Simple ways to make friends.

Maurits Kaptein
Eindhoven University of Technology
Den Dolech 2, HG 2.38
5600MB Eindhoven, the Netherlands
m.c.kaptein@tue.nl

Panos Markopoulos
Eindhoven University of Technology
Den Dolech 2, HG 2.38
5600MB Eindhoven, the Netherlands
p.markopoulos@tue.nl

Boris de Ruyter
High Tech Campus 34 -WB-5.27
5656 AE Eindhoven, the Netherlands
boris.de.ruyter@philips.com

Emile Aarts
High Tech Campus 34 -WB-7.51
5656 AE Eindhoven, the Netherlands
emile.aarts@philips.com

Abstract

This paper describes a study of the effects of mimicry and social praise when used by an artificial social agent. An experiment (N=50) is described which shows that social praise increases the perceived friendliness of a chat-robot, while, mimicry - displaying matching behavior - enhances the perceived intelligence of the robot. We advice designers to incorporate both mimicry and social praise in their persuasive systems. Different ways of implementing mimicry and praise in an ambient persuasive scenario are discussed.

Keywords: 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 07:30. 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 09.30 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 their digital training assistants since it would decrease obesities, coronary disease and diabetes, we all know this is tough on Sunday mornings.

Within the HCI community there is an increasing focus on developing persuasive technologies: Technologies intentionally designed to influence users' behavior [18]. These systems most typically target behaviors generally agreed upon as being positive: exercising more, smoking less, or complying with drug subscriptions (e.g., [34][26]).

The current study is an investigation of the effectiveness of what Thorndike already in 1920 called social intelligence [41]. More specifically we look at the effects of two manifestations of social intelligence namely *mimicry* and *praise*, which are both practically applicable in the Ambient Intelligence (AmI) scenario [1]. 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 [15]. This is especially relevant in the AmI scenario, since here embedded technology interacts with users to enhance their quality of life [17]. In this study we use a chat-robot to implement and test the effects of these two socially intelligent behaviors. The study presents a concrete example of enhancing the *social intelligence* of technology.

1.1 SOCIAL INTELLIGENCE IN AMI SYSTEMS

Systems like the digital wellness assistant are emerging. There is a recent shift towards the deployment of AmI technologies for *Wellbeing* and *Care* related applications [3]. Given the persuasive impact these technologies have on our daily lives [2], by shifting our attitudes and behavioral preferences, there is a need for important shifts in the

AmI paradigm. We need not only consider the system intelligence – the systems awareness and appropriate response to the actual context and behavior – of persuasive AmI applications but also their social intelligence: *Their appropriate reactions in situations where social conventions or user emotions and motives play an important role*.

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 [35]. Within this paradigm numerous experimental findings regarding socially intelligent human-to-human behavior have been replicated in system to human situations. It was shown for example, that humans like a digital actor more when they have personality traits in common [27]. Furthermore, preference and liking of teammates, as shown in early psychology research [30] seems to take place also in computer-human interaction [30]. The list of similarities between human-to-human interaction and human to digital actor interaction is long and its evidence overwhelming (See for example: [29][31][32]).

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 [15]; people generally say 'Yes' to the people they like [16] 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 two known effective social cues - acts of social intelligence - can be used to increase the perceived friendliness.

Next to perceived friendliness another well-known and effective persuasive cue is that of authority [16]: 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 in the field. 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. In this study we experimentally test whether the two social cues, mimicry and praise, affect the perceived intelligence and the perceived friendliness of a possible social actor: a chat-robot.

1.2 OUTLINE OF THE ARTICLE

In the remainder of this article we describe the setup and results of an experiment using a chat-robot which implements both mimicry and social praise and tests its effects on measures of perceived friendliness and perceived intelligence. We start by motivating our choice for mimicry and social praise, and explaining these social behaviors in more detail. In section three 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 section four we describe the results of our experiments and our method of analysis. The method of analysis, using nonparametric statistics, 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

It is frequently reported in social psychology literature how two social behaviors seem to influence both perceived friendliness and perceived intelligence of another. Mimicry – matching other people's behavior – and social praise – giving positive feedback on the social aspect of an interaction both increase overall evaluations of others. Since both mimicry and social praise are relatively easy to implement in a chat-robot setting, and since the effects of these two social cues are empirically shown in social sciences studies we decided to focus on these two cues. We briefly review the evidence for each of these social behaviors separately.

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 ten minutes. 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 [13]. This behavioral matching is called mimicking. While mimicking often occurs unconsciously 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 [39]. 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 [19]. For an overview of the effects of mimicry see [14].

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. Respondents that were being mimicked while interacting with an experimenter reported a smoother interaction than those not being mimicked [13]. Even when addressed in a more general fashion, respondents that were mimicked during an interaction scored higher on the "Inclusion of the Other" in the Self-scale [4] than respondents that were not mimicked. This showed that mimicking led to a closer feeling to others [5]. 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 [36].

Mimicking is thus a very powerful social cue, which leads to positive interpersonal impressions. Outside of the human to human context mimicry has also been shown to be effective also in human computer interaction: Prosodic mimicry of human utterances by a computer increases liking compared to similar utterances without prosodic mimicry [40]. We extend on this work by implementing time-based mimicry in a chat robot.

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. Whilst 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 remark above is a typical example of praise or flattery. And you probably are, just like anyone else, a "sucker for praise" [16]. Praise has been defined as *favorable interpersonal feedback* [8]. 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 [25]. There is a widespread believe 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 [6][23][43].

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 [9]. Furthermore, flattery has been shown to influence the receiver's perception of intelligence of the flatterer [33].

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 [35]. Finally, the effects of flattery seem more profound amongst woman than amongst men [11].

Given the role of praise in previous research we want to contribute to the literature to see if praise, delivered within the 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.

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 towards an artificial social agent: a chat-robot. To do so we set up a laboratory experiment in which respondents were asked to converse with a chat-robot for a maximum of ten minutes 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 authority based on intelligence leads to higher compliance in persuasive situations. Given the evidence from the studies of attitude and attitude change discussed above we hypothesize the following:

1. Mimicking displayed by a chat-robot will positively affect the perception of the chat-robot by humans.
2. Social praise given by a chat-robot will positively affect the perception of the chat-robot by humans.

To be more concrete; we expect the effects of mimicking and social praise to be measurable in the following domains: *Perceived friendliness* and *Perceived intelligence*.

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 ten minutes. 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 SUBJECTS

Fifty respondents took part in our experiment (27 males and 23 females). Respondents received €5 in gift coupons for their participation, which is the standard fee used at the Technical University of Eindhoven. Subjects were randomly assigned to one of the conditions of a 2 X 2 (No mimicry / mimicry X No praise / praise) between subjects factorial design. Subjects mainly consisted of undergraduate students. The average age of respondents was 23.83 years (SD = 5.09).

3.2 PROCEDURE

Prior to running the experiment a pilot study was conducted with five pilot subjects. 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 was run at the Psychology lab at the Technical University of Eindhoven, which is a lab with 10 sound isolated cubicles where participants can work individually using a PC. Respondents were assigned to a cubicle and followed the on-screen instructions on the screen that guided them trough the study.

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

Respondents were not informed about the different *mimicry* and *praise* conditions. Respondents 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. Respondents were told that they had a maximum of ten minutes to converse with Sara. Sara was introduced as being skilled in a number of topics, namely: *Sport*, *Geography*, *Politics* and *Artificial intelligence*. Respondents were told that their conversation would end automatically after ten minutes, however they could end their conversation whenever they wished. After the conversation respondents were asked to fill out a questionnaire. The exact questions asked are described in the materials section. Finally, respondents were notified that the experiment was over and were asked to leave the cubicle and notify the experimenter. After completion, respondents were debriefed and received the 5€ reward. They were instructed not to discuss the experiment with their classmates or friends.

3.2.1 Mimicry

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

1. In the *no-mimicry* condition Sara responded almost instantaneously (response times were shorter than 0.5 seconds) to any remark made by the respondent.
2. In the *mimicry* condition we recorded the time from the first keystroke of the respondent until the enter button was pressed or the send reply button was clicked. We reckoned this would capture respondent's response time excluding their reading time, which would heavily depend on the complexity and length of Sara's responses. Then, the response by Sara was delayed by giving a "*Wait for a response from Sara…*" message, and the response of Sara was presented in the same time it took for the respondent to type his or her answer.

We have to note here that the chosen manipulation of mimicry is not the only possible option. In face-to-face settings mimicry is normally defined as similarities in body language and posture these cues could not be used in a chat environment. Furthermore we believed that content wise mimicking; thus repeating parts of the respondent's sentences, could have profound negative effects when the natural language processing for this kind of mimicry was not implemented properly. Time-based mimicry presents a more straightforward approach which is easy to implement.

Other options of mimicry in human computer interaction would have emerged by the use of a 'talking head', an approach used for example by Kopp in the Max project [24], or by usage of an actual emotionally expressive robot like Kismet [10]. However, with this experiment we wanted to investigate effect of a very simple implementation of mimicry in relative simple human – system interactions. An effect found in such a setting in our view presents a greater motivation to enhance systems with simple social cues, even when the systems are not regarded as socially advanced by themselves.

3.2.2 Praise

The positive social feedback or praise condition was implemented as follows:

1. In the *no-praise* condition respondents conversed with Sara as implemented by an AJAX extension of the Program E php / ALICE implementation.
2. 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!*" and 5. "*I really like talking to you*". These sentences were presented embedded in Sara's answers right before the response from the ALICE system.

In our pilot we discussed the implementation of the praise condition with our respondents who remarked not to feel disturbed by the remarks. To further check for the unnaturality of the embedded remarks we added the open question "*What do you believe is the goal of this experiment?*" in our preceding questionnaire. This check was built in to prevent respondent biases because of prior expectations. None of the respondents 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. [42], which is an AIML interpreter. While implemented in PHP we extended the session management 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 respondent, never took more than 0.5 seconds in the *no-mimic* condition.

The questionnaire presented to respondents after the conversation consisted in 6 parts: Perceived friendliness, Perceived intelligence, Perceived connectedness, Remarks on the conversation, Individual Susceptibility to persuasion, the Ten Item Personality Inventory (TIPI) and Demographics. We describe each of these in more detail.

3.3.1 Perceived friendliness

Given the aim of the experiments the first questions after the conversation with Sara concerned the perceived friendliness of Sara. Respondents were asked to grade Sara's friendliness on a scale from 1 (very unfriendly) to 10 (very friendly). This corresponds to the Dutch high school grading system and as such is very natural for most of our respondents. Next to this grade respondents also filled out five items regarding Sara's friendliness on a scale from 1 (Totally disagree) to 7 (Totally agree). All respondents 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

Item 2 is phrased negatively and thus its coding was reversed in subsequent analysis.

3.3.2 Perceived intelligence

For perceived intelligence we used a similar approach as perceived friendliness. First, respondents were asked to grade Sara's intelligence on a ten-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

Item 2 is phrased negatively and thus its coding was reversed in subsequent analysis.

3.3.3 Perceived connectedness

Next to the hypothesized effects on friendliness and intelligence we added a measure of perceived connectedness. We were interested to see whether a higher friendliness score also let to a stronger perception of the bond between the user and the chat-robot [7].

Perceived connectedness was approached in a similar fashion as the friendliness and intelligence measures. Respondents were asked to grade how *emotionally connected they felt* to Sara on a ten-point scale. After the grading 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

3.3.4 Remarks on the conversation

After grading the friendliness, intelligence and connectedness we presented a number of open-ended questions. We asked respondents 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.

3.3.5 Individual differences in susceptibility to persuasion

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 [22]. 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 [16] 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.

3.3.6 Personality

Next to participants susceptibility to persuasion we also administered the TIPI: the Ten Item Personality Inventory [20]. 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.

3.3.7 Demographics

Finally, we included questions about respondents their 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 [11].

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

4 RESULTS

The first interesting finding in this research was the tendency of respondents to talk to Sara as long as possible. While the maximum conversation time was ten minutes, respondents were free to stop the conversation anytime they liked. However, 82% of respondents spent the full ten minutes conversing with Sara. In our opinion this showed respondents involvement in this study since it was clear to respondents 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. Respondents provided numerous helpful comments to improve Sara's conversational skills. Furthermore, given the elaborate answers and positive remarks respondents 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. In the next section we describe the relationships between our measurements and the possible confounds of susceptibility to persuasion and personality.

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 X 2 between subjects design using a nonparametric approach. Improper usage of parametric analysis can lead to serious errors and should thus be avoided [38][28][21]. For our analysis we use the concepts developed by Brunner and Munzel [12], and further elaborated upon by Shah and Madden [37].

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 can be 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 [37]. Results are presented using the estimated relative effect sizes ($\hat{p}$) – these being a convenient metric to depict nonparametric effects.

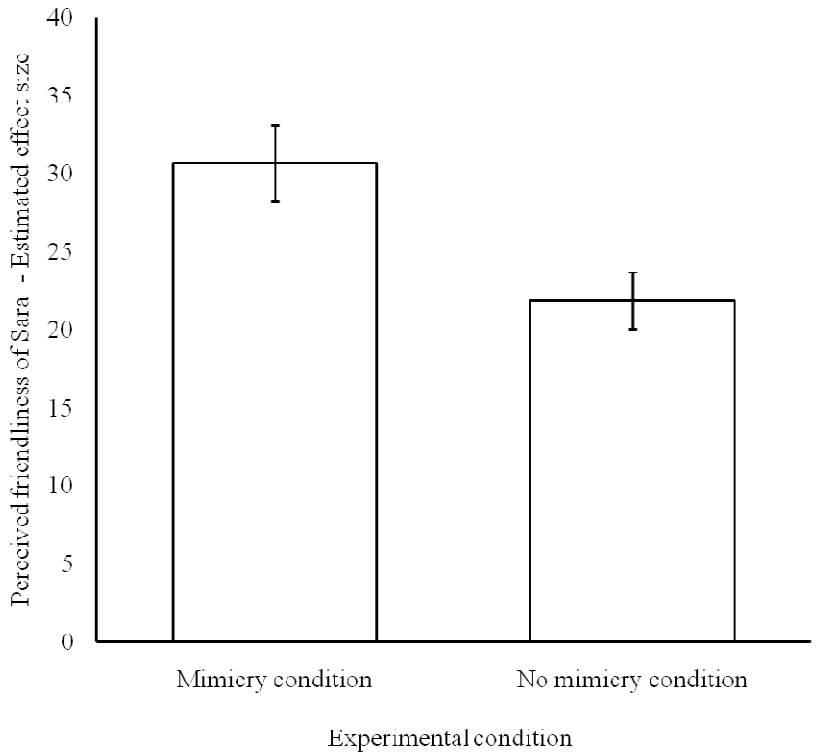
Since this way of analyzing the 2 X 2 experimental design that we used is (yet) uncommon in HCI – most researchers would use a parametric 2 X 2 mixed subject ANOVA – we want to state that the presented results have been checked for discrepancies with this frequently used method and that the effects reported upon are consistent over different methods. We thus have full confidence in the (internal) validity of the presented results. All analyses were done using a 2 X 2 X 2 (mimicry X praise X gender) model. Gender was incorporated since female

respondents 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 [11]. Effects of gender, or possible one and two way interactions, are not reported upon when not significant at a five percent level.

4.1.1 Effects of mimicry

Based on the social science literature we expected to find a positive effect of mimicry on both perceived intelligence as well as perceived friendliness. We first analyzed the scores provided by respondents on the 10 point scale (the grade) and then proceeded with analyzing the scores on the likert scales.

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



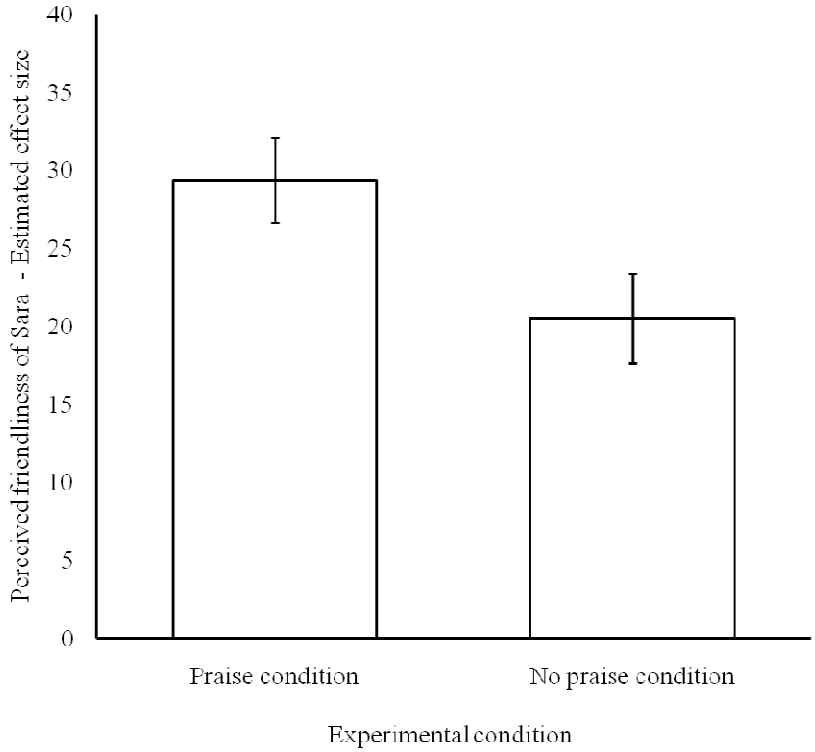
Contrary to our expectation we did not find a main effect of mimicry on the perceived friendliness of Sara ($\hat{\mu}_{mimic}$=28.22, $\hat{\mu}_{no\text{-}mimic}$= 21.67, F=2.72, p=0.111). Thus, mimicry did not directly influence the perceived friendliness of the chat-robot as measured on the 10 point 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 ($\hat{\mu}_{Female,\ mimic}$=25.29, $\hat{\mu}_{Female,\ no\text{-}mimic}$=18.06), but was not observed for males. This partly supports hypothesis one.

When looking at the ratings on the five 7-point likert scale items on perceived friendliness it was clear that these items properly measured one construct (Cronbach's Alpha=0.762). We thus computed a mean friendliness score for each of the participants. The 2 X 2 X 2 (mimicry X praise X gender) nonparametric analysis 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 does so only for females.

Analysis of the ten point perceived intelligence score confirmed the first hypothesis; we found a main effect of mimicry on the perceived intelligence scores ($\hat{p}_{mimic}$=30.64, $\hat{p}_{no\text{-}mimic}$=21.85, F=8.23, p<0.01) (See figure 1). Sara was perceived more intelligent when she displayed mimicry. When looking at the four 7-point likert scales for perceived intelligence (Cronbach's Alpha=0.706) 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.

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



Our results show that effects of mimicry on both intelligence and on friendliness were present. The positive effect of mimicry on intelligence confirms hypothesis one. Mimicry led to consistently higher ratings than no-mimicry. However, mimicry had a statistical significant positive effect on friendliness only for females.

Additionally we analyzed the effects of mimicry on connectedness. Even though the connectedness ratings in the mimicry condition, both on the 10 point grade as well as on the five 7 point scales (Cronbach's Alpha=0.891), were higher than in the no-mimicry condition, we did not find significant main effects (Ten point: F=1.63, p<0.21; Scale items: F=0.36, p<0.55). We expect that this finding is due to the relatively short conversation time (max. 10 minutes) which proves too short to get a feeling of connectedness with the chat-robot. This finding could also be due to the relatively weak implementation of mimicry which might also in the long run not have a profound effect on the perceived connectedness.

4.1.2 Effects of social praise

Hypothesis two was tested using the same 2 X 2 X 2 nonparametric analysis as described above. For perceived friendliness we confirmed hypothesis one; in the praise condition Sara was perceived more friendly than in the no-praise condition ($\hat{p}_{praise}$ =29.37, $\hat{p}_{no\text{-}praise}$ =20.53, F=4.94, p<0.05) as measured on the 10 point scale (See figure 2). The average of the five 7-point likert scale ratings 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). Due to the significant result for the ten point scale, and a similar trend on the 7 point likert scales we concluded that for perceived friendliness hypothesis two can be confirmed; the friendliness of Sara increases by the usage of praise during the conversation.

The perceived intelligence of Sara was not effected by the usage of social praise. Both for the 10 point grade, as well as for the averaged 7 point likert scale rating 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. Here again we believe that this is due to the relatively short conversation time, which prooved to short to build a connection between Sara and our participants (Ten point: F=1.43, p<0.24; Scale items: 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 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 score, and the friendliness and intelligence measures were explored using the computation of Spearman Rho's. This is the nonparametric equivalent of the Pearson correlation and as such indicates the strength of a linear relationship between two variables. Table 1 presents an overview of the relevant correlations for examination of the effects of personality and susceptibility.

Table 1. Overview of relationships between susceptibility and personality scores, and the intelligence and friendliness ratings. Significant correlations in **bold** face.

	Friendliness (10 point)	Friendliness scales	Intelligence (10 point)	Intelligence scales
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

4.2.1 Susceptibility to persuasion

The twelve item susceptibility questionnaire proved reliable (Cronbach's Alpha=0.698), which is in-line with previous empirical studies using this instrument [22]. Susceptibility to persuasion measures people's tendency to comply with persuasive requests which are cued by persuasive cues. The basic idea is that while for some people the sentence "*Buyers of this book also bought…*" heavily influences their decision making, while for others the effects are less immediate. We included this measure because we feel it reflects a general tendency of people to be influenced by external cues.

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 with a persuasive cue – leads to a higher friendliness score. This result is interesting, and would lead us to argue that people whom are more easily influenced also provide more positive ratings.

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 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 unaltered by the personality of the respondent.

5 DISCUSSION

In this study we showed that the socially intelligent behavior of a chat-robot influences its perceived friendliness and intelligence. The findings from the studies of social psychology in the field of attitude change 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. We contributed to the existing literature by empirically showing the effects of praise and mimicry on friendliness and intelligence in a controlled laboratory setting.

In this experiment it was clear that social praise like "*I really like our conversation*" led to a higher perceived friendliness of our chat-robot. Furthermore mimicry, in this case reflecting the response time of the robot, increased the perceived friendliness but did so only for women. These results showed that the perceived friendliness of digital actor is easily influenced by the use of social cues, and more research into the effectiveness of specific cues for specific situations and users is feasible. Utilizing the power of social cues to increase the persuasive powers of AmI systems is a promising approach. Praise and mimicry can be used to make friends, and friends are those you will listen to.

Next to the effects of praise and mimicry on perceived friendliness, we also showed that mimicry has a significant effect on the perceived intelligence of the robot. Even our rudimentary implementation of mimicry, copying the response times, had a significant effect on the perceived intelligence of Sara. Being able to show such an effect in a within subjects experiment shows the profound power of mimicry as a social cue. The absence of effects of praise on intelligence can be explained by one of the remarks given by respondents 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*". However, if this would have been the case for all respondents a logical consequence would have been a negative main effect of praise on intelligence which was absent in our experiment. We thus feel that to influence perceived intelligence praise needs to be delivered at the right moment in the conversational context however no detrimental effect of wrongly timed praise was found in this experiment.

Proper timing of praise would require a deeper level of social intelligence build into the system. We hypothesize that properly delivered timely praise will have its expected effects on the perceived intelligence of digital agents – even in this experiment the data clearly showed this trend. Engaging in the social act of praising or flattery can as such be a useful tool to increase the acceptance of AmI systems.

Table 2. Possible implementations of mimicry and praise for interactive systems.

Mimicry – implementations:
1. Using content wise mimicry and repetition of user phrases in ongoing communication.
2. Mimicking behavioral measures such as typing speed and style (chat interaction) or pitch and pitch variation (voice based interaction).
3. Mimicry of body language or posture: Approaching the user when he/she approaches the system.
Praise – implementations:
1. Providing content based praise based in ongoing communications,

reflecting on past communication instances.

2. Providing praise based on user performance instead of general conversation characteristics.
3. Responding appropriately to user generated instances of praise.

The overall effect of mimicry was smaller than expected, probably due to the operationalization of mimicry in this experiment. Since we only mimicked the response time of respondents 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 in a between subjects experiment based on such a small manipulation emphasizes the strength of mimicry as a social cue. When a system uses voice to communicate the implementation of mimicry can be much stronger, since mimicry can than also be performed based on the speech speed, pitch and the variations in pitch.

In hindsight we believe it was not surprising that we did not find any significant differences on perceived social connectedness to Sara. Most probably the ten-minute time slot was too short to create a 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 a higher connectedness 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 social intelligent acts within an embedded AmI setting.

For persuasive AmI systems to leverage the effects of praise and mimicry we propose a set of implementations of praise and mimicry. These can be found in table 2. Empirical validation of these implementations in a natural setting poses objectives for future research. We feel that the effects of properly implemented social cues can greatly enhance the compliance to persuasive systems.

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. 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 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.

REFERENCES

1 Aarts, E. H. L., Harwig, M., and Schuurmans, M. 2001. *Ambient Intelligence*. In Denning, P. (Ed.), The invisible future: the seamless integration of technology into everyday life. McGraw Hill.

2 Aarts, E. H. L., Markopoulos, P., and Ruyter, B. E. R. de. 2007. *The persuasiveness of ambient intelligence*. In M. Petkovic, W. Jonker (Eds.), Security, Privacy and Trust in Modern Data Management. Springer.

3 Aarts, E., and de Ruyter, B. 2009. New research perspectives on Ambient Intelligence, *Journal of Ambient Intelligence and Smart Environments*, 1, 5-14.

4 Aron, A. Aron, E. N., and Smollan, D. 1992. Inclusion of other in the Self Scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology*, 63, 596-612.

5 Baaren, R. B., Horgan, T. G., Chartrand, T. L., and Dijkmans, M. 2004. The Forest, the Tress, and the Chameleon: Context Dependence and Mimicry. *Journal of Personality and Social Psychology*, 86, 453-459.

6 Bandura, A. 1997. Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84, 191-215.

7 Baumeister, R. F., and Leary, M. R. 1995. The need to belong: Desire for Interpersonal Attachments as a Fundamental Human Motivation. *Psychological Bulletin*, 117, 497-529.

8 Baumeister, R. F., Hutton, D. G., and Cairns, K. J. 1990. Negative effects of praise on skilled performance. *Basic and Applied Social Psychology*, 11, 131-148.

9 Berscheid, E., and Walster, E. 1987. *Interpersonal attraction*. Addison-Wesley.

10 Breazeal, C. 2000. Sociable Machines: Expressive Social Exchange Between Humans and Robots. *Sc.D. dissertation*, Department of Electrical Engineering and Computer Science, MIT.

11 Burgoon, M., and Klingle, R. S. 1998. *Gender differences in being influential and/or influenced: A challenge to prior explanations*. In: Canary, D. J., and Dindia, K. (Eds.) Sex differences and similarities in communication: critical essays and empirical investigations of sex and gender interaction. Erlbaum.

12 Burnner, E., and Munzel, U. 2002. *Nichtparametrische Datenanalyse*. Springer, Heidelberg.

13 Chartrand, T. L., and Bargh, J. A. 1999. The Chameleon Effect: The Perception-Behavior Link and Social Interaction. *Journal of Personality and Social Psychology*, 76, 893-910.

14 Chartrand, T. L., Maddux, W. W., and Lakin, J. L. 2005. *Beyond the perception–behavior link: The ubiquitous utility and motivational moderators of nonconscious mimicry*. In R. Hassin, J. Uleman, and J. A. Bargh. (Eds.), Unintended thought 2: The new unconscious. New York: Oxford University Press.

15 Cialdini, R. 2001. *Influence, Science and Practice*. Allyn and Bacon.

16 Cialdini, R. 2004. The science of persuasion. *Scientific American Mind*.

17 De Ruyter, B. 2003. *User Centred Design*, In (Eds.): Aarts, Emile, and Stefano Marzano, The New EveryDay: Vision on Ambient Intelligence, 010 Publishers.

18 Fogg, B. J. 2003. Persuasive *Technology: Using Computers to Change what We Think and Do*. Morgan Kaufmann.

19 Garner, R. L. 2005. What is in a name: Persuasion perhaps. *Journal of Consumer Psychology*, 15.

20 Gosling, S. D., Rentfrow, P. J., and Swann, W. B., Jr. 2003. A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, 37, 504–528.

21 Kahler, E., Rogausch, A., Brunner, E., and Himmel, W. 2007. A parametric analysis of ordinal quality-of-life data can lead to erroneous results. *Journal of Clinical Epidemiology*. 61, 475-480.

22 Kaptein, M. C., Markopoulos, P., Aarts, E. H. L., and de Ruyter, B. 2009. Can you be Persuaded? Individual differences in susceptibility to persuasion. *Proceedings of Interact 2009*.

23 Koestner, R., Zuckerman, M., and Koestner, J. 1987. Praise, involvement, and intrinsic motivation. Journal of *Personality and Social Psychology*, 53, 383-390.

24 Kopp, S. and Wachsmuth, I. 2002. Model-based Animation of Coverbal Gesture. *Proceedings of Computer Animation*.

25 Lipnevich, A. A., and Smith, J. K. 2008. *Response to Assesment Feedback: The effects of Grades, Praise, and Source Information*. ETS RR-08-30.

26 Maheshwari, M., Chatterjee, S., and Drew, D. 2008. Evaluating the persuasiveness of "Just in Time" motivational messages for obesity management. *Lecture notes in computer science, 5033/2008, Springer*.

27 Moon, Y., and Nass, C. 1996. How "real" are computer personalities? Psychological responses to personality types in human-computer interaction. *Communication research*, 23, 651-674.

28 Munzel, U., and Bandelow, B. 1998. The use of parametric versus nonparametric tests in the statistical evaluation of ratings scales. *Pharmacopsychiatry*. 6, 222-224.

29 Nass, C., and Moon, Y. 2000. Machines and Mindlessness: Social responses to computers. *Journal of Social Issues*, 56, 81-103.

30 Nass, C., Fogg, B. J., and Moon, Y. 1996. Can computers be teammates? *International journal of human-computer studies*, 45, 669-678.

31 Nass, C., Moon, Y., Morkes, J., Kim, E. Y. and Fogg, B. J. 1997. Computers *are social actors: A review of current research*, in: Friedman, B. (Ed.) Human Values and the Design of Computer technology. CSLI Press.

32 Nass, C., Steuer, J. S. and Tauber, E. 1994. Computers as social actors. *Proceedings of the CHI Conference*, Boston, MA.

33 Pandey, J. and Kakkar, S. 1982. Supervisors' affects: Attraction and positive evaluation as a function of enhancement of others. *Psychological Reports*, 50, 479-546.

34 Räisänen, T., Oinas-Kukkonen, H. and Pahnila, S. 2008. Finding Kairos in Quitting Smoking: Smokers' Perceptions of Warning Pictures. *Lecture notes in computer science*, 5033/2008, Springer.

35 Reeves, B. and Nass, C. 1996. *The media equation: how people treat computers, television, and new media like real people and places*. Cambridge University Press.

36 Sandoval, V.A., and Adams, S.H. 2001. Subtle Skills for Building Rapport, *FBI Law Enforcement Bulletin*, 8, 1-9.

37 Shah, D. A., & Madden, L. V. 2004. Nonparametric Analysis of Ordinal Data in Designed Factorial Experiments. *Pythopathology*. 94, 33-60.

38 Singer, J. M., Poleto, F. Z., & Rosa, P. 2004. Parametric and nonparametric analyses of repeated ordinal categorical data. *Biomedical Journal*. 46, 460-473.

39 Suedfeldt, P., Bochner, S. and Matas, C. 1971. Petitioner's attire and petition signing by peace demonstrators. *Journal of applied psychology*, 1, 278-283.

40 Suzuki, N., Takeuchi, Y., Ishii, K., Okada, M. 2003. Effects of echoic mimicry using hummed sounds on human–computer interaction. *Speech Communication*, 40, 559-573.

41 Thorndike, E. L. 1920. Intelligence and its uses. *Harper's Magazine*, 140, 227-235.

42 Wallace, R. *The Anatomy of ALICE*. Retrieved from: http://www.alicebot.org/anatomy.html.

43 Weiner, B. Heckhausen, H. and Meyer, W. 1972. Causal ascriptions and achievement behavior: A conceptual analysis of effort and reanalysis of locus of control. *Journal of Personality and Social Psychology*, 21, 239-248