

Research Article

Does Humanity Matter? Analyzing the Importance of Social Cues and Perceived Agency of a Computer System for the Emergence of Social Reactions during Human-Computer Interaction

Jana Appel,¹ Astrid von der Pütten,¹ Nicole C. Krämer,¹ and Jonathan Gratch²

¹Department of Social Psychology: Media and Communication, University of Duisburg-Essen, Forsthausweg 2, 47057 Duisburg, Germany

²Institute for Creative Technologies, University of Southern California, 12015 Waterfront Drive Playa Vista, Los Angeles, CA 90094-2536, USA

Correspondence should be addressed to Astrid von der Pütten, astrid.von-der-puetten@uni-due.de

Received 31 March 2012; Revised 6 June 2012; Accepted 12 June 2012

Academic Editor: Kiyoshi Kiyokawa

Copyright © 2012 Jana Appel et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Empirical studies have repeatedly shown that autonomous artificial entities elicit social behavior on the part of the human interlocutor. Various theoretical approaches have tried to explain this phenomenon. The agency assumption states that the social influence of human interaction partners (represented by avatars) will always be higher than the influence of artificial entities (represented by embodied conversational agents). Conversely, the Ethopoeia concept predicts that automatic social reactions are triggered by situations as soon as they include social cues. Both theories have been challenged in a 2×2 between subjects design with two levels of agency (*low: agent, high: avatar*) and two interfaces with different degrees of social cues (*low: textchat, high: virtual human*). The results show that participants in the virtual human condition reported a stronger sense of mutual awareness, imputed more positive characteristics, and allocated more attention to the virtual human than participants in the text chat conditions. Only one result supports the agency assumption; participants who believed to interact with a human reported a stronger feeling of social presence than participants who believed to interact with an artificial entity. It is discussed to what extent these results support the social cue assumption made in the Ethopoeia approach.

1. Introduction

Since the computer found its way into private households, new standards for usability were necessary. Handling the computer had to become more easy and intuitive. Designing the computer to be and act more human-like seemed to be a good solution to improve human-computer interaction. This approach seems to have an impact on the way people interact with computers. Early psychological studies show that human-like interaction styles of computer interfaces had a greater impact on individuals' self-appraisals than machine-like interaction styles [1]. Subsequently, Nass and colleagues [2, 3] proved that people show similar social behavior during human-computer interaction (HCI) by systematically adapting studies from the field of human-human interaction (HHI) to HCI. In their CASA studies (computers are social actors) they could replicate many findings

from HHI, for example, that people show polite behavior towards computers [4, 5], use gender stereotypes for judging computers with female or male voices [6, 7], or reported a feeling of team spirit after being grouped in the same team with a specific computer [8]. Technological progress facilitates the development of computer interfaces with even more social cues like for instance virtual characters which are utilized in a variety of applications, for example, in games, application programs, or in the Internet. Many terms are in use to describe these characters: interface agent, embodied conversational agent, virtual assistant, autonomous agent, or avatar. The major difference with regard to the terms avatar and agent lies within the control of the virtual character. Bailenson and Blascovich [9] define an avatar as "a perceptible digital representation whose behaviors reflect those executed, typically in real time, by a specific human

being,” and an agent as “a perceptible digital representation whose behaviors reflect a computational algorithm designed to accomplish a specific goal or set of goals” [9, page 64].

Empirical results show that both types of representation can elicit social reactions on the side of the user, but opinions differ about the extent to which agents and avatars are able to elicit these social reactions. According to Blascovich et al. [10] the (perceived) agency of a virtual representation is crucial. While an avatar will always elicit social reactions because the user knows that the interaction partner is human whereas an agent will not do so unless it shows sufficiently realistic behavior. Indeed previous research found differing social influence regarding agency (e.g., [11–13]). Besides these two factors (agency and behavioral realism of the interactant) Blascovich et al.’s model of social influence in virtual environments also proposes that the social influence may be enhanced by the self-relevance of the interaction. Moreover, it is important to consider which behavioral-response system is targeted within an experimentation. When a virtual character shots a gun and produces a very loud sound, participants will most likely respond equally fearful, regardless whether they believe their counterpart is controlled by a human or an algorithm. If participants solved a cooperative task together with a virtual character, they might be more influenced by a positive feedback coming from the human-controlled avatar compared to the agent.

However, besides agency the number of social cues provided by the system is widely considered to be of great importance for the question whether people react socially. Recent studies about the impact of anthropomorphic virtual humans indicate that increasing the number of social cues leads to even stronger social reactions on the part of the user [13–20].

Although both approaches are very popular and well investigated, there were little attempts to systematically test both approaches against each other. In a previous study challenging both approaches, von der Pütten et al. [14] compared the importance of agency and behavioral realism of a virtual character and found the extent of displayed social behavior to be of greater importance than the participants’ knowledge about the virtual agent being introduced as an avatar or an agent. Although it is astonishing that the rather subtle variation of social cues in this study (virtual character with and without head nodding) had more impact than the knowledge of whether one interacted with a fellow human or a machine, it cannot be concluded that it is first and foremost social cues instead of agency that result in social reactions on the part of the user, because a multitude of social cues were present in all conditions, since in all conditions a human-like virtual agent was employed. The present study addresses this shortcoming by comparing the impact of agency on the one hand and a more explicit variation of social cues on the other hand. While agency is again varied by means of the instruction the participants are given, the number of social cues is varied by applying either a virtual figure (presented as either agent or avatar) or a text tool (presented as either textchat with a fellow human or as computer interface). In the following we give a detailed overview about previous work on this topic from which we derived our hypotheses.

2. Theoretical Background

As mentioned previously, there is empirical evidence that already conventional computers (using, e.g., command lines or graphical interfaces) elicit social behavior. The same is true for virtual characters in general, regardless whether they are avatars or agents. But opinions differ about which factor, agency, or number of human features/social cues are most crucial and in explaining how and why these social reactions happen.

2.1. The Effect of Social Cues. Nass and colleagues [3, 15] emphasize that people automatically treat artificial entities like real humans. They established the Ethopoeia term to describe these immediate automatic and unconscious reactions due to seemingly social characteristics of a computer. These basic social cues like speech, interactivity, or the filling of social roles [15] trigger social scripts and expectations, and humans cannot be prevented from reacting in social ways when they are confronted with social cues. This is seen as resulting from our evolutionary heritage; the human brain evolved in a time where only humans showed human social behavior. For instance, the usage of language was a definite sign of humanness, which made further cognitive efforts to assign an object into human or nonhuman category needless. The processing of situations and their social cues ensues mindless. Mindlessness [21–23] can be best understood as the failure to draw novel distinctions as no active information processing takes place. People adapt social scripts from human-human interaction (HHI) and use them in human-computer interaction (HCI), even though this behavior seems inappropriate. Indeed, Nass and colleagues have shown in numerous studies that mechanisms known from human-human communication also apply in HCI; people show politeness and cooperation towards computers and apply gender stereotypes and other rules of person perception [4–8]. Several of these computer as social actor studies were replicated with virtual agents (who besides speech, interactivity, and the fulfillment of social roles also often have a human-like appearance and show nonverbal behavior). For instance, Rossen et al. [24] showed that people apply ethnic stereotypes to agents. Caucasian medical students with a prejudice against Afro-Americans were found to show more empathetic verbal and nonverbal behavior towards an agent with a light skin tone than to an agent with a dark skin tone. The results from the politeness study [5] were replicated in an experiment with the virtual agent MAX [25]. Participants evaluated MAX more positively when it itself asked for their judgment compared to an evaluation via paper-and-pencil questionnaire.

Moreover, studies utilizing a large variety of virtual agents demonstrate that virtual humans draw attention just as real humans do [26], person perception was shown to be like that of real humans [27], cooperation and trust is fostered [28, 29], tasks are facilitated or inhibited by “social” presence of a virtual agent [30], and socially desirable behavior is triggered [29]. Several studies that compare plain text interfaces to virtual characters suggest that indeed the human-like appearance of the virtual agents is one factor

that increases social reactions towards the artificial entity. Sproull et al. [29] showed that participants preferred to interact and spend more time with a talking face compared to the interaction with a text interface. They referred to the mindless paradigm from Nass and colleagues [6] as explanation, why an obviously artificial talking face evoked stronger reactions than the text interface. This finding was affirmed by Krämer et al. [31] who found that participants who had the choice between a documentary about the life of Albert Einstein, a James Bond movie, or the daily TV listings were more likely to choose the socially desirable documentary when they were asked by an anthropomorphic interface agent compared to participants asked by a mere text-based or speech-based interface.

2.2. The Effects of Agency. Besides the social cues of a system, the second factor which is extensively discussed and considered to greatly influence the elicitation of social reactions is agency. Several researchers state that the perceived agency matters and impacts the elicitation of social reactions. Also, several studies present results that favor the importance of agency. For instance, Guadagno et al. [13] examined the effects of agency and behavioral realism on persuasion and found some supporting results for the importance of agency. Besides other results, participants in the avatar group experienced more social presence than subjects in the agent group. Hoyt et al. [12] demonstrated classic effects of social inhibition when participants were asked to perform a novel task. Task performance was inhibited when participants performed in front of two avatars, whereas performance was not inhibited in front of two agents or when performing the task alone. Conversely, they did not find effects of social facilitation when participants performed well-trained tasks in front of the avatars compared to performing alone or in the presence of agents. Also, other studies give evidence for the importance of agency but also do not yield consistent results; Aharoni and Fridlund [32] investigated the influence of the factor agency. Participants in their study interacted with a standard computer with prerecorded speech output. Participants believed that they were either interacting with a human interviewer or an artificial intelligent computer. The experimenters reported that participants used more silence fillers and smiled more while interacting with the human interviewer compared to the computer. However, the evaluation of the interviewer as well as the subjective emotional state of the participants was not affected by the factor agency. Surprisingly, also Nass and colleagues present results that show the influence of agency and are therefore not in line with their assumptions. In a study that analyzed the influence of humor on task performance [11], the perceived agency of the interaction partner and the level of humor (humor or no humor) were varied. Participants either believed to interact with a computer system or a human via text chat. The answers either contained humorous parts or no humor at all. The results were partially inconsistent with the assumptions drawn from the CASA paradigm. People in the HCI condition showed social reactions towards the computer, but those in the mocked CMC condition were even stronger than those in the HCI condition. Confronted

with these inconsistent results, Nass and colleagues searched for explanations: “In sum, although many results supported the SRCT [social responses to communication technologies] model, other findings point to a limitation in SRCT theory: HCI and CMC are not identical. In other words, the equating of HCI and human-human interaction (as proposed by [3]) is called into question here. Specifically, there is evidence for a conception of socialness as a continuum rather than a dichotomy” [11, page 423]. However, in a study about the in-group out-group phenomenon based on ethnicity Nass and colleagues were able to demonstrate results that affirmed their hypotheses [33]. Participants thought to either interact with an agent or with an avatar, whose skin color also was varied. There was no evidence for a significant difference due to the perceived agency. In line with the hypothesis, people evaluated the virtual character better, that seemed to have the same ethnicity. With a few exceptions (see [11]), the “human” conditions in these experiments have *not* elicited stronger social responses than the “computer” conditions.” [15, page 99]. Finally, Nowak and Biocca [34] found no effect of agency in a study about the influence of agency and anthropomorphism. Participants believed that they were interacting either with an agent or an avatar. Additionally, the degree of anthropomorphism was varied from no picture (control group), abstract eyes and mouth, (low anthropomorphism) to a realistic picture of a virtual character (high anthropomorphism). Agency showed no effects on the perceived degree of copresence or social presence, but participants reported increased social presence when confronted with a high anthropomorphic picture compared to a low anthropomorphic picture.

As neither the results from Blascovich and colleagues [12, 13] nor those from Nass and colleagues [11, 33] could clarify the role of the agency for the development of social reactions on the part of the user, a recent study [14] evaluated whether participants’ belief in interacting with either an avatar or an agent leads to different social effects. Von der Pütten and colleagues [14] used a design with two levels of agency (agent or avatar) and two levels of behavioral realism (showing feedback behavior versus showing no behavior). It could be found that the belief of interacting with either an avatar or an agent barely resulted in differences with regard to the evaluation of the virtual character or behavioral reactions, whereas higher behavioral realism affected both. It seems that behavioral realism plays an important role in human-agent interaction. Von der Pütten et al. see these results in line with the Ethopoeia concept: “The assumption that the more computers present characteristics that are associated with humans, the more likely they are to elicit social behavior [15, page 97] is confirmed in our experiment” [14, page 1647]. However, as noted above, the study does not allow for conclusions on the effects of social cues in general but only on the effects of the variation of a specific social cue, namely the presence of nonverbal feedback behavior. It has not yet been tested in comparison, what effects will be yielded by an agency manipulation if on the other hand the number of social cues (in terms of text-based interaction versus communication with a virtual character) is varied.

3. Hypotheses

- (H1) The social effects will be higher in the conditions with a presented virtual character as interaction partner (high number of social cues) than in the conditions with a presented text-based interface as interaction partner (low number of social cues). (The effect of social cues).
- (H2) The social effects will be higher in the conditions with an assumed avatar as interaction partner (high agency) than in the conditions with an assumed agent as interaction partner (low agency). (The effect of agency).

3.1. Experimental Design. To test our hypotheses, we used a 2×2 between subjects design with two levels of agency (*low: agent, high: avatar*) and two interfaces with different numbers of social cues (*low: textchat, high: virtual human*). Participants were randomly assigned to the conditions. Gender was distributed equally across conditions. During the interaction, the participants were asked three personal questions of increasing intimacy by the virtual character or by the textchat.

3.1.1. Factor Agency. For the factor agency, we varied the instruction given to the participants by the experimenter. Although, in fact, all of the questions asked by either the virtual character or the text chat were initiated manually by the experimenter in the room next door, half of them were led to believe that they interacted with a real person in another room (who was either mediated through the avatar or the chat system), and half of them were instructed that they interacted with a computer (which was either displayed as an agent or a text chat).

The agency was varied by giving different instructions to the participants. In the conditions with low agency (agent-condition), the participants were told that the interaction partner is a computer which uses an artificial intelligence program to talk to them.

In the conditions with high agency (avatar-condition), the participants were told that their interaction partner would be another participant in a room next door. Depending on the number of social cues, the instructions differed. When communicating with the virtual character, they were told that participant 2, the listener, would see a live video image of them and that they would see the avatar, whose movements copy the movements of Participant 2. If participant 2 moves his/her head or nods, the avatar would do the same at the same time. Because the virtual character did not provide any facial expressions, they were told that the avatar was not able to copy the facial expressions of participant 2. The instructions differed slightly when communicating with the text-based interface. Participants in the chat condition were led to believe that participant 2 posts questions through an instant messenger system. They were also told that participant 2 would see a live video image of them so that they would be able to use speech.

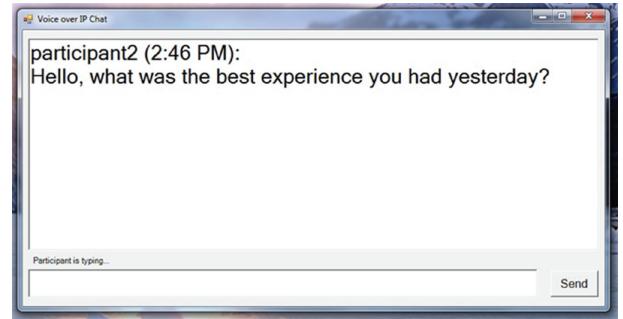


FIGURE 1: Mocked messenger interface for the high agency condition.

3.1.2. Factor Social Cues. The number of social cues of the system was varied by either using a virtual character or a text based interface. The virtual character possesses multiple human cues (such as a female voice, a human appearance, and the use of nonverbal feedback), while the only human cue of the text-based interface is the use of written language. The utilized systems will now be explained in detail.

3.1.3. Text-Based Interface. The text-based interface was applied because of its small number of social cues. In order to guarantee believability it was necessary to create two slightly different interfaces. In the condition with high agency, the participant had to be convinced of the existence of participant 2. Therefore we edited the appearance of the chat interface in order to make it more look like a real instant messenger system (Figure 1).

A text gadget with a send button was created, similar to those known from real text messenger systems. This text gadget was never used by the participants because they were asked to only answer verbally. Its only purpose was to simulate a fully functional messenger system. Second, a time delay was deployed before the appearance of the next question on the screen. This delay simulated the time necessary for the other person to type the next question. Additionally, during that 15–30 sec delay the status message “participant is typing...” was displayed. This should give the participant a feeling of real-time conversation. In the low agency condition a plain text window was used (Figure 2).

3.1.4. Virtual Character. For the condition with the virtual character we used the Rapport Agent, which was developed by Gratch et al. [16] at the Institute for Creative Technologies. The agent displays listening behaviors that correspond to the verbal and nonverbal behavior of a human speaker. The Rapport Agent has been evaluated in several studies [16–18] and has proven to be capable of creating the experience of rapport comparable with face-to-face conditions in certain contexts (e.g., storytelling, interview). To produce listening behaviors, the Rapport Agent first collects and analyzes the features from the speaker’s voice and upper-body movements via microphone and a Videre Design Small Vision System stereo camera, which was placed in front of the participants to capture their movements. Watson, an

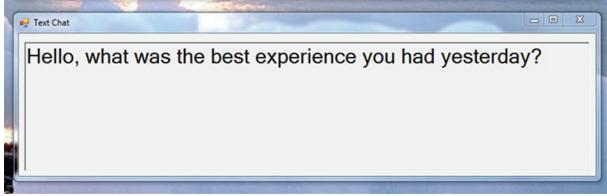


FIGURE 2: Plain text window for the low agency condition.



FIGURE 3: Female rapport agent.

image-based tracking library developed by Morency et al. [35], uses images captured by the stereo camera to track the participant's head position and orientation. Acoustic features are derived from properties of the pitch and intensity of the speech signal using a signal processing package, LAUN, developed by Grath et al. [16]. The Rapport Agent displays behaviors that show that the animated character is "alive" (eye blinking, breathing) and listening behaviors such as posture shifts and head nods are automatically triggered by the system corresponding to participants' verbal and nonverbal behavior. A female virtual character was used for this experiment (Figure 3).

3.1.5. Setting. All participants were asked to tell three short stories about their daily life. The whole conversation was fixed to five sentences. They appeared either as text on the text-based interfaces or were verbally asked by the virtual character. Therefore, five sentences were prerecorded with a female voice. All participants received the same questions. The following five sentences were used.

- (i) Okay, I'm ready.
- (ii) What was the most special experience for you yesterday?
- (iii) Which of your characteristics are you most proud of?
- (iv) What has been the biggest disappointment in your life?
- (v) Thank you. You're done.

3.2. Dependent Variables. As dependent variables, we assessed the participants' emotional state (PANAS) after the interaction, the person perception of the virtual character, the self-reported experience of social presence, and self-reported rapport. Besides these self-report measures, we also measured the following objective variables: the total number of words the participants used during the interaction and the percentage of pause fillers and interrupted words. We also carried out a qualitative analysis of the degree of self-disclosure. In the following, all measurements will be described in detail.

3.2.1. Quantitative Measurements. In the present study, we used the Positive And Negative Affect Scale [36] consisting of 20 items (e.g., strong, guilty, active, ashamed, etc.), which are rated on a 5-point Likert scale from "very slightly or not at all" to "extremely." The factorial analysis for the Positive And Negative Affect Scale resulted in three factors. The first factor, *Positive High-Dominance*, explains 32,53% of the variance (Cronbach's Alpha = .931). The second factor, *Negative High-Dominance*, explains 25,01% of the variance (Cronbach's Alpha = .876). The third factor, *Negative Low-Dominance*, explains 6,69% of the variance (Cronbach's Alpha = .759).

For the person perception (of the agent), we used a semantic differential with 27 bipolar pairs of adjectives (e.g., friendly-unfriendly, tense-relaxed), which are rated on a 7-point scale. The factor analysis for the person perception of the virtual character resulted in three factors. The first factor, *Negative Low-Dominance*, explains 39,14% of the variance (Cronbach's Alpha = .912). The second factor, *Negative High-Dominance*, explains 8,63% of the variance (Cronbach's Alpha = .839). The third factor, *Positive High-Dominance*, explains 6,69% of the variance (Cronbach's Alpha = .770).

Social presence [37] was measured by two scales: the social presence scale [19] with five items (e.g., "I perceive that I am in the presence of another person in the room with me") and the Networked Minds Questionnaire (NMQ; [38–40]). We concentrated on the following five aspects of the NMQ: empathy (with 4 items), mutual awareness (with 2 items), attention allocation (with 4 items), mutual understanding (with 3 items), and behavioral interdependence (with 4 items). All items from both scales were rated on a 7-point Likert scale.

To measure perceived rapport, we used a scale that had been developed for previous studies with the Rapport Agent. This scale contains ten items from the rapport construct by Tickle-Degnen [41], which were already in use in an experiment on the effects of nonverbal signal delay in telepsychiatry (see [42]). 19 ad hoc items were added, which proved to measure rapport in several studies [16–18]. The resulting 29 items were measured on an 8-point Likert scale. Because of the text chat conditions, another version of the Rapport Scale had to be prepared. Five items target the perceived embodiment of the interaction partner (e.g., "I watched the listener as I told the story"). These items had to be excluded in the text-based conditions because of the lack of embodiment. Therefore, the chat version of the Rapport Scale includes 24 items. Both versions of the Rapport Scale were used for a combined factor analysis.

TABLE 1: Main effect for agency on the social presence scale.

	Human-human interaction		Human-computer interaction		<i>F</i>	η^2	<i>P</i>
	μ	SD	μ	SD			
Social presence scale	4.182	1.392	3.236	1.396	10.870	.112	.001

The factor analysis for the self-reported rapport revealed four factors. The first factor, *Feelings and Self-Efficiency*, explains 37,3% of the variance (Cronbach's alpha = .926), the second factor, *Rapport and Connection*, 11,73% (Cronbach's alpha = .954), the third, *Attention Allocation*, 6,78% (Cronbach's alpha = .648%), and the fourth factor, *Embodiment*, explains 5,72% of the variance (Cronbach's alpha = .538).

3.2.2. Verbal Behavior. In addition, we analyzed the participants' verbal behavior. We counted the total amount of words, the amount of pause-fillers ("erm," "hm"), and the amount of broken words (e.g., "I was in the bib... library"). From the latter two, we calculated the percentage of speech disfluencies in relation to the total amount of words.

3.2.3. Qualitative Measurements. We conducted a qualitative analysis of the participants' answers to the questions asked by the virtual character. The first question ("What was the most special experience you had yesterday?") was excluded from the analysis because of too much variance due to the weekday on which they participated. When participants took part in the experiment on a Monday, they had more possibilities to report about their activities (i.e., on Sunday) than people who took part on a Thursday.

For the second question ("Which of your characteristics are you most proud of?"), we counted the number of characteristics the participants revealed. For the third question ("What has been the biggest disappointment in your life?"), we used a categorical coding scheme [43] with three categories: (1) no answer: the subject gives no answer or uses excuses to avoid an answer (e.g., "-um-... I don't know. I think I don't think I've had anything horrible happen to me yet. I'm lucky"); (2) low-intimacy answer: the disappointment (or unfulfilled wish) has not sustainably affected the private or business life of the subject (e.g., "I'd like to be wealthy so I think that's my biggest disappointment" or "-um-Not finishing tasks that I start or not following through with things I want to follow through with"); (3) high-intimacy answers: the disappointment (or unfulfilled wish) has sustainably affected the private or business life of the subject (e.g., "-The biggest disappointment in my life. -um- I would say probably the fact that -um- I never remarried -um- after being divorced for many years and -um- I'm starting now to feel a little disappointed about that fact that I didn't find another mate for myself"). The coding was processed by two coders. The interrater reliability showed substantial agreement (Cohen's Kappa = .810).

4. Procedure

Upon arrival, the participants were asked to read and sign informed consent forms. After completing a web-based

questionnaire [44] about their background including demographic data and the questionnaires of the explanatory variables, participants received a short introduction about the equipment and were given the instructions regarding their interaction partner and the task of the experiment (see above). Then, participants took a seat in front of a 30" screen, which displayed the interaction partner (virtual character or text chat). They were equipped with a headset with microphone. In order to assess the participants' verbal and nonverbal behavior, the whole session was videotaped. The camera was directed towards the participants and situated directly under the screen with the Rapport Agent in combination with the stereovision camera. Participants were instructed to wait until the systems starts, indicating readiness by a ping sound. They were asked three questions by the Rapport Agent with increasing intimacy. After the interaction, the participants completed the second web-based questionnaire. They were fully debriefed, given \$20, and thanked for their participation.

5. Participants

Ninety persons (49 females and 41 males) participated in the study. The mean age was 36.26 ($SD = 12.26$), ranging from 19 to 62 years. Participants were recruited via <http://www.craigslist.com/> from the general Los Angeles area and were compensated \$20 for one hour of their participation.

6. Results

We calculated MANOVAS with the two independent variables agency and number of social cues and the dependent variables: three PANAS factors, three person perception factors, four rapport factors, the social presence scale, the constructs empathy, attention allocation, mutual awareness, mutual understanding and behavioral interdependence from the NMQ, the total amount of words, the percentage of speech disfluencies, and the number of revealed characteristics.

We identified only one main effect for agency. A significant main effect on the dependent variable Social Presence Scale emerged by varying the perceived agency (Table 1). The feeling of social presence was more intense after communicating with the "other subject" via avatar or Textchat ($F(1;90) = 10.870$; $P = .001$; partial $\eta^2 = .112$) than after communicating with the computer.

On the other hand, several effects of the number of social cues emerged. Subjects in the virtual character conditions (high number of social cues) had a stronger feeling of social presence (factor Mutual Awareness, see Table 2) ($F(1;90) =$

TABLE 2: Main effect for social cues on the factor mutual awareness.

	Virtual character		Chat		F	η^2	P
	μ	SD	μ	SD			
Mutual awareness	2.728	1.223	3.693	1.095	15.207	.150	.000

TABLE 3: Main effect for social cues on the factor negative high-dominance.

	Virtual character		Chat		F	η^2	P
	μ	SD	μ	SD			
Negative high-dominance	-.217	.970	.226	.991	4.516	.050	.036

TABLE 4: Main effect for social cues on the factor attention allocation.

	Virtual character		Chat		F	η^2	P
	μ	SD	μ	SD			
Attention allocation	-.248	1.071	.302	.882	6.943	.075	.010

15.207; $P = .000$; partial $\eta^2 = .150$) than subjects in the text conditions (low number of social cues).

Also, a main effect for social cues with regard to the factor Negative High-Dominance of the Person Perception Scale emerged. Subjects in the virtual character conditions (high number of social cues) described the interlocutor less negative ($F(1;90) = 15.207$; $P = .000$; partial $\eta^2 = .150$, see Table 3) than subjects in the text-conditions (low number of social cues). A further significant main effect of social cues emerged on the dependent variable attention allocation (Table 4) of the Rapport Scale. Subjects in the virtual character conditions (high number of social cues) reported to pay more attention to the interaction partner ($F(1;90) = 6.943$; $P = .010$; partial $\eta^2 = .075$) than subjects in the text-conditions (low number of social cues). Low means are associated with paying more attention to the interaction partner. There were no effects with regard to the total amount of words, self-disclosure of information (number of characteristics), or percentage of speech disfluencies in relation to the total amount of words. Also, Chi-square tests did not reveal any effect with regard to the categories of question three. We also found no interaction effects of the factors agency and number of social cues.

7. Discussion

The main goal of this research was to empirically test which factors account for the emergence of social behavior in human-computer interaction. Two competing explanations were tested against each other: the agency and the number of social cues approach. To test the factor agency, we varied the factor agency by instructing the participants that they would either communicate with an artificial intelligence program or another real participant in the room next door. The other factor, number of social cues, was varied by using either a text-based interface (low number of social cues) or an animated character (high number of social cues). We used a wide range of dependent variables including quantitative and qualitative behavioral data, scales previously used within

the paradigm, and standardized psychological measures used for face-to-face interactions.

With regard to the agency assumption which basically states that there will be a difference between the assumed interaction with a computer or a human being—in the direction of real humans as interlocutor evoking stronger social reactions than computers—the results of this study do not provide a strong support for agency. Only one main effect for agency was found for the feeling of social presence which was more intense after communicating with the “other subject” via avatar or textchat than after communicating with the computer. However, this effect could also be a result of the wording of the Social Presence Scale, which explicitly asks for the presence of a real living person (e.g., “the person appears to be sentient (conscious and alive) to me”) and therefore provoked stronger reactions when participants were told that they interacted with a real person. Further studies investigating the agency assumption should use a revised version asking, for example, the interaction partner. Moreover, other measure should be considered like participants’ nonverbal behavior or psychophysiology. Besides this main effect, none of the other 19 dependent factors showed effects for agency. According to the social model of influence in virtual environments, it is especially surprising that we had no effects concerning the qualitative analysis of the verbal answers, since the interview situation was at least medium self-relevant, because participants were asked very personal and intimate questions. Moreover, answers to intimate questions can be regarded as a measure within a rather high-level behavioral-response system which should more easily lead to differences between the two agency conditions according to the model. Therefore, the hypothesis that *social effects will be higher in the conditions with an assumed avatar as interaction partner (high agency) than in the conditions with an assumed agent as interaction partner (low agency)* has to be rejected.

In contrast to the agency factor, we found several results supporting the assumption that the number of social cues displayed influences the strength of social reactions. For three

dependent variables main effects emerged which indicate that a human-like virtual character (high number of social cues) triggers stronger social reactions than a plain text-based interface (low number of social cues). Subjects in the virtual character conditions described the interlocutor less negative and had a stronger feeling of mutual awareness and reciprocal attention allocation than subjects in the text conditions. However, due to the relative small number of main effects for the independent variable social cues (three of in total twenty dependent variables) there is only partial support for hypothesis 1 which claimed stronger effects in those conditions with a higher number of social cues. Based on the results of von der Pütten et al. [14] and the results of the current study the assumption that “the more computers present characteristics that are associated with humans, the more likely they are to elicit social behavior” [15, page 97] seems to be reasonable.

8. Limitations

Although the amount of social cues can be regarded as a continuous variable, we used only two levels of social cues and treated it as a dichotomy variable for the purpose of our study. Future research should address this shortcoming by comparing more levels of social cues. Moreover, the two conditions differed not only in the amount of social cues but also in the modality, and this could have caused the differences in the attention participants paid to the system. Participants in the chat group were presented the questions as written text which stayed on screen until they finished giving their answer, while the participants in the virtual character condition were presented the questions as prerecorded sentences delivered by the virtual character. Thus it is possible that in the latter condition participants had to be more attentive to the character, because otherwise they would miss the question. In addition, unlike the chat window the virtual character presents continuous feedback by showing nonverbal responses to the participant’s verbal input. This also might bias the attention allocation in favor for the condition with the virtual character. This problem can be solved when providing continuous feedback for the chat condition as well by, for instance, showing the well-known rotating sandglass with the text “verbal input is being processed” or a red light indicating that recording and transmission of the signal is on air. As mentioned above, the Social Presence Scale probably was not adequate to use, because participants were biased by the wording of the scale. Future studies should also focus on the analysis of nonverbal behavior or psychophysiological measures which have been proven to be indicators of social presence (e.g., [19, 20]).

9. Conclusion

In sum, the results of the present study do not provide strong support for the assumption of the model of social influence in virtual environments [10] which claims general dissimilar social impact of agents and avatars due to their agency. The cognitive-mediated knowledge about the interlocutor’s

nature did not have a strong impact on the participants’ experience and behavior. The Ethopoeia concept by Nass and colleagues seems to be a better approach to explain the emergence of social reactions towards computers. Although not for all dependent variables, it could be shown that a high number of social cues provokes stronger social reactions on the part of the user. It has to be noted that this design did only consist of two levels of social cues, namely, high (humanoid appearance) and low (text) number of social cues. Additional studies should concentrate on systematically and gradually varying the level of social cues to analyze the relations in greater detail. This will provide a deeper insight into whether social reactions gradually become stronger along with a gradual increase of social cues. For example, it is still unclear whether single features, like the human appearance, the human voice, or the combination of all, triggered the social reactions. More research is needed to gain answers to these questions.

Acknowledgment

This research was supported by the National Scientific Foundation under grant # IIS-0916858 and the U.S. Army Research, Development, and Engineering Command. The content does not necessarily reflect the position or the policy of the U.S. Government, and no official endorsement should be inferred. This work was partially supported by a scholarship of the German Academic Exchange Service.

References

- [1] L. Quintanar, C. Crowell, J. Pryor, and J. Adamopoulos, “Human computer interaction: a preliminary social psychological analysis,” *Behaviour Research Methods & Instrumentation*, vol. 14, no. 2, pp. 210–220, 1982.
- [2] C. Nass, E. Tauber, and H. Reeder, “Anthropomorphism, agency and ethopoeia: computers as social actors,” in *Proceeding of the International Conference Companion on Human Factors in Computing Systems (CHI ’93)*, pp. 111–112, 1993.
- [3] B. Reeves and C. I. Nass, *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places*, Cambridge University Press, New York, NY, USA, 1996.
- [4] C. Nass, J. Steuer, and E. R. Tauber, “Computer are social actors,” in *Proceedings of the Conference on Human Factors in Computing Systems (CHI ’94)*, pp. 72–78, ACM Press, New York, NY, USA, April 1994.
- [5] C. Nass, Y. Moon, and P. Carney, “Are people polite to computers? Responses to computer-based interviewing systems,” *Journal of Applied Social Psychology*, vol. 29, no. 5, pp. 1093–1110, 1999.
- [6] C. Nass, Y. Moon, and N. Green, “Are machines gender neutral? Gender-stereotypic responses to computers with voices,” *Journal of Applied Social Psychology*, vol. 27, no. 10, pp. 864–876, 1997.
- [7] C. Nass, J. Steuer, L. Henriksen, and D. C. Dryer, “Machines, social attributions, and ethopoeia: performance assessments of computers subsequent to “self-” or “other-” evaluations,” *International Journal of Human-Computer Studies*, vol. 40, no. 3, pp. 543–559, 1994.

- [8] C. Nass, B. J. Fogg, and Y. Moon, "Can computers be teammates?" *International Journal of Human-Computer Studies*, vol. 45, no. 6, pp. 669–678, 1996.
- [9] J. N. Bailenson and J. Blascovich, "Avatars," in *Encyclopedia of Human-Computer Interaction*, pp. 64–68, Berkshire Publishing Group, 2004.
- [10] J. Blascovich, J. Loomis, A. C. Beall, K. R. Swinth, C. L. Hoyt, and J. N. Bailenson, "Immersive virtual environment technology as a methodological tool for social psychology," *Psychological Inquiry*, vol. 13, no. 2, pp. 103–124, 2002.
- [11] J. Morkes, H. K. Kernal, and C. Nass, "Effects of humor in task-oriented human-computer interaction and computer-mediated communication: a direct test of SRCT theory," *Human-Computer Interaction*, vol. 14, no. 4, pp. 395–435, 1999.
- [12] C. L. Hoyt, J. Blascovich, and K. R. Swinth, "Social inhibition in immersive virtual environments," *Presence*, vol. 12, no. 2, pp. 183–195, 2003.
- [13] R. E. Guadagno, J. Blascovich, J. N. Bailenson, and C. McCall, "Virtual humans and persuasion: the effects of agency and behavioral realism," *Media Psychology*, vol. 10, no. 1, pp. 1–22, 2007.
- [14] A. von der Pütten, N. C. Krämer, J. Gratch, and S.-H. Kang, "It doesn't matter what you are! explaining social effects of agents and avatars," *Computers in Human Behaviour*, vol. 26, no. 6, pp. 1641–1650, 2010.
- [15] C. Nass and Y. Moon, "Machines and mindlessness: social responses to computers," *Journal of Social Issues*, vol. 56, no. 1, pp. 81–103, 2000.
- [16] J. Gratch, A. Okhmatovskia, F. Lamothe, S. Marsella, M. Morales, R. J. van der Werf et al., "Virtual rapport," in *Proceedings of the 6th International Conference on Intelligent Virtual Agents*, Springer, Marina Del Rey, Calif, USA, 2006.
- [17] J. Gratch, N. Wang, J. Gerten, E. Fast, and R. Duffy, "Creating rapport with virtual agents," in *Proceedings of the 7th International Conference on Intelligent Virtual Agents (IVA '07)*, C. Pelachaud, J.-C. Martin, E. André, G. Chollet, K. Karpouzis, and D. Pelé, Eds., LNAI 4722, pp. 125–138, Springer, Paris, France, 2007.
- [18] J. Gratch, N. Wang, A. Okhmatovskia, F. Lamothe, M. Morales, and L.-P. Morency, "Can virtual humans be more engaging than real ones?" in *Proceedings of the 12th International Conference on Human-Computer Interaction: Intelligent Multimodal Interaction Environments. Part III (HCII '07)*, J. Jacko, Ed., LNCS 4552, pp. 286–297, Springer, Beijing, China, 2007.
- [19] J. N. Bailenson, J. Blascovich, A. C. Beall, and J. M. Loomis, "Equilibrium theory revisited: mutual gaze and personal space in virtual environments," *Presence*, vol. 10, no. 6, pp. 583–598, 2001.
- [20] J. N. Bailenson, J. Blascovich, A. C. Beall, and J. M. Loomis, "Interpersonal distance in immersive virtual environments," *Personality and Social Psychology Bulletin*, vol. 29, no. 7, pp. 819–833, 2003.
- [21] E. J. Langer, "Matters of mind: mindfulness/mindlessness in perspective," *Consciousness and Cognition*, vol. 1, no. 3, pp. 289–305, 1992.
- [22] E. J. Langer, *Mindfulness*, Addison-Wesley, Reading, Mass, USA, 1989.
- [23] E. J. Langer and M. Moldoveanu, "The construct of mindfulness," *Journal of Social Issues*, vol. 56, no. 1, pp. 1–9, 2000.
- [24] B. Rossen, K. Johnson, A. Deladisma, S. Lind, and B. Lok, "Virtual humans elicit skin-tone bias consistent with real-world skin-tone biases," in *Proceedings of the 8th International Conference on Intelligent Virtual Agents (IVA '08)*, H. Prendinger, J. Lester, and M. Ishizuka, Eds., LNAI 5208, pp. 237–244, Springer, Tokyo, Japan, 2008.
- [25] L. Hoffmann, N. C. Krämer, A. Lam-chi, and S. Kopp, "Media equation revisited: do users show polite reactions towards an embodied agent?" in *Proceeding of 9th International Conference on Intelligent Virtual Agents (IVA '09)*, Z. Ruttkay et al., Ed., LNAI 5773, pp. 159–165, Springer, Amsterdam, Netherlands, 2009.
- [26] D. M. Dehn and S. van Mulken, "The impact of animated interface agents: a review of empirical research," *International Journal of Human-Computer Studies*, vol. 52, no. 1, pp. 1–22, 2000.
- [27] G. Bente, N. C. Krämer, A. Petersen, and J. P. de Ruiter, "Computer animated movement and person perception: methodological advances in nonverbal behavior research," *Journal of Nonverbal Behavior*, vol. 25, no. 3, pp. 151–166, 2001.
- [28] S. Parise, S. B. Kiesler, L. Sproull, and K. Waters, "Cooperating with life-like interface agents," *Computers in Human Behavior*, vol. 15, no. 2, pp. 123–142, 1999.
- [29] L. Sproull, M. Subramani, S. Kiesler, J. H. Walker, and K. Waters, "When the interface is a face," *Human-Computer Interaction*, vol. 11, no. 2, pp. 97–124, 1996.
- [30] R. Rickenberg and B. Reeves, "The effects of animated characters on anxiety, task performance, and evaluations of user interfaces," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '00)*, pp. 49–56, April 2000.
- [31] N. C. Krämer, G. Bente, and J. Piesk, "The ghost in the machine. The influence of Embodied Conversational Agents on user expectations and user behaviour in a TV/VCR application," in *Proceedings of the International Workshop on Mobile Computing, Assistance, Mobility, Applications (IMC '03)*, G. Bieber and T. Kirste, Eds., pp. 121–128, Rostock, Germany.
- [32] E. Aharoni and A. J. Fridlund, "Social reactions toward people vs. computers: how mere labels shape interactions," *Computers in Human Behavior*, vol. 23, no. 5, pp. 2175–2189, 2007.
- [33] C. Nass, K. Isbister, and E.-J. Lee, "Truth is beauty: researching embodied conversational agents," in *Embodied Conversational Agents*, J. Cassell, J. Sullivan, S. Prevost, and E. Churchill, Eds., pp. 374–401, The MIT Press, Cambridge, Mass, USA, 2000.
- [34] K. Nowak and F. Biocca, "The influence of agency and the virtual body on presence, social presence and copresence in a computer mediated interaction," in *Proceedings of the 3rd International Presence Workshop*, Philadelphia, Pa, USA, 2001.
- [35] L. P. Morency, C. Sidner, and T. Darrell, "Towards context-based visual feedback Recognition for Embodied Agents," in *Proceedings of the Symposium on Conversational Informatics for Supporting Social Intelligence and Interaction*, pp. 69–72, AISB, Hatfield, UK, 2005.
- [36] D. Watson, A. Tellegen, and L. A. Clark, "Development and validation of brief measures of positive and negative affect: the PANAS scales," *Journal of Personality and Social Psychology*, vol. 54, no. 6, pp. 1063–1070, 1988.
- [37] J. Short, E. Williams, and B. Christie, *The Social Psychology of Telecommunications*, John Wiley & Sons, London, UK, 1976.
- [38] F. Biocca and C. Harms, "Defining and measuring social presence: contribution to the net-worked minds theory and measure," in *Proceedings of the 5th International Workshop on Presence*, F. R. Gouveia and F. Biocca, Eds., pp. 7–36, 2002.
- [39] F. Biocca, C. Harms, and J. K. Burgoon, "Toward a more robust theory and measure of social presence: review and suggested criteria," *Presence*, vol. 12, no. 5, pp. 456–480, 2003.

- [40] F. Biocca, C. Harms, and J. Gregg, "The networked minds measure of social presence: pilot test of the factor structure and concurrent validity," in *Proceedings of the 4th Annual International Workshop on Presence*, Philadelphia, Pa, USA, May 2001.
- [41] L. Tickle-Degnen and R. Rosenthal, "The nature of rapport and its nonverbal correlates," *Psychological Inquiry*, vol. 1, no. 4, pp. 285–293, 1990.
- [42] T. R. Manning, E. T. Goetz, and R. L. Street, "Signal delay effects on rapport in telepsychiatry," *Cyberpsychology and Behavior*, vol. 3, no. 2, pp. 119–127, 2000.
- [43] P. Mayring, *Einführung in die Qualitative Sozialforschung. Eine Anleitung zu qualitativem Denken*, Psychologie Verlags Union, Weinheim, Germany, 3rd edition, 1996.
- [44] D. Leiner, "ofb.msd-media.de," 2009, <http://ofb.msd-media.de/>.

