Benefits of Social Intelligence in Home Dialogue Systems

Privender Saini[1], Boris de Ruyter[2], Panos Markopoulos[3], and Albert van Breemen[2]

[1,3] User System Interaction Programme/ Industrial Design, Eindhoven University of
Technology, Den Dolech 2, 5600 MB Eindhoven, The Netherlands
{p.k.saini, p.markopoulos}@tue.nl
[2] Philips Research, Media Interaction, Prof.Holstlaan 4 (WY 2.01),
5656 AA Eindhoven, Netherlands
{boris.de.ruyter, albert.van.breemen}@philips.com

Abstract. This paper reports an exploration of the concept of social intelligence in the context of home dialogue systems for an Ambient Intelligence home. It reports a Wizard of Oz experiment involving a robotic interface capable of displaying several human social behaviors. Our results show that endowing a home dialogue system with some social intelligence can (a) create a positive bias in user's perception of technology in the environment, (b) increase user acceptance for the home dialogue system, and (c) trigger social behaviors of the user towards the home dialogue system.

1   Introduction

A defining characteristic of the vision of Ambient Intelligence is that humans will be confronted with an ever-increasing population of devices and applications embedded in our environments and in our everyday activities. There appear to be two prevalent and probably complementary views upon how humans will interact with such environments [14]: one is that interaction will be through a multitude of task specific information appliances [16]; the other is through a centralized user interface anticipating user needs through adaptivity and intelligence [17]. In the home domain, *home dialogue systems* are expected to fulfill this latter role by acting as intermediaries between systems embedded in the home and the people within them.

   This research examines the effect of endowing *social intelligence* to a home dialogue system; it sets out to verify whether designing such a system so that it is perceived as socially intelligent will lead to a more positive experience for the users. In particular we expect that users will not only have higher acceptance of the home dialogue system, but also of the technology embedded within the home environment.

2   Social Intelligence and User Experience

In its widest definition social intelligence is a person's ability to "…get along with people in general, social technique or ease in society, knowledge of social matters, susceptibility to stimuli from other members of a group as well as insight into the temporary moods and underlying personality traits of strangers." [25].

There is a large body of psychological literature on Social Intelligence. However, there does not seem to be a single list of characteristics or behaviors that constitute Social Intelligence. An important reason for this is that Social Intelligence is manifested in many different ways according to the context at hand. Socially intelligent behaviors may range from being nice and pleasant to interact with, admitting mistakes, displaying curiosity, to being able to read non-verbal cues of interlocutors, etc. This complex nature of social intelligence has two consequences. First, it means that there are multiple ways in which we can try to make a system exhibit social intelligence. Second, that it is not straightforward to assess whether a particular system or behavior is indeed socially intelligent.

In recent years, research into computational and robotic characters has demonstrated an increasing interest into the topic of social intelligence. Nevertheless, the benefits that social intelligence might offer to users are to this point only hypothesized. It is well known that people have a tendency to attribute human like properties to interactive systems [18]. Various works have shown that endowing on-screen computational characters with human-like behaviors, e.g., from eye-gaze to emotional facial expressions and from hand gestures to full body posture adjustments, may lead to enhancing the likeability of the character, trust in the character, satisfaction with interaction, naturalness of interaction, ease of use, efficiency of task completion, and closeness to human characteristics [4, 11, 22]. Similar effects have been found in research on social robots. Effects like the appropriateness of movements [19], joint attention [12], attention getting and cooperation elicitation from strangers [5], and invested effort in interaction task [3] have been measured, with the results in favor of the socially expressive robot. Such piecemeal results show a constant pattern in favor of the more 'colorful' and 'human-like' behaviors of the computational character.

Recent research has focused on examining the effect of singular factors conducive to social intelligence. This study examines what broader benefits could be brought upon the interaction experience by a more socially rich and coherent home dialogue system that is perceived to be socially intelligent. Such a holistic examination that would show the relevance and importance of social intelligence in the domain of human computer interaction has not been attempted before.

In this paper we report a controlled study into the effects of perceived social intelligence in a home dialogue system addressing the following questions:

- Will test participants be able to perceive the level of social intelligence implemented in the home dialogue system?
- What is the effect of bringing the concept of social intelligence into a home dialogue system on the perception of quality of the interactive systems (other than the home dialogue system) in the environment?
- Will the participant's acceptance for home dialogue systems increase if the concept of social intelligence is implemented into these systems?

In the following sections we present, a home dialogue system using a robotic interface, the iCat. We describe an experiment addressing the research questions raised above. The results of this experiment are discussed leading to conclusions in the final section.

3   The iCat as an Embodied Home Dialogue System

The home dialogue system used in our study takes the form of an "interactive Cat", or just iCat. The iCat is a research platform for studying social robotic user-interfaces. It is a 38 cm tall user-interface robot that is stationary, i.e. it has no mobility facilities (see figure 1). The robot's head is equipped with 13 standard R/C servos that control different parts of its face, such as the eyebrows, eyes, eyelids, mouth and head position. With this setup we are able to generate many different facial expressions that are needed to create an emotionally expressive character.
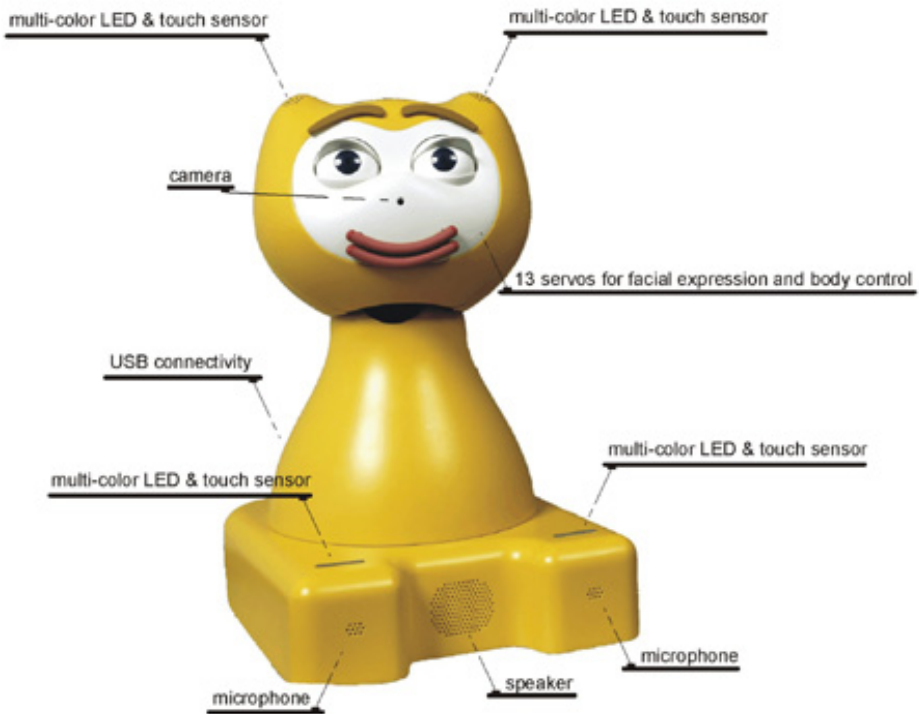

Fig. 1. Hardware setup of the home dialogue system "iCat"

A camera installed in the head of the iCat supports a range of computer vision capabilities, such as recognizing objects and faces. The iCat's foot contains two microphones to record the sounds it hears, perform speech recognition, and determine the direction of the sound source. By determining the direction of a sound source, the iCat can exhibit 'turn-to-speaker' behavior. Also, a loudspeaker is installed to play sounds and generated speech. Furthermore, the iCat is connected to an in-home network to control devices, e.g., light, VCR, TV, radio, and to access the Internet. Finally, touch sensors and multi-color LEDs are installed in the feet and ears to sense whether the user touches the robot and to communicate further information

encoded by colored light. For instance, the operation mode of the iCat (e.g. sleeping, awake, busy, listening, etc.) is encoded by the color of the LEDs in the ears.

4   Experiment

4.1   Participants

The experiment involved 36 paid participants, 14 women and 22 men, $\underline{M}$ = 28.7. 80 % of the people were either attending or had graduated from university. The rest had at least attended high school. Of these 20% all had office jobs. All participants were selected to have at least some basic experience with E-mail and Internet. Participants were recruited through an external agency.

4.2   Setting

The experiment took place at HomeLab, a dedicated laboratory that simulates a home, but which is equipped with an extensive observational infrastructure that allows testing of innovative technologies at home at a naturalistic setting [9].

Participants would be left with the iCat in the living room of HomeLab, while the experimenter would observe and control the experiment from the observation station of the HomeLab.

4.3   Design

We adopted a one-factor between-subjects design in which social intelligence was manipulated. There were two conditions. Participants were randomly assigned to one of the conditions. There were 18 participants in each condition.

Condition 1 can be called the 'social intelligence' condition. During this condition the robot talked (using synthesized speech) with lip synchronization. It blinked its eyes throughout the session and displayed facial expressions while exhibiting the following selected social intelligence aspects:

- *Listening attentively*: by looking at the participant when s/he is talking and occasional nodding of the head.
- *Being able to use non-verbal cues the participant displays*: responding verbally to repeated wrong actions of the participant by offering help.
- *Assessing well the relevance of information to a problem at hand*: by stating what went wrong, before offering the correct procedure to the user
- *Being nice and pleasant to interact with*: by staying polite, mimicking facial expressions (smile when participant smiles for example), being helpful.
- *Not ignoring affective signals from the user*: by responding verbally or by displaying appropriate facial expressions to obvious frustration, confusion, or contentment.
- *Displaying interest in the immediate environment*: the immediate environment being the participant and the equipment used in tasks, by carefully monitoring the person and the progress of the tasks.

- *Knowing the rules of etiquette*: by not interrupting the participant when s/he is talking.
- *Remembering little personal details about people*: addressing the participant by name, remembering login information, and passwords if asked.
- *Admitting mistakes*: by apologizing when something has gone wrong, but also when no help can be provided upon participant's request.
- *Being expressive*: by showing facial expressions while talking, if appropriate.
- *Thinking before speaking and doing*: by showing signs of thinking (with facial expression) before answering questions or fulfilling the participant's request.

Fig. 2. Snapshot from the test sessions showing respectively: the participant (LIV1), an overview of the living room (LIV2 and LIV3) and a close up of the iCat

The behaviors for the social intelligence condition were available as pre-programmed blocks. In the socially intelligent condition the experimenter would observe and listen to the participant and would type in responses for the iCat to utter. Further, the experimenter would initiate (press of a button) these pre-programmed social behaviors at appropriate moments in a Wizard of Oz fashion.

Condition 2 was the 'socially neutral' condition. The iCat did not display any facial expressions and did not blink its eyes. It talked and used lip-synchronization, but the

talking did not support the above listed aspects of social intelligence. It responded verbally only to explicit questions from the participant. The only self-initiated help was when the participants really got stuck (when after 12 minutes no progress was made) and could not continue without help.

We underline that contrary to studies listed in the previous section, we did not seek to assess the impact of each of the low level behaviors listed here. Rather, we hoped that their combination would lead the iCat to be perceived as socially intelligent and it is the impact of this perception that we aimed to assess.

4.4 Tasks

Participants were asked to perform 2 tasks. The first task was to program a DVD recorder to record 3 broadcast shows for the upcoming week. The iCat would sit on the table while participants performed the task. Participants could elicit help from the iCat by talking to it. The time allotted for this task was 10 minutes.

The second task was an online auction. An auction website was built for this purpose. The task was to register with the auction site and buy several items on a list. Participants could bid on the items, but that did not mean that they had bought them. Other people (in reality the experimenters) could out-bid them. For registration as a new user, the site required the participants' E-mail address (should be web-accessible). The E-mail was used in the experiment to notify participants when other people in the auction (in reality the experimenters) had outbid them on the listed items. In order to avoid having to check their E-mail for these outbid messages, participants could give iCat their E-mail details (login and password) if they wanted iCat to monitor their bid-on items. In other words: in asking iCat's help, participants would also need to entrust it with their E-mail password. The participants were allowed 20 minutes for this task.

4.5 Measures

A set of multiple measures was designed to test both the direct effects of the iCat's behaviors and the potential implicit "spillover" effects like satisfaction with the DVD recorder.

Social Behaviors Questionnaire (SBQ). In the absence of existing validated instruments to assess social intelligence in interactive systems a questionnaire was developed for the purpose of this study. Its purpose was to verify whether we succeeded in creating two separate conditions that the participants would rate differently in terms of social intelligence. The questionnaire (described in a separate publication) was built up of 5-point scales rating the degree of agreement (agree, somewhat agree, not agree or disagree, somewhat disagree, disagree) of participants to statements like:

- The robotic cat takes others' interests into account.
- The robotic cat does not see the consequences of things.
- The robotic cat says inappropriate things.
- The robotic cat is not interested in others' problems.
- The robotic cat tells the truth.

User Satisfaction Questionnaire (USQ). The USQ is an instrument developed previously in-house for assessing user satisfaction with consumer products [8]. It is a 54-item questionnaire that requires answers on a 5-point rating scale (agree to disagree). The USQ was used to assess the satisfaction with a DVD recorder that participants had to operate during the experiment.

The Unified Theory of Acceptance and the Use of Technology (UTAUT). [24] The UTAUT is a measure of technology acceptance. It was used with some adaptations for the home domain, to measure the extent to which participants expected to use iCat at home, if that were a possibility. It also requires participants to rate each item on a 5-point rating scale (agree to disagree).

In a post-experimental interview participants were asked some questions about what they thought abut their own performance on the auction task. The answers were scored by the experimenter on 5-point scale ranging from not happy to happy about performance.

Finally we recorded the number of times participants asked the robot general questions and the number of times they asked task-related questions. We also recorded the number of times that participants looked at the robot during the entire session.

4.6 Procedure

Participants were welcomed and it was explained that they were going to do two tasks. They were also told that while they did those tasks there would be a robot cat on the table that could be addressed if they needed or wanted its help. There would be times that the robot would initiate conversation when it thought it might be able to help. The instruction was such that emphasis was placed on completing the tasks, whereas interaction with iCat was secondary.

While participants did the tasks iCat was always sitting on the table facing them, available for input. When the opportunity arose iCat offered to monitor their E-mail account for outbids if the participants authorized it by giving it their password. The iCat was there to help in many other ways as well. If, for example, participants could not manage to register as a new user, the iCat could register on their behalf. This was done in both conditions when participants would not succeed in registering within 12 minutes. The iCat could also give information on the items that were offered in the auction. If authorized it could place bids for the participants.

After the experimental session the participants were taken to a separate room. They gave their first impressions of the experiment in an interview and filled in the questionnaires.

5 Results

Questionnaires. In light of the number of participants and the questionnaires having ordinal rating scales, Kruskal-Wallis one-way analyses of variance were performed to see whether responses on the three questionnaires differed from each other in the two conditions. An $\alpha = 0.05$ was set for all analyses.

SBQ. The difference between conditions for the SBQ was significant (χ^2 = 5.938, df = 1, p < 0.05). Inspecting the means (see table 1) indicated that our hypothesis was confirmed. Participants thought the socially intelligent iCat was indeed socially more intelligent than the neutral iCat.

USQ. The difference in user satisfaction with the DVD recorder was also significant (χ^2 = 4.294, df = 1, p < 0.05). Participants, who interacted with the socially intelligent iCat, were more satisfied with the DVD recorder.

UTAUT. The responses on the UTAUT also resulted in a significant difference (χ^2 = 9.633, df = 1, p < 0.05). Participants who worked with the neutral iCat were less inclined to want to continue working with iCat at home.

Table 1. Questionnaire results

Questionnaire	Socially intelligent	Neutral
SBQ	1.98	2.34
USQ	1.92	2.49
UTAUT	1.71	2.33

* Lower ratings mean socially more intelligent, more satisfied, and more likely to use technology.

Interviews. From the semi-structured interview it became clear that there is no difference between the two conditions on how participants evaluated the way they performed on the auction. On a scale from 1 (not happy about performance) to 5 (very happy about performance), the average for both conditions was 3.8 (χ^2 = 0.170, df = 1, p > 0.05). In response to the question: "*If you had iCat at home, what would you like it to do for you?*" many participants mentioned things like operating all their electronics and electrical appliances (such as lights, home heating system, household appliances, and home entertainment equipment). Many mentioned more privacy-sensitive tasks like having their E-mail checked for them, screening telephone calls, and Internet banking. The number of people that mentioned this was equal for both conditions.

However, when probed deeper there was a difference in the constraints posed before authorizing iCat to access personal information. Out of the 11 participants in the socially intelligent condition that would like iCat to handle their private tasks, six would like some user research data that tells them that it is safe and secure to use iCat for such tasks. The other five would use iCat without further evidence.

Participants in the neutral condition felt differently. There were 12 participants that would authorize iCat for personal tasks. 4 of them would like research evidence before using iCat. Only one participant would use iCat as is. But 7 participants were not sure if research was enough: they wanted to have further experience with iCat before allowing it more extensive access to their data. They stated that they would first give it small tasks. They would grant it full authorization only over the longer term and after proven success.

From the 36 participants, only 2 suspected the experiment was a Wizard-of-Oz.

Observations. Contrary to our expectations, there were no significant differences in the query behavior of subjects in the two conditions. The Wilcoxon-Mann-Whitney test was used to analyze the data. Participants asked questions to the iCat 13.6 times in the socially intelligent versus 11.1 in the neutral condition (Z = -0.954, p > 0.05). On average 4.9 questions were posed about items in the auction in the socially intelligent condition versus 3.2 in the neutral condition (Z = -0.486, p > 0.05). Participants looked at the iCat 11.6 times (average) in the social intelligence condition and 6.0 in the neutral condition (Z = -1.134, p > 0.05). In many cases participants looked at the robot in anticipation of an answer.

6 Discussion

The results from the SBQ confirm the distinctness of the experimental conditions that we wanted to create: Participants rated the socially intelligent iCat as more social than the neutral one. This corroborates the selection and the design of behaviors of the iCat to exhibit social intelligence.

The USQ also had a differential effect between the two conditions. Since the USQ was developed to test satisfaction with a consumer product after thorough interaction with that product and the DVD recorder task only consisted of exploring one function in a time frame of 10 minutes, the significant difference found between the two experimental conditions is remarkable.

The UTAUT was applied to the iCat and shows an explicit positive effect of the social intelligence manipulation.

There was no significant effect regarding the perceived auction performance; most participants in both conditions thought they did pretty well. The task of buying items was for most of them not a hard one. As such, many of them felt they did very well. Participants would have liked to delegate more chores to iCat and to ask it more questions. Most participants asked the iCat to monitor their items for bids from others (83% of participants). The only participants who were not very satisfied with how they had performed were those who in their daily lives do not spend much time on the web or on the computer.

Overall our impressions were that participants were more 'social' with the socially intelligent iCat. They were much more inclined to laugh, and (make small) talk, than with the neutral iCat. People generally talked more and more elaborately in the social intelligence condition. They were also more curious about the reasons the social robot said the things it said. For example, when asked which LCD monitor was a good one (to buy in the auction task) they were content with an answer. But in the socially intelligent condition they were also curious about how it knew this and why it was the best. They were also more inclined to ask about the iCat's opinion regarding other LCD monitors. They asked these questions politely and using full sentences. In the case of the neutral iCat, they were more inclined to take the suggestion of the best LCD monitor for what it was and not continue asking further. In cases that they did ask more, it was usually in shorter and more to-the-point command-like sentences than it was in the social condition.

Participants in the socially intelligent condition liked the fact that the robot was expressive in terms of facial expressions, that it nodded and shook its head in

response to their talking. Overall they agreed that it was only natural for the iCat to use its potential this way. However, participants in the neutral condition also liked iCat with its more neutral behavior. After all, they argued, it is a robot and it should not try or pretend to be anything other than that. Moving and facial expressions would only look like a poor attempt to seem alive and it would likely annoy and distract from whatever you are doing. This finding shows how hard it can be to imagine something you have not experienced. Neither group of participants could imagine the iCat being different than the way they had experienced it. This has important implications for future user-system interaction research. It is simply not helpful to just ask people what they like.

7   Conclusion

Participants in this study had the experience of a robot that could 'see' what the participant was doing, whether this was a task on the TV and DVD recorder or on the Internet. Although the participants were well aware of the level of invasion technology like this would have on their lives – should it be allowed into their homes – many of them welcomed this fact. Our study has shown that a few social behaviors in a robot may be sufficient to remove some discomfort that is brought about as domestic environments become richer in technology. One plausible explanation is that participants are more inclined to accept a single centralized interface for an Ambient Intelligence environment rather than a distributed set of separate products. Adding some thoughtful implementations of social intelligence to a perceptive robot that takes this role can make the robot easier to communicate with and more trusted by users.

Whereas most research on social robotic characters concerns the interaction with the robot as the focus of attention, this study has concentrated on the role of a robot as an intermediary, a home dialogue system. The interaction with the iCat was not the participants' priority. Despite its background function, the iCat and the behaviors it displayed had significant effects on satisfaction with the systems embedded in the home environment, their acceptance by users, and the sociability users exhibit towards the system. It was shown that social intelligence is not just important in the context of direct interaction with robotic or even screen characters, but it has relevance in systems that do not necessarily have a social function.

Concerning research in social intelligence of on-screen and robotic characters, this study has made a three-fold contribution. First we have shown how a collection of human-like behaviors can lead the character to be perceived as socially intelligent. Second, by means of the social behaviors questionnaire we developed, we have provided the means with which this perception can be evaluated. Third we have demonstrated the relevance of social intelligence as a concept for studying interaction between humans and computational characters. Researchers often assume this relevance implicitly; by showing how an increase in perceived social intelligence impacts people's perception positively of a system we can substantiate this assumption. On the basis of our results, it is possible for future research to explore the most effective ways to achieve social intelligence and to forge the links between lower level behaviors of robots or on-screen characters and the resulting experience of social intelligence.

References

1. Aarts, E, and Marzano, S (Editors) (2003), *The New Everyday: Vision on Ambient Intelligence*, 010 Publishers, Rotterdam, The Netherlands.
2. Banham, K. M. (1968). *Social competence inventory for adults : a social competence inventory for older persons*. Durham : Family Life Publications.
3. Bartneck, C. (2003). Interacting with an embodied emotional Character, *Proceedings of the DPPI2003 Conference*, Pittsburgh.
4. Bickmore, T., & Cassell, J. (2001). Relational agents: A model and implementation of building user trust. Proceedings. ACM CHI 2001, Seattle, Washington, 396-403.
5. Bruce, A., Nourbakhsh, I. R., &. Simmons, R. G. (2002) The Role of Expressiveness and Attention in Human-Robot Interaction. *ICRA 2002*, 4138-4142.
6. Cañamero, L., & Fredslund, J. (2001). I Show You How I Like You – Can You Read it in My Face? IEEE Transactions on Systems, *Man and Cybernetics – Part A: Systems and Humans, 31(5),* 454-459.
7. De Ruyter, B. & Aarts, E. (2004). Ambient Intelligence: visualising the future, *AVI 2004 - Advanced Visual Interfaces*, May 25-28, Gallipoli, Italy.
8. De Ruyter, B. & Hollemans, G. (1997). *Towards a User Satisfaction Questionnaire for Consumer Electronics: Theoretical Basis*. Eindhoven: Natuurkundig Laboratorium Philips Electronics N. V., NL – TN, 406/97.
9. De Ruyter, B., Markopoulos, P., Aarts, E., IJsselsteijn, W. (2004). Engineering the user experience, In: W. Weber, J. Rabaey, and E. Aarts (Eds): *Ambient Intelligence.* Springer-Verlag.
10. Fong, T., Nourbakhsh, I., & Dautenhahn, K. (2003). A survey of socially interactive robots. *Robotics and Autonomous Systems,* 42, 143-166.
11. Heylen, D., Van Es, I., Van Dijk, E. M. A. G. & Nijholt, A. (2004). Experimenting with the Gaze of a Conversational Agent. In J. van Kuppevelt, L. Dybkjaer and N.O. Bernsen (Eds.) *Natural, Intelligent and Effective Interaction in Multimodal Dialogue Systems.* Dordrecht: Kluwer Academic Publishers.
12. Imai, M., Ono, T., & Ishiguro, H. (2001) Physical relation and expression: joint attention for human-robot interaction. *Robot and Human Interactive Communication, 2001.* Proceedings. 512-517.
13. Keating, D.K. (1978). A search for social intelligence. *Journal of Educational Psychology*, 70, 218-233.
14. Markopoulos, P., (2004). Designing ubiquitous computer human interaction: the case of the connected family. In Isomaki, H., Pirhonen, A., Roast, C., Saariluoma, P., (Eds.) *Future Interaction Design*, Springer.
15. Moos, F. A., Hunt, K. T., Omwake, K. T. & Ronning, M. M. (1927). *Social Intelligence Test*. Washington, D.C.: The Center for Psychological Service.
16. Norman, D. A. (1998). The Invisible Computer. Cambridge,MA: MIT Press.
17. Pentland, A. (2000). Perceptual Intelligence. *Communications of the ACM*, 43(3), 35-44.
18. Reeves, B., and Nass, C. (1996). *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places*, Cambridge: Cambridge University Press.
19. Sidner, C. L., Kidd, C. D., Lee, C. H., & Lesh, N. B. (2004). Where to Look: A Study of Human-Robot Engagement*, ACM International Conference on Intelligent User Interfaces (IUI),* pp. 78-84, January 2004.
20. Strang, R., Brown, M. A., & Stratton, D. C. (1942). *Test of knowledge of social usage*, New York: Columbia University.

21. Sternberg, R.J., & Smith, C. (1985). Social intelligence and decoding skills in nonverbal communication, *Social Cognition*, 3, 168-192.

22. Thórisson, K. R. (1997). Gandalf: An Embodied Humanoid Capable of Real-Time Multimodal Dialogue with People. First ACM International Conference on Autonomous Agents, Marina del Rey, California, February 5-8, 536-7.

23. Van Breemen, A.J.N. (2004a). Bringing Robots to Life: Applying principles of animation to robots, *CHI2004 workshop Shaping Human-Robot Interaction*, Vienna.

24. Venkatesh, V,. Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User Acceptance of Information Technology: Towards a Unified View. *MIS Quaterly, 27(3)*, 425-478.

25. Vernon, P.E. (1933). Some characteristics of the good judge of personality. *Journal of Social Psychology*, 4, 42-57.