As Time Goes by: Long-term Evaluation of Social Presence in Robotic Companions

Iolanda Leite, Carlos Martinho, André Pereira and Ana Paiva

Abstract— **Given the recent advances in robot and synthetic character technology, many researchers are now focused on ways of establishing social relations between these agents and humans over long periods of time. Early studies have shown that the novelty effect of robots and agents quickly wears out and that people change their attitudes and preferences towards them over time. In this paper, we study the role of social presence in long-term human-robot interactions. We conducted a study where children played chess exercises with a social robot over a five week period. With this experiment, we identified possible key issues that should be considered when designing social robots for long-term interactions.**

I. Introduction

In the near future, we will live in a world populated with intelligent robots and virtual agents with the purpose of assisting us in everyday tasks. These agents usually have some human-like features in their embodiments and behave in ways people can understand and relate to. Since they will interact with us on a regular basis, they should be able to create and maintain social relationships with us so they can keep our engagement and become well accepted. However, existing work on artificial agents (with virtual or robotic embodiments) has mainly considered short-term interactions. After the novelty effect wears off, people usually lose interest and change their attitudes towards the agents [1, 2]. The main reasons behind this change are still unknown. Does the user's sense of presence, awareness or intimacy towards the agent fade away?

The term Social Presence was initially proposed by Short, Williams and Christie [3]. More recently, Biocca [4] proposed a definition of social presence oriented to human-computer interaction: "the amount of social presence is the degree to which a user feels access to the intelligence, intentions, and sensory impressions of another". Many studies regarding social presence are found in new forms of human-human communication such as computer conferencing [5]. But social presence is also used to measure individual's perception of a particular media, be it a virtual reality environment [6] or the interaction with a social robot [7]. Further, Lee and Nass argue that "social presence is the heart of all mediated vicarious experiences" and "people's social responses to media affect their feelings of social presence" [8].

The aim of our work is to find how one can keep the person's interest and motivation to interact with a social robot over long periods of time. Our hypothesis is that the feeling of social presence towards a particular agent motivates the user to maintain the interaction. So we argue that by evaluating user's perceived social presence over time, some indicators about what intelligent agents should have to engage users in long-term interactions can be retrieved. To better understand how social presence changes over time, we conducted a long-term experiment using "iCat, the Affective Chess Player", a system in which a social robot plays chess against a human opponent on a real chessboard. The results suggest that social presence decreases after five weeks of interaction, namely on three specific dimensions: attentional allocation, perceived affective and behavioural interdependence.

This paper is organized as follows. The next section contains a brief literature review on social robots and long-term interaction studies. Afterwards, we present the system that we used in the evaluation and describe the experiment. Finally, a discussion on the results is presented, followed by the conclusions and future work directions.

II. Related Work

The field of social robotics aims at developing intelligent robots that can communicate and interact with us, understand and even relate to us, in a personal way [9]. The application domains where social robots assist people is diverse, from health care [10] or tour guides [11], to robots whose sole purpose is to engage people in face to face interactions [12]. Although most part of these robots has been evaluated in practical scenarios with many users, only few experiments considered long-term interactions, i.e., the same user interacting with the system for several times.

One of the first long-term experiments with social robots was performed by Kanda et al. [1]. They performed a field trial evaluation for two weeks with elementary school Japanese students and two English-speaking interactive humanoid robots behaving as peer English tutors for children. The study revealed that the robot failed to keep most of the children's interest after the first week, mostly because the first impact created unreasonably high

Authors are with the IST – Technical University of Lisbon and INESC-ID, Av. Professor Cavaco Silva, 2780-990 Porto Salvo, Portugal (emails: iolanda.leite@tagus.ist.utl.pt, andre.pereira@tagus.ist.utl.pt, carlos.martinho@tagus.ist.utl.pt, ana.paiva@inesc-id.pt).

The research leading to these results has received funding from European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement nº 215554 and by two scholarships (SFRHBD/41358/2007 and SFRH/BD/41585/2007) granted by the "Fundação para a Ciência e Tecnologia".

expectations in the children. A longer study was carried out at Carnegie Mellon's University using Valerie, a "roboceptionist" [2]. Students and university visitors interacted with the robot over a nine month period. The results indicated that many visitors continued to interact daily with the robot, but over a certain period of time only few of them interacted for more than 30 seconds.

Some of the studies on long-term human-computer relationships are grounded on human social psychology theories, such as the work of Bickmore and Picard [13]. They developed a relational agent and evaluated it in a controlled experiment with approximately 100 users who were asked to interact daily with an exercise adoption system. After four weeks of interaction, the relational behaviors increased the participant's perceptions of the quality of the working alliance (on measures such as liking, trust and respect), when comparing the results with an agent without relational behaviors. Besides, participants interacting with the relational agent expressed significantly higher desire to continue interacting with the system.

The research on long-term interaction using artificial characters so far has been based on the analysis of the number of interactions, and how that number changes over time, given the assumption that people are free to interact (or not) with the agent. In this work we follow a different approach: by fixing the number of interactions (users interacted once a week during five consecutive weeks), we analyzed the differences on the user's mental model of the agent, particularly in terms of social presence.

III. iCat, the Affective chess player

The scenario that we used for the evaluation is composed by a social robot (Philip's iCat [14]), an electronic chessboard from DGT Projects [15] and a laptop where all the processing takes place (see Fig. 1).

The iCat acts as a game buddy, playing chess with children that already have some basic chess knowledge. The robot's affective behaviour is determined by the state of the chess game. It can display several facial expressions such as happy, sad or surprise, by moving its eyebrows, eyelids, mouth, neck and body.

A. Interaction model

The interaction starts with the iCat waking up. Afterwards, it invites the user to play. Users can play different chess exercises. After each user's move, the iCat performs an affective reaction to that move. Then, the robot asks the user to play its move in chess coordinates. When the user plays the move asked by the iCat, it sends a confirmation signal (a small utterance such as "ok, thank you" or a "nod" animation). If the user does not play the right move, there is also a set of "disapproval" animations and utterances. The game continues until one of the opponents checkmates the other.

B. System overview

The architecture behind the robot is separated in three different modules: game, emotion and animation. Following we provide a brief description of the modules (for more details please consult [16]).

1) Game module

The game module deals with the whole logic of the chess game. The chess engine was based on Tom Kerrigan's Simple Chess Program [17]. The engine is used not only to compute the moves of the iCat, but also to obtain a value that represents the evaluation of the game state after each user's move.

Fig. 1. System setup: a laptop, an electronic chessboard and the iCat robot.

2) Emotion Module

This module receives the evaluation values from the game module and updates the character's affective state. Inspired by Scherer's classification of affective states [18], the robot's affective state consists of two parts: *emotional reactions* and *mood*. Emotional reactions are triggered after every user's move, i.e., when a new board value is received from the game module. Despite being of short duration, they are quite explicit. On the other hand, *mood* represents a background affective state, less intense but always present.

Emotional reactions are often associated not only to the pure reaction to an event, but also take into account previous expectations built upon that event. Taking as an example a chess game, most people would feel much happier if they beat an opponent that they had never beat, than if they win against an opponent they won many times before. To model this kind of behaviour, we computed the emotional reactions using an anticipatory mechanism named *emotivector*. The emotivector is an anticipatory system that generates an affective signal resulting from the mismatch between the expected and the sensed values of the sensor to which it is coupled to [19]. The expected reward or punishment is computed, after each user's move, using the history of previous evaluation values, applying the moving averages prediction algorithm [20]. The same algorithm is used to calculate a confidence interval for the prediction, based on the history of previous errors.

For example, after three moves in the game, if the iCat has captured an opponent's piece (and the opponent has captured none), the iCat might be expecting a rewarding sensation on the user's next move. If the user then plays in agreement with what iCat was expecting (i.e., within the confidence interval), the elicited sensation would be an "expected reward", meaning "this is good, but I was already expecting that". But if the user plays a really bad move (e.g., by putting his/her queen in a very dangerous position), the elicited sensation would be a "stronger reward", which means "this was better than I was expecting".

Mood is represented as a valence variable that determines both the polarity (positive or negative) and the intensity of the mood. When the game starts the valence is neutral (i.e., zero). When a new evaluation value is received, it creates a new target value for the mood, towards which the valence gradually changes. Valence decays towards zero at a predefined decay rate in the absence of new moves played by the user.

3) *Animation Module*

The main role of the animation module is to manipulate the iCat's body parts to convey the affective states computed in the emotion module to the user. The iCat's embodiment can be animated by two different modalities: predefined animations (scripts containing a temporal sequence of a set of body parts and correspondent values) or direct manipulation (i.e., by setting the value of single body parts in real time).

The nine affective signals generated by the emotivector system were mapped into different predefined animations (*e.g.* sad, happy, surprise…). On the other hand, to convey mood, we used direct manipulation. Two different poses, each one corresponding to the limits of the valence space were defined, as depicted in Figure 2. The value of the body parts affected by mood is computed as an interpolation between those two body parameterizations.

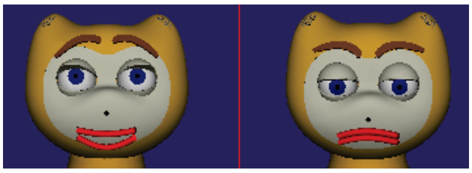

Figure 2. iCat's mood for extremes in the valence space (positive on the left, negative on the right).

The animation module also performs other predefined "idle" animations, like eye-blinking, looking to the sides randomly or nodding, to increase the believability and "life-like" appearance of the robot. Idle animations can only occur when no emotional reactions are being played, but mood coexists with them. In the presence of overlapping values, higher priority is given to emotional reactions, followed by mood and finally idle animations.

IV. LONG-TERM INTERACTION STUDY

From the studies we performed before, we realized that this scenario was well accepted by users. First [21], we evaluated the effects of the robot's affective behaviour in the user's perception of the game. The results indicated that the developed affective model enhanced the user's perception of the game. In another study [22], the iCat robot and a graphical version of the iCat were compared in terms of user's enjoyment. The experience was classified as more enjoyable by the users who played against the iCat physical robot.

Since most of the interactions in previous studies did not exceed one hour, we cannot claim that the user's attention to the robot was due to its behaviour rather than just a consequence of the novelty effect. In this study, we try to disambiguate those questions, and therefore our main objectives are: (1) to evaluate if user's perceived social presence changes over time; (2) identify which aspects of social presence are most affected.

A. *Method*

1) *Participants*

The experiment took place at a local chess club where every Saturday children between 5 and 15 years old take chess lessons from an instructor and play with each other. The class is composed of 7 children, even though we will only report the results of 5 of them (four males and one female) because the others missed more than one session.

Figure 3. User playing with the iCat.

None of the participants had interacted with the iCat or with any social robot before. Some of the younger subjects had limited reading comprehension.

2) *Procedure*

At the chess club, the robot, the electronic chessboard and the laptop were placed on a table. The subjects were seated in front of both the iCat and the chessboard like in a regular chess game, as Figure 3 shows.

A set of chess exercises was previously proposed to the chess instructor. He analyzed them and suggested modifications so that the difficulty of the exercises was adequate for each participant. In each session subjects played a different chess exercise against the iCat.

While the iCat was playing with one subject, the others would be watching the game or playing against each other, continuing their lessons. The idea was to integrate the robot in the group as one of their own. In this way, users were directly playing with the iCat and indirectly interacting with it during the remaining time.

The experiment was performed over five consecutive weeks. All the sessions were video recorded for further analysis. At the end of both the first and the last sessions children were asked to fill a questionnaire that measures social presence.

3) Measures

We measured social presence by questionnaires and by analyzing approximately a total of five hours of video. The social presence questionnaire was based on Harms and Biocca's [23] questionnaire, which conceptualizes social presence in six dimensions: **co-presence,** the degree to which the observer believes s/he is not alone; **attentional allocation**, the amount of attention the user allocates to and receives from an interactant; **perceived message understanding,** the ability of the user to understand the message from the interactant; **perceived affective understanding,** the user's ability to understand the interactant's emotional and attitudinal states; **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 **perceived behavioural interdependence,** the extent to which the user's behaviour affects and is affected by the interactant's behaviour. The social presence questionnaire was translated to the subjects' native language. We selected two items for each dimension that would be adequate for children (see Table 1). Subjects were asked to express their agreement or disagreement regarding each item on a five-point Likert scale (zero means "totally disagree" and five "totally agree").

The videos from the first, second and fifth week of five users were analyzed using ANVIL video annotation tool [24]. We annotated the parts of the video in which users were *looking at the iCat*, *looking sideways*, *talking to the iCat* and the *user's facial expressions*. We also distinguished the phase of the game in which users were looking at the iCat: *after the user's own move*, when the iCat performs an emotional reaction; *after playing the iCat's move*, when the user receives feedback from the robot, that confirms or disapproves his/her move, and while the user is *thinking*, when the iCat is performing idle behaviours (blinking and looking sideways) and its "face" reflects the mood.

B. Results and Discussion

1) Social Presence Questionnaire

According to the results in Table 1, in general, perceived social presence decreased after five weeks of interaction. Within the co-presence dimension, considering the Q1 means, there is slight evidence that users seem to notice the iCat less on the last week, which can be strengthened by the results of video observation. The amount of time that subjects looked at the iCat on the last interaction is lower than in the first ones. This may happen due to the novelty effect mentioned earlier, as none of the children had interacted with a social robot before. In spite of that, when asked if the iCat noticed them (Q2), all of them maintained their opinion. The turn-taking nature of the chess game may be the main cause for such result. Since the iCat reacts to children's moves and asks them to play its move, subjects may interpret those "reactive" behaviours as the robot noticing their presence.

Both the items regarding attentional allocation, Q3 and Q4, did not increase (on average decreased) after five weeks. From our observations at the chess club, when kids are playing with each other, they refer to previous games, explain theories behind certain moves and sometimes even make fun of each other. Some of these behaviours could be implemented in the iCat to increase the user's attention to the robot, especially the ones related to memory.

TABLE 1. SOCIAL PRESENCE QUESTIONNAIRE RESULTS

	1st week	5th week	↑	↔	↓
Co-Presence					
Q1. I noticed iCat.	4 [4;4]	3,75 [3;5]	1	1	2
Q2. iCat noticed me.	3,75 [2;5]	3,75 [2;5]	0	4	0
Attentional Allocation					
Q3. I remained focused on iCat.	3,5 [3;5]	2,75 [1;4]	0	2	2
Q4. iCat remained focused on me.	3,75 [3;5]	3,25 [2;4]	1	2	1
Perceived Message Understanding					
Q5. iCat's thoughts were clear to me.	3 [2;4]	3,25 [2;5]	2	1	1
Q6. My thoughts were clear to iCat.	3,25 [2;5]	2,75 [2;3]	0	3	1
Perceived Affective Understanding					
Q7. I could tell how iCat felt.	3 [2;4]	3 [1;4]	2	1	1
Q8. iCat could tell how I felt.	2,25 [1;3]	2,5 [1;4]	2	0	2
Perceived Affective Interdependence					
Q9. I was influenced by iCat's moods.	3,75 [3;5]	3 [2;4]	0	1	3
Q10. iCat was influenced by my moods.	3,5 [2;5]	2,75 [2;3]	1	1	2
Perceived Behavioural Interdependence					
Q11. My behaviour was tied to iCat's.	3,5 [2;5]	2,25 [1;4]	0	2	2
Q12. iCat's behaviour was tied to mine.	3,5 [2;4]	2 [1;4]	1	0	3

The table contains the means for the questionnaire items for the first and fifth week, with the corresponding minimum and maximum values between brackets. The last three columns contain the number of users who increased, maintained and decreased their ratings from the first to the fifth week.

In the perceived message understanding category, half the users claimed to know better what the iCat was thinking after the five week period (Q5). Even so, when they were asked if their thoughts were clear to the iCat, most of them maintained their position. This also happens for the perceived affective understanding dimension (Q7 and Q8). For instance, on the last weeks of interaction, when the iCat reacted sadly to a good move from the user, some of them talked to the robot: "I know you don't like that". Users felt that the iCat could not learn or adapt to their thoughts or affective states, which makes sense, given that the iCat does not have any mechanisms to adapt to different users.

The last two dimensions (perceived affective interdependence and perceived behavioural interdependence, from Q9 to Q12) were the ones whose means decreased the most after the long-term experiment. Over the weeks, iCat seems to be perceived much more as an automaton, behaving independently of how users feel or act, only reacting to their moves: subjects were expecting the iCat to behave more like a companion than a mere chess interface.

2) Video Observation

From all the aspects defined for annotation, *looking at the iCat* was the one with the largest number of annotations and also the one with more different results among different sessions. As such, it will be the main aspect of this discussion.

As one can see in Figure 4, the novelty effect apparently does not fade away immediately. On most subjects there are no substantial differences between the first and second weeks. However, the total time that subjects spent looking at the iCat on the last session is, on average, half the time they spent on the first one. These results are aligned with the ones obtained in the social presence questionnaire, especially with respect to the co-presence and the attentional allocation dimensions.

A detailed analysis of the average time that users were *looking at the iCat* in each phase of the game is shown in Figure 5. Even though the average time decreases in all phases, *after playing the iCat's move* and *thinking* were the phases that decreased the most over the weeks. In a*fter user's own move* phase, this fall was not so pronounced (between the first and second weeks the values remained roughly the same). Furthermore, in the questionnaire, subjects maintained their opinion regarding the understanding of iCat's behaviour (Q5 to Q8). One possible explanation is that, as the weeks go by, although subjects keep looking at the affective reaction performed by the iCat after their own move, they spend less and less time decoding its behaviour.

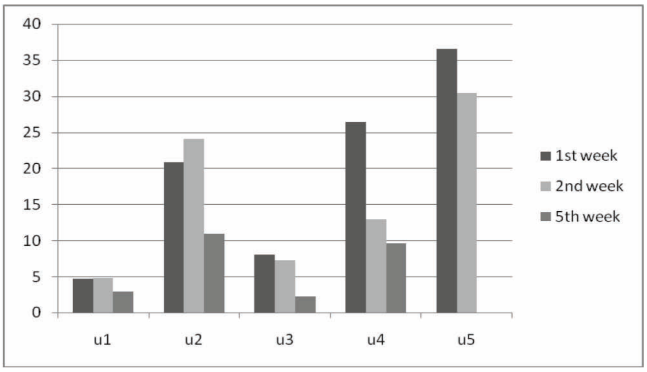

Figure 4. Total percentage of time[1] that each user spent looking at the iCat in the first, second and fifth weeks of interaction. U5 did not attend the chess club on the fifth week.

The remaining annotations were considerably more sporadic than *looking at the iCat.* For this reason, instead of presenting the quantitative results obtained from the annotation tool, we provide a brief qualitative discussion on each topic.

In the annotations for users *looking sideways*, no significant differences were found between the first and the last week of interaction. In most cases, users looked away from the iCat or the chessboard due to some external event at the chess club (e.g. someone arriving at the club). Still, in the last weeks some participants expressed signals of boredom after playing their move, while waiting for iCat's affective reaction and consequent move. Some of them looked away, but it was only for short periods of time.

Regarding *user talking to the iCat*, this annotation changed significantly among users and not so much over the weeks. Older subjects barely talked to the robot, but younger participants did. For example, when the iCat said "thank you for playing my move", young subjects replied "you're welcome". Over the weeks, younger subjects started talking to the iCat even when it was not their turn to play. On the first week this did not happened, probably because they were not so comfortable in the presence of the robot and the experimenters. It remains to be validated if this behaviour would continue over subsequent interactions.

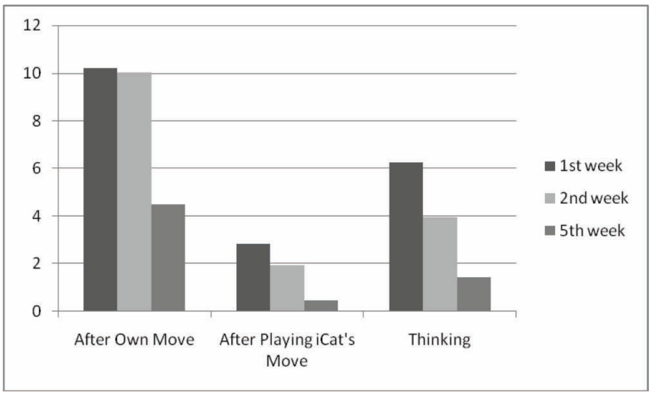

Figure 5. Average percentage of time[1] that users looked at the iCat in the first, second and fifth weeks, broken down by the phases of the game.

Concerning the *user's facial expressions*, we basically indentified two different types: the ones users displayed when they did not understand an iCat's affective reaction and the ones performed in the end of the game. The expressions users displayed to show misunderstanding about the iCat's reactions decreased over the weeks, which again indicates that over time the perceived message and affective understanding dimensions of social presence tend to improve (or at least remain the same). There were no substantial variations on the user's expressions in the end game though. Usually, when winning the game, users showed happy faces and when loosing they made a sad expression or showed no expression at all.

V. CONCLUSIONS AND FUTURE WORK

In this paper, we studied long-term interaction in terms of social presence. An experiment where users played chess with a social robot over a five week period was reported. Social presence was measured using a questionnaire and through video observation. The results of the questionnaire were reinforced by the analysis and annotation of 5 hours of video. The outcomes of the evaluation indicate that users' perception of social presence towards the iCat decreases after five weeks of interaction. We are aware that the results were obtained in a specific domain (a social robot in a chess

[1] Since all the exercises had different durations (from 5 to 20 minutes), we present the values as percentages of the total duration of the exercise, to be able to compare them between sessions.

game), and more experiments should be performed to see if the results can be generalized to other domains.

Our main contribution was the identification of the dimensions of social presence that decreased the most after five weeks of interaction. The questionnaire results indicate that attentional allocation, perceived affective and behavioural interdependence are the dimensions that decreased the most over time. These results were strengthened by video observation. Furthermore, the video analysis suggested that co-presence decreases over time as well. The identified dimensions are mainly related to robot's believability and user's attention to the system. We observed that the attention that users dedicate to the robot decreased significantly over the weeks, which suggests that new mechanisms and behaviours must be developed in order to maintain the engagement.

In summary, we concluded that the robot's current behaviour is not enough to create and maintain the perception of social presence after several interactions. Although it might appear believable and intelligent on the first impressions, *as time goes by*, users need more.

Several issues are still open for investigation. In the future, we intend to evaluate also the videos from the third and fourth weeks, in the attempt to understand if social presence decreases gradually or if there is a big collapse between two particular weeks, and if such decrease varies between different users or not. We also plan to implement new behaviours related to memory in the iCat, and perform a new evaluation. Examples of new behaviours that could sustain user's perceived social presence over the weeks would be the iCat being able to identify the users, remembering their names and the results of previous games they played together, and refer that during the game. Another possibility is the ability to conduct small talk [25], especially related to chess topics, which could increase the character's believability and credibility towards the user.

ACKNOWLEDGMENT

We thank "Clube de Xadrez da Mata de Benfica", the chess club where the experiments were conducted.

REFERENCES

[1] T. Kanda, T. Hirano, and E. Daniel, "Interactive Robots as Social Partners and Peer Tutors for Children: A Field Trial," *Human-Computer Interaction,* vol. 19, pp. 61-84, 2004.

[2] R. Gockley, A. Bruce, J. Forlizzi, M. Michalowski, A. Mundell, S. Rosenthal, B. Sellner, R. Simmons, K. Snipes, A. C. Schultz, and J. Wang, "Designing robots for long-term social interaction," in *IEEE/RSJ International Conference on Intelligent Robots and Systems - IROS'05*, 2005, pp. 1338-1343.

[3] J. A. Short, E. Williams, and B. Christie, *The social psychology of telecommunications*. New York: John Wiley & Sons, 1976.

[4] F. Biocca, "The cyborg's dilemma: Progressive embodiment in virtual environments," *Journal of Computer-Mediated-Communication,* vol. 3(2), 1997.

[5] L. Rourke, T. Anderson, D. R. Garrison, and W. Archer, "Assessing Social Presence In Asynchronous Text-based Computer Conferencing," *Journal of Distance Education,* vol. 14, pp. 51-70, 1999.

[6] C. Heeter, "Being there: the subjective experience of presence," *Journal of Presence: Teleoperation and Virtual Environments,* vol. 1, pp. 262-271, 1992.

[7] P. Schermerhorn, M. Scheutz, and C. R. Crowell, "Robot social presence and gender: do females view robots differently than males?," in *Proceedings of the 3rd ACM/IEEE international conference on Human robot interaction*, Amsterdam, The Netherlands, 2008, pp. 263-270.

[8] K.-M. Lee and C. Nass, "Social Psychological Origins of Feelings of Presence: Creating Social Presence With Machine-Generated Voices " *Media Psychology,* vol. 7, pp. 31-45, 2005.

[9] C. Breazeal, "Socially Intelligent Robots," *Interactions, ACM,* vol. 12, pp. 19-22, 2005.

[10] M. Montemerlo, J. Pineau, N. Roy, S. Thrun, and V. Verma, "Experiences with a Mobile Robotic Guide for the Elderly," in *Proceedings of the AAAI National Conference on Artificial Intelligence*, Edmonton, Alberta, Canada, 2002, pp. 587-592

[11] S. Thrun, M. Bennewitz, W. Burgard, F. Dellaert, D. Fox, D. Haehnel, C. Rosenberg, N. Roy, J. Schulte, and D. Schulz, "MINERVA: A second generation mobile tour-guide robot," in *IEEE International Conference on Robotics and Automation-ICRA* Detroit, 1999.

[12] C. Breazeal, "Emotion and Sociable Humanoid Robots," *International Journal of Human Computer Studies,* vol. 59, pp. 119-155, 2003.

[13] T. Bickmore and R. Picard, "Establishing and maintaining long-term human-computer relationships," *ACM Transactions on Computer-Human Interaction,* vol. 12, pp. 293-327, 2005.

[14] A. Breemen, X. Yan, and B. Meerbeek, "iCat: an animated user-interface robot with personality," in *Autonomous Agents and Multiagent Systems, AAMAS'05*, 2005, pp. 143-144.

[15] DGT Projects - Electronic Chessboard. http://www.dgtprojects.com/eboard.htm.

[16] I. Leite, A. Pereira, C. Martinho, and A. Paiva, "Are Emotional Robots More Fun to Play With?," in *IEEE RO-MAN 2008* Munich, Germany, 2008.

[17] Tom Kerrigan's Simple Chess Program. http://home.comcast.net/~tckerrigan/.

[18] K. R. Scherer, "Psychological models of emotion," in *The neuropsychology of emotion*, J. Borod, Ed. New York: Oxford University Press, 2000, pp. 137-162.

[19] C. Martinho and A. Paiva, "Using Anticipation to Create Believable Behaviour," in *Proceedings of the 21st National Conference on Artificial Intelligence and the 18th Innovative Applications of Artificial Intelligence Conference*, Boston MA, USA, 2006, pp. 175-180.

[20] NIST/SEMATECH e-Handbook of Statistical Methods. http://www.itl.nist.gov/div898/handbook/.

[21] I. Leite, C. Martinho, A. Pereira, and A. Paiva, "iCat: an affective game buddy based on anticipatory mechanisms," in *Proc. of 7th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2008)*, Estoril, Portugal, 2008, pp. 1229-1232.

[22] A. Pereira, C. Martinho, I. Leite, and A. Paiva, "iCat, the chess player: the influence of embodiment in the enjoyment of a game," in *Proc. of 7th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2008)*, Estoril, Portugal, 2008, pp. 1229-1232.

[23] C. Harms and F. Biocca, "Internal Consistency and Reliability of the Networked Minds - Measure of Social Presence," in *Seventh Annual International Workshop: Presence 2004*, Valencia, Universidad Politecnica de Valencia, 2004.

[24] M. Kipp, "Spatiotemporal Coding in ANVIL," in *Proceedings of the 6th international conference on Language Resources and Evaluation (LREC-08)*, 2008.

[25] T. Bickmore and J. Cassell, "Small Talk and Conversational Storytelling in Embodied Interface Agents," in *Proceedings of the AAAI Fall Symposium on Narrative Intelligence*, 1999, pp. 87-92.