

Two acts of social intelligence: the effects of mimicry and social praise on the evaluation of an artificial agent

Maurits Kaptein · Panos Markopoulos ·
Boris de Ruyter · Emile Aarts

Received: 26 February 2010 / Accepted: 9 September 2010 / Published online: 9 December 2010
© The Author(s) 2010. This article is published with open access at Springerlink.com

Abstract This paper describes a study of the effects of two acts of social intelligence, namely mimicry and social praise, when used by an artificial social agent. An experiment ($N = 50$) is described which shows that social praise—positive feedback about the ongoing conversation—increases the perceived friendliness of a chat-robot. Mimicry—displaying matching behavior—enhances the perceived intelligence of the robot. We advise designers to incorporate both mimicry and social praise when their system needs to function as a social actor. Different ways of implementing mimicry and praise by artificial social actors in an ambient persuasive scenario are discussed.

Keywords Ambient Intelligence · Attitude change · Persuasion · Friendliness · Chat-robot · Mimicry · Praise · Social intelligence

1 Introduction

You are woken up by an early alarm on a Sunday morning at 06:30. Your digital wellness assistant alerts you that you have not reached your target exercise level for this week, while all of your remote training mates have. Since your agenda is packed from 08.30, your wellness assistant has decided this would be a good time to exercise. On some Sundays you disagree...

While you are aware that training keeps you fit, and while for the common good it would be better if more people complied with such advice since it would potentially decrease the prevalence of obesity, coronary disease, and diabetes (Schroeder 2007), we all know this is tough on Sunday mornings.

In recent years, there is an increasing focus on developing persuasive technologies: Technologies intentionally designed to influence users' behavior (Fogg 2003). These systems most typically target behaviors generally agreed upon as being positive: exercising more, smoking less, or complying with drug subscriptions (e.g., Räsänen et al. 2008; Maheshwari et al. 2008).

The digital wellness assistant referred to in the scenario above represents a persuasive technology that embodies some of the defining characteristics of the *Ambient Intelligence scenario*: Embedded applications sense user's context and activities and provide personalized advice (Cialdini 2001). In the AmI world, distributed technologies are omnipresent and applications intelligently make sense of both user behaviors as well as the user context. Nevertheless, as suggested by Markopoulos et al. in (2005), research into Ambient Intelligence will greatly benefit by extending its scope beyond analytical intelligence into what Thorndike already in 1920 called social intelligence (Thorndike 1920).

M. Kaptein (✉) · P. Markopoulos
Eindhoven University of Technology, Den Dolech 2,
HG 2.38, 5600 MB Eindhoven, The Netherlands
e-mail: m.c.kaptein@tue.nl

P. Markopoulos
e-mail: p.markopoulos@tue.nl

B. de Ruyter
Philips Research The Netherlands,
High Tech Campus 34 -WB-4.053,
5656 AE Eindhoven, The Netherlands
e-mail: boris.de.ruyter@philips.com

E. Aarts
Philips Research The Netherlands,
High Tech Campus 34 -WB-7.051,
5656 AE Eindhoven, The Netherlands
e-mail: emile.aarts@philips.com

When creating systems like the *digital wellness assistant* both researchers and designers can benefit from the notion that computers function as social actors. Hence, when interacting with computers or digital agents people are inclined to treat them like they treat real people (Reeves and Nass 1996). This opens the door for persuasive technologies to use the same social influence strategies human persuaders do.

The current study is an investigation of the effectiveness of utilizing social intelligence (Thorndike 1920) for the design of ambient *persuasive* systems, which are Ambient Intelligence systems that have a persuasive purpose (Kaptein et al. 2010a). More specifically we look at the effects of two manifestations of social intelligence namely mimicry and social praise, in the Ambient Intelligence (AmI) scenario (Aarts et al. 2001; De Ruyter 2003). The effects of social praise and mimicry on the perceived friendliness and intelligence of a digital social actor are examined since an increased perception of friendliness and intelligence leads to higher compliance to persuasive requests and thus enhanced acceptance of the system (Cialdini 2001). In this study, we use a chat-robot to implement and test the effects of these two acts of social intelligence. The study presents a concrete example of assessing the effects of enhancing the social intelligence of technologies.

1.1 Social intelligence in persuasive AmI systems

Systems like the digital wellness assistant described in the introduction are emerging. There is a recent shift towards the deployment of AmI technologies for Wellbeing and Care related applications (Aarts and de Ruyter 2009). Given the persuasive impact these technologies have on our daily lives (Aarts et al. 2007)—by shifting our attitudes and behavioral preferences—there is a need for important shifts in the AmI paradigm. When designing AmI systems we should not only consider the system's intelligence—the systems awareness and appropriate response to the actual context and behavior. We should also consider the social intelligence of systems: their appropriate reactions in situations where social conventions or user emotions and motives play an important role (Markopoulos et al. 2005).

The quest of creating social intelligence in persuasive AmI applications benefits from the findings within the computers as social actors (CASA) paradigm started by Reeves and Nass (Reeves and Nass 1996). Within this paradigm, numerous experimental findings regarding socially intelligent human-to-human behaviors have been replicated in system-to-human situations. It was shown for example, that humans like a digital actor more when the actor manifests personality traits similar to their own (Moon and Nass 1996). Furthermore, preference and liking of teammates, as shown in early psychology research,

seems to take place also in computer–human interaction (Nass et al. 1996). The list of similarities between human-to-human interaction and human to digital actor interaction is long and its evidence overwhelming (See for example: Nass and Moon 2000; Nass et al. 1994, 1997).

Skeptics of the CASA paradigm have argued that the effects are most probably only applicable for new users of technology: No experienced user would ever ascribe social intelligence or the like to computers and act accordingly. Numerous experimental studies however show that these social effects are *larger* for computer literate than computer illiterate users (Nass and Moon 2000). This supports the idea that computers and digital actors will also in the future be treated as social actors. Thus, findings in social science can, at least partially, be used to support human–computer interactions.

1.2 Persuasive strategies

A specific area of social science that is of interest for the design of ambient persuasive systems is that of attitude and behavioral change. Investigators studying persuasion and compliance-gaining have varied in how they individuate persuasive strategies: Cialdini (2001) elaborates on six strategies at length, Fogg (2003) describes 40 strategies under a more general definition of persuasion, and others have listed over 100 (Rhoads 2007).

Although the large array of persuasive strategies studied can be confusing for both researchers and designers, multiple strategies have proven their effectiveness both in human–human as well as human–computer interaction. In this study investigating the effectiveness of using social intelligence for the design of ambient persuasive systems, we focus on two persuasive principles identified in the social science literature: *Liking*, and *Authority*.

1.2.1 Liking

For persuasive AmI applications to be adapted and to improve compliance to their requests, one key criterion is their perceived friendliness. Friendliness leads to higher compliance to persuasive requests (Cialdini 2001); people generally say 'Yes' to the people they like (Cialdini 2004) or perceive as friendly. We expect higher compliance rates to requests by ambient persuasive systems when the system is considered friendly. In this study, we explore how acts of social intelligence can be used to increase the *perceived friendliness* of a digital actor.

1.2.2 Authority

Next to liking another well-known and effective persuasive principle is that of authority (Cialdini 2004): People

comply with legitimate experts. When a human actor is perceived as an expert in a specific field people are inclined to follow advices given by the expert. We hypothesize that an increased perceived intelligence, both social as well as instrumental, of a digital actor leads to a more expert status. An increase in perceived intelligence would thus benefit the compliance rates to persuasive requests made by an AmI system. Thus, we also explore the effects of acts of social intelligence on the *perceived intelligence* of a digital actor.

1.3 Outline of the article

In the remainder of this article, we describe the setup and results of an experiment using an artificial social agent which implements two acts of social intelligence: mimicry and social praise. We evaluate the effects of these implementations on measures of perceived friendliness and perceived intelligence. We start by motivating our choice for the use of mimicry and social praise as experimental manipulations, and explaining these social behaviors in more details. In Sect. 3, we describe the setup of the experiment and we provide detailed explanations of the implementations of the independent variables as well as the operationalization of the dependent variables.

In Sect. 4, we describe the results of our experiments and our method of analysis. The method of analysis, using *nonparametric* statistics for factorial designs, is not the most commonly reported upon. We elaborate on our choice of this methodology and explain the underlying assumptions. Finally in the discussion section, we present the implications of our findings for the design of ambient persuasive systems, and we describe several alternative ways of implementing the effective social cues of praise and mimicry in ambient systems.

2 Mimicry and social praise

Given that computers are social actors, ambient persuasive technologies could use the same social influence strategies that humans do. To test whether simple acts of social intelligence as described by Thorndike (1920)—which are successful in human-to-human communication—are effective in a human–computer interaction scenario, we set out to implement two social strategies that might influence the perceived friendliness and the perceived intelligence of an artificial agent.

An examination of the social science literature raised our interest for implementations of mimicry—matching other people’s behavior—and social praise—giving positive feedback on the social aspect of an interaction. Both

mimicry and social praise are shown to have profound effects on both perceived friendliness as well as perceived intelligence in a human-to-human setting.

2.1 Mimicry

While in the midst of an important business negotiation you suddenly realize you have been shaking your leg intensely for the last 10 min. While you normally never shake your legs, you seem to have unconsciously copied the behavior of your negotiation partner.

Whenever people interact there is a natural tendency to match each other’s behavior (Chartrand and Bargh 1999). This behavioral matching is called mimicking. Mimicking often occurs unconsciously, and it has been shown to have profound effects on interpersonal behavior and attitudes.

In studies on the topic of attitude change, mimicking has been shown to make the mimicry, the one being mimicked, more susceptible to persuasive cues. As such it has been shown that mimicking displayed by interviewers asking to sign a petition led to higher response rates (Suedfeldt et al. 1971). Mimicking even works when it just concerns a similarity in name: People are more inclined to participate in a research project when the name of the researcher is similar to their own (Garner 2005). For an overview of the effects of mimicry see (Chartrand et al. 2005).

Besides these direct effects on persuasive requests, mimicking seems to have a number of more subtle effects on people’s attitudes towards each other. Empirical research has shown that mimicking leads to higher liking. Participants that were being mimicked while interacting with an experimenter reported a smoother interaction than those not being mimicked (Chartrand and Bargh 1999). Even when addressed in a more general fashion, participants that were mimicked during an interaction scored higher on the “Inclusion of the Other” in the Self-scale (Aron et al. 1992) than participants that were not mimicked. This shows that mimicking leads to a closer feeling to others (Baaren et al. 2004). The effects of mimicry are so profound, not just in laboratory settings but also in real life, that mimicry is part of common practice in neuro linguistic programming to establish rapport (Sandoval and Adams 2001).

Mimicking is thus a very powerful act of social intelligence, which leads to positive interpersonal impressions. Outside of the human to human context mimicry has also been shown to be effective in human– computer interaction: e.g. prosodic mimicry of human utterances by a computer increases liking compared to similar utterances without prosodic mimicry (Suzuki et al. 2003). We expect mimicry displayed by an artificial agent to have profound effects on peoples evaluations of the artificial agent.

2.2 Social praise

After lecturing to a group of undergraduate students sometimes one of them approaches you not just to ask a question, but to comment on the quality of the lecture. Apparently, it was one of the best lectures he ever attended. While there is no way to check whether indeed your lecture has been one of the best, or the student just mentions this to anyone, you will probably like the student. Most probably you will even believe he is one of the more intelligent students in the class.

The example above is typical of praise or flattery. And you probably are, just like anyone else, a “sucker for praise” (Cialdini 2004). Praise has been defined as favorable interpersonal feedback (Baumeister et al. 1990). Praise is a very common feature of interpersonal interaction and is frequently used to encourage people, to socialize, to integrate into groups, and to influence people (Lipnevich and Smith 2008). There is a widespread belief that praise alters the affective state of the recipient of the praise. Praise is believed to have beneficial effects on the receiver’s self-esteem, motivation and performance (Bandura 1997; Koestner et al. 1987; Weiner et al. 1972).

Empirical studies have shown several beneficial effects of the act of praising or flattering on the subject giving the praise. In a human-to-human context, flattery has been shown to increase liking of the receiver towards the flatterer (Berscheid and Walster 1987). Furthermore, flattery has been shown to influence the receiver’s perception of intelligence of the flatterer (Pandey and Kakkar 1982).

Similar findings have been obtained in human–computer interaction. Flattery has been shown to increase liking for the computer with whom one cooperates on a task. Furthermore, flattery led to a significant increase in evaluation of the performance of the computer on a specific task (Reeves and Nass 1996).

The effects of flattery on both perceived intelligence as well as friendliness seem more profound among women than among men (Burgoon and Klingle 1998). Studies have shown that woman are more susceptible to social praise and evaluate the praise giver more positively.

Given the role of praise in previous research, we want to contribute to the literature to see if praise, delivered within an ongoing communication between a human and a robot, can increase the acceptance of the robot. We hypothesize that praise for the interaction will lead to higher liking of the robot, and a higher perceived intelligence. In this way, praise can be used to enhance the effectiveness of persuasive AmI systems. We implement praise by inserting positive feedback into an ongoing conversation between a human and an artificial agent.

2.3 Hypothesis

To increase compliance to persuasive AmI systems, we set out to investigate the effects of mimicry and social praise on the attitudes toward an artificial social agent: a chat-robot. To do so, we set up a laboratory experiment in which participants were asked to converse with a chat-robot for a maximum of 10 min and the level of mimicry and social praise were varied.

We hypothesize that for chat-robots or other digital social actors to be persuasive they need to be perceived as friendly. Furthermore, we believe chat-robots will be more effective when they are perceived as intelligent, since increased perceived authority based on intelligence leads to higher compliance. Given the studies discussed above, we hypothesize that mimicry and social praise as used by a chat-robot will have positive effects on both the perceived friendliness as well as the perceived intelligence of a chat-robot.

3 Method

To test our hypothesis of the effects of mimicry and social praise on perceived intelligence and perceived friendliness, we set up an experiment in which subjects were asked to chat with a chat-robot for a maximum of 10 min. The chat-robot—named Sara—displayed praise, mimicry, both, or none of these. In this section, we describe the experimental setup in more detail.

3.1 Participants

Fifty Dutch college undergraduates took part in our experiment (27 males and 23 females). Participants received €5.-in gift coupons for their participation, which is the standard fee used at the Technical University of Eindhoven where the study was conducted. Participants were randomly assigned to one of the conditions of a 2 × 2 (No mimicry/mimicry × No praise/praise) between-subjects factorial design. Participants all indicated to be fluent in English—the language used by the chat-robot—and indicated moderate to high computer literacy. The average age of participants was 23.8 years (SD = 5.09). A between-subjects design was chosen since we expected large order effects and increased fatigue—and thus unreliable data later in the study—in the case of a within-subjects design.

3.2 Procedure

Prior to running the experiment, a pilot study was conducted with five pilot participants. This was done to test our implementations of the conditions and to make sure all

questions were easily understood. Both the pilot study as well as the experiment were run at the Psychology lab at the Technical University of Eindhoven, the Netherlands. The Psychology lab is a laboratory with 10 sound isolated cubicles where participants can work individually using a PC. Participants were assigned to one of the cubicles and followed the on-screen instructions that guided them through the study.

The first screen presented to participants was the informed consent form. Participants were told they were participating in a study to evaluate the implementation of a chat-robot named Sara. Participants were thus aware that they would be conversing with an artificial agent and *not* with a human.

Participants were not informed about the different mimicry and praise conditions. Participants were notified that their participation was voluntary and they could stop anytime they liked. Furthermore, we textually explained that the data gathered would be used for scientific purposes only.

After obtaining informed consent, the textual instructions introduced Sara. Participants were told that they had a maximum of 10 min to converse with Sara. Sara was introduced as a newly developed chat-robot that was skilled in discussing a number of topics, namely Sport, Geography, Politics and Artificial intelligence. Participants were also told that their conversation would end automatically after 10 min; however, they could end their conversation whenever they wished by clicking a button. During the conversation, a timer displayed the remaining conversation time and after 10 min participants automatically advanced to the next instruction section.

After the conversation, participants were asked to fill out a number of questionnaires. The exact questions asked are described in the materials section. Finally, participants were notified that the experiment was over and were asked to leave the cubicle and notify the experimenter. After completion, participants were debriefed and received the €5.-reward. They were instructed not to discuss the experiment with their classmates or friends.

3.2.1 Mimicry

Participants were randomly assigned to either the mimicry or no-mimicry condition. In our experimental setting, mimicry was operationalized in the following way:

- In the *no-mimicry* condition Sara responded almost instantaneously (response times were shorter than 0.5 s) to any remark made by the participant.
- In the *mimicry* condition, we recorded the time from the first keystroke of the participant until the “Enter”

button was pressed or the “Send reply” button was clicked. The reply was then delayed by the same time.

We reckoned our mimicry implementation would capture participant’s response time excluding their reading time, which would heavily depend on the complexity and length of Sara’s responses.

3.2.2 Praise

The positive social feedback or praise conditions were implemented as follows:

- In the no-praise condition, participants conversed with Sara as implemented by an AJAX extension of the Program E php / ALICE implementation—see Materials.
- In the praise condition, we presented a positive feedback message every ten request-response cycles.

The number ten was chosen since this was not overwhelming in the conversation but would still occur at least two times within every conversation with Sara as shown in our pilot study. The feedback that was presented was a random selection of one of the following sentences:

1. I really like our conversation a lot.
2. You are a very nice person to talk to.
3. Our conversation is very pleasurable. Thanks for talking to me!
4. You are such a kind person!
5. I really like talking to you.

These sentences were presented embedded in Sara’s answers right before the actual response from Sara.

In our pilot, we discussed the implementation of the praise condition with our participants who remarked not to feel disturbed by the remarks. To further check for suspicion of deceit or expectancy created by the embedded remarks, we added the open question “*What do you believe is the goal of this experiment?*” in one of the proceeding questionnaires. This check was built in to prevent participant biases because of prior expectations. None of the participants remarked anything about social feedback or praise, and thus, we are convinced that the remarks felt natural, given the human-chat-robot conversation: they did not disclose the experimental manipulations.

3.3 Materials

In this experiment, we used the Program E implementation of A.L.I.C.E. (Wallace 2009), which is an AIML—Artificial Intelligence Markup Language—interpreter.¹ While implemented in PHP we extended the session management

¹ The most recent version of Program E can be found at: <http://sourceforge.net/projects/program-e/>.

of the standard program E installation to enable an AJAX approach to manage the discussion. The front end of the application was done in HTML, CSS and JavaScript. This approach enabled us to implement the mimicking and praise conditions on the client site using JavaScript. We ran a standard installation of program E with a number of AIML libraries relating to the topics Sport, Geography, Politics and Artificial intelligence. As mentioned before the time from the AJAX HTTP-request from the client to the PHP server sending the response, and for this response to be rendered to the participant, never took more than 0.5 s in the no-mimic condition.

3.4 Questionnaires

The questionnaires presented to participants after the conversation assessed the following:

1. The perceived friendliness of Sara.
2. The perceived intelligence of Sara.
3. Participants perceived connectedness to Sara.
4. Remarks on the conversation.
5. Additional measures.

We describe each of these in more detail.

3.4.1 Perceived friendliness

Given the aim of the experiment, the first questions after the conversation with Sara concerned the perceived friendliness of Sara. Participants were asked to grade Sara's friendliness on a scale from 1 (very unfriendly) to 10 (very friendly). This 10-point scale corresponds to the Dutch high school grading system and as such is very natural for most of our Participants. Next to this grade, participants also filled out five items regarding Sara's friendliness on a scale from 1 (Totally disagree) to 7 (Totally agree). We chose to implement these two ways of measuring friendliness to improve the construct validity of our measure. If the obtained results are equal across both methods of measurement our confidence in the obtained results as a reflection of actual perceived friendliness is increased.

All participants rated their agreement to the following items:

1. Sara was friendly during our conversation
2. Compared to humans Sara's interaction style was unfriendly
3. If Sara was a real person I would consider her friendly
4. Compared to humans Sara was polite
5. I really liked Sara

To compute a final friendliness, index item 2 was reversed and an average of the 5 items was computed for each participant (Cronbach's $\alpha = 0.762$).

3.4.2 Perceived intelligence

For perceived intelligence, we used a similar approach as perceived friendliness. First, participants were asked to grade Sara's intelligence on a 10-point scale. Second, we presented the following items (7 point scale):

1. Sara was intelligent
2. Compared to humans Sara seemed dumb
3. If Sara was a real person I would consider her intelligent
4. Compared to humans Sara was smart

Again we computed a final perceived intelligence index. Item 2 was reversed, and for each participant, we computed an average score of the 4 items (Cronbach's $\alpha = 0.706$).²

3.4.3 Perceived connectedness

Next to measuring perceived friendliness and perceived intelligence—the constructs of core interest in this study—we added a measure of perceived connectedness (Van Bel et al. 2009). We were interested to see whether a higher friendliness score also led to a stronger perception of the bond between the user and the chat-robot (Baumeister and Leary 1995). Social connectedness is an emerging construct in the research literature and we wanted to see whether this measure of long-term bond was also influenced by the social intelligence manipulations.

Social connectedness is defined as the momentary experience of belongingness and relatedness with others (Van Bel et al. 2009; Kaptein et al. 2010c). Several attempts have been undertaken to assess this experience both quantitatively as well as qualitatively. In this experiment, social connectedness was measured using a similar approach as used for the perceived friendliness and perceived intelligence measures. Participants were first asked to grade how emotionally connected they felt to Sara on a 10-point scale. Next, the following items were presented on a seven-point scale:

1. I felt connected to Sara
2. Sara and I developed a bond during our conversation
3. I could connect to Sara
4. Sara shared my interest and ideas
5. I felt related to Sara

A social connectedness score was computed by averaging over the 5 items (Cronbach's $\alpha = 0.891$).

² A plausible reason for the lower Cronbach's α of the perceived intelligence scale as compared to the other two scales is the difference in the number of items. An increase in the number of scale items usually leads to an increase in Cronbach's α .

3.4.4 Remarks on the conversation

After grading the friendliness, intelligence and connectiveness of the chat robot, we presented a number of open-ended questions to participants. We asked participants to remark on the conversation, and to describe a typical good conversation. We also checked the understanding of the study by asking for an explanation of the purpose of the study. These items were added to the study to address possible suspicion of deceit or expectancy.

3.4.5 Additional measures

Next to the questions relating to Sara, we decided to gather a number of background measurements of the participants to be able to identify possible confounding relationships.

One of these measurements was participant's individual susceptibility to persuasive cues, as measured by the questionnaire presented in Kaptein et al. (2009). This is a twelve-item 7-point rated likert scale addressing the susceptibility to each of the six principles of persuasion as identified by Cialdini (2004) with two items. This scale has shown its predictive value in estimating participant's compliance to a persuasive request. We included this measure to be able to see whether individuals with higher susceptibility to persuasive cues would also be more influenced by the social cues of mimicry and praise. One overall susceptibility score was computed for each participant (Cronbach's $\alpha = 0.698$) (See Appendix).

Next to participants susceptibility to persuasion, we also administered the TIPI: the Ten Item Personality Inventory (Gosling et al. 2003). The TIPI represents a fast and convenient way to measure personality. While not elaborate we believed the TIPI scores could be used in our experiment to see if there were any confounding effects of participants' personalities on their judgments of friendliness and intelligence of Sara. The TIPI leads to a score on each of the 5 dimensions of the Big Five (Goldberg 1990): Extraversion, Agreeableness, Conscientiousness, Emotional Stability, and Openness to experiences.

Finally, we asked for participants age, gender, and living situation to enable us to control for possible confounds due to these characteristics. In our analysis, we especially focused on gender as a possible confound since gender differences for the effects of praise have previously been shown empirically (e.g. Burgoon and Klingle 1998).

All participants fully finished the study. The average completion time was 24 min (SD = 4.5).

4 Results

The first interesting finding in this research was the tendency of participants to talk to Sara as long as possible.

While the maximum conversation time was 10 min, participants were free to stop the conversation anytime they liked. However, 82% of participants spent the full 10 min conversing with Sara. In our opinion, this showed participants involvement in this study since it was clear to participants that there was no objective of the conversation and they could exit the conversation anytime they wanted to. Involvement became even more apparent when reading the answers to the open-ended questions. Participants provided numerous helpful comments to improve Sara's conversational skills. Furthermore, given the elaborate answers and positive remarks participants clearly seemed to enjoy participating in the study.

4.1 Main findings

In this section, we first describe the effects of mimicry on friendliness and intelligence, and then we describe the effects of social praise on these two dependent variables. Finally, we describe the relationships between our measurements and the possible confounds of susceptibility to persuasion and personality.

4.1.1 Method of analysis

Because one cannot assume that both the measures on the 10-point scale (i.e. "Was Sara friendly") nor the measures on the 7-point likert scales are of interval measurement level we choose to analyze our 2×2 between subjects design using a nonparametric approach. Improper usage of parametric analysis can lead to serious errors and should thus be avoided (Singer et al. 2004; Munzel and Bandelow 1998; Kahler et al. 2007; Kaptein et al. 2010b). For our analysis we use the concepts developed by Brunner and Munzel (2002), and further elaborated upon by Shah and Madden (2004) and Markopoulos et al. (2005).

In this nonparametric approach midranks—rank scores corrected for possible ties—are used to estimate the relative effect sizes of the different conditions. This approach has recently been extended to nonparametrically analyze complex experimental designs, and gives researchers the option of estimating effect sizes and computing interaction effects. For hypothesis testing, we use the Anova Type Statistic, as suggested by Shah and Madden (2004). Results are presented using estimated relative effect sizes (\hat{p})—these being a convenient metric to depict nonparametric effects.

Since this way of analyzing the 2×2 experimental design that we used is (yet) uncommon—most researchers would use a parametric 2×2 between-subject ANOVA—we want to state that the presented results have been checked for discrepancies with this frequently used method and that the effects reported upon are—in this particular case—consistent over the two different methods. We thus

have full confidence in the (internal) validity of the presented results. All analyses were done using a $2 \times 2 \times 2$ (mimicry \times praise \times gender) model. Gender was incorporated since female participants talking to Sara took part in a similar gender conversation while males took part in an opposite gender conversation and we wanted to control for possible effects of this difference between our experimental groups. Furthermore, different genders have responded differently to praise in previous research (Burgoon and Klinge 1998). Effects of gender, or possible one and two way interactions, are not reported upon when not significant at a five percent level.

4.1.2 Effects of mimicry

Contrary to our expectation, we did not find a main effect of mimicry on the perceived friendliness of Sara ($\hat{p}_{mimic} = 28.22$, $\hat{p}_{no-mimic} = 21.67$), $F = 2.72$, $p = 0.111$. Thus, mimicry did not directly influence the perceived friendliness of the chat-robot as measured on the 10-point scale. However, there was a significant interaction effect between gender and mimicry, $F = 5.60$, $p < 0.05$, which showed that the expected effect of mimicry was observed for females ($\hat{p}_{Female,mimic} = 25.29$, $\hat{p}_{Female,no-mimic} = 18.06$), but was not observed for males.

The $2 \times 2 \times 2$ (mimicry \times praise \times gender) non-parametric analysis of the computed perceived friendliness index showed the same results as obtained for the 10-point grade: There was no significant main effect of mimicry, $F = 0.06$, $p = 0.8026$ but there was a significant interaction between gender and mimicry, $F = 2.73$, $p < 0.05$. Here again the expected positive effect of mimicry was present for females, but not for males. Combining both of these results leads us to conclude that mimicry leads to an increase in perceived friendliness of a chat-robot, but has been shown to do so only for females.

Analysis of the 10-point perceived intelligence score confirmed the hypothesis; we found a main effect of mimicry on the perceived intelligence scores ($\hat{p}_{mimic} = 30.64$, $\hat{p}_{no-mimic} = 21.85$), $F = 8.23$, $p < 0.01$, (See Fig. 1). Sara was perceived more intelligent when she displayed mimicry. When looking at the perceived intelligence index we found a similar pattern; in the mimicry condition Sara is perceived as more intelligent than in the no-mimicry condition but this difference was not statistically significant, $F = 1.46$, $p = 0.24$. For perceived intelligence, we did not find the aforementioned interaction between mimicry and gender; however, a main effect of gender was found both on the 10-point grade, $F = 29.78$, $p < 0.01$, and on the 7-point rating scales, $F = 6.88$, $p < 0.05$. In both cases, females gave higher intelligence ratings than males.

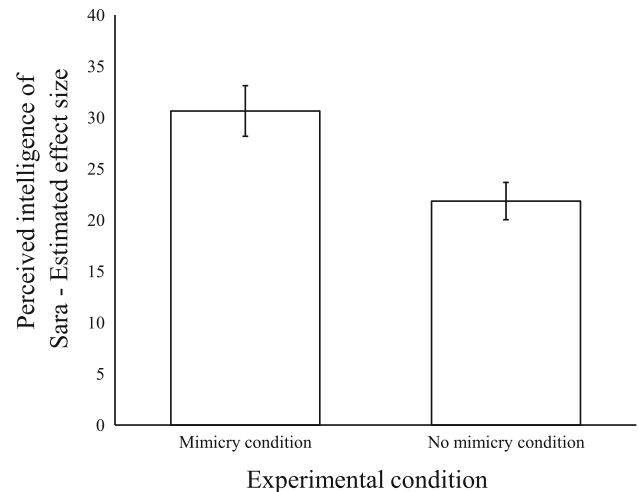


Fig. 1 Estimated relative effect sizes and standard errors for the Mimicry and No-Mimicry conditions on perceived intelligence of Sara

Additional to the analysis of perceived friendliness and praise we analyzed the effects of mimicry on connectedness. Even though the connectedness scores in the mimicry condition, both on the 10-point grade as well as on the five 7-point scales were higher than in the no-mimicry condition, we did not find significant main effects (Ten point: $F = 1.63$, $p < 0.21$; Index: $F = 0.36$, $p < 0.55$).

4.1.3 Effects of social praise

The effect of social praise was tested using the same $2 \times 2 \times 2$ nonparametric analysis as described above. For perceived friendliness, we confirmed our hypothesis: In the praise condition Sara was perceived more friendly than in the no-praise condition ($\hat{p}_{praise} = 29.37$, $\hat{p}_{no-praise} = 20.53$), $F = 4.94$, $p < 0.05$, as measured on the 10-point scale (See Fig. 2). The friendliness index indicated the same effect: the praise condition scoring higher than the no-praise condition. However, this effect was not significant at a five percent level, $F = 1.41$, $p = 0.25$.

The perceived intelligence of Sara was not affected by the usage of social praise. Both for the 10-point grade, as well as for the intelligence index we did not find a significant main effect of social praise, $F = 1.64$, $p = 0.21$; $F = 2.24$, $p = 0.15$.

As for the use of mimicry, no significant effects were found of praise on connectedness, 10-point scale: $F = 1.43$, $p < 0.24$; Index: $F = 3.20$, $p = 0.08$.

4.2 Additional findings

As mentioned in the method section, we included several measures of possible confounds in our experiment. The

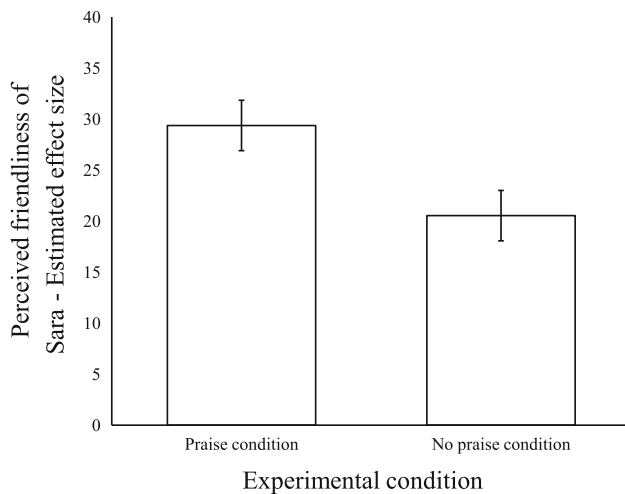


Fig. 2 Estimated relative effect sizes and standard errors for the Praise and No-Praise conditions on perceived friendliness of Sara

main possible confound—gender—was used as a control in the testing of our hypothesis. However, we also wanted to see whether possible effects of personality or susceptibility to persuasion on our dependent variables could be identified. Identification of such relationships would raise questions for follow-up research. The relationships between the possible confounds i.e. the personality scores and the susceptibility to persuasion scores, and the friendliness and intelligence measures were explored using the computation of Spearman Rho's. Table 1 presents an overview of the relevant correlations for examination of the effects of personality and susceptibility.

4.2.1 Susceptibility to persuasion

The individual susceptibility measure positively correlates with the friendliness and intelligence measures. However, this correlation is low to moderate, and is only significant for the friendliness scales. Here, a higher susceptibility score—thus, the participant is more inclined to comply to a message supported by an implementation of a persuasive

strategy—leads to a higher friendliness rating of the chat robot.

4.2.2 Personality findings

The personality dimensions as measured by the TIPI consist of Extraversion, Agreeableness, Conscientiousness, Emotional stability, and Openness to experiences. As is clear from the correlations in table one, none of these traits related to the friendliness or intelligence ratings of the chat-robot in our experiment. As such, personality traits of our participants did not influence the results of this experiment, and we thus, assume that the effects of mimicry and praise are relatively unaffected by the personality of the participant.

5 Discussion

In this study, we showed that socially intelligent behavior of a chat-robot influences its perceived friendliness and intelligence. Findings from social psychology can help us shape the social behavior of digital actors and increase their perceived friendliness and intelligence. Since friendliness and intelligence have profound effects on the perceived compliance to persuasive request in human-to-human communication, we believe that the social cues of mimicry and praise can be used to improve compliance to persuasive AmI systems—at least when the system functions as a social actor. We contributed to the existing literature by empirically showing the effects of praise and mimicry on friendliness and intelligence in a controlled laboratory setting.

We will discuss first the observed effects of mimicry on both perceived intelligence and friendliness, and next discuss possible implementations of mimicry outside the laboratory setting. Second, we will discuss the effects of social praise and its practical implementations. Finally, we will address the limitations of this study and give suggestions for future work.

Table 1 Overview of relationships between susceptibility and personality scores, and the friendliness and intelligence ratings—both the 10 point score and the index

	Friendly (10)	Friendly index	Intelligence (10)	Intelligence index
Susceptibility to persuasion	0.248	0.449	0.119	0.225
Extraversion	−0.116	−0.360	−0.201	−0.052
Agreeableness	−0.037	−0.032	0.089	0.028
Conscientiousness	0.063	0.260	0.163	0.102
Emotional stability	0.065	−0.148	0.027	0.072
Openness to experiences	0.079	0.125	−0.044	−0.066

Significant correlations in bold face

5.1 Mimicry

Our study showed a positive effect of mimicry—copying the response time of the user—on the perceived intelligence of an artificial agent. When the agent displayed mimicry participants rated the agent more intelligent. This effect was significant for the 10-point rating. The 5-item friendliness index showed the same trend. This increased perceived intelligence will probably—by utilizing authority as a persuasive strategy—increase the compliance to ambient persuasive systems that implement forms of mimicry.

Mimicry also increased the perceived friendliness of the artificial agent but did so only for women. As noted in the introduction responses to acts of social intelligence such as mimicry and praise have been previously shown to differ between males and females. However, the usage of mimicry had no adverse effect on the perceived friendliness of the artificial agent and as such mimicry can, in our opinion, safely be employed.

Overall, the effect of mimicry was somewhat smaller than we expected—mimicry had no effect on liking for males and its effect was not significant for the created indexes. This smaller than expected effect was probably due to the operationalization of mimicry in this experiment. Since only the response time of participants was mimicked the mimicry effects were small. We expect that bigger effects can be obtained when content wise mimicking is applied. However, our ability to show significant effects of mimicry for a number of dependent variables in a between subjects experiment based on our small manipulation emphasizes the strength of mimicry as a useful social cue in the Ambient Intelligence scenario.

We feel that implementations of mimicry can be made much stronger, especially when actual speech is employed by the artificial agent. In that case, mimicry can also be performed based on the speech speed, pitch and the variations in pitch. Other options of mimicry in human-computer interaction are possible when using a “talking head”. This approach has proven useful previously in the Max project (Kopp and Wachsmuth 2002). Physical emotionally expressive robots like Kismet (Breazeal 2000) extend the possible implementations of mimicry in human-computer interactions even further.

We suggest the following implementations of mimicry for usage in ambient persuasive systems to increase their perceived intelligence and thus leverage the persuasive principle of authority:

- Usage of content wise mimicry and repetition of user phrases in ongoing communication.
- Mimicking behavioral measures such as typing speed and style (chat interaction) or pitch and pitch variation (voice based interaction).

- Mimicry of body language or posture: Approaching the user when he/she approaches the system.

5.2 Social praise

In this study, social praise increased the perception of friendliness of the artificial agent. Simple remarks during the conversation led to more positive overall impressions. We believe that the increased perceived friendliness of the system—and a higher liking of the system—will lead to increased compliance due to the principle of Liking. As such, social praise is possibly useful to increase compliance to requests made by ambient persuasive systems.

We believe that a proper timing of praise would have strengthened the obtained effects. The absence of effects of praise on intelligence can be explained by similar argument and this is illustrated by one of the remarks given by a participant in the open-ended question section: “*I noticed Sara said she liked the conversation every now and then, however sometimes this was totally misplaced and this made her seem dumb*”. Although this did not show up as a negative main effect of mimicry on intelligence, we do feel that efforts should be devoted to implementing praise during natural conversation and properly timing the praise.

We suggest an exploration of the following implementations of social praise by ambient persuasive systems to increase liking and eventually increase compliance:

- Providing content based praise based in ongoing communications, reflecting on past communication instances.
- Providing praise based on user performance instead of general conversation characteristics.
- Responding appropriately to user-generated instances of praise.³

5.3 Limitations

A number of the effects of mimicry and social praise on perceived friendliness and perceived intelligence found in this study—while in the hypothesized directions—were rather small and this might seem confusing. We discuss two possible reasons for these small effects: The relatively short interaction time with the artificial agent, and the possible floor effect of the utilized indexes. We also address the null-effect on the social connectedness measures.

³ Several such instances occurred in the current study. Users would comment (e.g.) “You gave a funny reply!” and Sara would respond with “Do you think I am funny?” Instead, a “Thank you, I think you are funny to!” would probably be more appropriate.

We feel that one reason for the relatively small effects in this study is the relatively short conversation time (max. 10 min). While 10 min seemed like an appropriate time in our pretest the finding that almost all participants engaged in the conversation for the full 10 min might be an indicator this time was too short to really “get to know” the agent. As such, the implemented manipulations were not experienced for sufficient time to make a difference in the evaluation of the chat-robot. It would be interesting to see whether a prolonged conversation increases the effects of mimicry and social praise.

While we believe that the limited time leads to overall decreasing effects of the manipulations, we feel that the pattern of significant results for the 10-point scales accompanied by indicative but non-significant results for the indexes for both dependent variables warrants additional explanation. Overall, participants judgment of the chat-robot was relatively negative. In the Dutch grading system, the 10-point scale is effectively used from 4 to 9 in an educative setting. Any score lower than 6 represents a fail. Thus, a 2 or 3 on this scale for our Dutch participants reflected a very negative judgment. On the other hand, a score of 1 (totally disagree) on the 7-point scale probably seemed less harsh. As such, a floor effect prevented sufficient variation in the indexes to obtain significant results.

An alternative explanation for the observed effect of mimicry in this study is an effect of the delay in response time. Hence, the current finding need not necessarily be caused by mimicry per se, and could be an artifact of increased response time by the chat-robot. Delayed response time could lead a participant to believe that the chat-robot put more thought in the answer and thus a higher intelligence score was attributed. A further study which compares the current mimicry condition with a condition implementing a randomly delayed response time could be setup to test this distinction.

In hindsight, we believe it was not surprising that we did not find any significant differences on perceived social connectedness to Sara. Here also the 10-min time slot was too short to create an actual bond and more time and conversation is needed to build up a feeling of social connectedness. A longitudinal replication of this study would clarify this expected effect of praise and mimicry on long-term connectedness. Furthermore, a longitudinal study would also show whether the effects of mimicry and praise upon intelligence and friendliness are persistent over time.

We believe that ongoing communications with an actor that is perceived as friendly will lead to more bonding, and thus to a higher social connectedness score, than with an actor that is perceived as unfriendly. Overall it would be feasible, despite practical difficulties, to conduct more longitudinal studies of the effects of acts of social

intelligent within an embedded AmI setting. It is worthwhile to explore in which scenario’s the effects of social intelligence are indeed the same for HCI as for human-human interactions and in which cases they may be similar but not the same.

5.4 Future work

Our study confirms that endowing artificial agents with behaviors relating to social praise and mimicry can increase their perceived friendliness and perceived intelligence. These should in turn lead to higher persuasiveness of the agents. Further investigations are needed to verify that this is indeed the case and to examine (a) whether these effects carry over to usage situations outside the lab, and (b) what the impact is of repeated exposure to such social cues by artificial agents. Finally, it would be interesting to see how the findings in this study replicate across different cultures and backgrounds of participants.

The manipulations used in this experiment (varying the response time and providing positive comments on the conversation as such) were very simple but effective in increasing the overall opinion towards an artificial agent in a laboratory setting. An interesting challenge for future research is to develop subtler and more varied ways of mimicry and social praise that will be sensible to apply during real use in persuasive AmI systems.

6 Conclusion

This paper described an experiment which showed that *social praise*—positive feedback—increased the perceived friendliness of a chat-robot during an ongoing conversation. The study also showed that *mimicry*—displaying matching behavior—enhanced the perceived intelligence of the robot. Both of these acts of social intelligence led to improved evaluations of the chat-robot and should be considered in the design of AmI systems which take an active social role.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

Appendix: 12 item susceptibility questionnaire

1. When a family member does me a favor, I am very inclined to return this favor.
2. I always pay back a favor.
3. I believe rare products (scarce) are more valuable than mass products.

4. When my favorite shop is about to close, I would visit it since it is my last chance.
5. I always follow advice from my general practitioner.
6. When a professor tells me something I tend to believe it is true.
7. Whenever I commit to an appointment I do as I told.
8. I try to do everything I have promised to do.
9. If someone from my social network notifies me about a good book, I tend to read it.
10. When I am in a new situation I look at others to see what I should do.
11. I accept advice from my social network.
12. When I like someone, I am more inclined to believe him or her.

References

- Aarts EHL, Harwig M, Schuurmans M (2001) Ambient Intelligence. In: Denning P (ed) *The invisible future: the seamless integration of technology into everyday life*. McGraw Hill, New York
- Aarts EHL, Markopoulos P, de Ruyter BER (2007) The persuasiveness of ambient intelligence. In: Petkovic M, Jonker W (eds) *Security, privacy and trust in modern data management*. Springer, New York
- Aarts E, de Ruyter B (2009) New research perspectives on Ambient Intelligence. *J Ambient Intell Smart Environ* 1:5–14
- Aron A, Aron EN, Smollan D (1992) Inclusion of other in the Self Scale and the structure of interpersonal closeness. *J Pers Soc Psychol* 63:596–612
- Baaren RB, Horgan TG, Chartrand TL, Dijkmans M (2004) The Forest, the Tress, and the Chameleon: context dependence and mimicry. *J Pers Soc Psychol* 86:453–459
- Bandura A (1997) Self-efficacy: toward a unifying theory of behavioral change. *Psychol Rev* 84:191–215
- Baumeister RF, Leary MR (1995) The need to belong: desire for interpersonal attachments as a fundamental human motivation. *Psychol Bull* 117:497–529
- Baumeister RF, Hutton DG, Cairns KJ (1990) Negative effects of praise on skilled performance. *Basic Appl Soc Psychol* 11:131–148
- Berscheid E, Walster E (1987) *Interpersonal attraction*. Addison-Wesley, New York
- Breazeal C (2000) *Sociable machines: expressive social exchange between humans and robots*. Sc.D. dissertation, Department of Electrical Engineering and Computer Science, MIT, Cambridge
- Burgoon M, Klingle RS (1998) Gender differences in being influential and/or influenced: a challenge to prior explanations. In: Canary DJ, Dindia K (eds) *Sex differences and similarities in communication: critical essays and empirical investigations of sex and gender interaction*. Erlbaum
- Burner E, Munzel U (2002) *Nichtparametrische Datenanalyse*. Springer, Heidelberg
- Chartrand TL, Bargh JA (1999) The Chameleon effect: the perception-behavior link and social interaction. *J Pers Soc Psychol* 76:893–910
- Chartrand TL, Maddux WW, Lakin JL (2005) Beyond the perception-behavior link: the ubiquitous utility and motivational moderators of nonconscious mimicry. In: Hassin R, Uleman J, Bargh JA (eds) *Unintended thought 2: the new unconscious*. Oxford University Press, New York
- Cialdini R (2001) *Influence, science and practice*. Allyn and Bacon, Boston
- Cialdini R (2004) *The science of persuasion*. Scientific American Mind, London
- De Ruyter B (2003) User centred design. In: Aarts E, Marzano S (eds) *The new everyday: vision on Ambient Intelligence*, 010 Publishers
- Fogg BJ (2003) *Persuasive technology: using computers to change what we think and do*. Morgan Kaufmann, New York
- Garner RL (2005) What is in a name: persuasion perhaps. *J Consum Psychol* 15:108–116
- Goldberg LR (1990) An alternative “Description of personality”: the big-five factor structure. *J Pers Soc Psychol* 59(6):1216–1229
- Gosling SD, Rentfrow PJ, Swann WB Jr (2003) A very brief measure of the Big-Five personality domains. *J Res Pers* 37:504–528
- Kahler E, Rogausch A, Brunner E, Himmel W (2007) A parametric analysis of ordinal quality-of-life data can lead to erroneous results. *J Clin Epidemiol* 61:475–480
- Kaptein MC, Markopoulos P, Aarts EHL, de Ruyter B (2009) Can you be Persuaded? Individual differences in susceptibility to persuasion. In: *Proceedings of interact 2009*
- Kaptein M, Markopoulos P, de Ruyter B, Aarts E (2010a) Persuasion in Ambient Intelligence. *J Ambient Intell Human Comput Springer* 1(1):43–56
- Kaptein MC, Nass C, Markopoulos P (2010b) Powerful and consistent analysis of likert-type rating scales. In: *Proceedings of CHI 2010* (in press)
- Kaptein MC, Van Bel D, de Ruyter B, Markopoulos P, Aarts E (2010c) Social connectedness and compliance. *CHI 2010 workshop on social connectedness*
- Koestner R, Zuckerman M, Koestner J (1987) Praise, involvement, and intrinsic motivation. *J Pers Soc Psychol* 53:383–390
- Kopp S, Wachsmuth I (2002) Model-based animation of coverbal gesture. In: *Proceedings of computer animation*. IEEE Computer Society Press, Los Alamitos, pp 252–257
- Lipnevich AA, Smith JK (2008) Response to assessment feedback: the effects of grades, praise, and source information. *ETS RR-08-30*
- Maheshwari M, Chatterjee S, Drew D (2008) Evaluating the persuasiveness of “Just in Time” motivational messages for obesity management. *Lecture notes in computer science*, 5033/2008, Springer
- Markopoulos P, de Ruyter B, Saini P, van Breemen A (2005) Case study: bringing social intelligence into home dialogue systems. *ACM Interact* 12(4):37–44
- Moon Y, Nass C (1996) How “real” are computer personalities? Psychological responses to personality types in human-computer interaction. *Commun Res* 23:651–674
- Munzel U, Bandelow B (1998) The use of parametric versus nonparametric tests in the statistical evaluation of ratings scales. *Pharmacopsychiatry* 6:222–224
- Nass C, Moon Y (2000) Machines and mindlessness: social responses to computers. *J Soc Issues* 56:81–103
- Nass C, Fogg BJ, Moon Y (1996) Can computers be teammates. *Int J Human-Comput Stud* 45:669–678
- Nass C, Moon Y, Morkes J, Kim EY, Fogg BJ (1997) Computers are social actors: a review of current research In: Friedman B (ed) *Human values and the design of computer technology*. CSLI Press, Stanford
- Nass C, Steuer JS, Tauber E (1994) Computers as social actors. In: *Proceedings of the CHI conference*, Boston, MA
- Pandey J, Kakkar S (1982) Supervisors’ affects: attraction and positive evaluation as a function of enhancement of others. *Psychol Rep* 50:479–546
- Räsänen T, Oinas-Kukkonen H, Pahlila S (2008) Finding Kairos in quitting smoking: smokers’ perceptions of warning pictures. *Lecture notes in computer science*, 5033/2008, Springer

- Reeves B, Nass C (1996) *The media equation: how people treat computers, television, and new media like real people and places*. Cambridge University Press, Cambridge
- Rhoads K (2007) How many influence, persuasion, compliance tactics & strategies are there? (<http://www.workingpsychology.com/numbertactics.html>)
- Sandoval VA, Adams SH (2001) Subtle skills for building rapport. *FBI Law Enforc Bull* 8:1–9
- Schroeder S (2007) We can do better-improving the health of the american people. *N Engl J Med* 357:1221–1228
- Shah DA, Madden LV (2004) Nonparametric analysis of ordinal data in designed factorial experiments. *Pythopathology* 94:33–60
- Singer JM, Poleto FZ, Rosa P (2004) Parametric and nonparametric analyses of repeated ordinal categorical data. *Biomed J* 46:460–473
- Suedfeldt P, Bochner S, Matas C (1971) Petitioner's attire and petition signing by peace demonstrators. *J Appl Psychol* 1:278–283
- Suzuki N, Takeuchi Y, Ishii K, Okada M (2003) Effects of echoic mimicry using hummed sounds on human-computer interaction. *Speech Commun* 40:559–573
- Thorndike EL (1920) *Intelligence and its uses*. Harper's Mag 140:227–235
- Van Bel DT, Smolders KCHJ, IJsselsteijn WA, De Kort YAW (2009) Social connectedness: concept and measurement. In: *Proceedings of the 5th international conference on intelligent environments*, pp 67–74
- Wallace R (2009) The anatomy of ALICE. doi:10.1007/978-1-4020-6710-5_13
- Weiner B, Heckhausen H, Meyer W (1972) Causal ascriptions and achievement behavior: a conceptual analysis of effort and reanalysis of locus of control. *J Pers Soc Psychol* 21:239–248