent emotional behaviour throughout the game.

Have social memory. This guideline is of extreme importance in multiple interactions with the same participants. Although the implementation that we propose for an artificial opponent is prepared for coherently maintaining a relationship with a user throughout several interactions, our evaluation covered the span of only one interaction with each user. In the future, a longer-term evaluation should be performed. Nevertheless, participants still enjoyed and valued the robot's ability to call each participant by his/her name, and most participants in the socially present condition acknowledged that the robot established social relationships with them by remembering past actions in the game.

Simulate social roles common in board games. In the evaluation, more specifically in the socially present condition, the artificial opponent clearly expressed different social roles that were recognised by users. The robot expressed roles such as violator or dominator to players it did not like, for example, by threatening and attacking them more frequently. It frequently expressed the social roles of motivator and helper towards players it liked. By looking at participants' qualitative answers, we can also say that participants perceived the agent's social roles.

There is still much work to do before achieving a truly socially present artificial opponent or agent. However, by following guidelines and by taking inspiration from the case study that we have created, we believe that the next generation of artificial board game opponents or agents that are perceived as socially present can be created.

ACKNOWLEDGMENTS

This work was partially supported by national funds through FCT–Fundação para a Ciência e a Tecnologia, under project PEst-OE/EEI/LA0021/2013 and by the EU FP7 ICT-317923 project EMOTE.

REFERENCES

1. Adalgeirsson, S., and Breazeal, C. Mebot: a robotic platform for socially embodied presence. In *Proceeding of the ACM/IEEE international conference on Human-robot interaction* (2010), 15–22.

2. Biocca, F., Burgoon, J., Harms, C., and Stoner, M. Criteria and scope conditions for a theory and measure of social presence. *Presence: Teleoperators and virtual environments* (2001).

3. Biocca, F., Harms, C., and Burgoon, J. Toward a more robust theory and measure of social presence: Review and suggested criteria. *Presence: Teleoperators & Virtual Environments 12*, 5 (2003), 456–480.

4. Biocca, F., Harms, C., and Gregg, J. The networked minds measure of social presence: Pilot test of the factor structure and concurrent validity. In *4th annual International Workshop on Presence* (2001), 9–11.

5. Deshmukh, A., Castellano, G., Kappas, A., Barendregt, W., Nabais, F., Paiva, A., Ribeiro, T., Leite, I., and Aylett, R. Towards empathic artificial tutors. In *Proc. of the 8th ACM/IEEE int. conference on Human-robot interaction* (2013), 113–114.

6. Eriksson, D., Peitz, J., and Björk, S. Socially adaptable games. In *Proceedings of DiGRA Conference: Changing Views–Worlds in Play* (2005).

7. Garau, M., Widenfeld, H., Antley, A., Friedman, D., Brogni, A., and Slater, M. Temporal and spatial variations in presence: A qualitative analysis. In *Proc. of Int. Workshop on Presence* (2004), 232–239.
8. Gockley, R., Bruce, A., Forlizzi, J., Michalowski, M., Mundell, A., Rosenthal, S., Sellner, B., Simmons, R., Snipes, K., Schultz, A., et al. Designing robots for long-term social interaction. In *Intelligent Robots and Systems*, IEEE (2005), 1338–1343.
9. Harms, C., and Biocca, F. Internal consistency and reliability of the networked minds social presence measure. *Exploring the sense of presence* (2004), 246.
10. Heerink, M., Ben, K., Evers, V., and Wielinga, B. The influence of social presence on acceptance of a companion robot by older people. *Journal of Physical Agents 2*, 2 (2008), 33–40.
11. Heeter, C. Being there: The subjective experience of presence. *Presence: Teleoperators and virtual environments 1*, 2 (1992), 262–271.
12. Heider, F. *The Psychology of Interpersonal Relations*. Lawrence Erlbaum Associates, 1958.
13. Johansson, S. On using multi-agent systems in playing board games. In *Proceedings of the 5th International joint conference on Autonomous agents and multiagent systems*, ACM (2006), 569–576.
14. Karapanos, E., Zimmerman, J., Forlizzi, J., and Martens, J.-B. *User experience over time: an initial framework*. ACM, 2009, 729–738.
15. Kedzierski, J., Muszynski, R., Zoll, C., Oleksy, A., and Frontkiewicz, M. Emys-emotive head of a social robot. *IJ Social Robotics 5*, 2 (2013), 237–249.
16. Koda, T., and Maes, P. Agents with faces: The effect of personification. In *Robot and Human Communication*, IEEE (1996), 189–194.
17. Leite, I., Hajishirzi, H., Andrist, S., and Lehman, J. Managing chaos: models of turn-taking in character-multichild interactions. In *Proc. of the 15th ACM on Int. conference on multimodal interaction* (2013), 43–50.
18. Leite, I., Martinho, C., Pereira, A., and Paiva, A. As time goes by: Long-term evaluation of social presence in robotic companions. In *Robot and Human Interactive Communication*, IEEE (2009), 669–674.
19. Leite, I., Mascarenhas, S., Pereira, A., Martinho, C., Prada, R., and Paiva, A. "why can't we be friends?" an empathic game companion for long-term interaction. In *Intelligent Virtual Agents*, Springer (2010), 315–321.
20. Lombard, M., and Ditton, T. At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication 3*, 2 (1997), 0–0.
21. Martinho, C., and Paiva, A. Using anticipation to create believable behavior. In *Proceedings of the national conference on Artificial Intelligence*, vol. 21, Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999 (2006), 175.
22. Mori, M. The uncanny valley. *Energy 7*, 4 (1970), 33–35.
23. Pereira, A., Martinho, C., Leite, I., and Paiva, A. icat, the chess player: the influence of embodiment in the enjoyment of a game. In *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems - Vol 3* (2008), 1253–1256.
24. Pereira, A., Prada, R., and Paiva, A. Towards the next generation of board game opponents. In *FDG'11: Proceedings of the Sixth International Conference on Foundations of Digital Games* (June 2011).
25. Pereira, A., Prada, R., and Paiva, A. Socially present board game opponents. In *Advances in Computer Entertainment*, vol. 7624 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2012, 101–116.
26. Reeves, B. *The media equation: how people treat computers, television, and new media*. Stanford, Calif.: Center for the Study of Language and Information; Cambridge, 1996.
27. Ribeiro, T., and Paiva, A. The illusion of robotic life: principles and practices of animation for robots. In *Proceeding of the ACM/IEEE international conference on Human-robot interaction* (2012), 383–390.
28. Rogers, Y., and Lindley, S. Collaborating around vertical and horizontal large interactive displays: which way is best? *Interacting with Computers 16*, 6 (2004), 1133–1152.
29. Ryall, K., Morris, M., Everitt, K., Forlines, C., and Shen, C. Experiences with and observations of directtouch tabletops. In *Proceedings of IEEE TableTop the International Workshop on Horizontal Interactive Human Computer Systems* (2006), 89–96.
30. Short, J., Williams, E., and Christie, B. The social psychology of telecommunications.
31. Slater, M. Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences 364*, 1535 (2009), 3549.
32. Sodian, B., and Kristen, S. Theory of mind. *Towards a theory of thinking* (2010), 189–201.
33. Toney, A., and Thomas, B. Considering reach in tangible and table top design. In *Int. Workshop on Horizontal Interactive Human-Computer Systems* (2006).
34. Yankelovich, N., Levow, G., and Marx, M. Designing speechacts: Issues in speech user interfaces. In *Proc. of the SIGCHI conference on Human factors in computing systems* (1995), 369–376.
35. Yoshikawa, Y., Shinozawa, K., Ishiguro, H., Hagita, N., and Miyamoto, T. Responsive robot gaze to interaction partner. In *Robotics: Science and systems* (2006).